UNIVERSITÄT BONN

Hierarchical hub-filament structures of molecular gas: star formation and cluster evolution

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jianwen Zhou

aus

Chuzhou, China

Bonn, 08.05.2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer:	Prof. Dr. Pavel Kroupa
Gutachter:	Prof. Dr. Frank Bigiel
Tag der Promotion: 29.07.2025	
Erscheinungsjahr: 2025	

Summary

The first part of this thesis presents a comprehensive multi-scale investigation of hub–filament structures, examining their morphology, kinematics, and evolution from dense core (≈1000 AU) to clump (≈1 pc), molecular cloud (≈10–100 pc), and galaxy-cloud (≈1000 pc) scales. These observations reveal hierarchical hub–filament structures with self-similar properties spanning sub-parsec to kiloparsec scales, highlighting their crucial role in the star formation process. Within this hierarchy, a dense core serves as the hub within a clump, a clump functions as the hub within a molecular cloud, and a molecular cloud acts as the hub within a galaxy. Hub-filament structures form through the gravitational contraction of gas structures. Velocity gradient analyses suggest that gas inflows along filaments are driven by gravity. Molecular gas is organized into network structures formed by the gravitational coupling of multi-scale hub–filament structures, where local gravitational centers (hubs) serve as the primary sites of star formation.

The second part explores how tidal effects from surrounding material regulate gravitational collapse and influence the internal kinematics of gas structures across multiple scales. Tidal interactions are shown to significantly affect the velocity dispersion and density contrast of gas structures, emphasizing the critical role of the surrounding environment in their evolution.

The third part investigates the kinematics and dynamics of gas structures under stellar feedback across multiple scales. Significant correlations between velocity dispersion and column density suggest that velocity dispersion may partly originate from gravitational collapse, a conclusion further supported by the observed velocity gradients. During the formation and evolution of star clusters, although feedback from cloud to core scales will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, as well as changes in gravitational centers, but gravitational collapse remains ongoing.

The fourth part presents a comprehensive investigation of star cluster formation and evolution across broad time scales, from clumps to embedded clusters to open clusters. By combining observational analysis, numerical simulations, and theoretical modeling, these works reveal several key insights. Embedded clusters in ATLASGAL clumps exhibit burst-like star formation histories, with accretion luminosity playing a significant role in low-mass clumps. Star formation efficiency shows a strong anti-correlation with clump mass, averaging around 30%. Gas expulsion timescales significantly affect cluster dynamics, with all modes—fast, moderate, and slow—being observationally supported. The post-gas expulsion coalescence of embedded clusters within a common parental molecular cloud may be a dominant pathway for open cluster formation. Furthermore, the mechanisms of cluster formation and the initial gas expulsion processes are critical in the formation of black holes within open clusters.

The fifth part explores the characteristics of mass functions within star-forming regions. The relationship between core mass functions (CMFs) in high-mass star-forming regions and the initial mass function (IMF) is investigated. In clump complexes, the parameters of the clump mass function strongly correlate with cloud mass, and the presence of widespread substructures and mass segregation suggests gravitationally driven evolution. On galaxy–cloud scales, the integrated cloud-wide IMF (ICIMF) model is developed to connect the properties of molecular clouds with their internal stellar populations in

nearby galaxies.

In a hub-filament system, the hub is the primary site of star formation. A clump evolves into an embedded cluster, defining an IMF. A cloud consists of multiple clumps, leading to an ICIMF. Likewise, a galaxy contains numerous clouds, giving rise to an integrated galaxy-wide IMF (IGIMF). These hierarchical hub-filament structures (star-forming regions) correspond to hierarchically integrated IMFs. Extending this self-similar physical framework allows us to establish a connection between star formation in the Milky Way and nearby galaxies, linking star-forming regions with stellar populations across various scales.

太虚无形，气之本体；于太虚之中具有而未成乎形，气自足也，聚散变化，而其本体不为之损益。其聚其散，变化之客形尔。

——— 张载《正蒙·太和篇》

The Great Void (太虚) is formless; it is the fundamental substance of qi (气). Within the Great Void, qi exists but has not yet taken on a concrete form. Qi is self-sufficient. Whether it gathers or disperses, undergoing various transformations, its essence neither increases nor decreases. The gathering and dispersing of qi are merely temporary forms within the process of change.

——— Zhang Zai's Correcting the Ignorant: The Great Harmony Chapter

Man is but a reed, the most feeble thing in nature; but he is a thinking reed ..... A thinking reed. — It is not from space that I must seek my dignity, but from the government of my thought. I shall have no more if I possess worlds. By space the universe encompasses and swallows me up like an atom; by thought I comprehend the world.

——— Blaise Pascal's Pensées

Prelude

When I first became an undergraduate, I was torn between choosing philosophy and physics. In the end, I chose physics because I felt it had the power to explain things clearly. By the time I graduated, my interest in the origin of the universe and the nature of space-time led me to study gravity and cosmology. During the first year of my master's, I took many courses in astrophysics. However, these courses made me question theoretical research. The cosmology presented in textbooks felt somewhat unbelievable—rather than describing the actual universe, it seemed more like an abstract mathematical construct. This realization led me to shift my focus to observational research, aiming to reconstruct my understanding of the cosmos through data. As a result, I am trying to build a framework that integrates physics and objects across various timescales and spatial scales. The phenomena in the universe should be understood as part of a system rather than being considered in isolation. Within this framework, we can connect many aspects of the universe through causality, as well as temporal and spatial sequences, which can also be traced through multi-wavelength observations. Modern astronomy is divided into too many specialized branches, often isolated from one another. However, the objects we study are inherently interconnected within complex systems. Constructing a complete physical picture requires broad and comprehensive knowledge. I hope to expand the temporal and spatial scales of my research until it encompasses the entire universe. "To grasp the universe with thought"—this is a passion I have always carried. I want to understand the various phenomena of the universe and share that knowledge with others. Beyond the small confines of human society on Earth, there is an immensely vast world. Reflecting on life and existence within this grand cosmic backdrop allows us to consider our place in the universe and how we should find our footing within it.

Contents

1 Introduction **1**
1.1 Massive star and star cluster formation 1
1.2 Hub-filament structures 1
1.3 Tidal effects 2
1.4 Feedback 3
1.5 Formation and evolution of star clusters 5
1.6 Stellar and core mass functions 6
1.7 Objectives of the thesis 8
1.8 Structure of the thesis 10

2 Hierarchical hub-filament structures **11**
2.1 Data and observations 11
2.1.1 ALMA observations 11
2.1.2 APEX/LAsMA data 11
2.1.3 Other data 11
2.2 On clump scale 12
2.3 On cloud scale 15
2.3.1 Velocity gradients 15
2.3.2 Funnel structure 18
2.3.3 Large-scale inflow driven by gravity 18
2.4 From cloud to core scale 19
2.4.1 Connection of gas inflows at different scales 19
2.4.2 Infall+rotation motions 22
2.4.3 Gravitational focusing 22
2.5 On galaxy scale 23
2.5.1 Velocity component 23
2.5.2 Scaling relations 25
2.5.3 Velocity gradient 25
2.5.4 Deviation of the free-fall model 29
2.6 Decomposition of molecular clouds 30
2.6.1 Classification 30
2.6.2 Velocity gradient 32
2.6.3 ATLASGAL clumps and radio sources 34
2.6.4 Evolutionary sequence 34
2.6.5 Density contrast 35
2.6.6 Structure of molecular clouds 35
2.7 Decomposition of molecular gas in a galaxy 36
2.7.1 Classification 36

2.7.2 Velocity gradient . . . 39
2.7.3 Star formation . . . 39
2.7.4 Formation of hub-filament structures . . . 41
2.7.5 Evolutionary sequence of molecular clouds . . . 42
2.8 Gravitational network analysis . . . 43
2.8.1 Gravitational network . . . 44
2.8.2 Accretion or merger . . . 44
2.9 Summary . . . 46
2.9.1 On clump scale . . . 46
2.9.2 On cloud scale . . . 47
2.9.3 From cloud to core scale . . . 48
2.9.4 Decomposition of molecular clouds . . . 48
2.9.5 On galaxy scale . . . 49
2.9.6 Decomposition of molecular gas in a galaxy . . . 50
2.9.7 Gravitational network analysis . . . 50

3 Tidal effects . . . 52
3.1 Data . . . 52
3.2 On clump scale . . . 52
3.2.1 Tidal field . . . 53
3.2.2 Pixel-by-pixel computation . . . 53
3.2.3 Tide-regulated gravitational collapse . . . 55
3.3 On cloud scale . . . 57
3.3.1 Structural identification . . . 57
3.3.2 Tidal strength . . . 59
3.3.3 Velocity dispersion . . . 59
3.4 On galaxy-cloud scale . . . 60
3.4.1 Density contrast distribution . . . 60
3.4.2 Shear and tidal fields of the galaxy . . . 61
3.4.3 Tidal effects of neighboring material . . . 63
3.4.4 Baryonic matter . . . 63
3.4.5 Comparison of different tidal calculations . . . 64
3.4.6 Tidal effects . . . 65
3.5 Summary . . . 65
3.5.1 On clump scale . . . 65
3.5.2 On cloud scale . . . 66
3.5.3 On galaxy-cloud scale . . . 66

4 Feedback . . . 68
4.1 On cloud scale . . . 68
4.1.1 Virial parameter . . . 69
4.1.2 Scaling relation . . . 70
4.1.3 Feedback from HII regions . . . 71
4.1.4 The Heyer relation in feedback . . . 72
4.1.5 Cloud disruption and collapse under feedback . . . 73
4.2 On clump scale . . . 74
4.2.1 Evolutionary stages . . . 74

4.2.2 Statistics of physical quantities . . . 75
4.2.3 Scaling relations . . . 76
4.2.4 Velocity gradient . . . 77
4.2.5 Relative importance of feedback processes . . . 77
4.2.6 The role of feedback . . . 78
4.3 Summary . . . 79
4.3.1 On cloud scale . . . 79
4.3.2 On clump scale . . . 80

5 Formation and evolution of star clusters . . . 82
5.1 Data and sample . . . 82
5.1.1 Embedded clusters . . . 82
5.1.2 Clumps . . . 82
5.1.3 Open clusters . . . 83
5.2 Embedded clusters in ATLASGAL clumps . . . 83
5.2.1 Synthetic embedded clusters . . . 83
5.2.2 Accretion luminosity . . . 84
5.2.3 Age distribution . . . 85
5.2.4 Critical clump mass . . . 85
5.3 Star formation efficiency of ATLASGAL clumps . . . 85
5.3.1 Accretion . . . 86
5.3.2 Star formation efficiency . . . 88
5.4 Isolated evolution from clumps to embedded clusters . . . 90
5.4.1 Cluster mass . . . 90
5.4.2 Self-similar mass-size relations . . . 91
5.4.3 Isolated evolution . . . 92
5.5 N-body simulations . . . 93
5.5.1 Individual cluster simulation . . . 94
5.5.2 Slow and moderate gas expulsions . . . 96
5.5.3 Star formation efficiency . . . 96
5.5.4 Procedure . . . 96
5.5.5 Simulation results analysis . . . 96
5.5.6 Cluster coalescence simulation . . . 97
5.5.7 Simulation results analysis . . . 101
5.6 Expanding very young star clusters . . . 102
5.6.1 Evolution of cluster radii in N-body simulations . . . 102
5.6.2 Mass–radius relation . . . 103
5.6.3 Expansion of embedded clusters in HII-clumps . . . 105
5.7 Origin of open clusters . . . 105
5.7.1 Mass and radius . . . 106
5.7.2 Separations of open clusters . . . 106
5.7.3 Mergers at different stages . . . 107
5.7.4 Post-gas expulsion coalescence . . . 108
5.8 Evolution of the $m_{\rm max} - M_{\rm ecl}$ relation . . . 108
5.8.1 N-body simulations . . . 108
5.8.2 Evolutionary patterns of the $m_{\rm max} - M_{\rm ecl}$ relation . . . 109
5.8.3 The $m_{\rm max} - M_{\rm cluster}$ relation in observation . . . 110

5.8.4 Comparison of observation and simulation . . . 110
5.8.5 Individual cluster simulation vs. cluster coalescence simulation . . . 117
5.9 Constraining the gas expulsion modes of embedded clusters . . . 117
5.9.1 Kolmogorov-Smirnov test . . . 117
5.9.2 Possible causes of slow expansion . . . 119
5.10 Black hole formation in open clusters . . . 121
5.10.1 Cluster evolution . . . 121
5.10.2 Black hole formation . . . 122
5.11 Summary . . . 124
5.11.1 Embedded clusters in ATLASGAL clumps . . . 124
5.11.2 Star formation efficiency of ATLASGAL clumps . . . 124
5.11.3 Isolated evolution from clumps to embedded clusters . . . 125
5.11.4 Expanding very young star clusters . . . 126
5.11.5 Origin of open clusters . . . 126
5.11.6 Evolution of the $m_{\rm max} - M_{\rm ecl}$ relation . . . 127
5.11.7 Constraining the gas expulsion modes of embedded clusters . . . 128
5.11.8 Black hole formation in open clusters . . . 128

6 Mass functions in star-forming regions . . . 130
6.1 Data . . . 130
6.1.1 Molecular cloud, clump and core catalogs . . . 130
6.1.2 Star cluster and stellar association catalogs . . . 130
6.1.3 CO (2-1) cube and 21 μm map . . . 131
6.2 From CMF to IMF . . . 131
6.2.1 Observed CMFs . . . 131
6.2.2 Clumps and embedded clusters . . . 132
6.2.3 Initial stellar populations . . . 133
6.2.4 Comparison of synthesized and observed CMFs . . . 133
6.3 Clump populations in molecular clouds . . . 136
6.3.1 Clump mass function in observation . . . 138
6.3.2 Analytical model of the clump mass function . . . 139
6.3.3 Spatial distribution of clumps . . . 140
6.3.4 Mass segregation . . . 141
6.4 The ICIMF theory . . . 141
6.4.1 Correlations between molecular clouds and their internal stellar populations . . . 142
6.4.2 Embedded stellar populations . . . 144
6.4.3 The ICIMF theory . . . 146
6.4.4 Cluster complex rather than single cluster . . . 148
6.5 Summary . . . 150
6.5.1 From CMF to IMF . . . 150
6.5.2 Clump populations in molecular clouds . . . 150
6.5.3 The ICIMF theory . . . 150

7 Conclusion and outlook . . . 152

Bibliography . . . 155

Acknowledgements . . . 181

Scientific contributions 182

Original papers of Chapter.2 . . . . . **185**

Original papers of Chapter.3 . . . . . **235**

Original papers of Chapter.4 . . . . . **252**

Original papers of Chapter.5 . . . . . **281**

Original papers of Chapter.6 . . . . . **347**

Nomenclature

Physical and astronomical constants and units

The values used in this dissertation are based on CODATA recommended values 2018 (**CODATA2018**).[1]

Speed of light in vacuum	c	$= 299\,792\,458\ \mathrm{m\,s^{-1}}$
Electron charge	e	$= 1.602\,176\,634 \times 10^{-19}\ \mathrm{C}$ (coulombs)
Electron mass	m_e	$= 9.109\,383\,7015(28) \times 10^{-31}\ \mathrm{kg}$
Gravitational constant	G	$= 6.674\,30(15) \times 10^{-11}\ \mathrm{m^3\,kg^{-1}\,s^{-2}}$
Boltzmann constant	k_B	$= 1.380\,649 \times 10^{-23}\ \mathrm{J\,K^{-1}}$
Astronomical unit	1 AU	$= 1.495\,978\,707 \times 10^{11}\ \mathrm{m}$
Parsec	1 pc	$= 648\,000/\pi$ AU
Light-year	1 ly	$= 9\,460\,730\,472\,580\,800$ m
Nominal solar mass parameter	$(\mathcal{GM})^\mathrm{N}_\odot$	$= 1.327\,1244 \times 10^{20}\ \mathrm{m^3 s^{-2}}$
(**Prsa2016**)	$T_\odot$	$= (\mathcal{GM})^\mathrm{N}_\odot / c^3 \approx 4.925\,490\,947\,641\ \mu\mathrm{s}$
Solar mass (derived)	$\mathrm{M}_\odot$	$= (\mathcal{GM})^\mathrm{N}_\odot / G \approx 1.998\,410(45) \times 10^{30}\ \mathrm{kg}$

[1] *Committee on Data of the International Science Council*, see `http://physics.nist.gov/constants`.

Frequently used acronyms

ALMA	Atacama Large Millimeter/submillimeter Array
ATOMS	ALMA Three-millimeter Observations of Massive Star-forming regions
IRAS	Infrared Astronomical Satellite
APEX	Atacama Pathfinder Experiment
LAsMA	Large APEX sub-Millimeter Array
GILDAS	Grenoble Image and Line Data Analysis Software
GLIMPSE	Galactic Legacy Infrared Mid-Plane Survey Extraordinaire
ATLASGAL	APEX Telescope Large Area Survey of the Galaxy
PHANGS	Physics at High Angular resolution in Nearby GalaxieS
JWST	James Webb Space Telescope
HST	Hubble Space Telescope
MUSE	Multi Unit Spectroscopic Explorer
MYStIX	Massive Young Star-Forming Complex Study in Infrared and X-ray
IMF	Initial Mass Function
ICIMF	Integrated Cloud-wide IMF
IGIMF	Integrated Galaxy-wide IMF
ISM	Interstellar Medium
LTE	Local Thermodynamic Equilibrium
FOV	Field-Of-View
SFH	Star Formation History
SFE	Star Formation Efficiency
PMS	Pre-Main Sequence
LMC	Large Magellanic Cloud
MSFR	Massive Star-Forming Region
PPV	Position–Position–Velocity
GMC	Giant Molecular Cloud
SNR	Supernova Remnant
IRDC	Infrared-Dark Cloud
SED	Spectral Energy Distribution
OC	Open cluster
EC	embedded cluster
ZAMS	Zero Age Main Sequence
MST	Minimum Spanning Tree
KinMS	Kinematic Molecular Simulation
UC-HII	Ultracompact HII
moment 0	integrated intensity map
moment 1	intensity-weighted velocity map

CHAPTER 1

Introduction

1.1 Massive star and star cluster formation

Our understanding of massive star and star cluster formation remains incomplete in the present day. Most and perhaps all the observed stars were formed in embedded clusters (Kroupa, 1995a; Kroupa, 1995c; C. J. Lada & E. A. Lada, 2003; Kroupa, 2005; Megeath et al., 2016; Dinnbier et al., 2022), introducing considerable complexity in the identification and analysis of individual young stellar objects. Unlike their lower-mass counterparts, massive stars progress to the main sequence while still deeply embedded in their natal clumps, rendering all initial stages invisible even when observed at mid-infrared wavelengths (not all). Adding to the challenge, massive stars are both rare and undergo rapid evolution. Typically, the sites where they form are situated at much greater distances compared to their low-mass counterparts, further complicating observational efforts (Urquhart et al., 2022; Wells et al., 2022).

Fortunately, our understanding of massive star and cluster formation has seen significant advancements, thanks to the combination of Galactic plane surveys and high-angular resolution images obtained through submillimeter facilities (Motte et al., 2018b; Beuther et al., 2025). Notably, surveys like HiGAL (Molinari et al., 2010) and ATLASGAL (Schuller et al., 2009) have played a crucial role by providing unbiased samples of dense clumps, identified by their dust emission. Dense clumps serve as the initial stages of massive star and cluster formation. Investigations utilizing extensive and statistically representative clump samples can offer valuable perspectives on the effectiveness of transforming molecular gas into stellar clusters (Elia et al., 2021; Urquhart et al., 2022; Molinari et al., 2025). This approach enables us to evaluate how the physical environment influences this intricate process, providing a deeper understanding of the kinematics and dynamics involved.

1.2 Hub-filament structures

Giant molecular clouds (GMCs) are widely acknowledged as the primary reservoirs of gas that drive star formation, serving as the cradles for nearly all stars. Understanding the properties of GMCs is therefore essential for unraveling the intricate link between gas and star formation in galaxies (Kennicutt & N. J. Evans, 2012; Schinnerer & Leroy, 2024). Analyzing the dynamical interaction between density enhancements in molecular clouds and gas motion in their surrounding environment provides insight into the formation of hierarchical structures in molecular clouds (McKee & Ostriker, 2007; Dib et al., 2012; Motte et al., 2018b; Henshaw et al., 2020). Velocity and density fluctuations are discussed in detail in Henshaw et al. (2020), who detected ubiquitous velocity fluctuations across all spatial scales and galactic environments and discovered oscillatory gas flows with wavelengths ranging from 0.3-400

pc that are coupled to regularly spaced density enhancements that probably formed via gravitational instabilities (Henshaw et al., 2016; B. G. Elmegreen et al., 2018). Furthermore, the spatial correlation between density enhancements and velocity gradient extrema may indicate convergent motion due to gravitational collapse (Hacar & Tafalla, 2011; Hacar et al., 2016; S. D. Clarke et al., 2016; Misugi et al., 2019; J.-W. Zhou et al., 2022; J. W. Zhou et al., 2023). These studies may distinguish between competing concepts for high-mass star formation, such as monolithic collapse of turbulent cores in virial equilibrium (McKee & Tan, 2003a; Krumholz et al., 2007), competitive accretion in a protocluster environment through Bondi-Hoyle accretion (Bonnell et al., 1997; Bonnell et al., 2001), turbulence-driven inertial-inflow (Padoan et al., 2020a) and gravity-dominated global hierarchical collapse (Vázquez-Semadeni et al., 2009; Ballesteros-Paredes et al., 2011a; Hartmann et al., 2012; Vázquez-Semadeni et al., 2017; Vázquez-Semadeni et al., 2019), as discussed in J.-W. Zhou et al. (2022).

Filamentary structures are ubiquitous in high-mass star-forming molecular clouds. High-resolution observations show that density enhancements are organized in filamentary gas networks, especially in hub-filament systems. A collapsing cloud with converging filaments forms a hub-filament system. Hub-filament systems are known as a junction of three or more filaments. Filaments have lower column densities compared to the hubs, but filaments show much higher aspect ratio than the hubs (Myers, 2009; Schneider et al., 2012). In such systems, converging flows are funneling matter into the hub through the filaments. Therefore, cores embedded in denser clumps can prolong the accretion time for growing massive stars due to the sustained supply of matter from the filamentary environment. Numerical simulations of colliding flows and collapsing clumps reveal velocity gradients along the dense filamentary streams converging toward the hubs (P. Wang et al., 2010; Gómez & Vázquez-Semadeni, 2014; R. J. Smith et al., 2016; Padoan et al., 2020a). In observations, velocity gradients along filaments are often interpreted as evidence for gas inflow along filaments (Kirk et al., 2013; T. Liu et al., 2016a; Yuan et al., 2018; G. M. Williams et al., 2018; H.-R. V. Chen et al., 2019; M. C.-Y. Chen et al., 2020; Pillai et al., 2020; J.-W. Zhou et al., 2022).

Many case studies have suggested that hub-filament systems are the birth cradles of high-mass stars and clusters (Peretto et al., 2013; Henshaw et al., 2014; Q. Zhang et al., 2015; T. Liu et al., 2016a; Yuan et al., 2018; Lu et al., 2018; Issac et al., 2019; Dewangan et al., 2020; X.-L. Liu et al., 2021; H.-L. Liu et al., 2022; Kumar et al., 2020; J.-W. Zhou et al., 2022; H.-L. Liu et al., 2023; F.-W. Xu et al., 2023; J. W. Zhou et al., 2023; Yang et al., 2023; J. W. Zhou et al., 2024f). Statistically, Kumar et al. (2020) identified the candidate hub-filament systems in the Milky Way inner disc by searching for filamentary structures around 35 000 clumps in the Hi-GAL catalogue (Elia et al., 2017). ≈11% of the clumps are the candidates of hub-filament systems. The column densities of hubs are found to be enhanced from a factor of approximately two (pre-stellar sources) up to about ten (proto-stellar sources). High-mass stars preferentially form in the density-enhanced hubs of hub-filament systems.

1.3 Tidal effects

One prevailing paradigm about star formation is that the collapse of gas is driven by gravity, balanced by processes such as turbulence and magnetic fields (McKee & Ostriker, 2007). Based on this picture, the virial parameter (Bertoldi & McKee, 1992) is always used to determine whether a structure is in a state of gravitational collapse. However, as pointed out in Ramırez-Galeano et al. (2022) and G.-X. Li (2024), these studies focus on the role of gravity in individual, isolated parts, neglecting long-range gravitational interactions. Diverse manifestations of gravitational effects on gas within molecular clouds were unveiled in G.-X. Li (2024). Gravity influences cloud evolution in various ways. Within dense regions, it facilitates fragmentation and collapse. Outside these regions, it can suppress the low-density

gas from collapsing through extensive tidal forces and drive accretion.

Gravity is a long-range force. A local dense structure evolves under its self-gravity, but as a gravitational center, its gravity can also affect neighboring structures. At the same time, it also experiences the external gravity from neighboring material. The tidal and gravitational fields are mutually interdependent. The multi-scale gravitational coupling of gas structures implies extensive tidal interactions between them, which may prevent gravitational collapse of these structures.

1.4 Feedback

High-mass stars (>8 $M_\odot$) have a profound impact on the evolution of the interstellar medium (ISM). Throughout their short lifetimes ($\approx 10^6$ yr), radiation-driven stellar winds from high-mass stars create HII regions in molecular clouds (Zinnecker & Yorke, 2007; Molinari et al., 2014; Motte et al., 2018b). High-mass stars end their lives in the form of supernovae, whose explosions can release $\approx 10^{51}$ erg of energy near-instantaneously. Shocks from expanding HII regions and supernova remnants (SNRs) can accelerate and heat their surrounding gas, and add turbulence to the gas. The combined effects of multiple feedback mechanisms may play a major role in determining star formation efficiency and the initial mass function in molecular clouds (Dib et al., 2010; Krumholz et al., 2014). In Krumholz et al. (2014), feedback mechanisms are divided into three categories: (1) momentum feedback, including protostellar outflows and (probably) radiation pressure; (2) explosive feedback, including winds from hot main sequence stars, photoionizing radiation, and supernovae; and (3) thermal feedback, with non-ionizing radiation as the main form.

In observations, Lopez et al. (2014) assessed the role of stellar feedback at scales ≈10–100 pc toward 32 HII regions in the Large and Small Magellanic Clouds. They employed multiwavelength data to examine several stellar feedback mechanisms and found that the ionized gas feedback dominates over the other mechanisms in all of the sources. Similar methods were also used in Olivier et al. (2021), but for 106 deeply embedded HII regions with radii < 0.5 pc, they concluded that the dust-processed radiation dominates in 84% of the samples. In simulations, massive stellar feedback, including ionizing radiation, stellar winds, and supernovae (Matzner, 2002; Dale et al., 2012; Rogers & Pittard, 2013; Dale et al., 2014; Rahner et al., 2017; M. C. Smith et al., 2018; Lewis et al., 2023), can suppress star formation and destroy the natal cloud. However, whether stellar feedback promotes or suppresses star formation remains controversial. The simulations of Dale et al. (2014) examined the effects of photoionization and momentum-driven winds from O-stars on molecular clouds, and found that feedback has little effect on denser and more massive clouds, which mainly destroy clouds with a lower density and mass. Dib et al. (2013) showed, using semi-analytical models in which gas-expulsion from clumps is driven by stellar wind feedback originating from both B and O stars, that the gas expulsion timescale depends on the clump mass and the coupling efficiency of the wind with the surrounding gas. The observations of Pabst et al. (2019) and Henshaw et al. (2022) suggest that at early stages energy-winds could dominate, even if this is questioned by theoretical works (Lancaster et al., 2021a; Lancaster et al., 2021b). It is also possible for the feedback to be positive by triggering new star formation. In Deharveng et al. (2005), Zavagno et al. (2006), and H.-L. Liu et al. (2017), using data at various wavelengths, they identified and studied several objects in which star formation appears to have been triggered at the borders of HII regions. M. A. Thompson et al. (2012) studied a large sample of infrared bubbles and estimated that the fraction of massive stars in the Milky Way formed by triggering could be between 14% – 30%. Kendrew et al. (2012) found that ≈22% of massive young stars may have formed as a result of feedback from expanding HII regions.

W49A is one of the most massive and luminous young star-forming regions in the Milky Way. As

presented in Rugel et al. (2019), it is more likely that only limited parts of W49A were affected by feedback from the central stellar cluster, while stars in the outer parts of W49A formed independently. Moreover, all feedback models used in Rugel et al. (2019) predict recollapse of the shell after the first star formation event, which means that feedback from the first formed cluster is therefore not strong enough to disperse the cloud. As presented in Watkins et al. (2019a), stellar feedback from O stars clearly modifies the structure of the larger-scale clouds, but without much effect on the dynamical properties of the assembled dense gas. Previous work on the G305 region observed with the Large APEX sub-Millimeter Array (LAsMA) 7 beam receiver on the Atacama Pathfinder Experiment 12 meter submillimeter telescope (APEX) found that strong stellar winds drive turbulence in the G305. Feedback has triggered star formation through the collect and collapse mechanism in the G305 (Mazumdar et al., 2021b; Mazumdar et al., 2021a). The dense molecular gas structures inside the cloud serve as star-forming sites and their physical states directly determine the star-formation capability of the molecular cloud under feedback. A basic question is how the dense gas structures survive and maintain star formation activity in a strong feedback environment, which depends on the relative strengths of the gravity and turbulence therein.

The relative importance of turbulence and gravity in massive star-forming regions is a long and widely debated topic, with different views leading to different physical pictures of massive star formation, such as the turbulent-core model (McKee & Tan, 2003a; Krumholz et al., 2007), the competitive-accretion model (Bonnell et al., 1997; Bonnell et al., 2001), the inertial-inflow model (Padoan et al., 2020a), and the global hierarchical collapse model (Vázquez-Semadeni et al., 2009; Ballesteros-Paredes et al., 2011a; Hartmann et al., 2012; Vázquez-Semadeni et al., 2017; Vázquez-Semadeni et al., 2019). Larson's laws claim that, in molecular clouds, the velocity dispersion σ scales proportionally to the scale R, and molecular clouds are approximately in virial equilibrium, with a mostly uniform column density. The Larson relation ($\sigma_v \propto R^{0.5}$) is generally used to emphasize the importance of turbulence in molecular clouds, and turbulence acts to sustain the clouds against gravitational collapse (R. B. Larson, 1981; Solomon et al., 1987; M. H. Heyer & Brunt, 2004; Mac Low & Klessen, 2004; McKee & Ostriker, 2007; Hennebelle & Falgarone, 2012). M. Heyer et al. (2009) generalized the Larson relation by extending the Larson ratio $L \equiv \sigma_v/R^{0.5}$ with the surface density Σ of Galactic GMCs, which gives $\sigma_v/R^{0.5} \propto \Sigma^{0.5}$. Subsequently, the Heyer relation was used to emphasize the importance of gravity in molecular clouds (Ballesteros-Paredes et al., 2011b; Traficante et al., 2018a; Traficante et al., 2018b; Ballesteros-Paredes et al., 2018; Vázquez-Semadeni et al., 2019; Ballesteros-Paredes et al., 2020). In particular, in the high-column-density portions of star-forming regions, such as clumps or cores, the Heyer relation always performs better than the Larson relation (Ballesteros-Paredes et al., 2011b; Camacho et al., 2016; Traficante et al., 2018b), suggesting strong gravity at relatively small scales in molecular clouds. Ibáñez-Mejıa et al. (2016) showed that the Heyer relation cannot be reproduced without self-gravity in simulations of the ISM with SN-driven turbulence. In contrast, purely SN-driven turbulence in the ISM generates the Larson relation.

Regarding the explanation of the Heyer relation, M. Heyer et al. (2009) claimed that it is consistent with the clouds being in virial equilibrium, as it follows directly from the condition $2E_\mathrm{k} = E_\mathrm{g}$, where $E_\mathrm{k} = M\sigma^2/2$ and $E_\mathrm{g} = 3GM^2/5R$. Ballesteros-Paredes et al. (2011b) further pointed out that the scaling is also consistent with the clouds undergoing free-fall, in which case $E_\mathrm{k} = |E_\mathrm{g}|$. However, the differences between the effects of free-fall and virial equilibrium in the $\sigma_v/R^{0.5}$ versus Σ diagram are smaller than the typical uncertainty of the observational data (Ballesteros-Paredes et al., 2011b), and are therefore difficult to distinguish. Both explanations involve only gravitational and kinetic energy, which may be a workable approximation for relatively isolated molecular clouds, but is often too simplistic for substructures inside a molecular cloud, especially for a cloud affected by feedback. When the substructures can self-gravitationally collapse, they may decouple from the surrounding environment

(Peretto et al., 2023). If not, the exchange of energy with the surrounding environment will break the conversion between gravitational potential energy, E_g, and kinetic energy, E_k, of the structures, and thus violate the Heyer relation.

1.5 Formation and evolution of star clusters

Until a decade ago, the fields of stellar dynamics and chemical abundances on the one hand and star cluster formation on the other were operating almost independently with the stellar kinematics and abundances community focusing on the long-term dynamical evolution of stellar clusters (Aarseth, 1974; D. C. Heggie et al., 2006; Lamers et al., 2008; Portegies Zwart et al., 2010a; Spurzem et al., 2018), while the star-formation community has directed its attention toward the dynamics of interstellar gas and the impacts of star formation and stellar feedback (Bate, 1998; Dib et al., 2008; R. J. Smith et al., 2009; Dale, 2017; Motte et al., 2018b; J. W. Zhou et al., 2024g). As highlighted in Krumholz et al. (2019), maintaining this division is no longer tenable. Advances in observational capabilities and large surveys of the populations of young and less young clusters in the Milky Way now allow us to trace the origins of field stars back to their birthplaces, or at least reconstruct some of the now-dissolved structures in which they originated (Wright et al., 2024; Cantat-Gaudin et al., 2024). To guide such reconstructions, input from both theoretical and observational perspectives within the star-formation community is essential.

Massive star-forming regions (MSFRs) represent a major mode of star formation within the Galaxy (C. J. Lada & E. A. Lada, 2003; Zinnecker & Yorke, 2007; Motte et al., 2018c; Dib, 2023a). Young stellar clusters within these regions can serve as progenitors for open clusters. Measurements of the internal structure and dynamics of young clusters and star-forming regions are necessary to fully understand the process of their formation and dynamical evolution. Despite their importance, the properties of embedded clusters and their relationship to open clusters remain unclear. The mechanisms driving the fragmentation of gas that forms these clusters, as well as the subsequent dynamics of their stars, are not yet well understood (Megeath et al., 2016; Dib, 2023a; J. W. Zhou et al., 2024f). Although embedded clusters are the likely precursors to open clusters, C. J. Lada & E. A. Lada (2003) discovered that only 7% of embedded clusters survive the gas dispersal phase. In N-body simulations of Kroupa (2001) and Brinkmann et al. (2017), massive clusters containing many O stars most likely survive as open clusters. However, low-mass embedded clusters or groups dissolve quickly, partly owing to the loss of their residual gas.

Recent observational progress has highlighted the crucial role of early expansion in the evolution of young star clusters. By leveraging Gaia DR2 and DR3 data in combination with multi-object spectroscopy, studies have revealed expansion processes occurring in very young clusters within star-forming regions through both case studies (Wright et al., 2019; Cantat-Gaudin et al., 2019; Kuhn et al., 2020; Lim et al., 2020; Swiggum et al., 2021; Lim et al., 2022; Mužić et al., 2022; Das et al., 2023; Flaccomio et al., 2023) and statistical works (Kuhn et al., 2019; Della Croce et al., 2024; Wright et al., 2024; Jadhav et al., 2024). Bastian & Goodwin (2006) compared the luminosity profiles of young massive clusters, such as M82-F, NGC 1569-A, and NGC 1705-1, with N-body simulations of clusters that had experienced rapid gas expulsion, resulting in the loss of a substantial portion of their mass. They found the observed profiles to closely agree with the simulations. Goodwin & Bastian (2006) analyzed the differences between luminosity-based and dynamical mass estimates for young massive clusters. Their results revealed significant discrepancies, which they attributed to the possibility that many young clusters are not in virial equilibrium due to undergoing violent relaxation following gas expulsion. Additionally, these authors noted that the increasing core radii observed in young clusters within the Large Magellanic Cloud and Small Magellanic Cloud can be well explained as an effect of rapid gas loss. In Kroupa et al.

(2001) and Banerjee & Kroupa (2013), Banerjee & Kroupa (2014) and Banerjee & Kroupa (2015), direct N-body modeling of realistic star clusters with a monolithically formed structure and undergoing residual gas expulsion consistently reproduced the characteristics of several well-observed very young star clusters, namely the Orion Nebula Cluster (ONC) and the clusters R136 and NGC 3603, thereby shedding light on the birth conditions of massive star clusters. Similar simulations were also carried out in Banerjee & Kroupa (2017) to explain how young massive star clusters attain their current shapes and sizes from an initially compact embedded morphology. These authors found that young massive clusters cannot expand to their current sizes if only two-body relaxation, stellar mass loss and dynamical heating through initial binaries are considered. They also confirmed that a substantial residual gas expulsion with ≈30% star formation efficiency can adequately swell the newborn embedded clusters.

The question of whether newborn stars in star clusters are drawn randomly from the initial mass function (IMF) is critical for various fields of astrophysics. The stochastic (or random) sampling approach posits that each star's formation in an embedded cluster is an independent event, with its mass being randomly selected from a probability density function (PDF), effectively treating the IMF as a PDF. Conversely, the optimal sampling theory (Kroupa et al., 2013; Schulz et al., 2015) suggests a deterministic relationship between the mass of an embedded cluster and the mass of each individual star within it, due to a self-regulation star formation process in the cluster (Kroupa, 2015; Yan et al., 2023). A key factor in determining whether a star cluster should be modeled using stochastic sampling or optimal sampling from the IMF is the presence of a significant relationship between the mass of the most massive star ($m_{\rm max}$) and the total mass of the embedded cluster in stars ($M_{\rm ecl}$). If a statistically significant correlation exists between $m_{\rm max}$ and $M_{\rm ecl}$, it would suggest that physical processes, such as self-regulation during star formation within the cluster, are at play. Actually, the $m_{\rm max} - M_{\rm ecl}$ relation has been well established analytically (B. G. Elmegreen, 2000; Weidner & Kroupa, 2004; Oey & C. J. Clarke, 2005), observationally (R. B. Larson, 2003; Weidner & Kroupa, 2006; Weidner et al., 2010; J. W. Zhou et al., 2024b) and numerically (Bonnell et al., 2003; Bonnell et al., 2004; Peters et al., 2011). Yan et al. (2023) provided a deterministic tool for sampling stellar mass that reproduces the observed $m_{\rm max} - M_{\rm ecl}$ relation and aligns well with the luminosities of star-forming clumps (also refer to J. W. Zhou et al. 2024d).

1.6 Stellar and core mass functions

Galaxies act as stellar factories in the universe, producing stars through the gravitational collapse of the densest regions within molecular clouds. Molecular clouds are widely distributed within galaxies and serve as localized sites for star formation. Each molecular cloud typically contains multiple clumps, which are the local sites of star formation within the cloud and the precursors of embedded star clusters (Miville-Deschênes et al., 2017; E. Rosolowsky et al., 2021; Urquhart et al., 2022; Yan et al., 2017; J. W. Zhou et al., 2024d; J. W. Zhou et al., 2024b). As presented in Motte et al. (2018b), Vázquez-Semadeni et al. (2019), Kumar et al. (2020), Henshaw et al. (2020), J.-W. Zhou et al. (2022), J. W. Zhou et al. (2023), J. W. Zhou et al. (2024f), J.-W. Zhou & T. Davis (2024), J. W. Zhou et al. (2024a) and J. W. Zhou & G.-X. Li (2025), the molecular gas from galaxy to dense core scales presents multi-scale/hierarchical hub-filament structures. Like dense core as the hub in clump, clump as the hub in molecular cloud, molecular cloud can be the hub in galaxy. Hubs at different scales as local gravitational centers can continuously accrete surrounding diffuse gas, thereby accumulating mass. The hierarchical structure of molecular gas in a galaxy determines the spatial distribution of stellar populations (Gouliermis et al., 2015; D. M. Elmegreen et al., 2014; Grasha et al., 2017; Chevance et al., 2023). Young clusters with different ages constitute a cluster population within the parent molecular cloud. The evolution of the

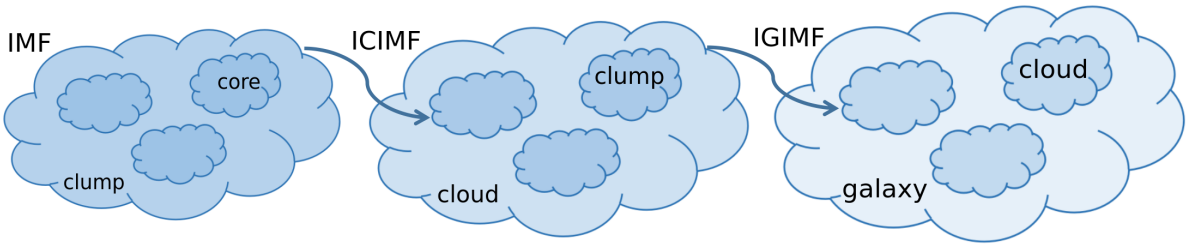


Figure 1.1: Correspondence between hierarchical gas structures (star-forming regions) and hierarchically integrated IMFs. "ICIMF" and "IGIMF" are the integrated cloud-wide IMF and the integrated galaxy-wide IMF, respectively. In the hierarchical hub–filament structures, a dense core (≈1000 AU) serves as the hub within a clump (≈1 pc), a clump functions as the hub within a molecular cloud (≈10–100 pc), and a molecular cloud acts as the hub within a galaxy. Clumps are localized high-density structures within a molecular cloud and serve as the primary sites of star formation within the cloud. Cores are assumed to be overdense structures with respect to their parental clump.

molecular cloud is closely related to the formation and evolution of the internal star clusters. The star clusters become fully exposed after the molecular cloud is dispersed by feedback (J.-G. Kim et al., 2016; Kruijssen et al., 2019; Chevance et al., 2022; Watkins et al., 2023; J. W. Zhou et al., 2024e).

Most, if not all, observed stars are believed to form within embedded clusters (Kroupa, 1995a; Kroupa, 1995c; C. J. Lada & E. A. Lada, 2003; Kroupa, 2005; Megeath et al., 2016; Dinnbier et al., 2022). Based on this viewpoint, the integrated galaxy-wide initial mass function (IGIMF) theory has been developed to calculate galaxy-wide IMFs, starting with the IMF that is observationally constrained at the star cluster scale (Kroupa & Weidner, 2003; Kroupa et al., 2013; Yan et al., 2017; Jerabkova et al., 2018; Dib, 2022; Haslbauer et al., 2024). The IMFs assembled in star-forming regions combine to create the IMF of the entire galaxy. This process involves first generating a series of embedded star clusters, then producing a series of stars for each embedded cluster from the IMF, and finally summing the stars from all clusters to construct the galaxy-wide IMF. Star clusters first form in molecular clouds. Molecular clouds, serving as intermediaries between star clusters and the galaxy, should also be incorporated into the IGIMF theory. Molecular clouds and their internal star clusters could be quantitatively characterized using a localized IGIMF theory, which can be called the integrated cloud-wide initial mass function (ICIMF) theory. Fig.1.1 illustrates the correspondence between hierarchical gas structures (star-forming regions) and hierarchically integrated IMFs, providing a depiction of their relationship.

A dense core is a localized fragment or overdensity within a molecular cloud, representing a local minimum in gravitational potential and serving as the precursor to an individual star or stellar system (P. André et al., 2014). A prestellar core is a dense core that remains starless while being gravitationally bound. In star-forming regions within the solar neighborhood that produce solar-type stars, the prestellar core mass function (CMF) closely resembles the IMF (Alves et al., 2007; Enoch et al., 2008; P. André et al., 2014; Könyves et al., 2015; Di Francesco et al., 2020). This similarity has led to the hypothesis that the IMF's shape may be directly inherited from the CMF. However, the Milky Way includes both low-mass and high-mass star-forming regions. A thorough analysis of large statistical samples of massive protoclusters is essential to observationally determine whether the origin of the IMF is truly independent of cloud properties. Thanks to observations with the Atacama Large Millimeter/submillimeter Array (ALMA), numerous studies have now examined CMFs in high-mass star-forming regions (Motte et al., 2018a; M. Liu et al., 2018; Cheng et al., 2018; Sanhueza et al., 2019; Lu et al., 2020; Sadaghiani et al., 2020; O'Neill et al., 2021; Pouteau et al., 2022; Nony et al., 2023; Louvet et al., 2024), revealing that the CMFs in these regions are flatter than the ones in nearby, low-mass star-forming regions. The excess of high-mass cores in these regions indicates a top-heavy CMF, which may challenge the universality of the IMF. In the literature, the CMF fitting for high-mass star-forming regions typically does not differentiate

between core types (such as prestellar or protostellar). However, the differentiation was done in Sanhueza et al. (2019) and Nony et al. (2023). Sanhueza et al. (2019) derived a power law index of -2.17±0.10 for the prestellar CMF, which is slightly shallower than the Salpeter IMF. In Nony et al. (2023), on the other hand, the prestellar CMF has a power law index of -2.46±0.15. The protostellar CMF in Nony et al. (2023) is significantly shallower than the Salpeter value, with an index of -1.96±0.09. Therefore, the observed CMF of high-mass star-forming regions may not be robust.

It is now widely accepted that the majority of stars both form and exist on the main sequence with at least one stellar companion; "single stars are the exception, not the rule" (see Offner et al. (2023) and references therein). Goodwin & Kroupa (2005) argued that the short ($< 10^5$ yr) dynamical decay times of nonhierarchical multiple systems imply that most stars must form in binary systems or dynamically stable hierarchical higher-order systems to explain the observed very high fraction of binaries in about 1 Myr old star-forming regions. Multiple star systems are formed during the earliest stages of star formation. Observations have offered evidence that the multiplicity fraction of protostars is higher than that of pre-main-sequence and main-sequence stars (Looney et al., 2000; X. Chen et al., 2013), indicating that most systems initially form as multiples. Observations of multiplicity in the youngest stellar systems are key to understanding their origins (Izquierdo et al., 2018; Olguin et al., 2021; Olguin et al., 2022; Tobin et al., 2022; S. Li et al., 2024). As is shown in Fig.1 of Offner et al. (2023), the likelihood of massive protostars forming as multiple systems may approach nearly 100%. The multiplicity of protostars is tightly correlated with the origin of the IMF or the transformation from the CMF to the IMF. Despite the high multiplicity fraction among protostars, it is often assumed that a single core will produce only one star. A core at low resolution may actually consist of multiple smaller-scale cores, which should form a multiple system rather than a single star. Assuming a one-to-one mapping between cores and stars artificially shifts the cores to the higher-mass range, potentially making the CMF appear top-heavy. Apart from the multiplicity, the mass conversion efficiency from core to star ($\epsilon_{\rm core}$) will also have a significant impact on the transformation between the CMF and the IMF. For the cores in star-forming regions within the solar neighborhood, the values of $\epsilon_{\rm core}$ are $\approx 20\%-40\%$ (Alves et al., 2007; Könyves et al., 2010; Könyves et al., 2015; Megeath et al., 2016).

1.7 Objectives of the thesis

(1) To deepen our understanding of high-mass star formation in hub-filament systems, it is essential to obtain kinematic information at scales down to ≈0.1 pc, comparable to studies of nearby clouds. This necessitates high spatial resolution observations of massive filaments. However, due to the field-of-view (FOV) limitations of ALMA, high-resolution single-dish observations (e.g., APEX/LAsMA) are indispensable for tracing gas flows on molecular cloud scales. Such observations provide a crucial link between ALMA studies of star formation on clump scales (several parsecs) in the Milky Way and those on molecular cloud complex scales (several hundred parsecs) in nearby galaxies. By combining high-resolution APEX/LAsMA and ALMA observations, we may directly reconstruct the hierarchical hub-filament structures and track gas inflows from cloud to dense core scales. ALMA's high-resolution CO imaging data enable systematic studies of molecular clouds in nearby galaxies. The picture of self-similar/multi-scale hub-filament structures and filamentary accretion feeding central regions from cloud to dense core scales may be further extended to galaxy-cloud scales. Based on the evidence from the kinematics, we may directly reconstruct multi-scale hub-filament structures and investigate the evolution of gas structures leading to star formation. The molecular gas can be viewed as a network structure formed by gravitational coupling of multi-scale hub-filament structures. Based on this picture, we may construct gravitational networks and analyze both the gravitational interactions and the physical

properties of the gas structures using geometry-based network metrics.

(2) The ISM in galaxies ranges from diffuse gas to regions of localized clumping, followed by a further increase in density due to gravitational collapse and accretion, ultimately leading to star formation. A gas structure is inevitably impacted by the cumulative tidal interactions with surrounding material, which can impede gravitational collapse and influence star formation. Given the diffuse and intricate nature of the ISM, performing a full tidal calculation becomes quite complex. A practical approach involves deriving the gravitational potential from the 3D density distribution and calculating the tidal field from that. In observational studies, however, we typically only have access to the 2D projected surface density map. Fortunately, with the improved distance measurements provided by the Gaia mission, we can now derive the 3D density distribution of the ISM through advanced dust mapping techniques. By utilizing these 3D dust maps, we can calculate and break down the tidal field. This allows us to explore methods for estimating the tidal forces acting on molecular clouds and evaluate how these tides impact the physical properties of these clouds.

(3) High-mass stars have a significant influence on the evolution of the ISM. The combined effects of various feedback mechanisms are thought to play a key role in regulating star formation efficiency and shaping the initial mass function in molecular clouds. However, whether stellar feedback fosters or inhibits star formation remains a topic of debate. Dense molecular gas structures within clouds serve as sites of star formation, and their physical states directly influence the cloud's ability to form stars in the presence of feedback. Different stages of these dense gas structures are dominated by distinct feedback processes. By comparing their physical properties, we can potentially infer the relative importance of these feedback mechanisms and explore how dense gas structures endure and sustain star formation in environments with strong feedback.

(4) Young stellar clusters within massive star-forming regions are potential progenitors of open clusters. To fully understand the processes of cluster formation and their dynamical evolution, it is essential to study the internal structures and dynamics of these young clusters and star-forming regions. Despite their significance, the properties of embedded clusters and their relationship to open clusters are not yet fully understood. The mechanisms behind the fragmentation of gas that leads to the formation of these clusters, as well as the subsequent dynamics of their stars, remain poorly understood. ATLASGAL clumps with HII regions (HII-clumps) represent the final stage of embedded cluster formation within the clumps. They are crucial for bridging the gap between stellar kinematics, abundances, and star formation studies, as protoclusters or embedded clusters in HII-clumps mark the initial stage of gas-free star clusters. By combining observations and simulations, we may reconstruct the evolutionary pathways from clumps to embedded clusters and eventually to open clusters.

(5) In a hub-filament system, the hub serves as the primary site for star formation. One of the key objectives in star formation research is to understand the origin of the initial mass function. A clump, which forms an embedded cluster, essentially defines an IMF. Therefore, clumps should be considered the fundamental units for studying the origin of the IMF. Hierarchical gas structures (star-forming regions) correspond to hierarchically integrated IMFs. By extending and drawing analogies from this self-similar physical framework, we may establish connections between star formation in the Milky Way and nearby galaxies. Moreover, molecular clouds in the Milky Way offer an opportunity to study their internal structures and stellar populations in detail. Research on the interstellar medium and stellar populations within the Milky Way can serve as a valuable template for analyzing similar structures in nearby galaxies.

1.8 Structure of the thesis

In Chapter.2, we present observational evidence of hierarchical hub-filament structures spanning multiple scales, from dense cores to clumps, molecular clouds, and up to galaxies. Molecular gas is organized into network structures formed by the gravitational coupling of these multi-scale hub-filament structures. These studies highlight the interplay of gravitational forces across molecular gas structures spanning multiple scales. In Chapter.3, we further investigate the widespread tidal interactions and find that tidal effects can suppress gravitational collapse and influence star formation. In Chapter.4, we examine the role of feedback in the formation and evolution of star clusters. This process involves both the destruction and formation of structures, along with shifts in gravitational centers, yet gravitational collapse persists throughout. As an extension on the time scale, Chapter.5 focuses on bridging the formation and evolution of star clusters. We explore the evolutionary pathways leading from clumps to embedded clusters and ultimately to open clusters. Chapter.6 aims to connect star formation in the Milky Way with that in nearby galaxies, linking star-forming regions (hub-filament structures) to stellar populations across various scales. We also discuss the transition from core mass functions to the initial mass function.

In this thesis, each chapter comprises several papers (published or submitted), with each section corresponding to one paper. The selected content from these papers is organized according to the chapter's central theme, and only the parts relevant to that theme are included.

CHAPTER 2

Hierarchical hub-filament structures

The content presented in this chapter are reproduced from the publications by J.-W. Zhou et al. (2022) (Sec.2.2), J. W. Zhou et al. (2023) (Sec.2.3), J. W. Zhou et al. (2024f) (Sec.2.4), J.-W. Zhou & T. Davis (2024) (Sec.2.5), J. W. Zhou et al. (2024a) (Sec.2.7), J. W. Zhou & G.-X. Li (2025) (Sec.2.8) and a submitted work (Sec.2.6). This chapter mainly presents a picture of multi-scale/hierarchical hub-filament structures from dense core to clump to molecular cloud to galaxy scales by observations.

2.1 Data and observations

2.1.1 ALMA observations

We used ALMA data from the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Project ID: 2019.1.00685.S; PI: Tie Liu). The ATOMS sources were initially selected from a complete and homogeneous CS J=2-1 molecular line survey toward IRAS sources with far-infrared colors characteristics of UC HII regions (Bronfman et al., 1996). As discussed in T. Liu et al. (2020), the ATOMS sources are distributed in very different environments of the Milky Way, and are an unbiased sample of the protoclusters with the strongest CS J=2-1 line emission located in the inner Galactic plane (Faúndez et al., 2004). The data reduction and the details of the 12 m array and 7 m array ALMA observations were summarized in T. Liu et al., 2020; J.-W. Zhou et al., 2022.

2.1.2 APEX/LAsMA data

For the G333 molecular cloud complex, we mapped a $3.4^{\circ}\times1.2^{\circ}$ area centered on $(l, b) = (332.33^{\circ}, -0.29^{\circ})$. The 7 pixel LAsMA receiver on the APEX telescope was used to observe the $J = 3 - 2$ transitions of ^{12}CO ($\nu_{\rm rest} \approx 345.796$ GHz) and ^{13}CO ($\nu_{\rm rest} \approx 330.588$ GHz) simultaneously in the upper and lower sideband, respectively. More details about the receiver are given in Mazumdar et al. (2021b). The data were calibrated using the three load chopper wheel method, and reduced using the GILDAS package[1]. More details can refer to J. W. Zhou et al. (2023). The final noise levels of the ^{12}CO (3–2) and ^{13}CO (3–2) data cubes are ≈0.32 K and ≈0.46 K, respectively.

2.1.3 Other data

We used images at 3.6, 4.5, 5.8, and 8.0 μm, obtained by the *Spitzer* Infrared Array Camera (IRAC), as part of the GLIMPSE project (Benjamin et al., 2003). The images of IRAC were retrieved from the *Spitzer*

[1] `http://www.iram.fr/IRAMFR/GILDAS`

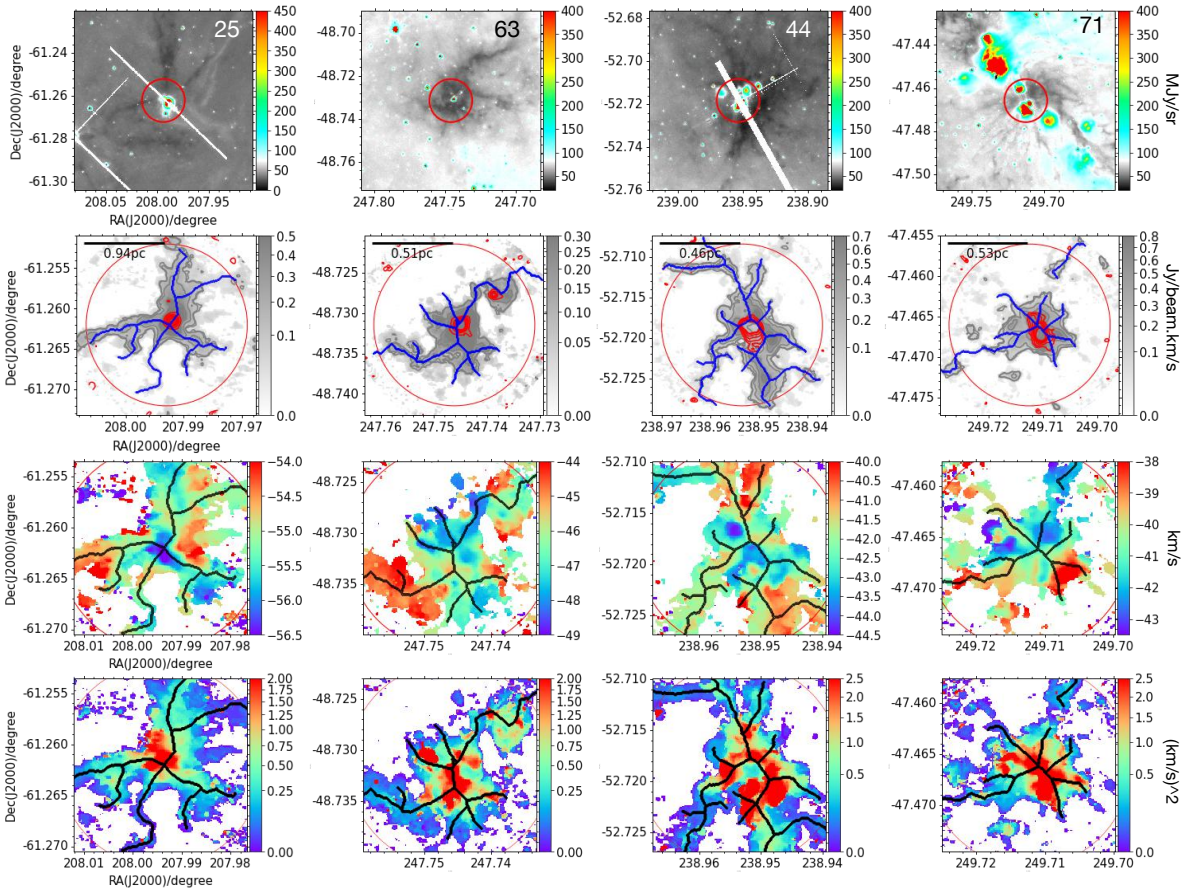


Figure 2.1: Hub-filament systems in four exemplar sources. The first row shows the colour images of *Spitzer* 8 μm emission. The source numbers in the upper-right corners are from the Table.A1 of T. Liu et al. (2020). The second row displays the filaments (blue lines) identified in the moment 0 maps (background) of $H^{13}CO^+$ J=1-0. The red contours represent 3 mm continuum emission, with lowest contour 5σ. The third and fourth rows show the moment 1 and 2 maps of $H^{13}CO^+$ J=1-0 with filament skeletons shown in black lines. The radius of red circles in the maps is 36″, which corresponds to the FOV of the ALMA observations.

Archive and the angular resolutions of images are better than 2″. We used ATLASGAL+Planck 870 μm data (ATLASGAL combined with Planck data), which are sensitive to a wide range of spatial scales at a resolution of ≈ 21″ (Csengeri et al., 2016). We used the combined 12m+7m+TP PHANGS-ALMA CO (2−1) data cubes to investigate gas kinematics and dynamics in spiral galaxies. We also used the James Webb Space Telescope (JWST) 7.7 and 21 μm maps from the PHANGS-JWST survey and the H_α emission map from the PHANGS-MUSE survey. Overviews of the PHANGS-ALMA, PHANGS-MUSE and PHANGS-JWST surveys' science goals, sample selection, observation strategy, and data products are described in Leroy et al. (2021b), Leroy et al. (2021a), Emsellem et al. (2022) and J. C. Lee et al. (2023), respectively. All the data are available on the PHANGS team website [2].

2.2 On clump scale

To deepen our understanding of high-mass star formation in hub-filament systems, it is necessary to obtain kinematic information on scales down to ≈0.1 pc, similar to studies in nearby clouds (P. André

[2] https://sites.google.com/view/phangs/home

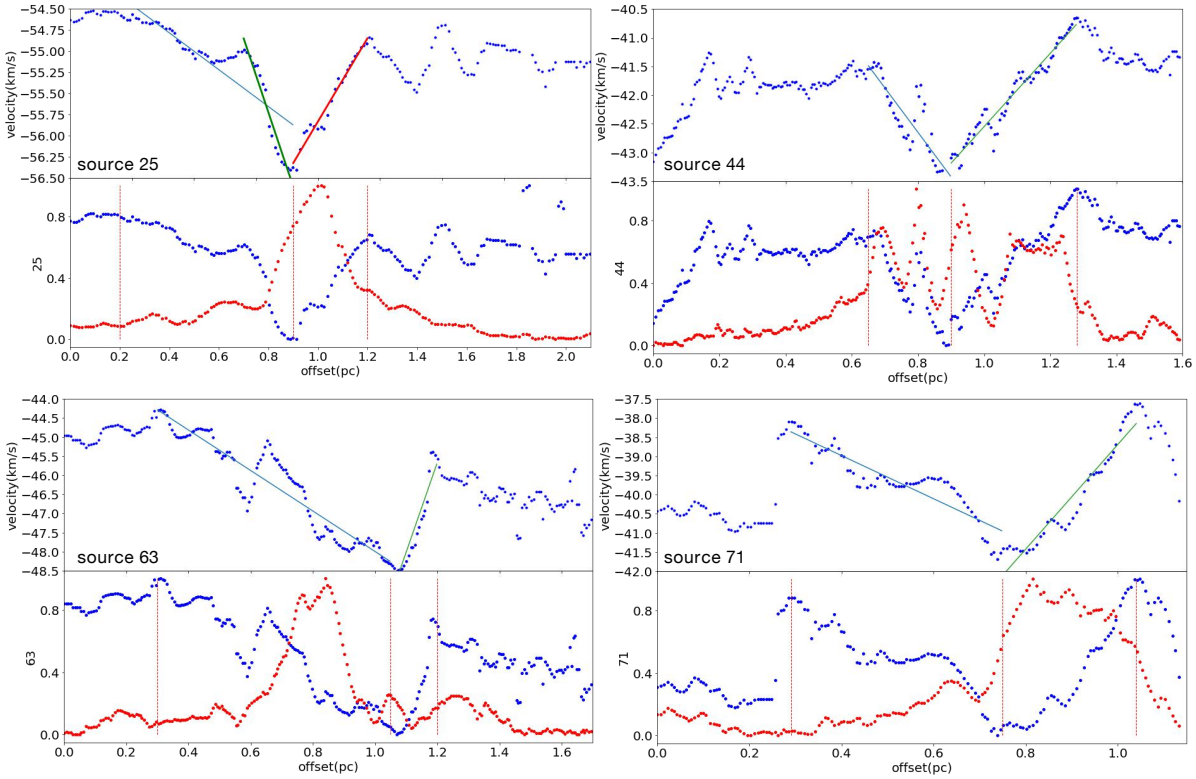


Figure 2.2: The distributions of velocity and intensity along the longest filament in four exemplar sources. Upper panel for each source: Velocity gradients are fitted in the ranges defined by the red vertical dashed lines, and straight lines show the linear fitting results; Lower panel for each source: Blue and red dotted lines show the normalized velocity and intensity, respectively.

et al., 2014; P. André et al., 2016). The work presented in this section is reproduced from the publication by J.-W. Zhou et al. (2022). This work was completed during my master's degree. However, for the sake of the physical image's completeness, and because a series of subsequent works are based on this project and extend it to larger scales using similar methods, it is necessary to first provide a detailed introduction to this work.

We used the FILFINDER algorithm (Koch & E. W. Rosolowsky, 2015a) to identify filaments from the integrated intensity (moment 0) maps of $H^{13}CO^+$ J=1-0. Fig. 2.1 presents the maps of four well-defined hub-filament systems. These hub-filament systems are connected to larger filamentary structures revealed by extinction in *Spitzer* 8 μm images (see top row of panels in Fig. 2.1). Although their velocity fields are complicated, most hub-filament systems show clear velocity gradients, as shown in the intensity-weighted velocity (moment 1) maps of $H^{13}CO^+$ J=1-0 (third row of Fig. 2.1). In addition, the velocity dispersion of $H^{13}CO^+$ J=1-0 increases toward the hub regions (see bottom panels of Fig. 2.1).

ATOMS sources are at various distances from 0.4 to 13.0 kpc , enabling us to investigate filaments from core scale to clump/cloud scale under a uniform angular resolution ($\approx 2''$) with a large sample (T. Liu et al., 2020). The length distribution of the longest filaments in these sources ranges from 0.086 pc to 4.87 pc with a mean value of 1.61 pc and a median value of 1.35 pc, indicating that filaments as well as hub-filament systems can exist not only in small-scale (≈0.1 pc) dense cores but also in large-scale clumps/clouds (≈1-5 pc), and probably up to 10 pc, considering projection effects and the limited FOV of ALMA.

Fig. 2.2 shows the intensity-weighted velocity (moment 1) and integrated intensity of $H^{13}CO^+$ J=1-0 line emission along the longest filaments in four exemplar sources. Velocity and density fluctuations

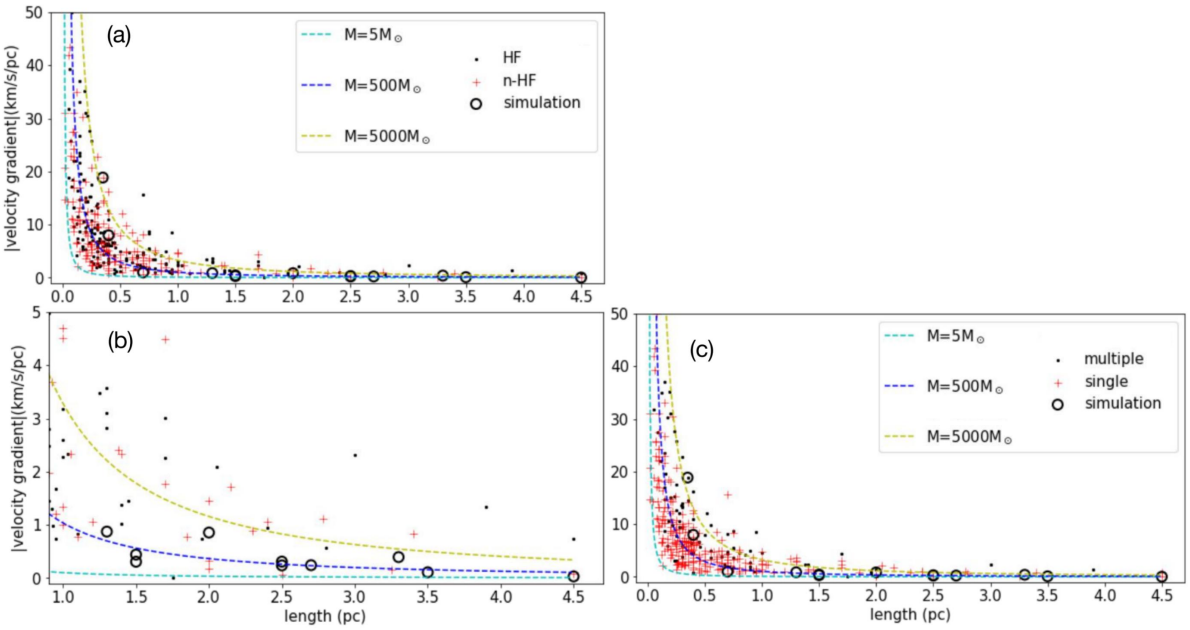


Figure 2.3: Velocity gradient versus the length over which the gradient has been estimated for all sources. (a) Non-hub-filament systems (red crosses) and hub-filament systems (black dots) are plotted severally. The color lines show free-fall velocity gradients for comparison. For the free-fall model, cyan, blue and yellow lines denote masses of $5M_{\odot}$, $500M_{\odot}$ and $5000M_{\odot}$, respectively. Black circles represent the velocity gradients fitted by simulation data; (b) A blow-up of the region with lengths greater than 1 pc in panel (a); (c) Same as in panel (a), but sources showing singly-peaked profiles (red crosses) and multiple-peaked profiles (black dots) in $H^{13}CO^{+}$ line emission in hub regions are plotted separately.

along these filaments are seen, which are likely caused by dense structures embedded in these filaments. The density and velocity fluctuations along filaments may indicate oscillatory gas flows coupled to regularly spaced density enhancements that probably form via gravitational instabilities (Henshaw et al., 2020). On the other hand, one can see clear velocity gradients along the longest filaments. We firstly estimate two velocity gradients between velocity peaks and valleys at the two sides of the strongest intensity peaks of $H^{13}CO^{+}$ emission (i.e., the center of the gravity potential well), as marked by the vertical dashed lines in Fig. 2.2, and ignored local velocity fluctuations. We also derive additional velocity gradients over smaller distances around the strongest intensity peaks of $H^{13}CO^{+}$ emission for some very long filaments (such as source 25 in the lower panels of Fig. 2.2).

Fig. 2.3 shows velocity gradients as a function of the filament lengths over which the gradients have been estimated. There are no significant difference in velocity gradients among filaments with or without hubs. This may indicate that the physical drivers that generate the kinematics, such as filamentary accretion at different scales, are the same for filaments in various systems. We find that velocity gradients are very small at scales larger than 1 pc relative to small scales (<1 pc), probably hinting for the existence of inertial inflow at large scales driven by turbulence (Padoan et al., 2020b), or by large-scale gravitational collapse (Gómez & Vázquez-Semadeni, 2014; Vázquez-Semadeni et al., 2019). This is evidenced by Fig. 2.3(b), which shows that the variation of velocity gradients on large scales more or less follow the trend at small scales. Below ≈1 pc, velocity gradients dramatically increase as filament lengths decrease, indicating that gravity dominates gas inflow at such small scales where gas is being accumulated in dense cores or hub regions. The large velocity gradients, however, could be overestimated due to the complicated line emission profiles in hub regions.

The prevalent hub-filament systems found in proto-clusters favors the pictures advocated by either

global hierarchical collapse or inertial-inflow scenarios of high-mass star formation, which emphasize longitudinal flow along filaments. However, here we argue that gas infall at scales smaller than ≈1 parcsec is likely dominated by the gravity of the hub, while velocity gradients are very small at scales larger than ≈1 pc, probably suggesting the dominance of either pressure-driven inflow, which may be driven either by large-scale turbulent motions (inertial inflow) or large-, cloud-scale collapse, or else the self-gravity of the filament.

Hub-filament structures can exist at various scales from 0.1 parsec to several parsecs in very different Galactic environments. Below the 0.1 pc scale, slender structures similar to filaments have also been detected in the surroundings of high-mass protostars (H. B. Liu et al., 2015; Maud et al., 2017; Izquierdo et al., 2018; X. Chen et al., 2020; Sanhueza et al., 2021). Therefore, self-similar hub-filament systems and filamentary accretion seem to exist at all scales (from several thousands au to several parsecs) in high-mass star forming regions. This paradigm of hierarchical hub-filament like systems should be considered in any promising model for high-mass star formation.

2.3 On cloud scale

In J.-W. Zhou et al. (2022), we measured the velocity gradients around the hubs. The statistical analysis found that velocity gradients are very small at scales larger than ≈1 pc, probably suggesting the dominance of pressure-driven inertial inflow, which can originate either from large-scale turbulence or from cloud-scale gravitational contraction. Below ≈1 pc, the velocity gradients dramatically increase as the filament lengths decrease, indicating that the hub or core gravity dominates gas infall on small scales. Due to the FOV limitation of the ALMA observation, our previous work was restricted to scales up to about 5 pc. In this work, following a similar method as described in J.-W. Zhou et al. (2022), we generalized the results from clump-core scales to cloud-clump scales. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2023).

2.3.1 Velocity gradients

Following the method described in J.-W. Zhou et al. (2022), we used the moment 0 maps of ^{13}CO (3−2) to identify filaments in the G333 complex using the FILFINDER algorithm (Koch & E. W. Rosolowsky, 2015a). The skeletons of identified filaments overlaid on the moment 0 maps of the ^{13}CO (3-2) line emission are shown in Fig.2.4.

The kinematical features of hub-filament systems can be revealed by the velocity gradients along the filaments. As done in J.-W. Zhou et al. (2022), the same method was used to derive the velocity gradients from the LAsMA data. Then the fitted velocity gradients derived from the LAsMA and ALMA data were combined for a statistical analysis, as shown in Fig.2.6.

Fig.2.5 shows the moment 1 and moment 0 maps of the ^{13}CO (3−2) line emission along the skeletons of selected filaments with intense velocity and density fluctuations along these filaments. Generally, the peaks of the density fluctuations are associated with ATLASGAL clumps, as shown in Fig. 2.4. The density and velocity fluctuations along the filaments may indicate converging gas flows coupled to regularly spaced density enhancements that probably form via gravitational instabilities (Henshaw et al., 2020). Fig. 2.5 shows the typical V-shape velocity structure around the intensity peaks, which we fit as velocity gradients on either side of the intensity peak marked by the straight lines. For the long filaments, except for the local velocity fluctuations or gradients, they also exhibit velocity gradients on larger scales, which were fit by ignoring the local velocity fluctuations. Generally, large-scale velocity gradients are associated with several intensity peaks.

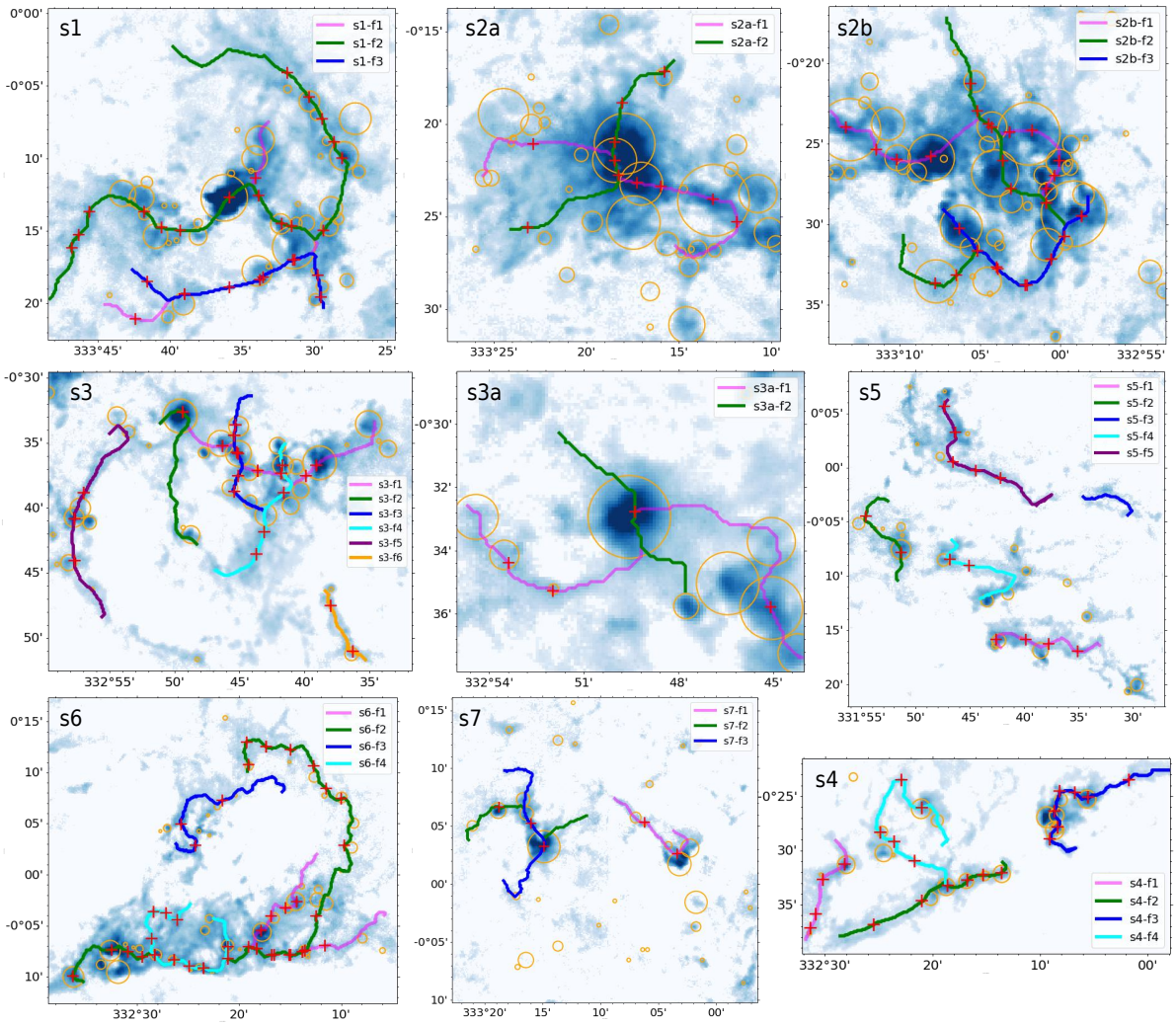


Figure 2.4: Background images showing the moment 0 maps of ^{13}CO (3-2). The lines in color present the filament skeletons. The orange circles show the ATLASGAL clumps. The size of the circle represents the clump radius. The red plus marks the intensity peak of ^{13}CO (3-2) emission in Fig. 2.5.

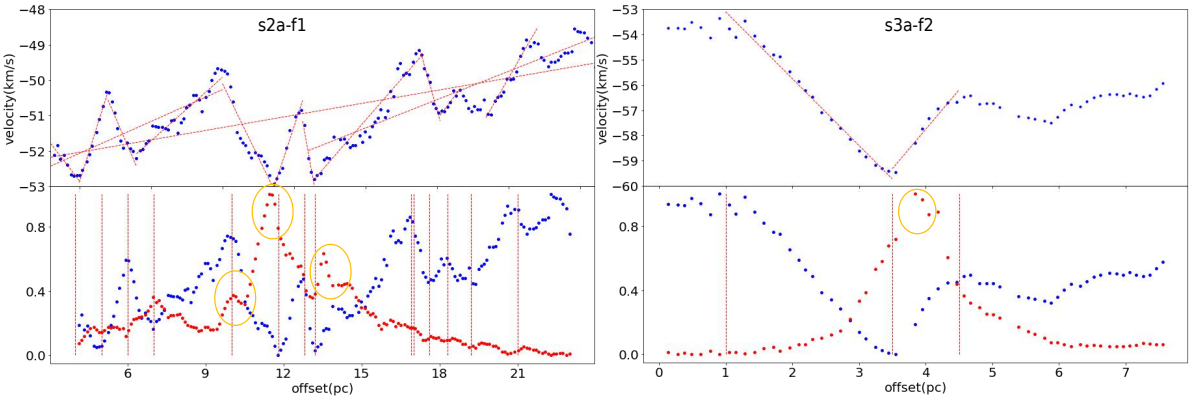


Figure 2.5: Two selected filaments illustrate the fitting of the velocity gradient. Upper panel: Velocity gradients are fit in the ranges defined by the vertical dashed red lines, and straight lines show the linear fitting results. Lower panel: Dotted blue and red lines show the normalized velocity and intensity, respectively. The orange circles mark the intensity peaks of ^{13}CO (3−2) emission associated with ATLASGAL clumps.

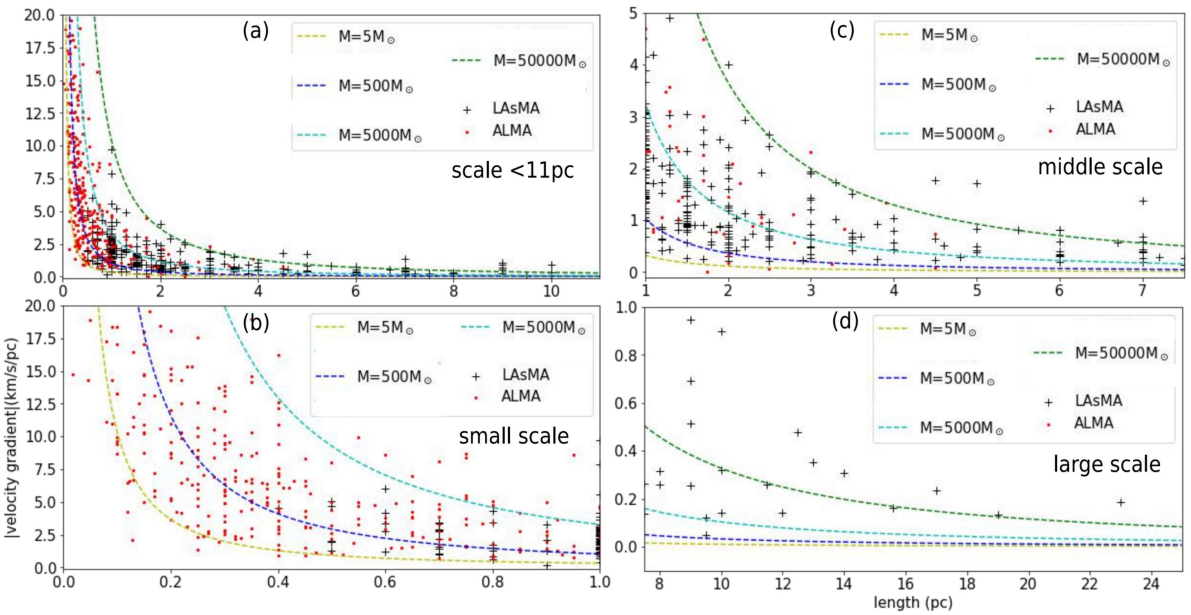


Figure 2.6: Statistical analysis of all fitted velocity gradients. (a) Velocity gradient vs. the length over which the gradient has been estimated for all sources in the ATOMS survey (red dots) and current LAsMA observation (black plus). The color lines show the freefall velocity gradients for comparison. For the freefall model, yellow, blue, cyan, and green lines denote masses of 5 $M_\odot$, 500 $M_\odot$, 5000 $M_\odot$, and 50000$M_\odot$, respectively. Panels (b), (c), and (d): Zoomed maps with lengths < 1 pc (small scale), ≈ 1 – 7.5 pc (medium scale), and > 7.5 pc (large scale) in panel (a).

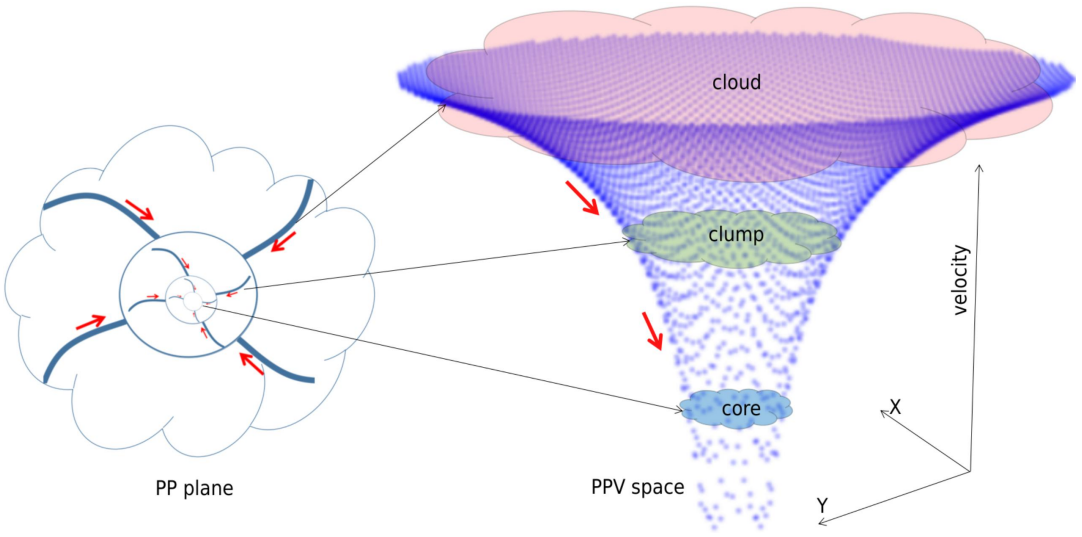


Figure 2.7: Schematic diagram of the funnel structure in PPV space and on the PP plane. The PP plane shows the multiscale hub-filament structure in a molecular cloud, which can turn into a funnel structure in PPV space. The red arrow represents the gas inflow.

2.3.2 Funnel structure

When we compare the velocity gradients of the molecular gas to iron filings, the distribution of the velocity gradients may to some extent reflect the shape of the force field that plays the dominant role in the molecular clouds, just as the iron filings can reflect the shape of a magnetic field. Below, we investigate the distribution of the molecular velocity gradients in position-position-velocity (PPV) space. In the case of gravitational freefall onto a central point mass, velocities change as $v \propto r^{-1/2}$ ($r = \sqrt{x^2 + y^2}$). In the case of a collapse that is driven by the gravity of a gaseous mass with a certain density profile, the radial dependence of the infall speed depends on the slope of the density profile (Gómez et al., 2021). In either case, this results in a funnel morphology of the velocity field in PPV space, as illustrated in Fig.2.7. The exterior cloud velocities show only gentle variations (a small velocity gradient that funnels material from the cloud to the clump), while the velocities in the interior part near the center change dramatically with scales (a large velocity gradient that funnels material from the clump to the core). We note that the schematic in Fig.2.7 is oversimplified. A molecular cloud generally contains many clumps, and a clump also contains many cores. We choose only one clump and one core to demonstrate the gas inflow and the variation in the velocity gradient from the molecular cloud scale to the core scale.

Moreover, the V-shape velocity structure around the intensity peaks indicates accelerated material inflowing toward the central hub (Gómez & Vázquez-Semadeni, 2014; Kuznetsova et al., 2015; Hacar et al., 2017; Kuznetsova et al., 2018; J.-W. Zhou et al., 2022), which is shown everywhere in Fig. 2.5. The V-shape velocity structure can be treated as the projection of the funnel structure from PPV space to the PV plane. Hence, the funnel structures in PPV space can be an effective probe for the gravitational collapse motions in molecular cloud.

2.3.3 Large-scale inflow driven by gravity

In this work, we have investigated gas motions on sufficiently large scales. The longest filament was ≈50 pc. We obtained results similar to previous small-scale ALMA observations. We fit a wide range of values of the velocity gradients, as shown in Fig. 2.6(a). The velocity gradients are small at large scales, while they become significantly larger at small scales ($\lesssim$ 1 pc), as is the case in the ATOMS survey (J.-W. Zhou et al., 2022). For both ALMA and LAsMA data, most of the fitted velocity gradients are concentrated on the scales ≈ 1 pc, as shown in Fig.2.6. This is consistent with the fact that 1 pc is considered the characteristic scale of massive clumps (Urquhart et al., 2018). In Fig.2.6, the variation in the velocity gradients with the scales is comparable with the expectations from gravitational freefall, in the sense that the gradient decreases smoothly with increasing scale. The most frequent velocity gradient around 1 pc is ≈ 1.6 km s^{-1} pc^{-1}. Assuming freefall,

$$\nabla V_{free} = -\frac{d}{dR}\sqrt{\frac{2GM}{R}} = \sqrt{\frac{GM}{2R^3}}, \tag{2.1}$$

we estimate that the kinematic mass corresponding to 1 pc is ≈ 1190 $M_\odot$, which is comparable with the typical mass of clumps in the ATLASGAL survey (Urquhart et al., 2018). Thus the clump may be a gravity-dominated collapsing object, which is also consistent with the survey results that most Galactic parsec-scale massive clumps seem to be gravitationally bound, regardless of how evolved they are (T. Liu et al., 2016b; Urquhart et al., 2018; I. Evans N. J. et al., 2021).

As shown in Fig.2.6, the velocity gradients we fit to the LAsMA and ALMA data agree with each other over the range of scales covered by ALMA observations in the ATOMS survey (< 5 pc). The variations in velocity gradients at small scales (< 1 pc), medium scales (≈ 1-7.5 pc), and large scales (> 7.5 pc) are

consistent with gravitational freefall with central masses of $\approx 500\,M_\odot$, $\approx 5000\,M_\odot$, and $\approx\ 50000\,M_\odot$. This means that the velocity gradients on larger scales require a higher mass to be maintained, which is also revealed by the funnel structure shown in Fig.2.7. Higher masses imply larger scales, that is to say, the larger-scale inflow is driven by the larger-scale structure, which may be the gravitational clustering of smaller-scale structures, consistent with the hierarchical structure in molecular clouds and the gas inflow from large to small scales. Thus, the large-scale gas inflow may also be driven by gravity, with the molecular clouds in G333 in a state of global gravitational collapse, which is consistent with the argument that these molecular clouds act as cloud-scale hub-filament structures. Moreover, the funnel structure of the velocity field shown in PPV space also tends to support the large-scale gravitational collapse of the molecular clouds in the G333 complex.

2.4 From cloud to core scale

In this work, we focus on a local region G333-s3a in the G333 complex, called G332.83-0.55 hereinafter, which is marked in Fig.1 of J. W. Zhou et al. (2023). also shown in Fig.2.8. G332.83-0.55 presents a typical hub-filament morphology at ≈10 pc. As a cloud-scale hub-filament system, its hub corresponds to the ATLASGAL clump AGAL332.826-00.549. High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation allowed us to trace the gas kinematics and gas inflows from cloud to clump scales in G332.83-0.55. We also zoomed in on the hub region of G332.83-0.55 using high-resolution ALMA HCO^+ (1-0) and $H^{13}CO^+$ (1-0) observations to investigate gas motions from clump to core scales. By combing LAsMA and ALMA observations, we could directly recover the hierarchical or multi-scale hub-filament structures and gas inflows in G332.83-0.55. In addition, we were also able to characterize the details of the funnel structure at different scales. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024f).

2.4.1 Connection of gas inflows at different scales

Moment maps of various molecular lines are shown in Fig.2.8. The moment 0 maps of ^{13}CO (3−2), $H^{13}CO^+$ J=1-0, and HCO^+ J=1-0 emission show a typical hub-filament morphology. The FILFINDER algorithm (Koch & E. W. Rosolowsky, 2015a) was used to identify and characterize filaments traced by $H^{13}CO^+$ (1-0) and ^{13}CO (3−2) emission. The skeletons of the identified filaments, overlaid on the moment 0 maps in Fig.2.9, are highly consistent with the gas structures. We selected the filaments through the redshifted and blueshifted peaks (LAsMA-f3 and ALMA-f3), and extracted the velocity and intensity along the filaments, as shown in Fig.2.10.

Fig.2.11 displays the position-velocity (PV) diagrams along the dashed lines in panels (b1) and (d1) of Fig.2.8. For $H^{13}CO^+$ J=1-0 emission, feedback from the central UC-HII region identified in C. Zhang et al. (2023) has destroyed the surrounding molecular gas structure (Fig.2.8(d)). Thus, the diagram looks broken, but it still has a clear double-peak velocity structure, indicating that the surrounding gas is still infalling onto the protocluster. In Fig.2.11(b), the central part of the PV diagram has a similar morphology to Fig.2.11(a). Large-scale gas structures with clear velocity gradients are connected to the central velocity jump. Correspondingly, large-scale filamentary structures connected to the central clump can be seen in Fig.2.8(a) and (b); small-scale filamentary structures down to dense core scales are also presented in Fig.2.8(c) and (d). These features indicate that cloud-scale gas inflows extend down to core scales. The inheritance of gas flows on large scales to small scales is clearly visible in Fig.2.10(c).

Fig.2.12(a) displays the PPV map of ^{13}CO (3−2) emission decomposed by the fully automated Gaussian decomposer `GAUSSPY+` (Lindner et al., 2015; Riener et al., 2019). We can see a funnel structure that can be a probe of gravitational collapse, as discussed in J. W. Zhou et al. (2023). At small scales, we

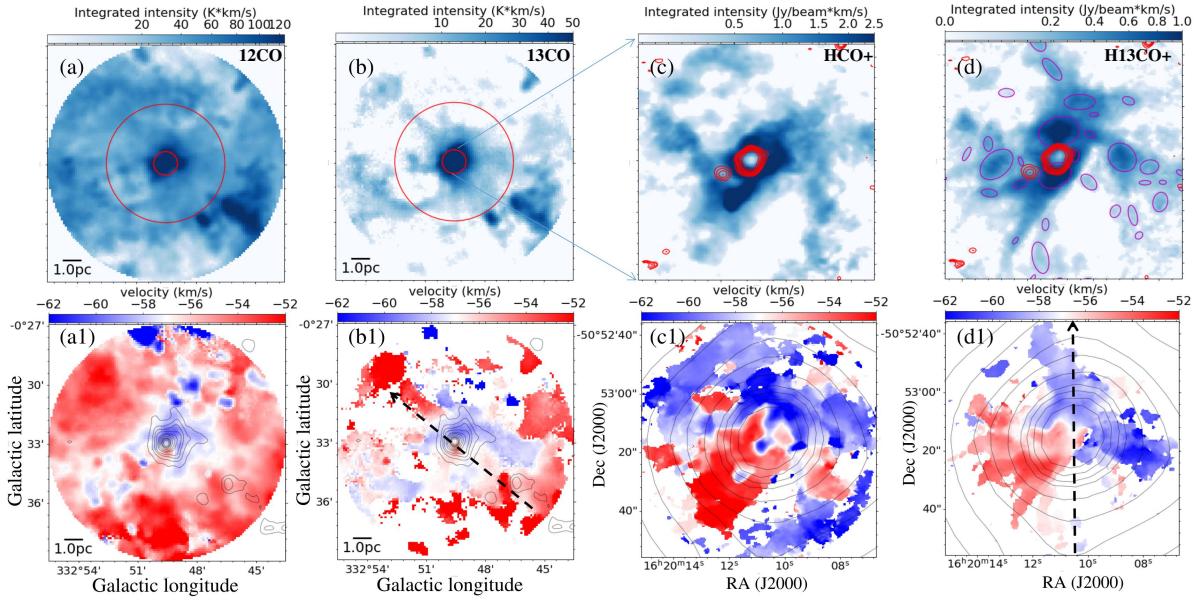


Figure 2.8: First and second columns: Moment maps of LAsMA ^{12}CO (3−2) and ^{13}CO (3−2) emission. The radius of the entire region is 0.1^{o} (≈6.28 pc). The big red circle marks the region with a 0.05^{o} (≈3.14 pc) radius, and the small red circle shows the region with a 0.01^{o} (≈0.63 pc) radius. Third and fourth columns: Moment maps of ALMA 7m+12m HCO^{+} J=1-0 and H^{13}CO^{+} J=1-0 emission. The radius of the entire region is 0.01 (≈0.63 pc). Red contours in panels (c) and (d) show the 3mm continuum emission. Black contours in the second row show ATLASGAL 870 μm continuum emission. Dashed black lines in panels (b1) and (d1) show the paths of the PV diagrams in Fig.2.11. Ellipses in panel (d) represent the identified dendrogram structures.

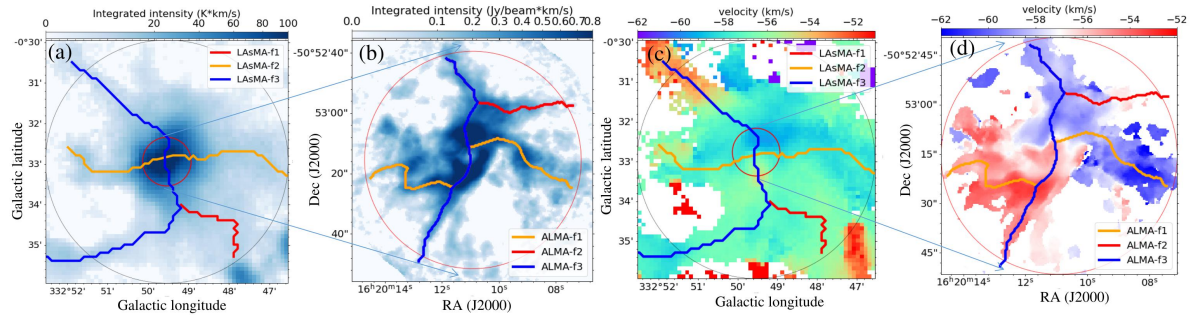


Figure 2.9: Hierarchical hub-filament structures in G332.83-0.55. (a) and (b) Moment 0 maps of LAsMA ^{13}CO (3−2) and ALMA H^{13}CO^{+} (1-0) emission. (c) and (d) Moment 1 maps of LAsMA ^{13}CO (3−2) and ALMA H^{13}CO^{+} (1-0) emission. The hub region marked by a red circle in panels (a) and (c) are enlarged in panels (b) and (d). More detailed structures can be seen in panels (b) and (d), thanks to the high resolution of ALMA observation. Lines in color present the filament skeletons identified by the FILFINDER algorithm.

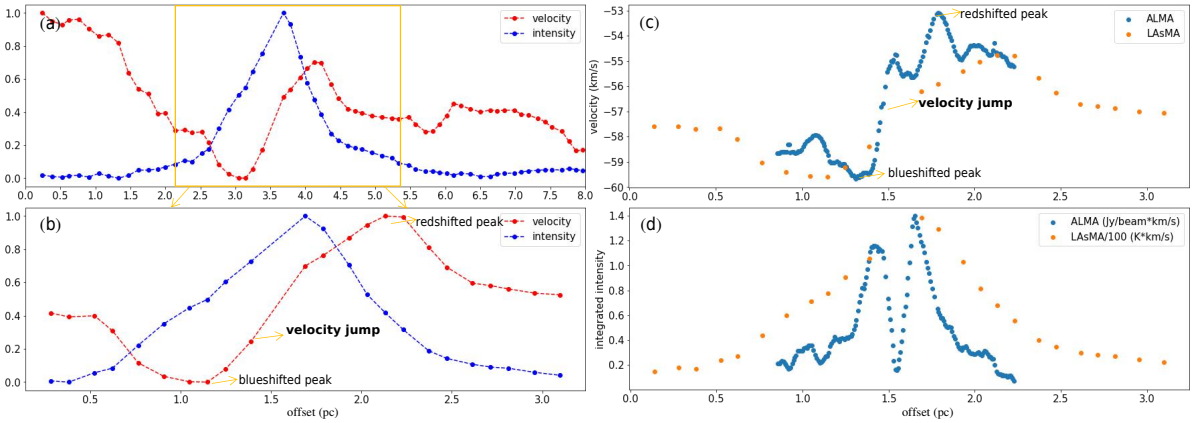


Figure 2.10: (a) Normalized velocity and intensity along the filament LAsMA-f3 in Fig.2.9(a); here we can see a blueshifted V-shape velocity structure. (b) Segment of panel (a) around the central hub; here we can see a V+Λ-shape velocity structure. (c) and (d) Velocity and intensity along the segment of LAsMA-f3 in panel (b) compared to the velocity and intensity along the filament ALMA-f3 in Fig.2.9(b).

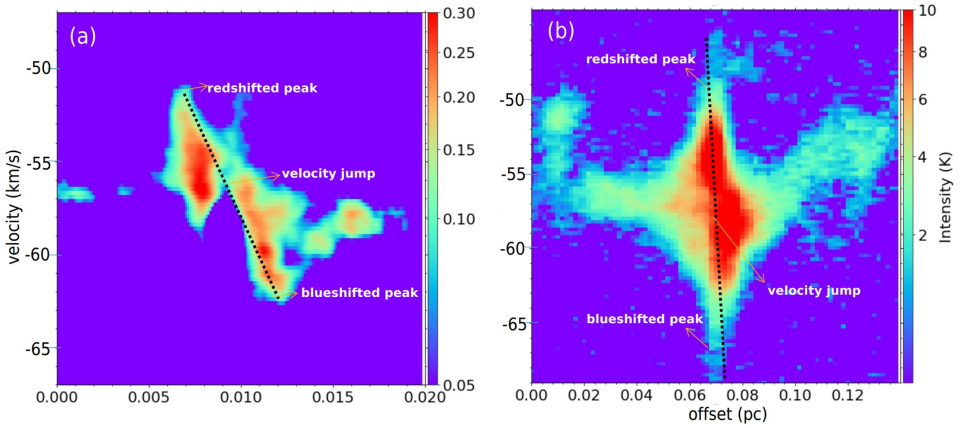


Figure 2.11: PV diagrams of ALMA $H^{13}CO^+$ (1-0) and LAsMA ^{13}CO (3−2) along the paths marked in Fig.2.8 (d1) and (b1).

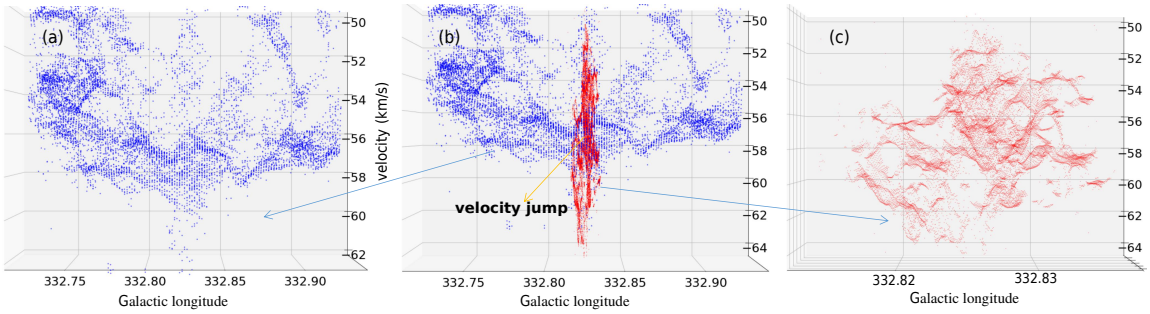


Figure 2.12: Linkage of large- and small-scale velocity fields in PPV space. (a) Velocity field revealed by LAsMA data ^{13}CO (3−2) on large scales decomposed by `GAUSSPY+`. (b) Combined velocity fields traced by ^{13}CO (3−2) and $H^{13}CO^+$ (1-0) at different scales, i.e., panel (a) plus panel (c). (c) Velocity field revealed by ALMA data $H^{13}CO^+$ (1-0) on small scales decomposed by `GAUSSPY+`.

can roughly identify a velocity jump comprised of the redshifted and blueshifted peaks in Fig.2.9(c), which is clearly shown in Fig.2.9(d), thanks to the higher resolution of the ALMA data.

To conclude, G332.83-0.55 shows at least two levels of hub-filament structures. Fig.2.8(a) and (b) present the first level from cloud to clump scales. The second level from clump to core scales is shown in Fig.2.8(c) and (d). In a statistical study of hub-filament systems in J.-W. Zhou et al. (2022), we found filaments are ubiquitous in proto-clusters, and hub-filament systems are very common from dense core scales (≈0.1 pc) to clump and cloud scales (≈1-10 pc). These observed features are reproduced in G332.83-0.55 well.

The structures marked by small red circles in Fig.2.8(a) and (b) present an approximately spherical morphology due to the limited resolution in single-dish observation, usually referred to as a "clump". However, in the eyes of ALMA, the spherical clump displays anisotropic and filamentary accretion structures (hub-filament) in Fig.2.8(c) and (d). Thus, relative to the traditional "infall envelope", filamentary accretion and inflow structures may be more realistic in high-mass star-forming clumps.

2.4.2 Infall+rotation motions

In Fig.2.8(c) and (d), on small scales, $H^{13}CO^+$ J=1-0 and HCO^+ J=1-0 emission present spiral-like structures, indicating the rotation of gas around the protocluster. PV diagrams in Fig.2.11 also reveal rotation motions. The PV diagrams show two peaks separated by several km/s, and they are symmetrically located with respect to the gravitational center. The double-peaked feature on the PV map is often seen in cluster-forming clumps, and this feature can only be reproduced when the clump has both infall motion and rotation at the same time, thus strongly implying that the clump is collapsing with rotation (Ohashi et al., 1997; Shimoikura et al., 2016; Shimoikura et al., 2018; Sanna et al., 2019; S.-Y. Liu et al., 2020; Shimoikura et al., 2022). Around the gravitational center, we can see the blueshifted and redshifted extrema at the end of the velocity jump in Fig.2.10, which is comparable with Fig.2 of Sakai et al. (2014). In the schematic diagram of an infalling and rotating envelope model for IRAS 04368+2557 (Sakai et al., 2014), the shift of blueshifted and redshifted peaks can be used to estimate the radius of the centrifugal barrier (r_c), where the rotation velocity ($V_{rotation}$) has its maximum value. According to equations (1) and (2) in Sakai et al. (2014), we can estimate the mass of the gravitational center according to r_c by

$$M_* = \frac{V_{rotation}^2 r_c}{2G}. \tag{2.2}$$

According to the velocity difference and the separation between blueshifted and redshifted peaks in Fig.2.10(c), we roughly estimate $V_{rotation,ALMA}$ ≈3.29 km/s and $r_{c,ALMA}$ ≈ 0.21 pc. Then the mass of the gravitational center (i.e., the protocluster) is ≈264 $M_\odot$.

As a further test, following the method described in Kauffmann et al. (2008), the mass can be computed by dust emission. A mass ≈264 $M_\odot$ corresponds to the flux of 3mm dust emission ≈0.12 Jy. This value is ≈3% of the total flux of the 3mm continuum emission in ALMA observations. As discussed in E. Keto et al. (2008) and C. Zhang et al. (2023), 3mm continuum emission of UC HII regions in high-resolution interferometric observations should be mostly free–free emission rather than from dust, and thus the proportion of ≈3% for the dust emission is reasonable.

2.4.3 Gravitational focusing

In Fig.2.11(b) and Fig.2.12(b), on large scales, the funnel structure is consistent with the schematic diagram in Fig.9 of J. W. Zhou et al. (2023). However, it looks as though the bottom of the funnel splits into two parts in Fig.2.11(b) and Fig.2.12(b). A velocity jump appears in the center comprised of the

redshifted and blueshifted components, which may be attributed to the infalling and rotating motions around the protocluster. If we turn over the redshifted component and let it merge with the blueshifted component, then we obtain the funnel structure in Fig.9 of J. W. Zhou et al. (2023).

Similar to lights converging to the focus, now the gas inflows are converging to the gravitational center, that is, the hub region of G332.83-0.55. Thus, there is a vivid "gravitational focusing" process, which also provides a clue as to the formation of the hub-filament structure. As shown in the Fig.1 of J. W. Zhou et al. (2023), the entire G333 complex has a large-scale shell structure. G332.83-0.55 is a fragment on the shell, which should have a sheet configuration initially. Then the subsequent gravitational focusing process yields the hub-filament structure. In Fig.2.8(b), the hub-filament structure is decoupling from the surrounding gas environment due to gravitational collapse.

2.5 On galaxy scale

In previous works, we have investigated gas dynamics and kinematics in molecular clouds from cloud to core scales. The main task of this work is to generalize the above physical pictures to galaxy-cloud scales, now the smallest unit is the molecular cloud itself. In J.-W. Zhou et al. (2022), dense core is the hub in clump. In J. W. Zhou et al. (2023), clump is the hub in molecular cloud. Here we treat molecular cloud as the hub in galaxy. Self-similar or hierarchical/multi-scale hub-filament structures and filamentary accretion feeding the central regions exist from molecular cloud to dense core scales (Motte et al., 2018b; Vázquez-Semadeni et al., 2019; Kumar et al., 2020; J.-W. Zhou et al., 2022; J. W. Zhou et al., 2023), this picture will be extended to galaxy-cloud scales in this work. The work presented in this section is reproduced from the publication by J.-W. Zhou & T. Davis (2024).

2.5.1 Velocity component

In this work, NGC5236 and NGC4321 were selected for the following reasons: (1) They are face-on galaxies. The inclination angles of NGC5236 and NGC4321 are 24^{o} and 38.5^{o} (Lang et al., 2020), respectively. Thus, we can ignore the projection effect in velocity gradient fitting; (2) They have strong CO (2−1) emission, presenting large-scale continuous structures (strong spiral arms), thus we can trace galaxy-scale gas motions; (3) NGC5236 has similar morphology to NGC4321, but it is about 3 times closer than NGC4321, thus we can resolve more detailed structures.

Before studying gas kinematics, we need to ascertain the distribution of gas components in the galaxy. We utilized the fully automated Gaussian decomposer `GAUSSPY+` (Lindner et al., 2015; Riener et al., 2019), designed to break down intricate spectra of molecular lines into multiple Gaussian components. For the CO (2−1) data cube of NGC4321, 50071 spectra were fitted by `GAUSSPY+`, 98% spectra are single component, 2% spectra have two velocity components. 323059 spectra were fitted for NGC5236, 87% spectra show single component, 11.5% spectra have two components. The reduced χ^2 values output in the fitting of NGC4321 and NGC5236 are around 1, which means that both of them get a good fit. Generally, multiple velocity components (>1 velocity component) mainly concentrate on the central region of the galaxy. In this work, we mainly focus on the gas kinematics on the spiral arms. Therefore, the analysis can be carried out based on the moment maps.

In Fig.2.13, we can see ubiquitous velocity fluctuations, which may be attributed to gravitational instability, as discussed below. In the central region of the galaxy, there is almost no velocity fluctuations, clear velocity fluctuations are mainly distributed on the spiral arms, revealing the interconnection between velocity fluctuations and star formation activities.

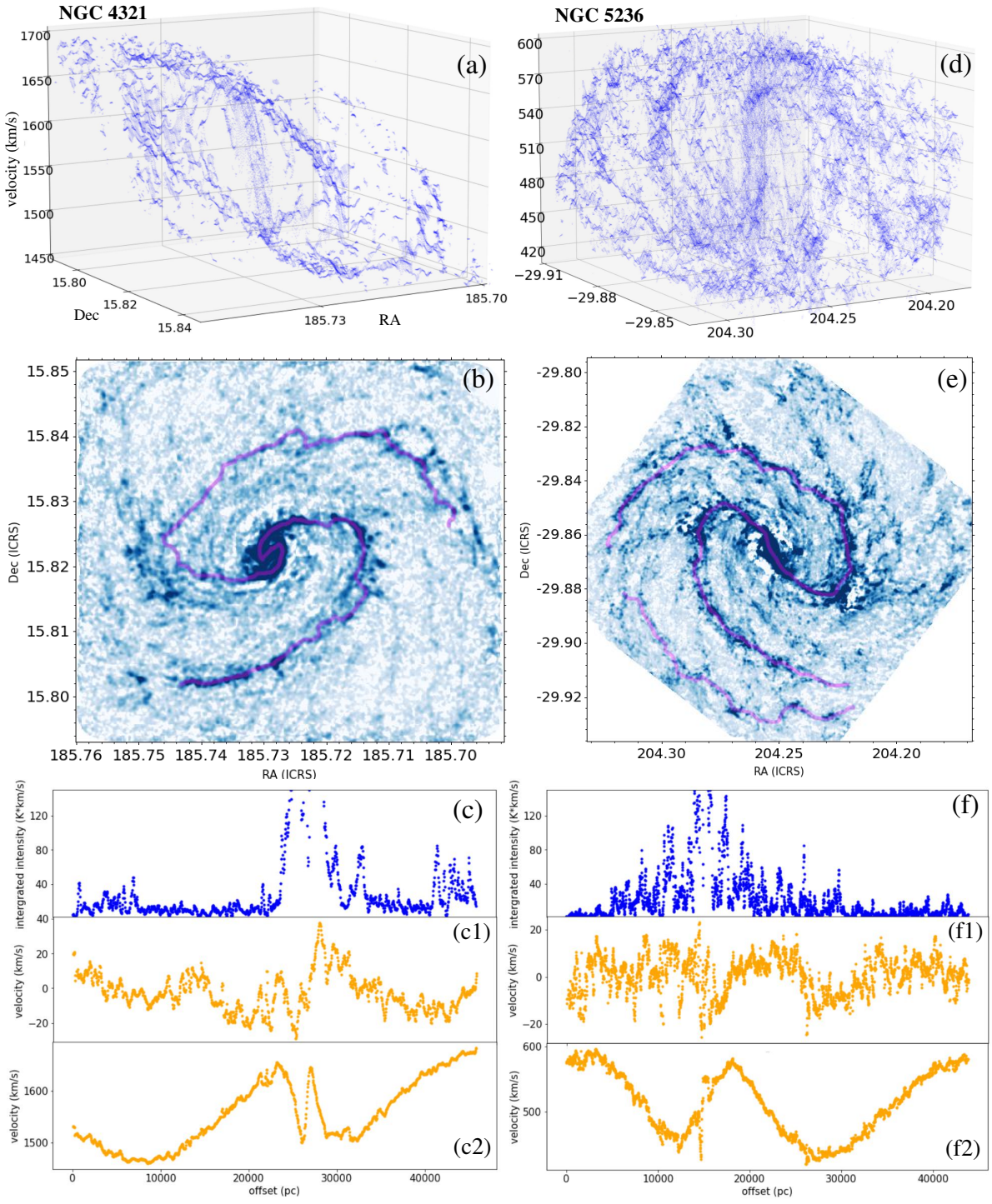


Figure 2.13: An overview of gas kinematics in NGC4321 and NGC5236. (a) and (d) The velocity distribution of the entire galaxies in PPV space decomposed by GAUSSPY+; (b) and (e) The main filaments of NGC4321 and NGC5236 identified by FILFINDER algorithm overlaid on the Moment 0 maps; (c) and (f) The integrated intensity distributions extracted from the moment 0 maps along the main filaments in panels (b) and (e); (c2) and (f2) The velocity distributions extracted from the moment 1 maps along the main filaments in panels (b) and (e); (c1) and (f1) The velocity residual distributions along the main filaments in panels (b) and (e) after subtracting the large-scale velocity gradients due to the galaxy rotation.

2.5.2 Scaling relations

We conducted a direct identification of hierarchical (sub-)structures based on the 2D integrated intensity map of CO (2−1) emission by the dendrogram algorithm. As described in E. W. Rosolowsky et al. (2008), the dendrogram algorithm decomposes intensity data into hierarchical structures called leaves and branches (see Sec.4.1.1 for more details). Subsequently, we extracted the average spectrum of each identified structure to delve into its velocity components and gas kinematics. The large-scale velocity gradients due to the galaxy rotation will contribute to the non-thermal velocity dispersion. Before extracting the average spectra of the identified structures, we subtracted the large-scale velocity gradients in PPV data cube based on the constructed gas dynamical model. According to their averaged spectra, the structures with absorption features were eliminated firstly. Following J. W. Zhou et al. (2024e), we only consider type1 (single velocity component) and type3 (blended velocity components) structures in this work.

The scaling relations provide insights into the physical states of the structures. Although the identified structures in NGC5236 are much more than that in NGC4321, the scaling relations in Fig.2.14 are comparable. Due to the greater distance, some faint structures were filtered out in NGC4321, which can be seen in NGC5236. In Fig.2.14, the faint or low-density structures produce dispersive tails, similar to the Fig.14 of J. W. Zhou et al. (2024g) for the G333 molecular cloud complex in the Milky Way.

In Fig.2.14(g)-(i), it appears that only branch structures demonstrate a discernible correlation between velocity dispersion and scale. Clumps or cores characterized by high column density exhibit greater velocity dispersion relative to those with lower column density, a phenomenon attributed to gas motions associated with gravitational collapse (Ballesteros-Paredes et al., 2011b; Traficante et al., 2018a; Ballesteros-Paredes et al., 2018; S. Li et al., 2023; J. W. Zhou et al., 2024g). As depicted in Fig.2.14(d)-(f), the positive correlations observed between velocity dispersion and column density suggest a gravitational origin for the velocity dispersion. In the context of pure free-fall, an anticipated $\sigma - N * R$ relation would exhibit a slope of 0.5. However, in Fig.2.14(a)-(c), the slopes of the $\sigma - N * R$ relations are notably shallower than 0.5, suggesting a deceleration from the expected behavior of pure free-fall gravitational collapse.

2.5.3 Velocity gradient

The gravitational collapse of the structures has been revealed by the scaling relations. This section gives a more detailed discussion of the gravitational collapse from the perspective of the velocity gradient. The FILFINDER algorithm (Koch & E. W. Rosolowsky, 2015a) was used to identify and characterize filaments in the galaxy. Following the method described in J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023), we used the moment 0 maps of CO (2−1) emission from the data products of the PHANGS-ALMA survey to identify filaments. Fig.2.13(b) and (e) display the skeletons of the identified filaments superimposed on the moment 0 maps. In this work, we only consider the filaments along the spiral arms (the main filaments), where have strong CO (2−1) emission, which allows us to study large-scale continuous gas motions in the galaxy.

In J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023), the filaments traverse multiple local hub-filament structures, characterized by fluctuations in velocity and density along the filaments. These local dense structures, acting as gravitational centers, facilitate the accretion of surrounding diffuse gas, ultimately leading to the formation of local hub-filament structures. In the PPV space, a hub-filament structure can be represented as a funnel structure, as demonstrated in Fig.9 of J. W. Zhou et al. (2023). The gradient of the funnel profile serves as an indicator of the gravitational field's strength. Consequently, examining the local velocity gradients along the filaments provides insights into the strength of local

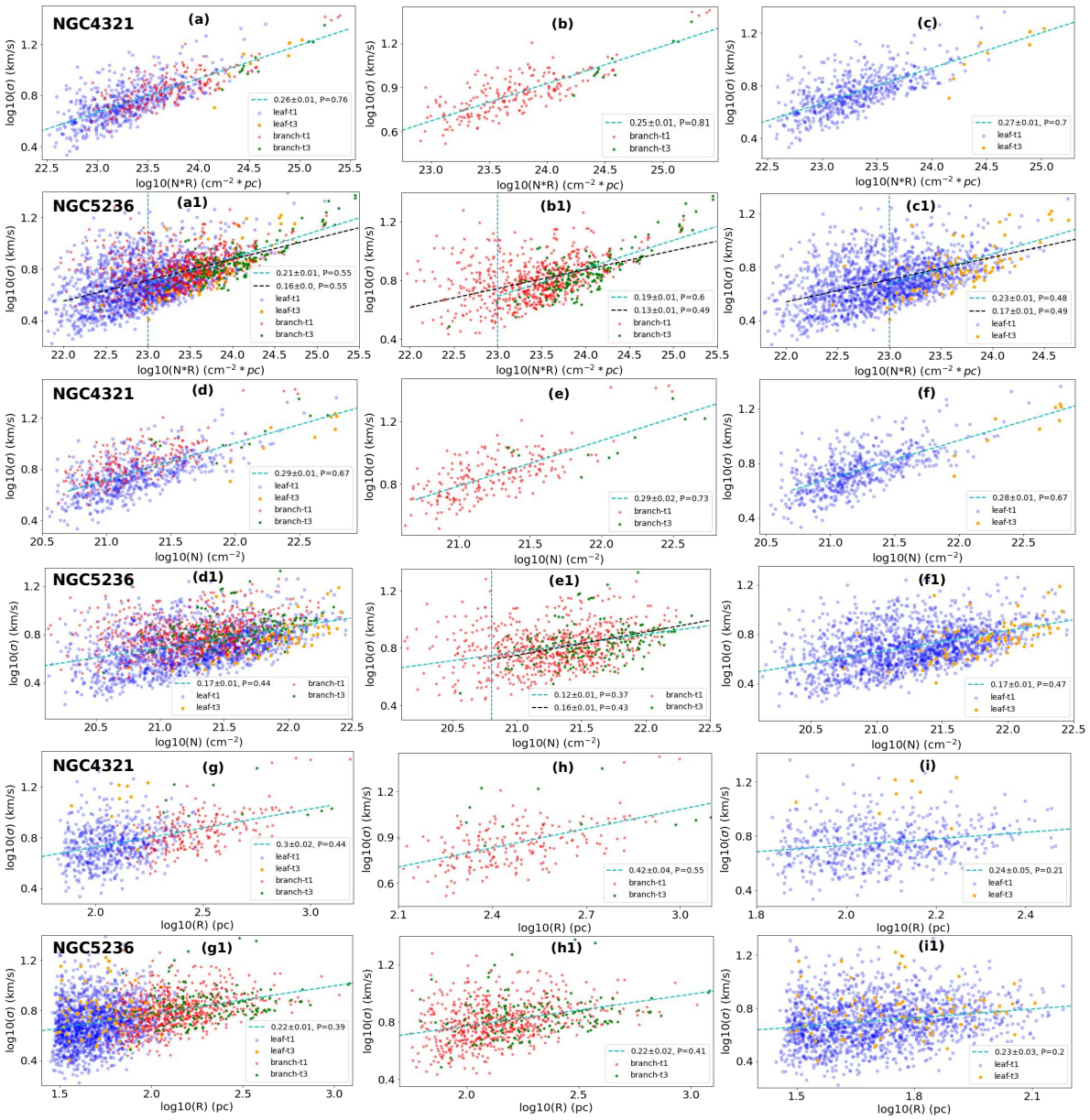


Figure 2.14: Scaling relations of leaf and branch structures of NGC 4321 and NGC 5236. (a)-(c) and (a1)-(c1) $\sigma - N * R$; (d)-(f) and (d1)-(f1) $\sigma - N$; (g)-(i) and (g1)-(i1) $\sigma - R$. σ, R and N are the velocity dispersion, effective radius, and column density of each structure, respectively. Here, "P" represents the Pearson coefficient. Dashed vertical line marks the boundary of two fittings.

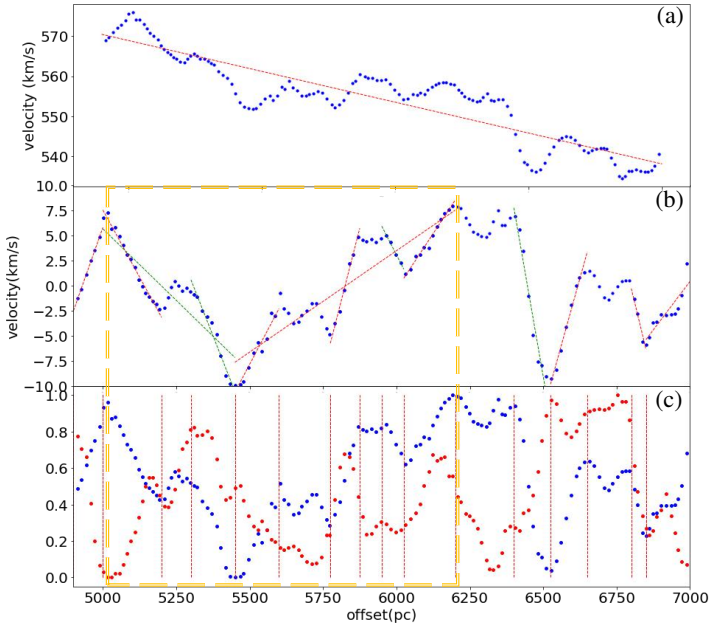


Figure 2.15: A segment of NGC5236's main filament, used to demonstrate the velocity gradient fitting. (a) Linear fitting of the large-scale velocity gradient due to the bulk motion; (b) Velocity field of the segment in panel (a) after subtracting the fitted large-scale velocity gradient. Local velocity gradients are fitted in the ranges defined by the red vertical dashed lines in panel (c), and straight lines show the linear fitting results. The color-coding of straight lines are meaningless, but it is convenient for the reader; (c) Blue and red dotted lines show the normalized velocity and intensity, respectively. Orange box marks a gravitational coupling structure.

gravitational fields. In Fig.2.13, although the velocity fluctuations at small scales can be seen everywhere, they are hidden in global large-scale velocity gradients due to the galaxy rotation. To derive the local velocity fluctuations, we must first subtract the large-scale velocity gradients. We take two methods to solve this issue.

Fig.2.13 (c2)&(c) and (f2)&(f) show the intensity-weighted velocity and integrated intensity of CO (2−1) line emission along the main filaments with intense velocity and density fluctuations. A segment of NGC5236's main filament was used to demonstrate the velocity gradient fitting in Fig.2.15. Here, we eliminated the large-scale velocity gradient by employing a simple linear function for modeling. Subsequently, we subtracted this model from the velocity field obtained from the moment 1 map along the main filament. As a comparison, we also removed the bulk motion or the galaxy rotation by creating gas dynamical models using the Kinematic Molecular Simulation (KinMS) package of T. A. Davis et al. (2013) based on the moment 1 map. Then we used the moment 1 map to subtract the created model, and obtained the velocity residual map. Finally, we extracted the velocity fluctuations from the velocity residual map rather than the moment 1 map.

As shown in Fig.2.16, two methods give consistent fitting results. In Fig.2.17, we can find the same gas kinematic modes presented in J. W. Zhou et al. (2023) and J. W. Zhou et al. (2024e). The variations in velocity gradients at both small and large scales (with a boundary around 500 pc) are roughly comparable with expectations from gravitational free-fall, with central masses of $\approx 10^5$–$10^{6.5}$ $M_\odot$ and $\approx 10^6$–$10^{7.5}$ $M_\odot$. This implies that sustaining velocity gradients on larger scales requires correspondingly larger masses, and larger masses imply larger scales, suggesting that the larger-scale inflow is driven by the larger-scale structure which may arise from the gravitational clustering of smaller-scale structures, in harmony with the presence of hierarchical or multi-scale hub-filament structures within the galaxy and the gas inflows from large to small scales. In the orange box marked in Fig.2.15(b) and (c), multiple

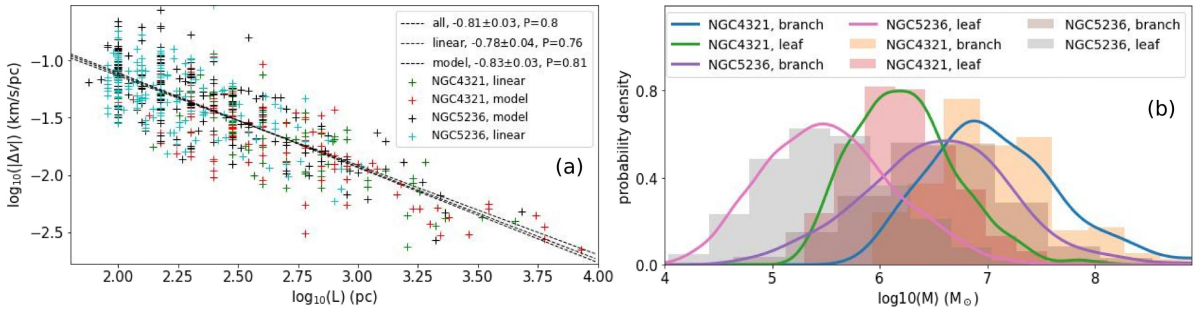


Figure 2.16: (a) The correlation between the fitted velocity gradient and the scale; (b) The mass distribution of the identified structures.

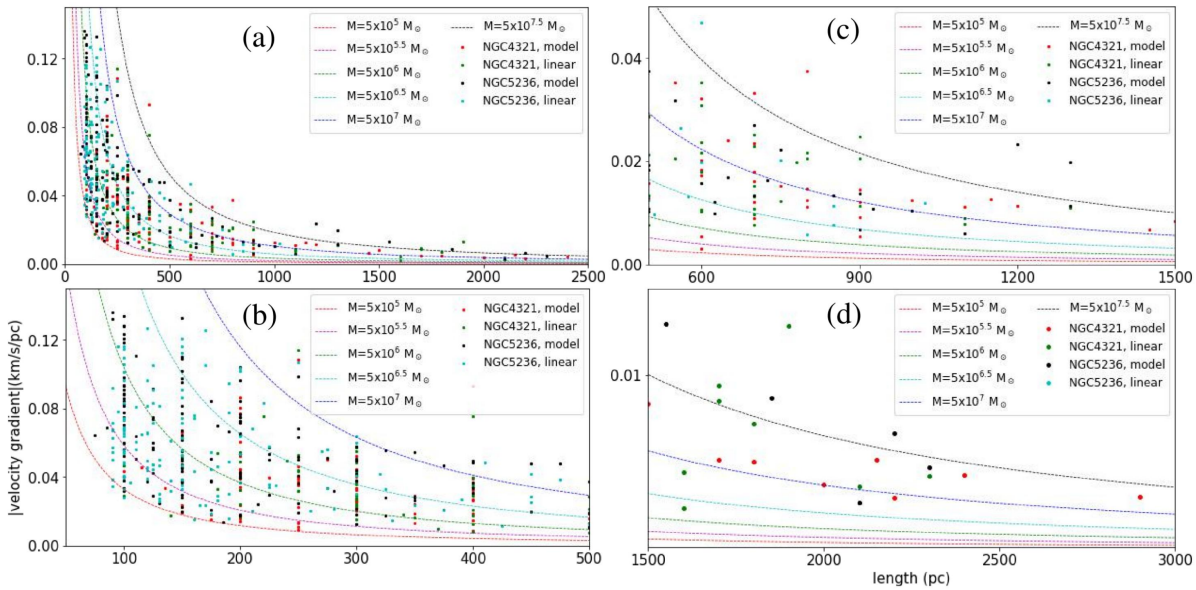


Figure 2.17: Statistical analysis of all fitted velocity gradients. (a) Velocity gradient vs. the length. The color lines show the freefall velocity gradients for comparison. For the freefall model, red, magenta, green, cyan, blue and black lines denote masses of 10^5 $M_\odot$, $10^{5.5}$ $M_\odot$, 10^6 $M_\odot$, $10^{6.5}$ $M_\odot$, 10^7 $M_\odot$ and $10^{7.5}$ $M_\odot$, respectively. Panels (b), (c), and (d): Zoomed maps with lengths < 500 pc (small scale), $\approx 500 - 1500$ pc (medium scale), and > 1500 pc (large scale) in panel (a).

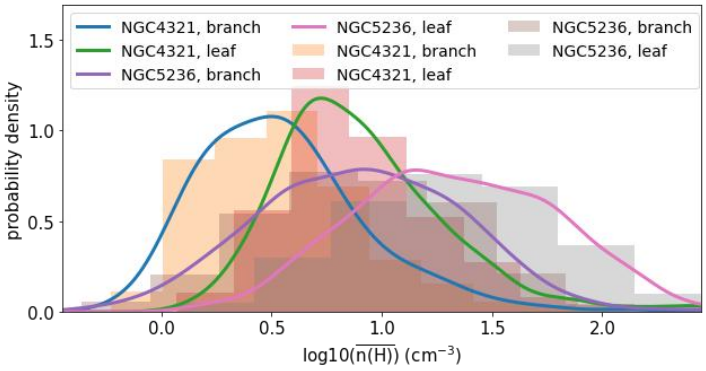


Figure 2.18: The average number density distribution of the identified clouds.

peaks are coupled together to form a gravitational potential well on larger scale, and each peak itself is also a local gravitational potential well.

In Fig.2.16(a), almost all measured velocity gradients can be fitted in the mass range $\approx 10^5$–10^7 $M_\odot$, which is consistent with the mass distribution of the identified structures shown in Fig.2.16(b), indicating that local dense structures and their complex as gravitational centers will accrete the surrounding diffuse gas and then produce the observed velocity gradients at different scales.

2.5.4 Deviation of the free-fall model

We can do a deeper analysis for the fitted velocity gradients with a simple model. Assuming free-fall, in equation.2.1, velocity gradient is more sensitive to scale. Thus we only fit the correlation between velocity gradient and scale in Fig.2.16(a). In equation.2.1, $\nabla v \propto R^{-1.5}$, however, the linear fitting in Fig.2.16(a) gives $\nabla v \propto R^{-0.8}$, thus ≈2 times smaller than the slope of the free-fall model, also indicating the slowing down of a pure free-fall gravitational collapse.

Gas kinematics on galaxy-cloud scales clearly deviate from the free-fall model. The deviation may come from measurement or calculation biases. We only measured the projection of the realistic velocity vector and hence acceleration, thus tends to underestimate the observed acceleration. Moreover, we only considered the molecular line CO (2−1), which only traces part of the baryonic matter in the galaxy, thus underestimated the mass of the gravitational center. The shallower slope in Fig.2.16(a) may have physical correlation with the flat rotation curve of the galaxy. Gas motions on galaxy-cloud scales may still couple with the galactic potential, a comprehensive model should account for the interplay between motions within the galactic potential and the self-gravitational potential of the cloud (Dobbs & Pringle, 2013; Meidt et al., 2018; Meidt et al., 2020; Utreras et al., 2020). In the model proposed by Meidt et al. (2020), the transition to free-fall collapse can only happen once the gas has completely decoupled from the galactic potential. However, it is not clear down to which spatial scales gas motions remain dynamically relevant to galactic potential or start decoupling with galactic potential.

Tidal forces have been advocated in previous works to restrict or trigger star formation. Previous studies by Ballesteros-Paredes et al. (2009b) and Ballesteros-Paredes et al. (2009a) have shown the impact of tidal forces arising from an effective Galactic potential on molecular clouds. These forces have the potential to either compress or disrupt molecular clouds, influencing the overall star formation efficiency. Thilliez et al. (2014) examined the stability of molecular clouds within the Large Magellanic Cloud (LMC) against shear and the galactic tide. However, their findings indicate that star formation in the LMC is not impeded by either tidal or shear instability. Moreover, in Ramırez-Galeano et al. (2022), tidal stresses from neighbouring molecular cloud complexes may increase interstellar turbulence, rather

than the galactic potential.

Gravity is a long-range force. A local dense structure evolves under its self-gravity, but as a gravitational center, its influence can also affect neighboring structures. At the same time, it also experiences the gravitational pull from other nearby structures. The tidal and gravitational fields are mutually interdependent. As described above, the hierarchical/multi-scale gravitational coupling of gas structures means the extensive tidal interactions between them. Whether the galactic potential has effect on molecular clouds and their complexes or not, molecular clouds should be affected by the cumulative tidal interactions of nearby material, which may inhibit gravitational collapse, the growth of instabilities, or star formation within the cloud. The influence of tidal effects on the physical properties of gas structures is discussed in Chapter.3.

In addition to the factors discussed above, magnetic fields may have a significant impact on the gas kinematics of molecular clouds as suggested by simulations (J.-G. Kim et al., 2021; Seifried et al., 2020; Ibáñez-Mejıa et al., 2022; Ganguly et al., 2023) and observations (Crutcher et al., 2010; H.-B. Li & Henning, 2011; Crutcher, 2012; Stephens et al., 2022; Ngoc et al., 2023; Rawat et al., 2024). Especially, in simulations, the diffuse gas with a number density less than 100 cm^{-3} in the envelopes of molecular clouds may be upheld against gravitational collapse due to magnetic support (Ibáñez-Mejıa et al., 2022; Ganguly et al., 2023). As shown in Fig.2.18, most of structures have the average number density $\overline{n(H)}$ < 100 cm^{-3}. The supportive role of the magnetic field may be a significant factor contributing to the deviation from free-fall motions, and further detailed investigation is warranted in future research.

2.6 Decomposition of molecular clouds

Kinematically, the results in J. W. Zhou et al. (2023) have revealed the multi-scale hub-filament structures in the G333 complex. Numerous structures within the G333 complex exhibit the characteristic kinematic properties of hub-filament systems. For instance, the intensity peaks, which represent the hubs, are linked to converging velocities, indicating that the surrounding gas is flowing towards these dense regions. Based on the evidence from the kinematics, the main goal of this work is to directly recover the multi-scale hub-filament structures in the G333 complex and the G331 giant molecular cloud (GMC) using the high-resolution ^{13}CO (3−2) data in LAsMA observation.

2.6.1 Classification

The most likely hub candidates are the intensity peaks. Following the procedure described in J. W. Zhou et al. (2024g), we directly identify the intensity peaks based on the integrated intensity map of ^{13}CO (3−2) emission by the dendrogram algorithm. The dendrogram algorithm decomposes intensity data into hierarchical structures called leaves and branches. For hub-filament systems, hubs are essentially structures with higher density compared to the surrounding diffuse gas. As local gravitational centers, these hubs can accrete the surrounding diffuse gas, forming hub-filament structures. The filamentary structures are actually gas inflows converging towards the hubs. A hub-filament structure essentially consists of a gravitational center and gas inflows converging towards it. Therefore, all identified intensity peaks (i.e. leaf structures) are thought to be potential hubs. Around the intensity peaks, we extended the spatial ranges to investigate the filamentary structures connected with them. As shown in the first row of Fig.2.19, extending to 2.5 times the hub size (the effective radius of the corresponding leaf structure) can well recover a relatively complete hub-filament morphology. However, the hub-filament structures are identified directly according to the integrated intensity map of the full velocity range. Therefore, the overlapping of different velocity components is inevitable, as shown in Fig.3 of J. W. Zhou et al. (2023).

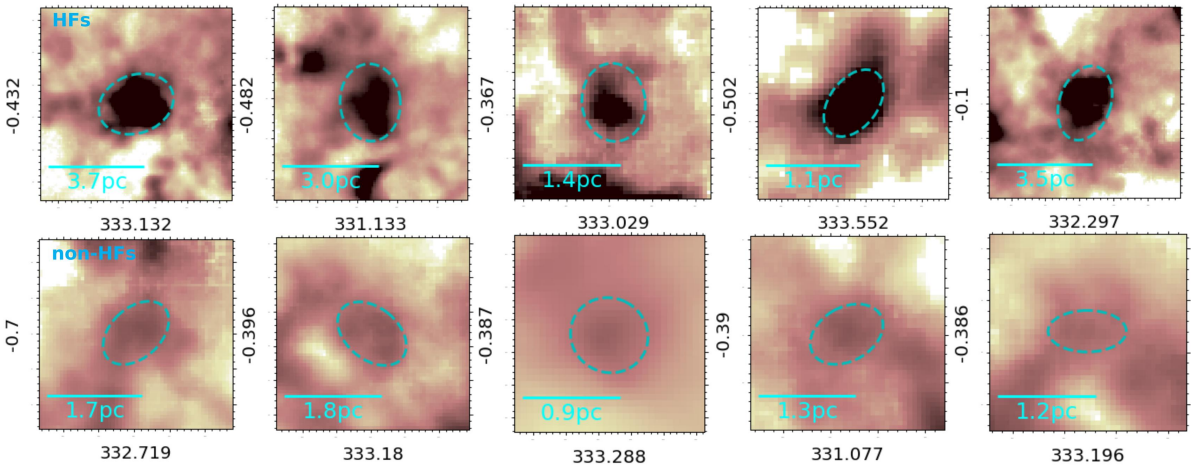


Figure 2.19: Some examples of HFs and non-HFs. The dashed ellipses represent the equivalent ellipses of the identified leaf structures by the dendrogram.

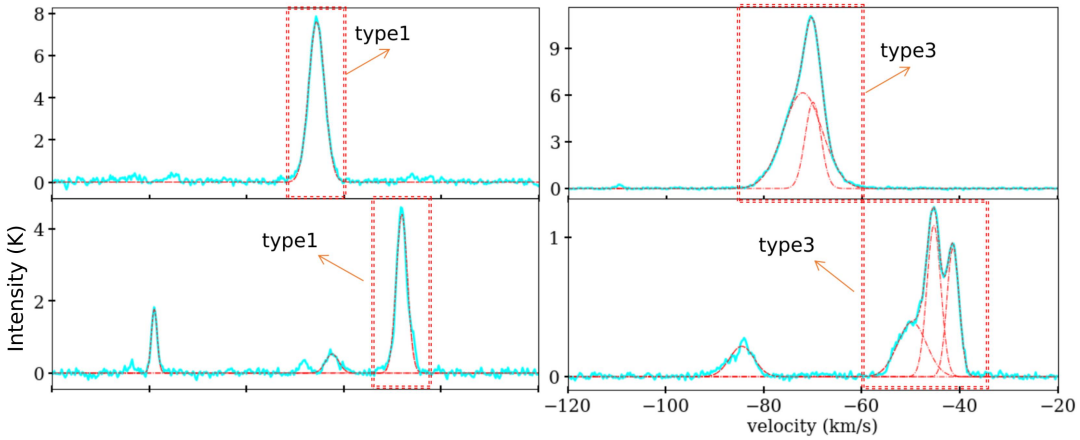


Figure 2.20: Typical ^{13}CO (3−2) line profiles of the identified structures. Red dashed box marks the limited velocity range of each structure.

To address this issue, we extracted the average spectra for all structures in the 2.5 times extended regions to decompose their velocity components.

The velocity components within the same structure that exhibit large velocity differences are generally independent of each other, which can be attributed to projection effects. Thus, for each structure, we mainly focus on the main emission marked by the red dashed box in Fig.2.20. As shown in Fig.2.20(a) and (c), the main emission can be one velocity component or multiple blended velocity components. Using the names in J. W. Zhou et al. (2024g), they are called type1 and type3 structures. For type2 structures with separated velocity components (such as Fig.2.20(b)), we only consider the dominant component, thus they are reduced to type1 structures. In this work, we consider type3 structures to be independent entities, similar to type1 structures, but characterized by more complex gas motions.

Restricted to the fixed velocity ranges of all structures (2.5 times extended) illustrated in Fig.2.20, their moment 0 maps were reproduced. In hub-filament systems (HFs), hubs exhibit substantially higher densities compared to filaments. Based on this, we categorized the structures into two groups: those with a distinct high-density central region and those without. The two categories comprise 195 and 243 structures (non-HFs), respectively. For the 195 structures featuring a distinct high-density center, we further identified the associated filaments. Structures in which filaments converge at the

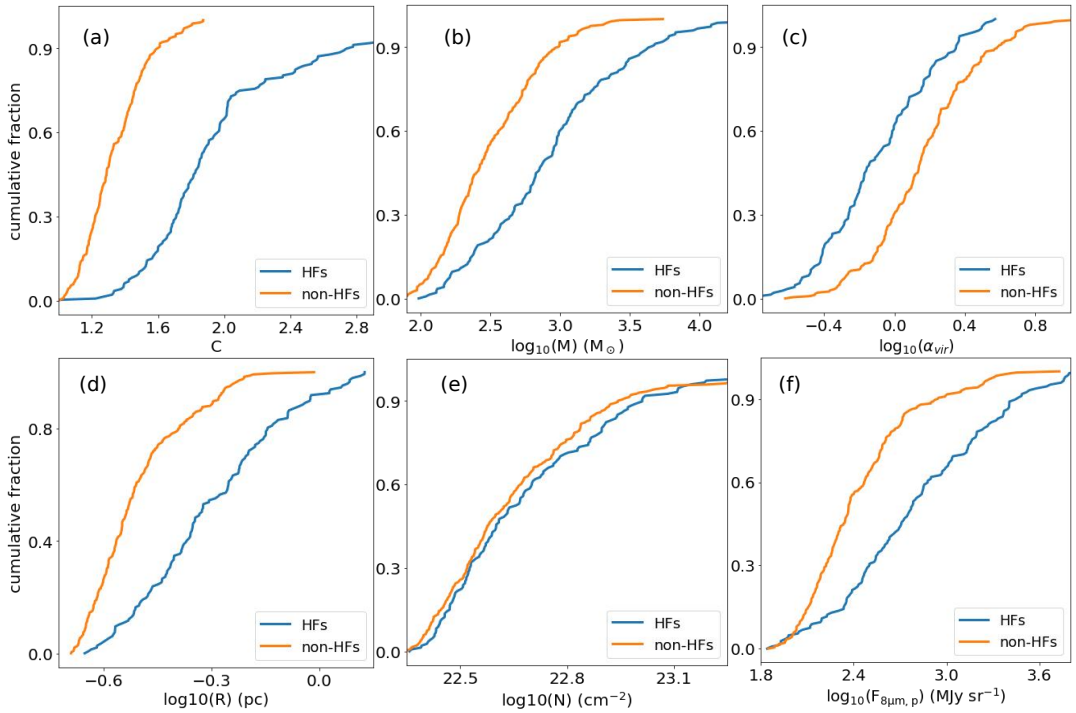


Figure 2.21: Physical properties of HFs and non-HFs. (a) Column density contrast; (b) Mass; (c) Virial ratio; (d) Effective radius; (e) Average column density; (f) Peak intensity of *Spitzer* 8 μm emission.

high-density center were classified as hub-filament systems, totaling 148 structures (HFs). A small number of remaining structures, which were difficult to classify, were excluded to maintain the clarity of comparison in the subsequent analysis. Fig.2.19 illustrates examples of non-HFs and HFs. In Fig.2.21, we present a comparison of the physical properties of HFs and non-HFs, revealing significant differences between the two groups.

2.6.2 Velocity gradient

As done in J. W. Zhou et al. (2023), we also use the FILFINDER algorithm (Koch & E. W. Rosolowsky, 2015b) to characterize the filamentary structures around the hub. In J. W. Zhou et al. (2023), we focused on large-scale filaments within molecular clouds. Along these filaments, we observed clear fluctuations in both velocity and intensity, with each intensity peak indicating a local hub. In this study, we began by identifying local intensity peaks (i.e. leaf structures), and subsequently searched for hub-filament structures centered around these intensity peaks.

A kinematic feature of hub-filament structures is that the gas flow along the filament converges towards the central hub, resulting in a measurable velocity gradient along the filament. Since we are now studying each local hub-filament structure, we follow the same analysis presentend in J.-W. Zhou et al. (2022) to fit the velocity gradients around the intensity peaks. In Fig. 2.23(a), the velocity gradients fitted in those local hub-filament structures are quite comparable to those presented in J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023). Almost all of them can be fitted with masses in the range $\approx 100-10000$ $M_\odot$, which is consistent with the mass distribution of leaf structures shown in Fig.2.21(b), indicating that local dense structures as gravitational centers can accrete the surrounding diffuse gas and produce the observed velocity gradients. Measurements of velocity gradients provide evidence for gas inflow within these structures, which can serve as kinematic evidence that these structures can be regarded as

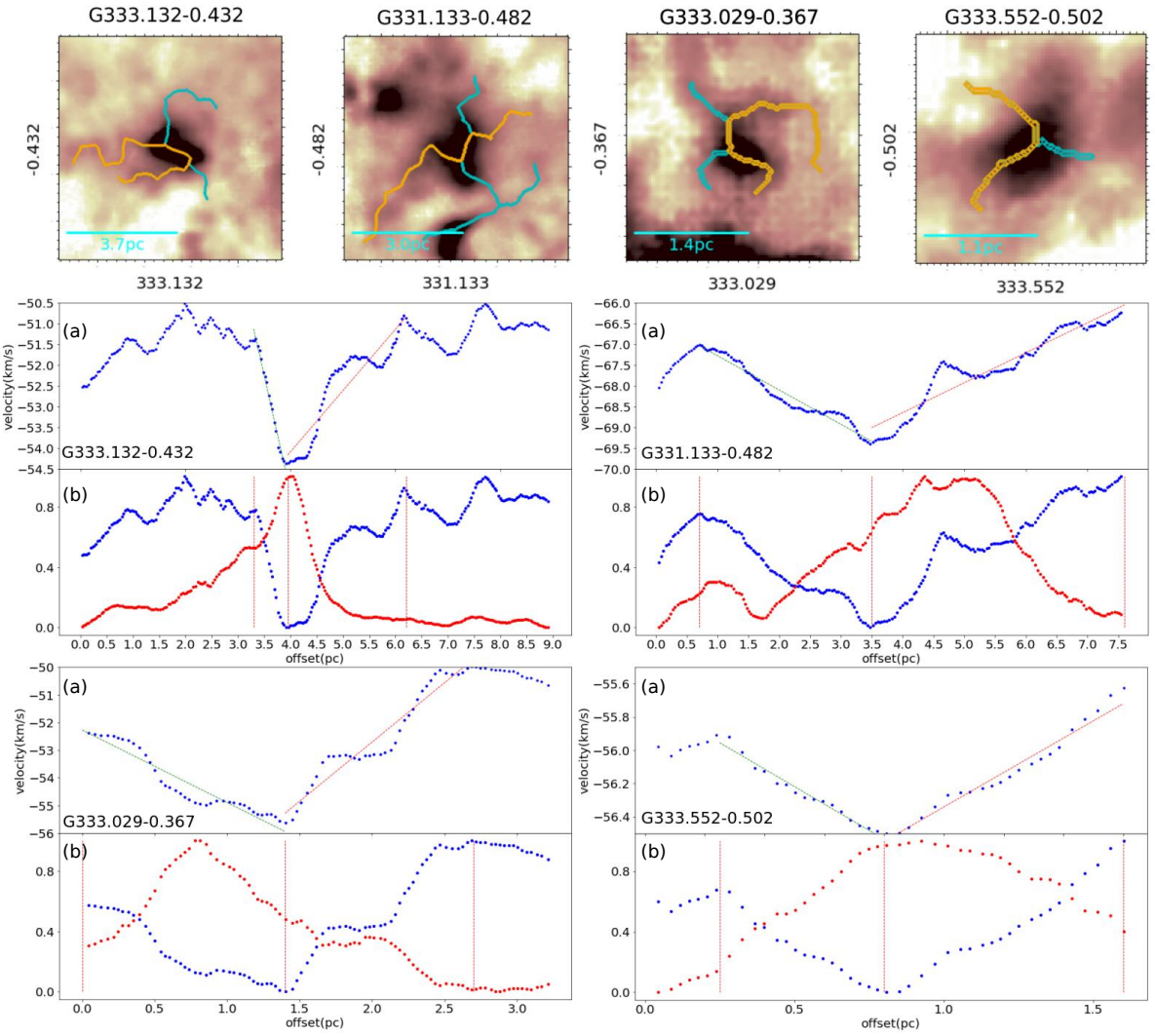


Figure 2.22: Four HFs structures used to demonstrate the velocity gradient fitting. In the first row, the background is the integrated intensity map of ^{13}CO (3−2), orange and cyan lines are the filaments identified by the FILFINDER algorithm. The two rows below show the velocity and intensity profiles extracted along the orange filaments. In panel (a), velocity gradients are fitted in the ranges defined by the red vertical dashed lines in panel (b). Straight lines show the linear fitting results. In panel (b), blue and red dotted lines show the normalized velocity and intensity, respectively.

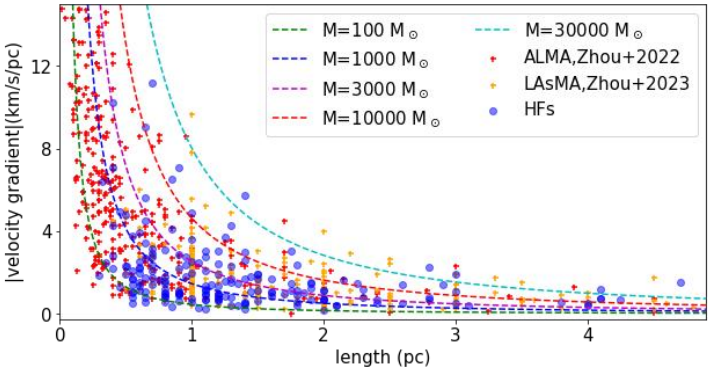


Figure 2.23: Velocity gradient versus the length over which the gradient has been fitted. Red and orange "+" represent the velocity gradients fitted in J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023), respectively. The dashed lines show free-fall velocity gradients for comparison. For the free-fall model, from left to right, the lines denote masses of 100 $M_\odot$, 1000 $M_\odot$, 3000 $M_\odot$, 10000 $M_\odot$ and 30000 $M_\odot$, respectively.

hub-filament systems.

2.6.3 ATLASGAL clumps and radio sources

The physical properties of the ATLASGAL clumps (Urquhart et al., 2022) associated with leaf structures can reveal their evolutionary stages. If the distance between the central coordinates of a clump and a leaf structure is less than the effective radius of the leaf structure, we associate the clump with the leaf structure. In total, 438 leaf structures are associated with 293 ATLASGAL clumps. The fraction of hubs associated with ATLASGAL clumps is approximately 0.83 for HFs and about 0.32 for non-HFs. We also match leaf structures with the radio sources identified in the CORNISH-South survey (Irabor et al., 2023). Approximately 29% of hubs in HFs are associated with radio sources, compared to only around 4% in non-HFs. The ratio of HFs associated with clumps and radio sources is much higher than that of non-HFs, indicating that HFs are more evolved than non-HFs. In Fig.2.21(f), the *Spitzer* 8 μm peak intensity of HFs is significantly higher than that of non-HFs. The luminosity-to-mass ratios, L/M, of the clumps can be used to trace the evolutionary process of star formation (Saraceno et al., 1996; Molinari et al., 2008; T. Liu et al., 2016b; Stephens et al., 2016; Urquhart et al., 2018). The clumps associated with HFs have significantly higher L/M ratios compared to those associated with non-HFs, which also suggests that HFs are more evolved than non-HFs.

2.6.4 Evolutionary sequence

In Fig.2.24, we compared the physical properties of two types of non-HFs, i.e. without and with associated ATLASGAL clumps. The structures associated with ATLASGAL clumps (more evolved) have higher density contrast, larger mass and lower virial ratio. We also compared the physical properties of two types of HFs, i.e. without and with associated radio sources. The structures associated with radio sources (more evolved) have higher density contrast, larger mass and lower virial ratio. As discussed above, HFs is more evolved than non-HFs. In Fig.2.21, compared with non-HFs, the more evolved HFs have higher density contrast, larger mass and lower virial ratio. Therefore, there may be an evolutionary sequence from non-HFs to HFs. In terms of spatial distribution, HFs and non-HFs are mixed together, sharing the same physical environment. In other words, the differences between HFs and non-HFs are not caused by their physical environments. If there is an evolutionary relationship between them, these differences can be naturally explained.

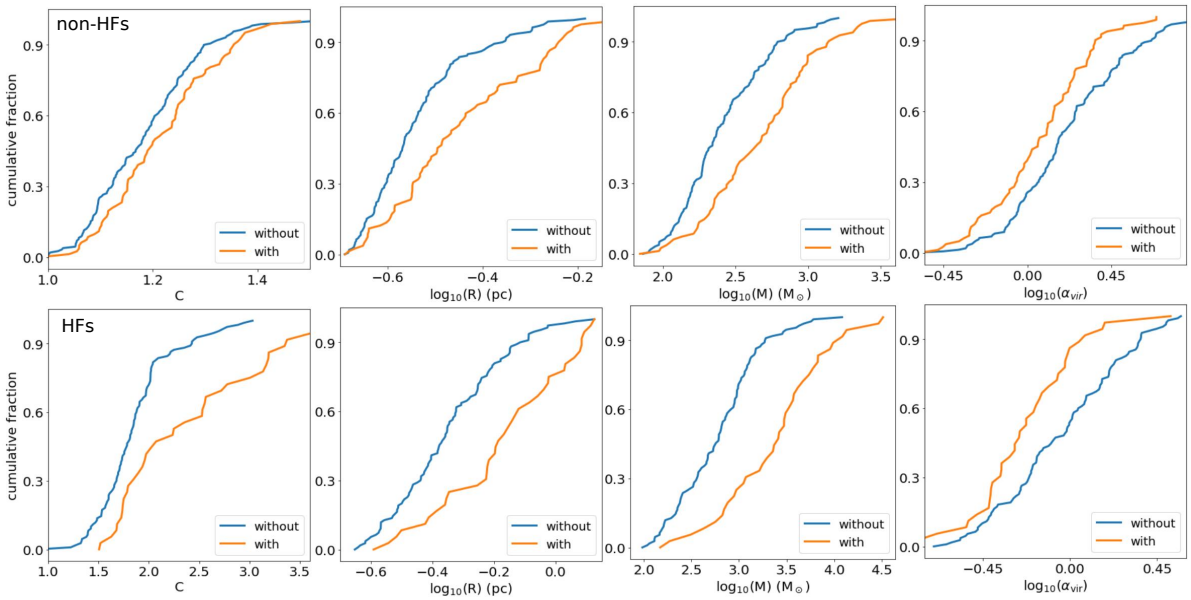


Figure 2.24: The first row shows the physical properties of non-HFs without and with associated ATLASGAL clumps. For left to right, the physical parameters are the column density contrast, effective radius, mass and virial ratio, respectively. The second row shows the physical properties of HFs without and with associated radio sources.

2.6.5 Density contrast

As illustrated in Fig.2.19, the presence of a distinct central hub indicates a significant density contrast between the hub and the surrounding diffuse gas. We define the density contrast, C, as the ratio of the average column density in the hub region, $N_{\rm hub}$, to the average column density within an elliptical ring surrounding the hub, which has a width equal to the size of the hub (the effective radius of the leaf structure), $N_{\rm a}$,

$$C = N_{\rm hub}/N_{\rm a}. \tag{2.3}$$

As expected, in Fig. 2.21(a), HFs display the highest density contrast. Moreover, HFs also possess the largest masses and the lowest virial ratios. As shown in Fig.2.21(e), although HFs have a larger density contrast, the average density of non-HFs is almost comparable to that of HFs. Therefore, non-HFs are not necessarily low-density structures. It is just that their density distribution is relatively uniform. Gas structures may undergo a gravitational focusing process starting from an initially relatively uniform density distribution, ultimately forming hub-filament structures. This process has been investigated in detail for the fragment G332.83-0.55 (one of HFs in this work) in J. W. Zhou et al. (2024f), where a strong gravitational center plays a crucial role in shaping the hub-filament morphology. According to the possible evolutionary sequence from non-HFs to HFs, currently, non-HFs lack a distinct gravitational focusing process that would result in significant density contrast. Thus, their density distribution is relatively uniform. Therefore, density contrast effectively measures the extent of gravitational collapse and the strength of the gravitational center of the structure, which definitively shape the hub-filament morphology. In terms of reflecting the development and evolutionary stage of a structure, density contrast is more crucial than the density itself.

2.6.6 Structure of molecular clouds

Considering the evolutionary sequence from non-HFs to HFs, non-HFs can be regarded as potential HFs, since they may evolve into HFs in the future. In this work, the molecular clouds are decomposed

into hundreds of HFs and potential HFs with different scales. Combined with the kinematic evidence in J. W. Zhou et al. (2023) and J. W. Zhou et al. (2024g), we suggest that molecular clouds are network structures formed by the gravitational coupling of multi-scale hub-filament structures. Hub-filament structures serve as the basic constituent units of molecular clouds. The nodes in the networks are the hubs, and they are the local gravitational centers and the main star-forming sites. Actually, clumps in molecular clouds are equal to the hubs.

This network picture can naturally explain the results in J. W. Zhou et al. (2024g) and J. W. Zhou et al. (2024e): In star-forming regions, spanning from cloud to core scales, initial-stage feedback has minimal impact on the physical properties of embedded dense gas structures. While feedback at these scales can disrupt the original cloud complex, the substructures within it can reorganize into new configurations governed by gravity around new gravitational centers. This dynamic process involves both structural destruction and formation, shifts in gravitational centers, but gravitational collapse is always ongoing (see Chapter.4 for more details). The network structure of molecular clouds implies that the nodes or local dense structures as local hubs or local gravitational centers are relatively independent of each other. Although some nodes in the cloud evolve more rapidly, forming HII regions, due to their relative independence with the neighboring nodes, early feedback from them does not significantly impact the physical properties of the neighboring nodes. Therefore, early feedback mainly reshapes the network structure of molecular clouds, representing a topological deformation rather than a destruction of the molecular cloud, as illustrated in Fig.8 of J. W. Zhou et al. (2024e).

2.7 Decomposition of molecular gas in a galaxy

Kinematic analyses presented in J. W. Zhou et al. (2023) and J.-W. Zhou & T. Davis (2024) demonstrate the presence of multi-scale hub-filament structures within molecular clouds and spiral galaxies. The results notably show that intensity peaks, acting as hubs, are correlated with converging velocities, suggesting that surrounding gas inflows are directed towards these dense regions. Based on the kinematic evidence, the main goal of this work is to directly recover the multi-scale hub-filament structures in the spiral galaxy NGC 628 (M74 or the Phantom Galaxy). NGC 628 was selected for the following reasons: (1) It is a face-on galaxy. The inclination angle of NGC 628 is only 8.9 degree (Lang et al., 2020). Therefore, this facilitates the identification of the hub-filament structures embedded within the galaxy; (2) High-resolution CO data and the galaxy's relatively near distance can clearly reveal the hub-filament structures; (3) It is covered by multi-wavelength observations. As discussed below, high-resolution 21 μm and H_α emission are crucial to determine the evolutionary states of CO structures. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024a).

2.7.1 Classification

We conducted a direct identification of hierarchical (sub-)structures based on the 2D intensity maps (CO J=2−1, 21 μm and H_α emissions) using the dendrogram algorithm. For the identified molecular gas structures, we extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. Large-scale velocity gradients from the galaxy's rotation contribute to the non-thermal velocity dispersion. To address this, we removed the bulk motion due to the rotation of galaxy by creating gas dynamical models using the KinMS package (T. A. Davis et al., 2013). Then, following the procedure described in J. W. Zhou et al. (2024g), we fitted the averaged spectra of 773 leaf structures individually using the `GAUSSPY+` (Lindner et al., 2015; Riener et al., 2019) algorithm.

From the line profile, we can fix the velocity range of each structure. Then, the moment 0 map of each structure was reproduced in the corresponding velocity range to eliminate the overlap of potential

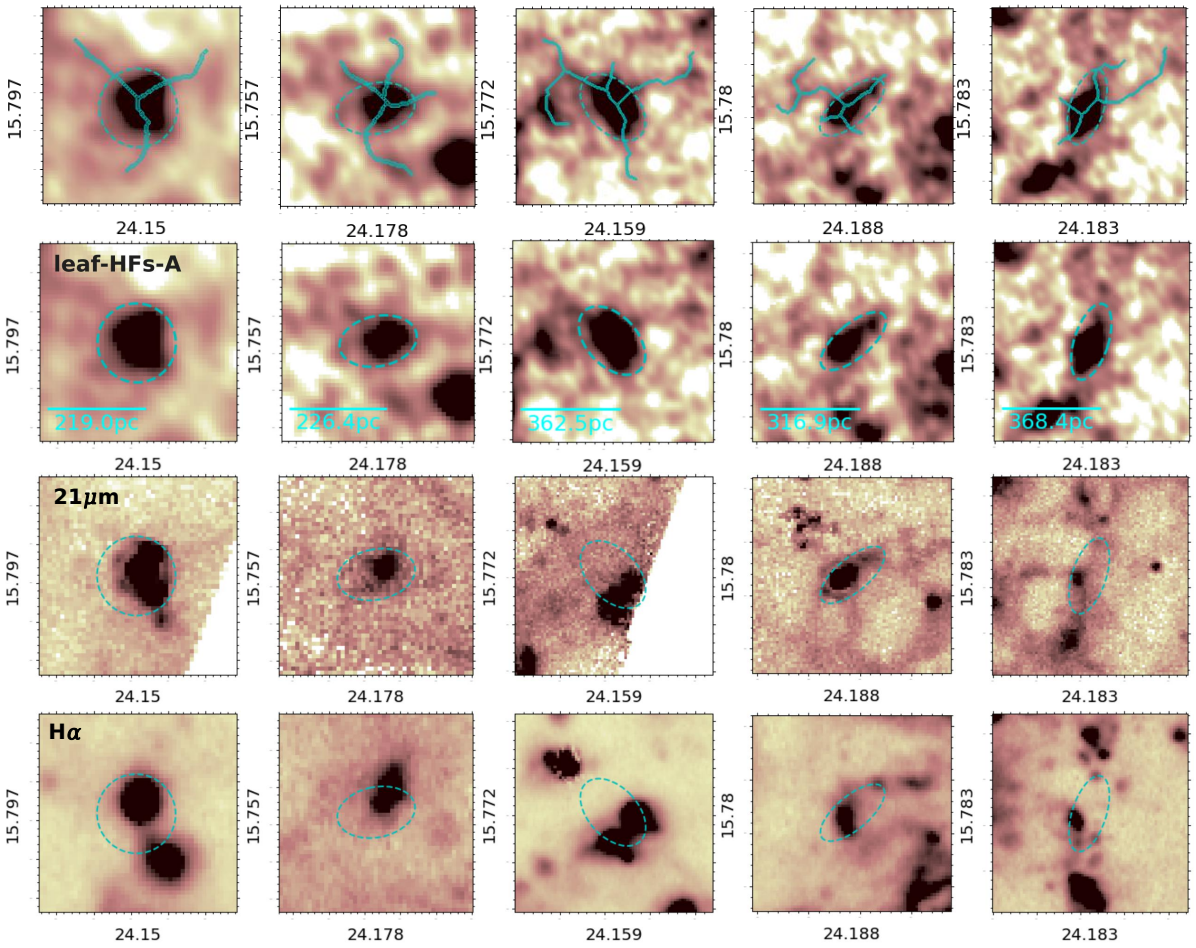


Figure 2.25: Some examples of leaf-HFs-A. CO emission (the first and second rows), 21 μm emission (the second row) and H_α emission (the third row). In the first row, the filaments identified by the FILFINDER algorithm are overlaid on the moment 0 maps.

incoherent velocity components. All identified intensity peaks (leaf structures) of CO (2−1) emission are thought to be potential hubs. Around these intensity peaks, we extended the spatial ranges to investigate the filamentary structures connected with them. All leaf structures were classified by eye into three categories only based on their morphology, i.e. leaf-HFs-A, leaf-HFs-B and leaf-HFs-C. The numbers of structures in the three categories are 234, 181 and 358. Some examples in the three categories are shown in Fig. 2.25, Fig. 2.26 and Fig. 2.27.

From these maps, we can observe that leaf-HFs-C do not exhibit clear central hubs, meaning that the density contrast between the hub and the surrounding diffuse gas is not pronounced. In contrast, leaf-HFs-A and leaf-HFs-B feature distinct central hubs or high-density central regions, characteristic of hub-filament structures. Compared to leaf-HFs-B, leaf-HFs-A demonstrate more defined hub-filament morphology, with filamentary diffuse gas surrounding their central hubs. The hubs of leaf-HFs-A are more outstanding. While the boundary between leaf-HFs-A and leaf-HFs-B might be less distinct, the difference between leaf-HFs-A and leaf-HFs-C is significant. Therefore, in the subsequent discussion, we focus on the comparison between leaf-HFs-A and leaf-HFs-C. The initial classification was only based on the morphology, because other differences between the structures were unknown. In the subsequent analysis, we can see that the physical properties of the structures in three categories are also significantly different. Therefore, the morphological differences are the result of certain physical processes shaping

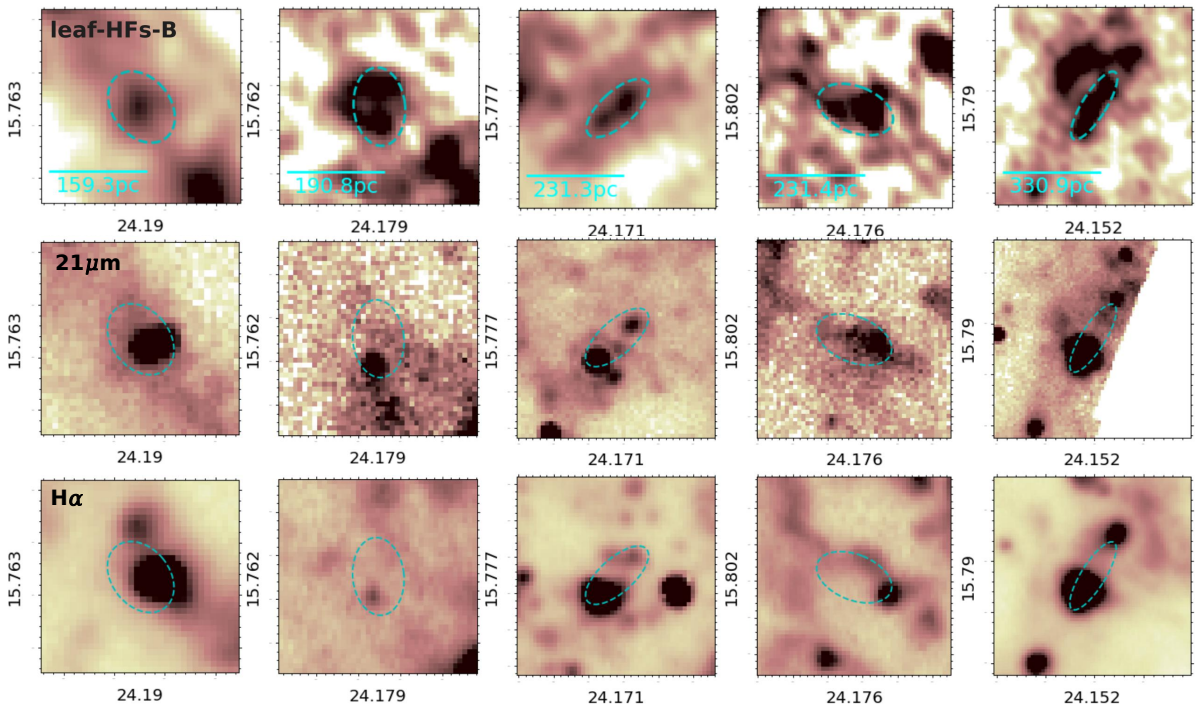


Figure 2.26: Same as Fig.2.25, but for leaf-HFs-B.

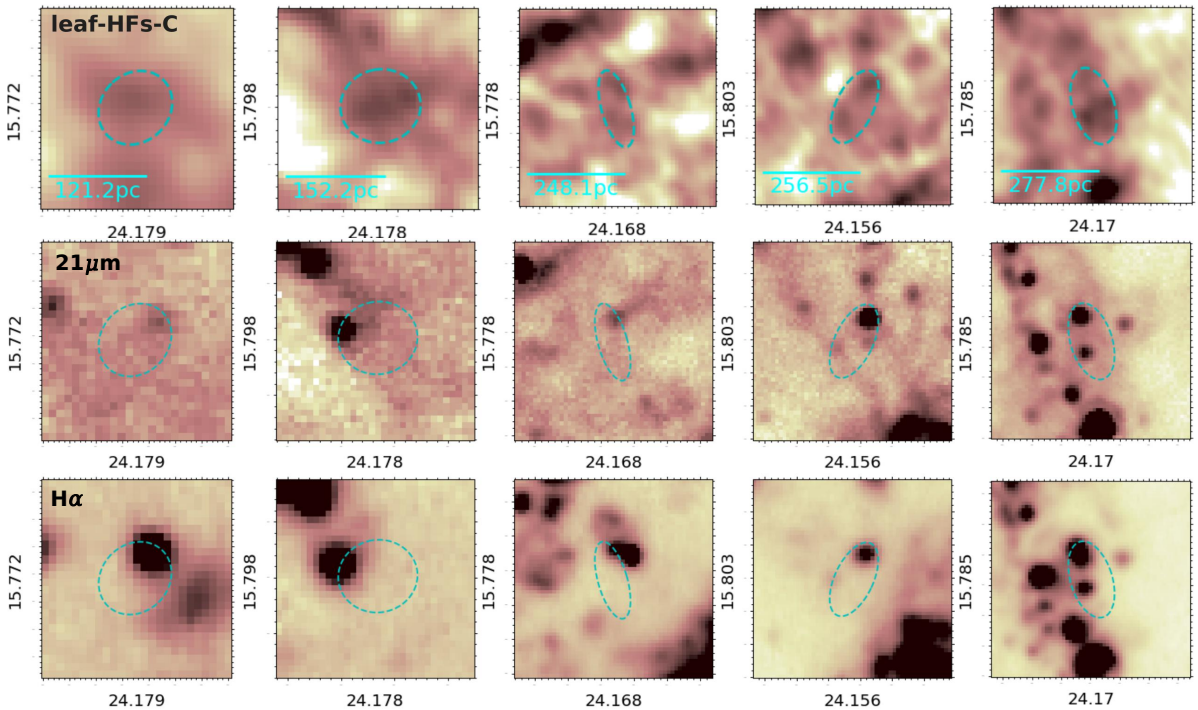


Figure 2.27: Same as Fig.2.25, but for leaf-HFs-C.

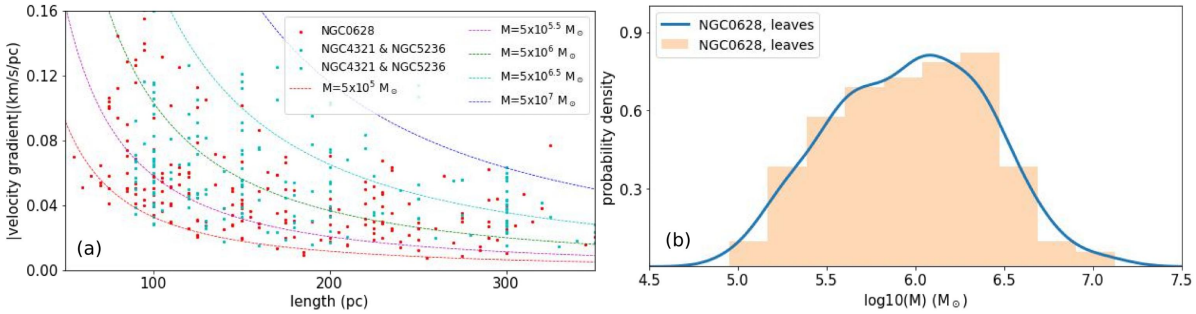


Figure 2.28: Statistical analysis of all fitted velocity gradients. (a) Velocity gradient vs. the length. The color lines show the freefall velocity gradients for comparison. For the freefall model, red, magenta, green, cyan, and blue lines denote masses of 10^5 $M_\odot$, $10^{5.5}$ $M_\odot$, 10^6 $M_\odot$, $10^{6.5}$ $M_\odot$ and 10^7 $M_\odot$, respectively; (b) The mass distribution of leaf structures.

them and are not coincidental.

2.7.2 Velocity gradient

As done in J.-W. Zhou & T. Davis (2024), we also use the FILFINDER algorithm to characterize the filamentary structures around the hub. In J.-W. Zhou & T. Davis (2024), we only focused on large-scale filaments along the spiral arms. Clear fluctuations in velocity and density were observed along these filaments. Each individual intensity peak reveals a local hub. In this work, we first identified the local intensity peaks (leaf structures), and then searched for hub-filament structures centered around these intensity peaks as potential hubs. In Fig. 2.25, the identified filamentary structures align well with the background. However, due to the resolution limitations, the filamentary structures do not appear very extended.

A kinematic feature of hub-filament structures is that the gas inflow along the filament converges towards the central hub, resulting in a measurable velocity gradient along the filament. Since we are now studying each local hub-filament structure, we follow the same analysis presentend in J.-W. Zhou et al. (2022) to fit the velocity gradients around the intensity peaks. In Fig. 2.28(a), the velocity gradients fitted in NGC 628 are quite comparable to those in NGC 4321 and NGC 5236 presented in J.-W. Zhou & T. Davis (2024). The measured velocity gradients predominantly fit within the mass range of $\approx$ 10^5-10^7 $M_\odot$. This range aligns with the mass distribution of leaf structures (molecular clouds) shown in Fig. 2.28(b). This consistency suggests that local dense structures act as gravitational centers, accreting surrounding diffuse gas and thereby generating the observed velocity gradients. Measurements of velocity gradients provide evidence for gas inflow within these structures, offering kinematic support for classifying them as hub-filament systems.

2.7.3 Star formation

We consider the corresponding 21 μm or H_α structures to be associated with the CO structure if their separations are less than the effective radius of the CO structure. Finally, there are 313 21 μm structures (37%, 313/841) and 324 H_α structures (28%, 324/1142) associated with CO structures, where 185 CO structures have both strong 21 μm and H_α emissions. Limited to the FOV of 21 μm observations, the total numbers of leaf-HFs-A, leaf-HFs-B and leaf-HFs-C are 163, 119, and 251, respectively. The proportions of leaf-HFs-A, leaf-HFs-B and leaf-HFs-C associated with both 21 μm and H_α structures are 58% (95/163), 37% (44/119) and 18% (46/251), indicating that leaf-HFs-A are more evolved than

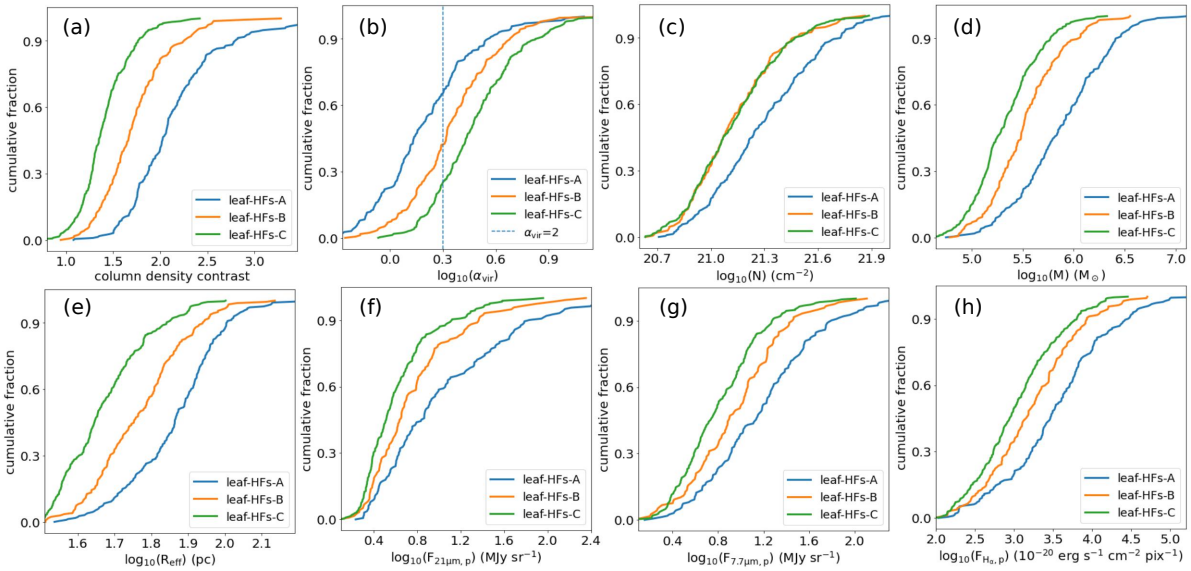


Figure 2.29: Physical parameters of three types of structures. (a) Column density contrast; (b) Virial ratio. The vertical dashed line marks the position of $\alpha_{vir} = 2$; (c) Mass; (d) Column density; (e) Scale (the effective radius of the CO structures); (f) Peak intensity of 7.7 μm emission; (g) Peak intensity of 21 μm emission; (h) Peak intensity of H$_\alpha$ emission.

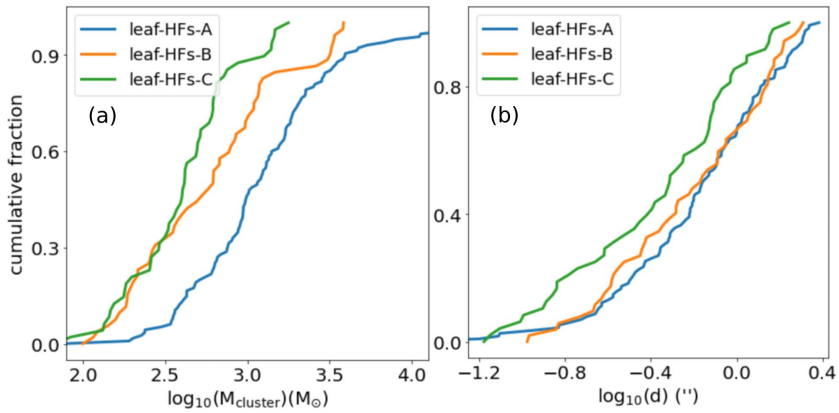


Figure 2.30: (a) The mass distribution of embedded star clusters; (b) The spatial separations between the intensity peaks of CO and associated 21 μm structures.

leaf-HFs-C. We also calculated the peak intensities of 7.7 μm, 21 μm and H$_\alpha$ emissions for each CO structure. As shown in Fig. 2.29 (f)-(h), all emissions gradually increase from leaf-HFs-C to leaf-HFs-A.

We further examined the physical properties of embedded star clusters within molecular clouds. Following the prescriptions described in Leroy et al. (2021b), the 21 μm data can be used to calculate the local star formation rate (SFR) surface density. Kruijssen & Longmore (2014) and Kruijssen et al. (2018) have developed a statistically robust method to convert the observed spatial separations between cold gas and SFR tracers into their underlying timescales. This approach was employed by J. Kim et al. (2023) to study NGC 628 using emission maps of CO, 21 μm, and H$_\alpha$ emission maps which respectively trace molecular clouds, embedded star formation, and exposed star formation. This analysis yielded systematic constraints on the duration of the embedded phase of star formation in NGC 628. They defined the duration of the embedded phase of star formation as the time during which CO and 21 μm emissions are found to be overlapping, $t_{\mathrm{fb},21\mu\mathrm{m}}$. In contrast, the "heavily obscured phase" refers

to the period when both CO and 21 μm emissions are present without associated H_α emission, $t_{\rm obsc}$. Finally, they obtained $t_{\rm fb,21\mu m} \approx 5.1^{+2.7}_{-1.4}$ Myr and $t_{\rm obsc} \approx 2.3^{+2.7}_{-1.4}$ Myr. In this work, we directly take the typical values $t_{\rm fb,21\mu m} = 5.1$ Myr and $t_{\rm obsc} = 2.3$ Myr. Therefore, the estimate is quite rough, given the significant uncertainty regarding the duration time. For 185 CO structures associated with both 21 μm and H_α emissions, the duration time should be ≈2.3−5.1 Myr. The spatial separations (d) between the intensity peaks of CO and 21 μm structures can be used to estimate the age of the embedded star clusters in each molecular cloud. Assuming $t_0 = 2.3$ Myr, the age of the embedded star clusters is

$$t = t_0 + \frac{t_{\rm fb,21\mu m} - t_{\rm obsc}}{d_{\rm max} - d_{\rm min}} \times d, \tag{2.4}$$

where $d_{\rm max}$ and $d_{\rm min}$ are the maximum and minimum spatial separations for all 185 clouds. By combining the SFR of each cloud, we estimated the total mass of the corresponding embedded star clusters. Fig. 2.30(a) shows the mass distribution of embedded star clusters. The cluster mass of leaf-HFs-A is much larger than those of leaf-HFs-B and leaf-HFs-C. The results here are quite consistent with the findings in Hassani et al. (2023). For the JWST 21 μm compact source population identified in Hassani et al. (2023), using spectral energy distribution (SED) fitting, they found that the 21 μm sources have SEDs consistent with stellar population masses of $10^2 < M_*/M_\odot < 10^{4.5}$. They also found a range of mass-weighted ages spanning 2–25 Myr, though most objects are < 8 Myr.

In Fig. 2.30(b), the spatial separations between the intensity peaks of CO and 21 μm structures of leaf-HFs-A are also larger than leaf-HFs-C. Since the spatial separation can measure the evolutionary timescales of molecular clouds, this is further evidence supporting that leaf-HFs-A is more evolved than leaf-HFs-C. However, the sample size here is very limited. The numbers of leaf-HFs-A, leaf-HFs-B and leaf-HFs-C are 95, 44 and 46, respectively. More samples are necessary to quantitatively assess the correspondence between the separations and the evolutionary stages of molecular clouds.

2.7.4 Formation of hub-filament structures

The hub-filament morphology could be shaped by either gravitational collapse (Vázquez-Semadeni et al., 2019) or strong shocks due to turbulence (Padoan et al., 2020b). In the second case, there is no anticipated consistency between the velocity gradients and the masses of the hubs, while in the first case, a relationship between the velocity gradients and the hub masses is naturally expected. Thus, the results presented above clearly support the scenario in which the hub-filament structures form through gravitational contraction of gas structures on galaxy-cloud scales, which is also revealed in J.-W. Zhou & T. Davis (2024). Similar findings on cloud-clump scales and clump-core scales are presented in J. W. Zhou et al. (2023) and J.-W. Zhou et al. (2022), also see Motte et al. (2018b), Vázquez-Semadeni et al. (2019) and Kumar et al. (2020) and references therein. The kinematic characteristics of gas structures on galaxy-cloud scale are very similar to those on cloud-clump and clump-core scales. Therefore, a picture emerges where dense cores act like hubs within clumps, clumps act like hubs within molecular clouds, and molecular clouds act like hubs in spiral galaxies. The molecular gas from galaxy to dense core scales presents multi-scale/hierarchical hub-filament structures. Gas structures at different scales in the galaxy may be organized into hierarchical systems through gravitational coupling. The hierarchical hub-filament structures are also present in the galaxy cluster (hub)–cosmic web (filament) picture, reflecting the self-similarity of the structural organization in the universe. The fundamental reason may be that structures at different scales in the universe primarily evolve under the influence of gravity.

Hubs are essentially structures with higher density compared to the surrounding more diffuse gas. As local gravitational centers, these hubs seem to accrete the surrounding diffuse gas, forming hub-filament structures. The filamentary structures seem to be associated with gas flows converging towards the hubs.

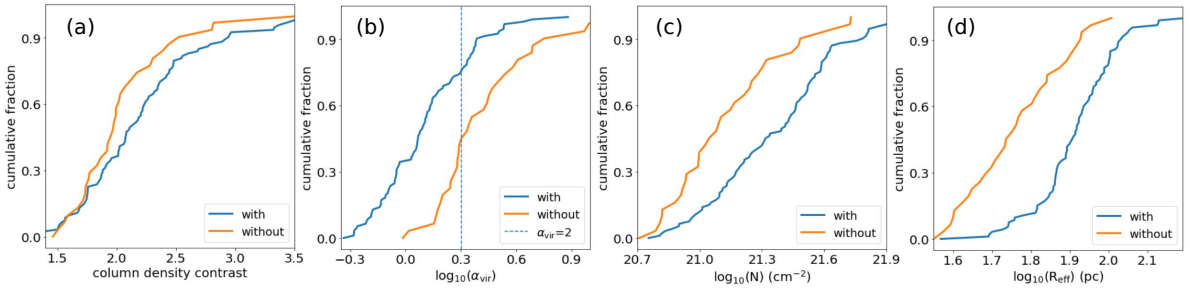


Figure 2.31: Comparison of leaf-HFs-A with and without 21 μm and H$_\alpha$ emission. (a) Column density contrast; (b) Virial ratio. The vertical dashed line marks the position of $\alpha_{vir} = 2$; (c) Column density; (d) Scale (the effective radius).

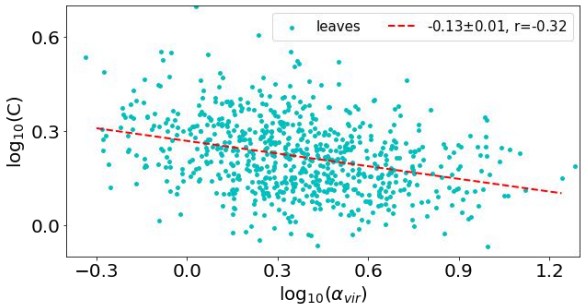


Figure 2.32: Correlation between the density contrast and the virial ratio for all leaf structures. "r" is the Pearson correlation coefficient.

A hub-filament structure essentially consists of a gravitational center and the gas flow converging towards it. This could be considered as a fundamental structural type in the ISM. In short, as long as there are local dense structures within the relatively diffuse ISM, hub-filament structures will inevitably form.

Finally, we note that although the measured velocity gradients support the gravitational collapse of gas structures on galaxy-cloud scales, the collapse is much slower than a pure free-fall gravitational collapse. In the free-fall, the velocity gradient ∇v and the scale R satisfy $\nabla v \propto R^{-1.5}$. However, Fig.2.28(a) only gives $\nabla v \propto R^{-0.9}$. For NGC 4321 and NGC 5236, J.-W. Zhou & T. Davis (2024) obtained a similar slope, i.e. $\nabla v \propto R^{-0.8}$. As discussed in J.-W. Zhou & T. Davis (2024), the deviation from the free-fall may come from the measurement biases, the coupling with the galactic potential, the tidal forces from the galaxy or the neighboring structures, or the turbulence and magnetic field supports (Chevance et al., 2023).

2.7.5 Evolutionary sequence of molecular clouds

In Fig. 2.29(a) and (e), leaf-HFs-C have significantly smaller density contrasts and scales. However, leaf-HFs-C and leaf-HFs-B show comparable column density distributions in Fig. 2.29(d), although leaf-HFs-B have larger density contrasts. Therefore, leaf-HFs-C are not necessarily low-density structures. It is just that the density distribution in leaf-HFs-C structures is relatively uniform. The results of this work seem to suggest an evolutionary sequence among molecular clouds. First, we focus solely on leaf-HFs-A. We compared two types of leaf-HFs-A, i.e. without and with 21 μm and H$_\alpha$ emissions. As shown in Fig. 2.31, the structures with 21 μm and H$_\alpha$ emissions present larger density contrasts, lower virial ratios, higher column densities and larger scales (larger masses). In Fig. 2.29, from leaf-HFs-C to leaf-HFs-A, we can see the same trend for each physical quantity. Leaf-HFs-A without and with 21

μm and H_α emissions represent structures from two different evolutionary stages. Leaf-HFs-C and leaf-HFs-A may have the same relation.

A typical characteristic of hub-filament structures is that the hub region has significantly higher density compared to the surrounding filamentary structures, i.e. the density contrast. The hub-filament morphology is shaped by gravitational collapse. Hub-filament structures only become apparent when gravitational collapse has progressed to a certain extent. In this work, the identified dense CO structures are potential hubs, which act as local gravitational centers. Leaf-HFs-A have significantly greater mass than leaf-HFs-C, so they serve as stronger gravitational centers. Leaf-HFs-A also present much better hub-filament morphology than leaf-HFs-C, which reflects the extent of gravitational collapse. Leaf-HFs-C have not yet exhibited a pronounced gravitational collapse process that would lead to a substantial density contrast. From leaf-HFs-C to leaf-HFs-A, gravitational collapse will increase the density contrast of the structures. Thus, the density contrast effectively gauges the degree of gravitational collapse and the depth of the gravitational potential well within a structure, which decisively shape the hub-filament morphology.

Fig.2.32 shows a clear correlation between the density contrast and the virial ratio. As expected, the density contrast decreases as the virial ratio increases. However, the classical virial analysis only considered the gravitational potential energy and internal kinetic energy may not completely reflect the true physical state of the identified structures. A more comprehensive virial analysis should include additional physical mechanisms, such as external pressure (Spitzer, 1978; B. G. Elmegreen, 1989; Kirk et al., 2017; S. Li et al., 2020; Sun et al., 2020; J. W. Zhou et al., 2024g), tidal forces (Ballesteros-Paredes et al., 2009b; Ballesteros-Paredes et al., 2009a; Ramırez-Galeano et al., 2022; G.-X. Li, 2024; J. W. Zhou et al., 2024f) and magnetic fields (Dib et al., 2007; Crutcher et al., 2010; H.-B. Li & Henning, 2011; Crutcher, 2012; Seifried et al., 2020; Ibáñez-Mejıa et al., 2022; Ganguly et al., 2023). Local dense structures are embedded within larger gas environments. Their interactions with the surrounding environment may significantly impact their physical state. After calculating a more accurate virial ratio, the correlation in Fig.2.32 should be stronger.

As the masses and scales of the structures also increase from leaf-HFs-C to leaf-HFs-A, the results of this work suggest that the structures inevitably accrete matter from their surrounding environment during the evolution. It is consistent with the velocity gradient analysis and the results presented in J.-W. Zhou & T. Davis (2024) for NGC 5236 and NGC 4321. In these spiral galaxies, there are clear velocity gradients around intensity peaks, and the variations in velocity gradient across different scales suggest a gradual and consistent increase in velocity gradient from large to small scales, indicative of gravitational collapse and gas inflow at different scales. Local dense structures act as local gravitational centers, naturally accreting matter from the surrounding diffuse gas environment and thereby accumulating mass.

2.8 Gravitational network analysis

Molecular gas in spiral galaxies is structured into a network of interconnected systems through the gravitational coupling of multi-scale hub-filament structures. Molecular clouds act as nodes within these hub-filament networks, serving as local gravitational centers and primary sites for star formation. In this work, we employ network-based analytical methods to gain further insights into the formation and evolution of molecular clouds in the spiral galaxy NGC 628. The work presented in this section is reproduced from the publication by J. W. Zhou & G.-X. Li (2025).

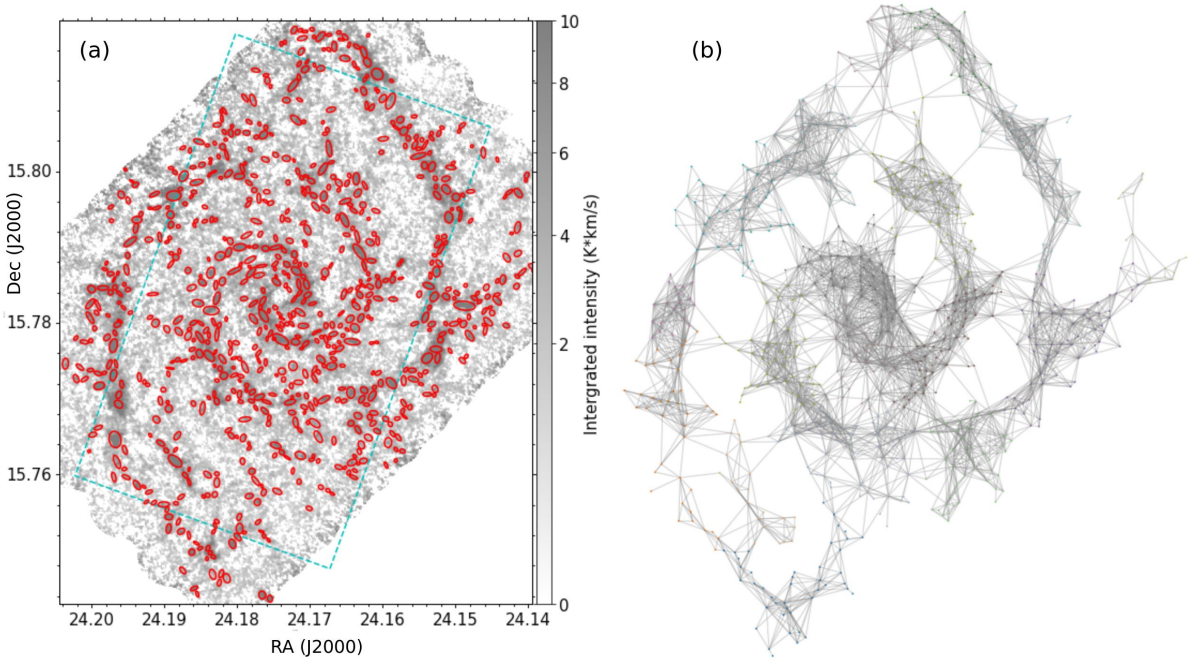


Figure 2.33: (a) Background is the broad-masked integrated intensity map of CO (2−1) emission. Red ellipses represent leaf structures identified by the dendrogram algorithm in J. W. Zhou et al. (2024a). Cyan box marks the field-of-views of the JWST 21 μm observation; (b) The gravitational network constructed in Sec.2.8.1, with molecular clouds as nodes.

2.8.1 Gravitational network

Considering the picture of multi-scale/hierarchical hub-filament structures in spiral galaxies revealed by observations, we can construct a gravitational network with molecular clouds as nodes all over the galaxy NGC 628. The elements of a network include nodes, edge weights and connection rules. The edge weights are defined by the gravitational force between nodes, with greater weights indicating stronger gravitational interactions. The connection rules stipulate that nodes separated by a distance greater than 1 kpc are not connected, as this work focuses only on locally coupled structures. The distance is calculated in a 2D plane using Euclidean distance according to the central coordinates of nodes. We tested different values of the separation threshold and found that a threshold below 1 kpc (e.g., 500 pc) mainly results in discrete local networks, preventing the formation of the overall network shown in Fig.2.33. A larger threshold (e.g., 1500 pc) leads to more connections, which eliminate locally coupled structures (local networks) but do not affect the correlations between node geometric parameters and node mass presented in the next section. Furthermore, a gravitational threshold is defined as the median gravitational force calculated for all node pairs with separations less than 1 kpc. Nodes are connected only if their gravitational force exceeds this threshold to outstand the main gravitational interactions between nodes. Fig.2.33 presents the constructed network, revealing some localized cloud clusters.

2.8.2 Accretion or merger

The physical properties of molecular clouds in NGC 628 have been calculated and discussed in J. W. Zhou et al. (2024a). Now, with each cloud being considered a node, we can calculate the geometric parameters of each node and explore the correlation between geometric and physical properties of molecular clouds. In other words, having visualized the gravitational interactions between molecular clouds through the

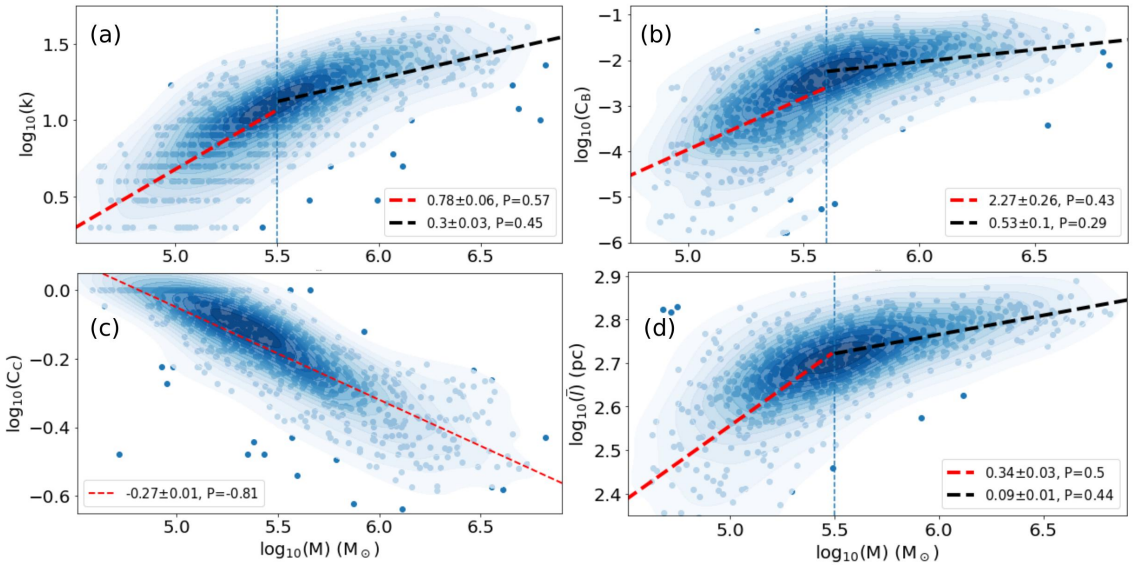


Figure 2.34: Correlation between the geometric parameters of nodes and their masses. (a) The degree (k); (b) The betweenness centrality (C_B); (c) The clustering coefficient (C_C); (d) The average physical path length ($\bar{l}$) between each node and its neighbors. The background contour in each panel shows the probability density (KDE) of the scatter points. Vertical dashed lines mark the inflection points of the segmented fit.

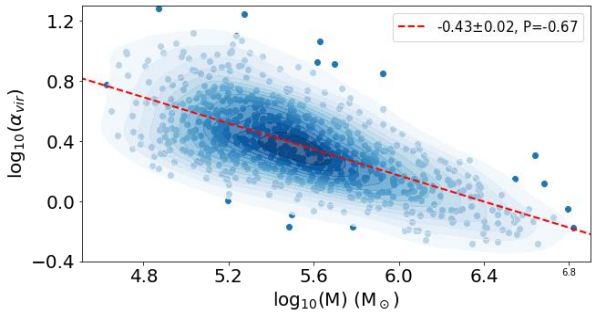


Figure 2.35: Correlation between the virial ratio and mass of molecular clouds. The background contour shows the probability density (KDE) of the scatter points.

gravitational network, we can now characterize both these interactions and the physical properties of the clouds using geometry-based network metrics.

In a network, the degree of a node (k_i) represents the number of other nodes directly connected to this node. It is a fundamental attribute for measuring the importance of a node. Particularly in a gravitational network, nodes with high degrees may represent massive objects that exert significant gravitational influence on their surroundings. Fig.2.34(a) shows a strong positive correlation between node mass and degree.

The betweenness centrality of a node ($C_{B,i}$) measures the extent to which this node acts as an intermediary within the network. Nodes with high betweenness centrality have greater influence in the network. Massive clouds, serving as local hubs and gravitational centers, naturally exhibit greater influence, leading to higher values of C_B, as shown in Fig.2.34(b).

The clustering coefficient of a node ($C_{C,i}$) is an important metric for measuring the degree of connectivity among its neighbors. It reflects the tightness of the local network structure around a node. In a gravitational network, a high clustering coefficient indicates that the neighbors of a node have stronger

gravitational relationships with one another, forming a highly interconnected subnetwork. Surprisingly, as shown in Fig.2.34(c), high-mass molecular clouds exhibit a lower clustering coefficient. We also calculate the average physical path length ($\overline{l_i}$) between each node and its neighbors. In Fig.2.34(d), $\bar{l}$ increases with the node's mass, which means an anti-correlation between $\bar{l}$ and C_C. This could be due to high-mass nodes connecting to more distant nodes, which increases the average distance to their neighbors. Alternatively, it might be that there are simply fewer nodes near high-mass nodes, which also accounts for the lower C_C observed in these high-mass nodes.

These results suggest that high-mass clouds, during their formation and evolution, deplete nearby gas through accretion or merger, resulting in more isolated network characteristics. Molecular clouds are inherently irregular gas structures, so describing them as being merged or accreted essentially conveys a similar idea. Initially, molecular clouds may exhibit a relatively uniform distribution. Over time, molecular gas accumulates toward gravitational centers, and the high-mass clouds that form eventually decouple from their surroundings due to their stronger self-gravity. As shown in Fig.2.35, the virial ratio of molecular clouds decreases as the molecular cloud mass increases.

To further verify the above scenario, now we focus on two different types of molecular clouds classified in J. W. Zhou et al. (2024a), i.e. leaf-HFs-A and leaf-HFs-C (A-type and C-type). According to the evolutionary sequence proposed in J. W. Zhou et al. (2024a), A-type clouds are more evolved than C-type clouds. However, C-type clouds as nodes have higher $C_{\rm C}$ and smaller $\bar{l}$. This indicates that, compared to A-type clouds, C-type clouds are generally located in tighter gravitational subnetworks, with shorter distances to other molecular clouds, which makes them more prone to accretion or merging as they evolve. As shown in J. W. Zhou et al. (2024a), the more evolved A-type clouds have significantly greater mass and lower virial ratio than C-type clouds. In Fig.2.34(d), for the relation $M \propto (\bar{l})^{\alpha}$, low-mass clouds exhibit a significantly steeper slope ($\alpha \approx 0.34$) compared to high-mass clouds ($\alpha \approx 0.1$), suggesting that low-mass clouds are more susceptible to accretion or merging during evolution, consistent with the above results.

2.9 Summary

2.9.1 On clump scale

The physical properties and evolution of hub–filament systems in a large sample of protoclusters observed in the ATOMS survey are studied. Moment 0 maps (integrated intensity maps) of the $H^{13}CO^+$ J=1–0 emission and the FILFINDER algorithm are used to identify filaments in the ATOMS sources. Filaments are found to be nearly ubiquitous in protoclusters. With a strict definition, 49% of the sources are classified as hub–filament systems.

Hub–filament structures are shown to exist not only in small-scale (≈0.1 pc) dense cores but also in large-scale clumps/clouds (≈1–10 pc), suggesting that self-similar hub–filament systems across various scales play a crucial role in massive star and star cluster formation throughout different Galactic environments. The proportion of hub–filament systems decreases with increasing dust temperature and luminosity-to-mass ratio, indicating that stellar feedback from HII regions gradually disrupts hub–filament systems as protoclusters evolve. In clumps containing HII regions, hub–filament systems may have been destroyed by stellar feedback as the HII regions expand.

The longest filaments in the ATOMS sources exhibit clear velocity gradients. The approximately symmetric distribution of positive and negative velocity gradients strongly indicates the presence of converging gas inflows along filaments. At scales larger than ≈1pc, velocity gradients are very small, possibly reflecting the dominance of pressure-driven inertial inflow originating from large-scale turbulence or cloud-scale gravitational contraction. Below ≈1pc, velocity gradients increase dramatically

as filament lengths decrease, implying that gravity from the hub or core governs gas infall at such small scales. Assuming the velocity gradients are caused by gas inflow, a linear correlation is identified between the gas inflow timescale and the filament length. These observations imply that high-mass stars in protoclusters may acquire the majority of their mass via longitudinal inflow along filaments.

Any promising models for massive star formation should include self-similar hub-filament systems and filamentary accretion at all scales (from several thousand au to several parsecs).

2.9.2 On cloud scale

The gas kinematics of hub–filament structures in the G333 complex were investigated using ^{13}CO (3–2) emission. The G333 complex contains three primary velocity components, which are well distinguished by average spectra and the velocity distribution in position–position–velocity (PPV) space. The velocity distribution was derived through decomposition of ^{13}CO (3–2) emission using the `GAUSSPY+` algorithm. The overall G333 complex appears to be fragmented into several independent molecular clouds. Within these, some substructures display typical hub–filament morphology, accompanied by characteristic kinematic signatures of hub–filament systems. Disrupted morphologies observed in some infrared-bright structures suggest that stellar feedback is actively reshaping star-forming regions.

Filaments in the G333 complex were identified from integrated intensity maps of ^{13}CO J=3–2. Velocity and intensity data along the filament skeletons were extracted from moment maps. A clear correlation between velocity and density fluctuations was observed, and velocity gradients around intensity peaks were fitted accordingly. The scale-dependent variation in velocity gradients reveals a funnel-like morphology of the velocity field in PPV space. This funnel structure likely results from accelerated inflow toward central hubs and the gravitational contraction of star-forming clouds or clumps. The commonly observed V-shaped velocity pattern along filaments is consistent with a projection of this funnel structure onto the position–velocity (PV) plane, indicating that such features may serve as effective tracers of gravitational collapse motions within molecular clouds.

Gas motions were traced over sufficiently large spatial scales, with the longest filament extending ≈50 pc. Despite this large scale, results remain consistent with those from small-scale ALMA observations. A typical velocity gradient on a 1 pc scale is ≈1.6 km/s/pc. Under the assumption of a freefall model, gravitational accelerations toward each hub were predicted. Observed accelerations show good agreement with these predictions, although uncertainties remain due to limitations in observing only a single component of the acceleration and in the estimated hub masses. Nonetheless, the findings provide strong evidence for gravitational acceleration of material inflowing along filaments into hubs.

The velocity gradients derived from LAsMA and ALMA data are consistent across overlapping scales (less than 5 pc). On larger scales, inflows appear to be driven by more extensive structures, supporting the existence of a hierarchical molecular cloud framework with gas funneled from large to small scales. This large-scale inflow is gravity-driven, suggesting that the G333 molecular clouds are undergoing global gravitational collapse. The funnel-like velocity structure observed in PPV space further reinforces this interpretation.

Despite the complexity introduced by substructures and local fluctuations in velocity and density, the overall large-scale configuration exhibits a loosely organized funnel structure. This suggests that hub–filament systems across multiple scales may be integrated into a hierarchical macroscopic system, connected through gravitational centers at different scales. In short, the variation in velocity gradients across scales suggests a funnel-like structure in the velocity field within PPV space. This is indicative of a smooth, continuously increasing velocity gradient from large to small scales, and it is thus consistent with gravitational acceleration.

2.9.3 From cloud to core scale

G332.83–0.55 is a subregion within the G333 complex. High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3–2) observations reveal a clear hub–filament morphology on scales of ≈10 pc, with the central hub corresponding to the ATLASGAL clump AGAL332.826–00.549. Higher-resolution ALMA observations using HCO^+ (1–0) and $H^{13}CO^+$ (1–0) further confirm the presence of a hub–filament structure within the central region of G332.83–0.55. This indicates that the subregion contains at least two hierarchical levels of hub–filament systems. The combination of high-resolution single-dish and interferometric observations enables the tracing of gas inflows from cloud to core scales.

The G332.83–0.55 region is undergoing gravitational collapse with rotation. Velocity fields derived from both LAsMA and ALMA data show distinct velocity discontinuities near the gravitational center—identified as a protocluster—with redshifted and blueshifted peaks. Based on the observed velocity difference and spatial separation between the two peaks, the rotational velocity is estimated to be ≈3.29 km s^{-1}, and the centrifugal barrier radius is ≈0.21 pc. The mass of the central protocluster is consequently estimated to be ≈264 $M_\odot$, consistent with the mass derived from 3 mm dust continuum emission.

2.9.4 Decomposition of molecular clouds

The G333 complex and the G331 GMC were decomposed into multi-scale hub–filament (HF) structures using high-resolution ^{13}CO (3–2) data from LAsMA observations. Intensity peaks as potential hubs were identified from the integrated intensity map of ^{13}CO (3–2) emission, utilizing the dendrogram algorithm. Around intensity peaks, average spectra were extracted within regions extended by a factor of 2.5 to decompose the velocity components and determine velocity ranges. Final identification of hub–filament systems was performed using cleaned moment 0 maps constrained to these velocity ranges, effectively excluding artifacts resulting from the overlap of uncorrelated velocity components.

Compared to non-HFs, the identified HFs exhibit significantly higher density contrasts, greater masses, and lower virial ratios. The FILFINDER algorithm was employed to identify and characterize filaments within the HFs. Velocity and intensity profiles along the filament skeletons were extracted, and velocity gradients around the intensity peaks were fitted. These velocity gradients display scale-dependent variations that approximately follow a freefall model over a mass range of ≈100–10000 $M_\odot$, aligning with the mass distribution of local dense gas structures. This suggests that local dense structures as gravitational centers accrete surrounding diffuse gas and produce the observed velocity gradients. Velocity gradient measurements provide evidence of gas inflow within these structures, offering kinematic support that they are hub-filament systems.

The proportion of HFs associated with clumps and radio sources is markedly higher than that of non-HFs. Additionally, HFs exhibit significantly higher *Spitzer* 8 μm peak intensities compared to non-HFs. The luminosity-to-mass ratios of ATLASGAL clumps associated with HFs are also higher. These results indicate that HFs represent a more evolved phase in the star formation process. Within non-HFs, those structures associated with clumps display higher density contrasts, greater masses, and lower virial ratios than those lacking such associations. Similarly, among HFs, structures linked to radio sources show enhanced density contrast, larger mass, and lower virial ratios compared to those without radio counterparts. Collectively, these findings support the existence of an evolutionary sequence progressing from non-HFs to HFs.

Despite the higher density contrasts in HFs, the average densities of non-HFs are comparable with that of HFs, indicating that non-HFs are not inherently low-density structures. Instead, their relatively uniform density distributions suggest the absence of a pronounced gravitational focusing process. This

implies that non-HFs currently lack a strong central gravitational potential capable of inducing significant contrast. Therefore, the density contrast effectively measures the extent of gravitational collapse and the strength of the gravitational center of the structure that definitively shape the hub-filament morphology. In terms of reflecting the development and evolutionary stage of a structure, density contrast is more crucial than density itself.

Molecular clouds are decomposed into hundreds of hub–filament structures spanning a range of spatial scales. Integrating these findings with the kinematic analyses from J. W. Zhou et al. (2023) and J. W. Zhou et al. (2024g), it is proposed that molecular clouds constitute network-like structures formed through gravitational coupling among multi-scale HFs. In this framework, hub–filament systems act as the fundamental building blocks of molecular clouds. The hubs as local gravitational centers represent the "nodes" of the network and are the primary sites of star formation. Clumps observed within molecular clouds are effectively equivalent to these hubs.

The networked nature of molecular clouds suggests that these hubs operate with a degree of independence. While some hubs evolve more rapidly, forming HII regions, their relatively isolated positions within the network prevent early feedback effects from significantly altering the physical conditions of neighboring hubs. This conclusion is supported by the findings in J. W. Zhou et al. (2024g) and J. W. Zhou et al. (2024e).

2.9.5 On galaxy scale

The kinematics and dynamics of gas structures on galaxy-cloud scales were investigated in two spiral galaxies, NGC 5236 (M83) and NGC 4321 (M100), using the CO (2–1) line. Hierarchical gas structures were directly identified based on the 2D integrated intensity map of CO (2–1) emission. The average spectrum for each structure was extracted to analyze its velocity components and gas kinematics. To account for large-scale velocity gradients due to galaxy rotation, which contribute to the non-thermal velocity dispersion, the large-scale velocity gradients were subtracted from the PPV data cube using a constructed gas dynamical model before extracting the average spectra of the identified structures.

Scaling relations among velocity dispersion (σ), effective radius (R), and column density (N) were examined across all structures. It was found that $\sigma - N * R$ relation consistently exhibits a stronger correlation than either $\sigma - N$ or $\sigma - R$ relations. The observed correlations between velocity dispersion and column density suggest a potential connection to gravitational collapse, supported by the measured velocity gradients. However, the slopes of the $\sigma - N * R$ relations were significantly shallower than the anticipated value of 0.5, indicating a deceleration from the behavior characteristic of a pure free-fall gravitational collapse.

The FILFINDER algorithm was employed to identify and characterize filaments within the galaxy using integrated intensity maps. Velocity and density fluctuations along these filaments were analyzed by fitting local velocity gradients around intensity peaks, with global large-scale velocity gradients attributed to the galaxy's rotation removed beforehand. Statistical analysis of the fitted velocity gradients on galaxy-cloud scales revealed the same gas kinematic modes observed on cloud-clump and clump-core scales.

The variations in velocity gradients at both small and large scales (with a boundary around 500 pc) are roughly comparable with expectations from gravitational free-fall, with central masses of $\approx 10^5$–$10^{6.5}$ $M_\odot$ and $\approx 10^6$–$10^{7.5}$ $M_\odot$, respectively. This suggests that maintaining velocity gradients on larger scales requires larger masses, and that larger masses correspond to larger scales. Therefore, the larger-scale inflow is likely driven by the larger-scale structure, which may arise from the gravitational clustering of smaller-scale structures. This is consistent with the presence of hierarchical or multi-scale hub-filament structures within the galaxy, with gas inflows occurring from large to small scales.

In the freefall model, the velocity gradient and scale satisfy $\nabla v \propto R^{-1.5}$. However, in the observations, $\nabla v \propto R^{-0.8}$, further indicating a slowing down of the pure free-fall gravitational collapse.

2.9.6 Decomposition of molecular gas in a galaxy

The spiral galaxy NGC 628 was decomposed into multi-scale hub-filament structures using the CO (2–1) line map. Intensity peaks, which are potential hubs, were identified based on the integrated intensity map of CO (2–1) emission. For all intensity peaks, the average spectra for each structure were extracted to decompose their velocity components and define their velocity ranges. The final identification of hub-filament structures was based on cleaned moment 0 maps, which were constructed by restricting to the fixed velocity ranges to exclude potential overlap of uncorrelated velocity components. In practice, these steps were largely unnecessary, as nearly all structures exhibited a single velocity component. All identified structures were classified into three categories: leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C. For leaf-HFs-C, the density contrast between the hub and surrounding diffuse gas is not pronounced. Both leaf-HFs-A and leaf-HFs-B have clear central hubs, with leaf-HFs-A exhibiting the best hub-filament morphology, the highest density contrast, the largest mass, and the lowest virial ratio.

The FILFINDER algorithm was employed to identify and characterize filaments within the 185 leaf-HFs-A structures. Local velocity gradients were fitted around intensity peaks, after removing the global large-scale velocity gradients attributed to the galaxy's rotation. The velocity gradient measurements provided evidence of gas inflow within these structures, serving as kinematic support for their classification as hub-filament structures. Leaf-HFs-C are not necessarily low-density structures; rather, their density distribution is relatively uniform. An evolutionary sequence from leaf-HFs-C to leaf-HFs-A may exist, with leaf-HFs-C lacking a distinct gravitational collapse process that would generate significant density contrast.

The number of associated 21 μm and Hα structures, as well as the peak intensities of 7.7 μm, 21 μm, and Hα emissions, decrease from leaf-HFs-A to leaf-HFs-C. Additionally, the spatial separations between the intensity peaks of CO and 21 μm structures are larger in leaf-HFs-A compared to leaf-HFs-C. These observations suggest that leaf-HFs-A are more evolved than leaf-HFs-C. Combining the local star formation rate (SFR) derived from JWST 21 μm emission with the timescale revealed by the spatial separation between CO and 21 μm emissions yields mass estimates for embedded star clusters in molecular clouds, which are comparable to those obtained from spectral energy distribution (SED) fitting. As expected, the cluster mass of leaf-HFs-A is much larger than those of leaf-HFs-B and leaf-HFs-C.

A clear correlation between density contrast and virial ratio was observed, with the density contrast decreasing as the virial ratio increases. The density contrast effectively measures the extent of gravitational collapse and the depth of the gravitational potential well of the structure that shapes the hub-filament morphology.

Together with the kinematic evidence presented in J.-W. Zhou & T. Davis (2024), a coherent picture emerges in which molecular gas in spiral galaxies is structured into a network through the gravitational coupling of multi-scale hub–filament systems, consistent with the global hierarchical collapse scenario (Vázquez-Semadeni et al., 2019). Molecular clouds act as nodes within the hub–filament network, serving as local gravitational centers and primary sites of star formation.

2.9.7 Gravitational network analysis

Observations reveal that molecular gas in spiral galaxies forms a network of interconnected systems through the gravitational coupling of multi-scale hub-filament structures. Based on this physical picture, a gravitational network is constructed for the galaxy NGC 628. By visualizing the gravitational interactions

between molecular clouds in this network, both the gravitational interactions and the physical properties of the clouds are characterized using geometry-based network metrics. In the gravitational network, a strong correlation exists between the geometric and physical properties of the nodes (i.e. clouds). Metrics such as node degree (k), betweenness centrality (C_B), clustering coefficient (C_C), and average physical path length ($\bar{l}$) show strong correlations with node mass.

High-mass molecular clouds exhibit lower C_C and longer $\bar{l}$, indicating that high-mass nodes tend to have fewer neighbors. Additionally, a positive correlation between node mass and the distance to the nearest neighbor is observed. For subnetworks consisting of a node and its neighbors (node degree ≥ 20), nodes with greater mass and longer $\bar{l}$ are more significantly displaced from the geometric center of the subnetwork. These findings suggest that high-mass clouds, during their formation and evolution, deplete nearby gas through accretion or merger, resulting in more isolated network characteristics. This is consistent with the observed decrease in the virial ratio of molecular clouds as their mass increases.

For molecular clouds at different evolutionary stages classified in J. W. Zhou et al. (2024a) (i.e. A-type and C-type), a comparison reveals that C-type clouds, which are less evolved, have higher C_C and shorter $\bar{l}$ compared to A-type clouds. This indicates that C-type clouds are generally located in tighter gravitational subnetworks, with closer proximity to other molecular clouds. Consequently, they are more prone to accretion or merging during evolution. For the relation $M \propto (\bar{l})^{\alpha}$, the slope for low-mass clouds ($\alpha \approx 0.34$) is significantly steeper than that for high-mass clouds ($\alpha \approx 0.1$). This also indicates that low-mass clouds are more susceptible to accretion or merging as they evolve.

CHAPTER 3

Tidal effects

The content presented in this chapter are reproduced from the publications by J. W. Zhou et al. (2024f) (Sec.3.2), J. W. Zhou (2025) (Sec.3.3), and J. W. Zhou & Dib (2025) (Sec.3.4). The studies in Chapter.2 demonstrate the multi-scale gravitational coupling between molecular gas structures. In this chapter, we will focus on analyzing the tidal interactions between them.

3.1 Data

Edenhofer et al. (2024) offers the most accurate 3D dust map to date for the solar neighborhood within 1.25 kpc, employing a novel Gaussian process prior approach to reduce the impact of fingers-of-god artifacts. This 3D map was generated using stellar distance and extinction estimates from X. Zhang et al. (2023) derived from *Gaia* spectra. There are 12 samples from their inferred 3D dust extinction distribution. Our calculation is based on the posterior mean of their reconstruction, which is the average of these 12 samples. Using the method described in O'Neill et al. (2024), we converted the differential extinction in the 3D map to the number density of hydrogen nuclei, n, independent of phase ($n = n_{\mathrm{H_I}}+2n_{\mathrm{H_2}}$). Then the volume density is, $\rho = 1.37 m_{\mathrm{p}} n$, where m_{p} represents the mass of a proton, and the factor 1.37 accounts for the contribution of helium to the total mass, based on cosmic abundance ratios relative to hydrogen.

We utilized the full position-position-velocity (l-b-v) cube of CO [1] from Dame et al. (2001). The spatial range of this cube covers all Galactic longitudes and $\pm\ 30^0$ in Galactic latitude. The velocity range is $\pm$ 320 km s^{-1}.

3.2 On clump scale

A local dense structure evolves under its own gravity, but as a gravitational center, it also exerts influence on its surroundings while simultaneously being affected by the gravitational forces of nearby structures. The gravitational and tidal fields are inherently interconnected. This hierarchical or multi-scale gravitational coupling among gas structures means the widespread presence of tidal interactions between them. A gas structure is inevitably impacted by the cumulative tidal interactions from the surrounding material. Such interactions may hinder gravitational collapse, suppress instability growth, and inhibit star formation within it. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024f).

[1] `https://lweb.cfa.harvard.edu/rtdc/CO/`

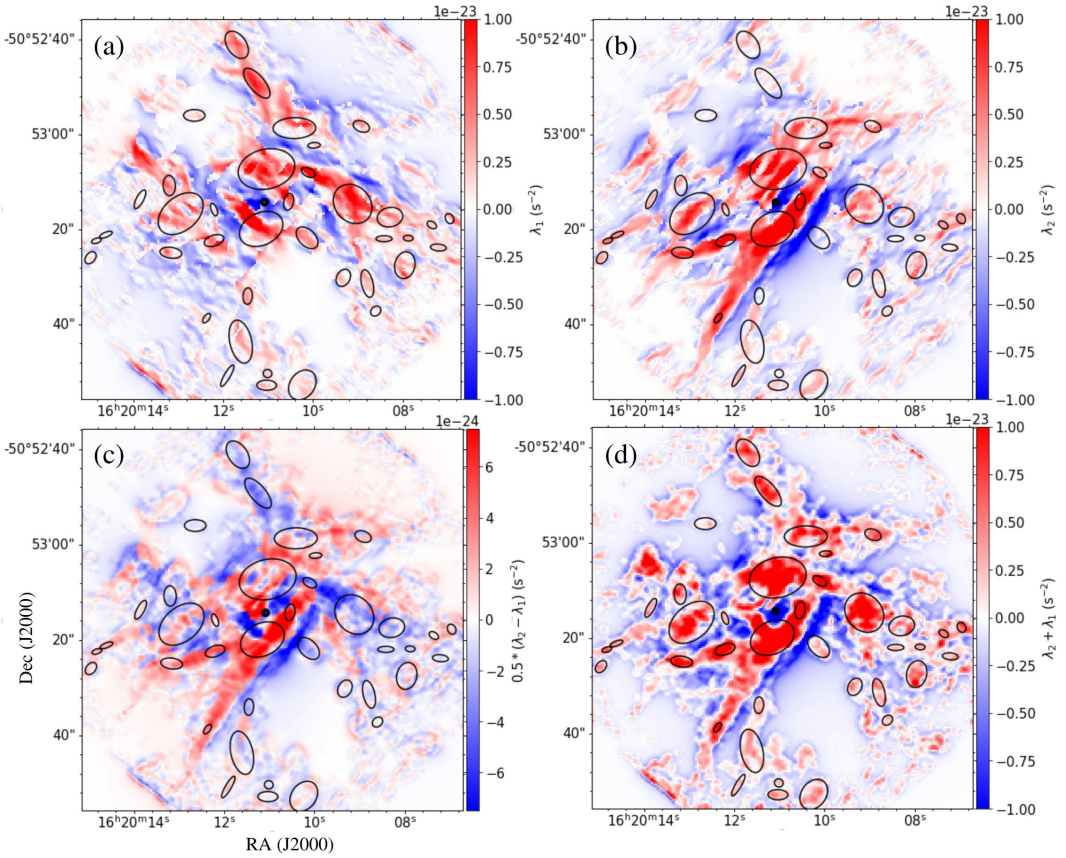


Figure 3.1: Different components of the projected tidal field in the FOV of the ALMA observation.

3.2.1 Tidal field

To fully consider the tidal interactions, we directly derived the tidal field according to the density distribution. Only the projected surface density map is available in the observations. To align with other projected physical quantities derived from the observations, we also calculated the 2D projected tide. We note that 3D calculations can refer to Ganguly et al. (2024a) and G.-X. Li (2024).

Calculating the gravitational potential field $\phi(x, y)$ based on the surface density map $\Sigma(x, y)$ has been discussed in Gong & Ostriker (2011) and He et al. (2023). As described in J. W. Zhou et al. (2024f), the tidal tensor has only two eigenvalues in the case of 2D, that is to say λ_i (i=1,2). The tidal tensor of a structure includes both self-gravity and external tides, that is, $\mathbf{T} = \mathbf{T}_{\rm ext} + \mathbf{T}_{\rm int}$. For three limiting cases, the external tides can be measured by λ_1, λ_2, and $(\lambda_2 - \lambda_1)/2$, and they are called $\mathrm{T_{d,1}}$, $\mathrm{T_{d,2}}$, and $\mathrm{T_{d,3}}$, respectively. Fig.3.1 displays different components of the tidal tensor.

3.2.2 Pixel-by-pixel computation

We directly identified hierarchical gas structures based on the 2D integrated intensity map of $\mathrm{H^{13}CO^+}$ (1−0) emission, and then extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. We derived the column density by combining $\mathrm{H^{13}CO^+}$ (1−0) and $\mathrm{HCO^+}$ (1−0) emissions using the same method described in J. W. Zhou et al. (2024e) by the local thermodynamic equilibrium (LTE) analysis. Then the masses and virial ratios of all leaf structures were calculated.

Below, we explain how we tested the tidal tensor analysis described above using an original method.

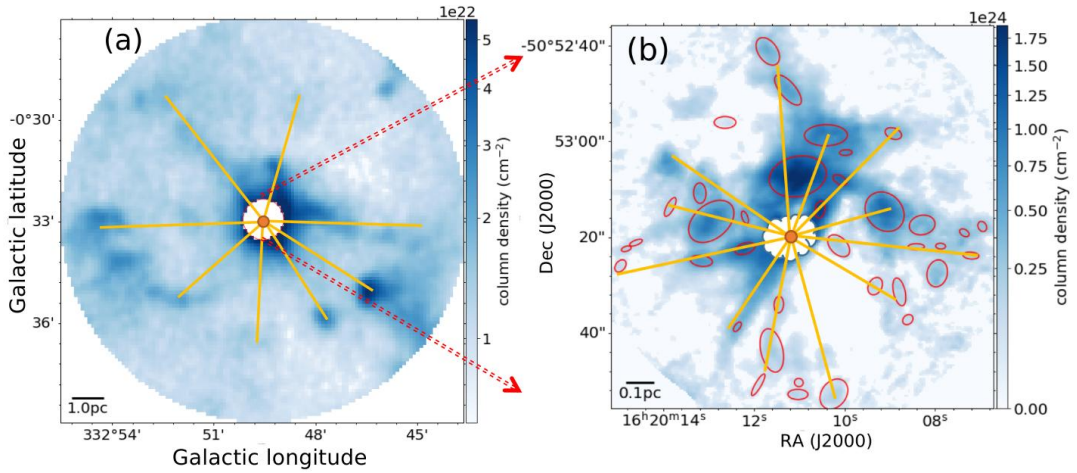


Figure 3.2: Computed structure first masked from the density map. The external tides from all outside material sustained by a point in the computed structure was calculated pixel by pixel and is illustrated by orange lines. (a) Background representing the column density map derived from ATLASGAL+Planck 870 μm data. The material in the FOV of the ALMA observation (white region) was masked. (b) Background representing the column density map derived from $H^{13}CO^+$ J=1-0 and HCO^+ J=1-0 lines from the ALMA observation by the LTE analysis.

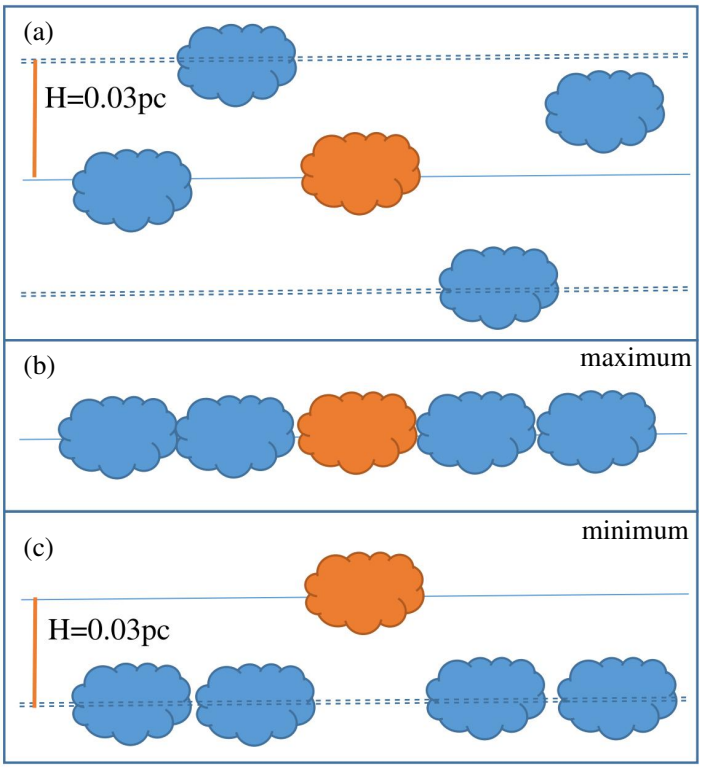


Figure 3.3: Edge-on view of G332.83-0.55. The calculated gas structure is in red, and all external material apart from the calculated structure is in blue. (I) The gas distribution in G332.83-0.55 can be everywhere. (II) In 2D computation, all molecular gases are assumed to be concentrated on the same plane. Thus, we obtained a maximal estimate for the tidal strength. (III) Taking a characteristic scale height of the molecular gas as 0.03 pc, all external material apart from the calculated structure were moved vertically to 0.03 pc away. Thus we obtained a minimal estimate of the tidal strength.

For a structure A, located at a distance R from the center of another structure B, of mass M, the tidal strength sustained by structure A due to the external gravity from structure B is

$$T \approx \frac{2GM}{R^3}. \tag{3.1}$$

The total tidal strength at a pixel measures the total deformation due to the external tides at that point. A gas structure is neither a rigid body nor a point mass, but it is a parcel of gas with malleability. Most gas structures are also not spherically symmetric, and their morphologies can be highly complex. Although the external gravitational forces from all directions may cancel out within a structure, the resulting tidal forces inside the structure do not. Therefore, we used the scalar superposition to measure the overall tidal strength in the structure from the external gravity.

Using the mask of the dendrogram structure, we divided the density map into two parts, that is, the structure itself and all other materials excluding this structure. For the external tides from all outside material sustained by a point in structure A, a simple calculation involves using equation.3.1 to directly calculate the tidal strength at that point pixel by pixel, as illustrated in Fig.3.2(b). The average tidal strength of all the pixels in structure A can then represent the average tidal strength sustained by the structure.

As illustrated in Fig.3.2(b), the above computation is limited to a 2D plane. In equation.3.1, the tidal strength is very sensitive to the distance. If the external structure and the calculated structure are not in the same plane, the 2D calculation underestimates the distance between them and thus overestimates the tidal strength (maximum computation, T_{max}). However, we only have the 2D projection in the observation. The 3D structure of the molecular gas is unknown, but we can consider an extreme case. Using the average effective radius (≈0.03 pc) of all identified structures as the characteristic scale height of G332.83-0.55, all external material apart from the calculated structure was moved vertically to 0.03 pc away, as illustrated in Fig.3.3(c). This is unrealistic, given the continuity of the matter distribution in a cloud. As a limiting case, the estimated tidal strength should be smaller than the actual value (minimal computation, $T_{min,s}$).

The other minimal computation ($T_{min,v}$) is the vector superposition of tides, which allows the tidal interactions to directly cancel each other out. Using the calculated point in the target structure as the coordinate origin to build a 2D coordinate system, all external tides were decomposed to the coordinate axes. Then we calculated the overall tidal strength at the calculated point after performing vector superposition along each axis.

The dendrogram structures are embedded in a larger-scale gas environment. The material outside of the FOV of ALMA observation may also produce considerable external tides (T_{out}) on small-scale dendrogram structures. Using the H_2 column density derived from ATLASGAL+Planck 870 μm in J. W. Zhou et al. (2024g), the matter in the FOV of ALMA observation was first masked, as shown in Fig.3.2(a). Then we calculated the external tides at each pixel in the FOV. The averaged tidal strength of all pixels is $\approx 1.4 \times 10^{-25}$ s^{-2}, which is very small and can be ignored. As shown in J. W. Zhou et al. (2024f), the components $T_{d,1}$, $T_{d,2}$, and $T_{d,3}$ are comparable. Especially, they are also comparable with $T_{min,s}$ and $T_{min,v}$. As expected, T_{max} is significantly larger than others. For $T_{d,1}$, $T_{d,2}$, $T_{d,3}$, $T_{min,s}$, and $T_{min,v}$, $T_{min,v}$ is slightly larger than others and the value of $T_{min,s}$ is in the middle. As a compromise, we use the value of $T_{min,s}$ to represent the external tides in the subsequent analysis.

3.2.3 Tide-regulated gravitational collapse

The average strength of the external tides ($T_{min,s}$) of a structure was used to measure the deformation of the structure due to the external gravity. The effective velocity dispersion due to the external tides of the

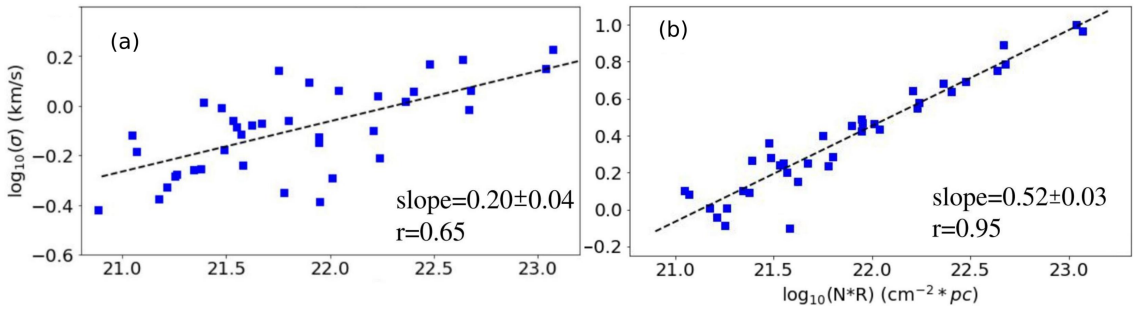


Figure 3.4: (a) $\sigma - N * R$ relation, where σ is the local gas velocity dispersion. (b) $\sigma_s - N * R$ relation, where $\sigma_s = \sqrt{\sigma_t^2 + \sigma^2}$. We note that "r" represents the Pearson correlation coefficient.

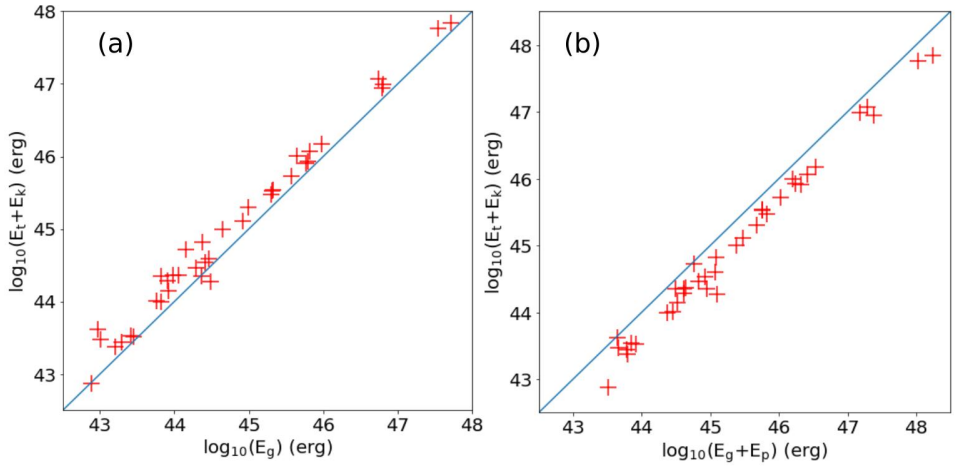


Figure 3.5: Comparison between the structural energies, i.e., the gravitational potential energy E_g, the kinetic energy E_k, the tidal energy E_t, and the external pressure energy E_p.

structure is, $\sigma_t^2 \approx T_{\mathrm{min,s}} * R^2$, and the corresponding energy

$$E_{\mathrm{t}} \approx \frac{1}{2} M * T_{\mathrm{min,s}} * R^2. \tag{3.2}$$

The gravitational potential energy E_{g} and internal kinetic energy E_{k} were calculated in the same way as presented in J. W. Zhou et al. (2024g).

As shown in Fig.3.4(a), the slowing down of a pure free-fall gravitational collapse as revealed by the shallow $\sigma - N * R$ relation means that there are some physical processes working against gravitational collapse, such as tidal tearing from external gravity. The effect of external tides on the gravitational collapse of gas structures can also be reflected in the scaling relation. After adding the contribution of the tidal interactions to the observed velocity dispersion, the effective velocity dispersion $\sigma_{\mathrm{s}} = \sqrt{\sigma_{\mathrm{t}}^2 + \sigma^2}$, then the $\sigma_{\mathrm{s}} - N * R$ relation was fitted again. As shown in Fig.3.4(b), now the slope is close to 0.5, and the correlation also gets stronger. Thus, the deformation caused by external tides may effectively slow down the pure free-fall gravitational collapse of gas structures. However, in Fig.3.5(a), the sum of E_{k} and E_{t} is notably larger than E_{g}.

The velocity dispersion is a manifestation of gas perturbations due to many physical processes, including gravitation collapse, tidal tearing, and turbulence. Thus, E_{k} should be overestimated. Moreover, when $E_{\mathrm{k}} + E_{\mathrm{t}} > E_{\mathrm{g}}$, it does not mean that the structure cannot be gravitationally bound. A structure is

deformed by external tides from the surrounding material, and at the same time, the surrounding material also exerts a confining pressure on the structure.

Previous studies have suggested that the external pressure provided by the larger-scale molecular gas might help to confine dense structures in molecular clouds (Spitzer, 1978; McKee, 1989; B. G. Elmegreen, 1989; Ballesteros-Paredes, 2006; Kirk et al., 2006; Dib et al., 2007; C. J. Lada et al., 2008; Pattle et al., 2015; Kirk et al., 2017; S. Li et al., 2020; J. W. Zhou et al., 2024g). External pressure can have various origins; here we mainly consider the external pressure from the ambient cloud for each decomposed structure using

$$P_{\rm cl} = \pi G \bar{\Sigma} \Sigma_r, \tag{3.3}$$

where $P_{\rm cl}$ is the gas pressure, $\bar{\Sigma}$ is the mean surface density of the cloud, and $\Sigma_{\rm r}$ is the surface density at the location of each structure (McKee, 1989; Kirk et al., 2017). The external pressure energy can be calculated as

$$E_{\rm p} = -4\pi P_{\rm cl} R^3. \tag{3.4}$$

As shown in Fig.3.5(b), after considering the external pressure, $E_{\rm k} + E_{\rm t} < E_{\rm g} + E_{\rm p}$. The external tide tries to tear up the structure, but the external pressure on the structure prevents this process from occurring. The counterbalance between the external tide and external pressure hinders the gravitational collapse of the structure, which can also slow down a pure free-fall gravitational collapse. Gravity can play multiple roles in cloud evolution. Inside dense regions, it drives fragmentation and collapse, and outside these regions, it can suppress the collapse of low-density gas through extensive tidal forces and drive accretion (G.-X. Li, 2024). This work confirms that tidal interactions can significantly influence the dynamics and kinematics of gas structures and regulate their gravitational collapse.

3.3 On cloud scale

Due to the diffuse and complex morphology of ISM, a complete tidal calculation would be complex. One should derive the gravitational potential distribution from the 3D density distribution and then calculate the tidal field according to the gravitational potential, as presented in G.-X. Li (2024). In observations, we can usually only obtain the 2D projected surface density map. However, building on the accurate distances enabled in the *Gaia* era, we have been able to infer the 3D density distribution of the ISM by the 3D dust mapping technique (Rezaei Kh. et al., 2018; B. .-. Chen et al., 2019; Green et al., 2019; Lallement et al., 2019; Hottier et al., 2020; Leike et al., 2020; Edenhofer et al., 2024). In this work, we first calculated and decomposed the tidal field using the 3D dust map, then explored methods to estimate the tidal strength within the clouds induced by external gravity. Finally, we assessed the impact of external tides on the physical properties of molecular clouds. The work presented in this section is reproduced from the publication by J. W. Zhou (2025).

3.3.1 Structural identification

Based on the 3D density distribution of hydrogen nuclei derived from the 3D dust map, Fig.3.6(a) shows the surface density map of hydrogen nuclei projected on the *l*-*b* plane. We applied the dendrogram algorithm to identify local dense structures based on this surface density map. Then the CO (1-0) cube was utilized to derive the velocity dispersion of the identified structures. Since the CO (1-0) cube is in the Galactic coordinates (*l*-*b*-*v*), we also used the 3D dust map provided in the Galactic coordinate system (*l*-*b*-*d*)[2]. This allows both cubes to be analyzed by the same way, as they have the same data structure.

[2] https://zenodo.org/records/10658339

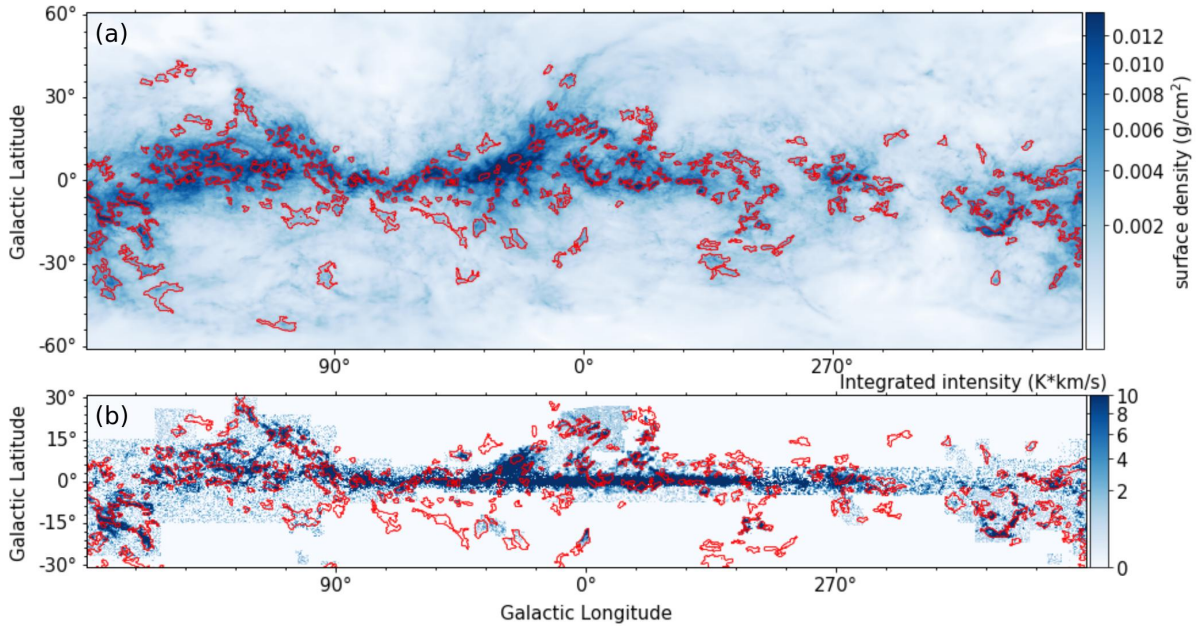


Figure 3.6: Structural identification. (a) The surface density of hydrogen nuclei on the l-b plane. Red contours are the structures identified by the dendrogram algorithm; (b) The integrated intensity map of CO (1-0) line emission. Red contours are the same with panel (a).

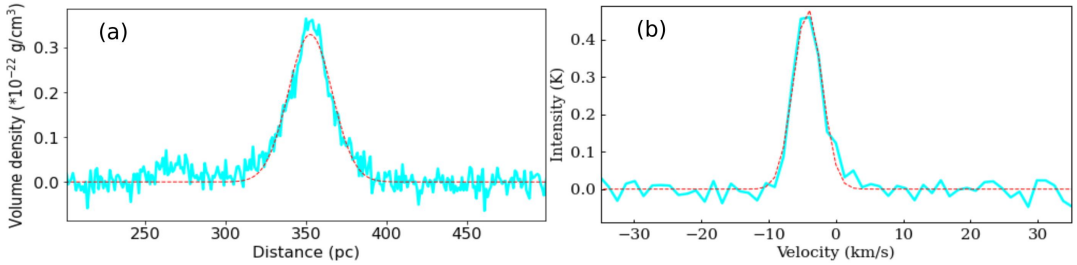


Figure 3.7: The average "spectra" (cyan lines) of a single structure extracted from the l-b-v and l-b-d cubes, respectively. (a) The averaged intensity distribution along the velocity axis; (b) The averaged density distribution along the distance axis. Red dashed lines represent the Gaussian fitting.

As shown in Fig.3.6(a), the structures identified by the dendrogram correspond well to the surface density distribution. In Fig.3.6(b), CO and dust also show a good correspondence in their distribution, but CO emission is primarily concentrated in the Galactic plane.

Using the mask of each identified structure in Fig.3.6(a), we extracted the average "spectra" according to the l-b-v and l-b-d cubes. Specifically, for each identified structure, we averaged the intensity or density values across all pixels within the masked region and analyzed the intensity distribution along the velocity axis and the density distribution along the distance axis, as shown in Fig.3.7. This approach allows us to investigate the characteristic velocity and distance distributions of the structures, providing insights into their kinematics and spatial properties.

If an identified structure is a single structure without overlap, we would expect both types of "spectra" to exhibit a single peak, as shown in Fig.3.7. In total, 90 structures satisfy this condition. From the "spectra", we can obtain the systematic velocity, the velocity dispersion (σ_g), the distance and the thickness of each structure. The thickness is the full width at half maximum (FWHM) of the line profile, displayed in Fig.3.7(a). The dendrogram algorithm approximates the morphology of each structure as an ellipse. Using the long and short axes (l_1 and l_2) of the effective ellipse and the half-thickness (l_3), we

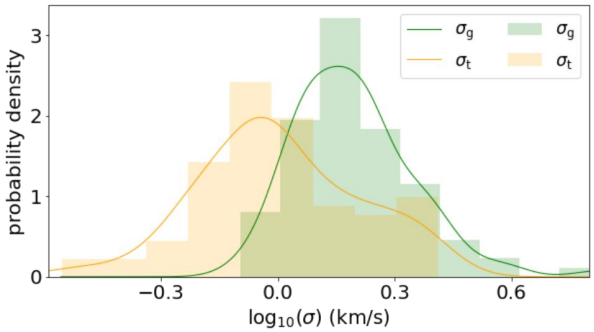


Figure 3.8: The KDE (kernel density estimation) and the histogram of the velocity dispersion of the clouds. σ_t is the equivalent velocity dispersion derived from the average tidal strength. σ_g is the velocity dispersion measured from CO emission.

can estimate a 3D effective radius for each structure, $R \approx (l_1 * l_2 * l_3)^{1/3}$.

3.3.2 Tidal strength

As presented in J. W. Zhou (2025), for a subregion of the 3D map, based on the 3D density distribution, the gravitational potential was computed by solving the Poisson equation, then the tidal tensor was derived from the gravitational potential. For a cloud, its tidal tensor can be divided into two parts (internal and external) due to the matter inside and outside the cloud. Different decomposition methods result in four possible maximum eigenvalues ($|\lambda_{\rm ext,max}|$) of the external tidal tensor ($\mathbf{T}_{\rm ext}$). An independent and original pixel-by-pixel computation method was used to select the optimal $|\lambda_{\rm ext,max}|$. For the eigenvalues, we found that the optimal $\mathbf{T}_{\rm ext}$ shows the same formalism with the tidal tensor of a point-mass. Therefore, for the external tides exerted by surrounding material at a point within a cloud, the external material can be approximated as a collection of point masses, each generating its own tidal field. The combined effect of the external tides is equivalent to the superposition of the tidal fields produced by all these point masses. Essentially, this approach is analogous to the pixel-by-pixel computation.

Our current computational resources are insufficient to directly calculate the tidal tensor of the entire 3D map. Therefore, we adopted the pixel-by-pixel computation approach to calculate the average tidal strength ($\lambda_{\rm pp}$) of all structures in the subsequent analysis. This method only considers the structure being calculated and its neighboring matter each time. Typically, accounting for matter within 8–10 times the structure's radius is sufficient, as tidal strength decreases rapidly with distance.

3.3.3 Velocity dispersion

We estimated an equivalent velocity dispersion for the clouds according to the average tidal strength, as done in J. W. Zhou et al. (2024f). The tidal strength defined in this work is the tidal acceleration per unit distance (Stark & Blitz, 1978). Thus, for a cloud with radius R and mass M, $\lambda_{\rm pp} * R$ is the tidal acceleration. The velocity dispersion of a gas structure measures the intensity of internal gas motion within the structure. The tides induced by external gravity pervade the interior of the gas structure. We propose that these ubiquitous gravitational differences serve as the driving force for the motion of gas parcels within the structure. The work done by tidal forces is converted into kinetic energy, which corresponds to the equivalent velocity dispersion, σ_t, i.e.

$$\frac{1}{2}M\sigma_t^2 \approx \int_0^R M\lambda_{\rm pp} R\, dR. \tag{3.5}$$

Then, we have

$$\sigma_t \approx \sqrt{\lambda_{\rm pp} * R^2}. \tag{3.6}$$

Since σ_t is derived from the average tidal strength over the entire structure, it is isotropic and can be directly compared to the observed velocity dispersion of the structure (σ_g). As shown in Fig.3.8, σ_t is roughly comparable with σ_g in magnitude, which means that tidal effects from external gravity can significantly contribute to the velocity dispersion of the clouds. A relatively smaller value of σ_t is also reasonable, as other physical processes (such as gravitational collapse and feedback activities) also contribute to the overall velocity dispersion of the clouds.

3.4 On galaxy-cloud scale

In J. W. Zhou et al. (2024a), we decomposed the molecular gas in the spiral galaxy NGC 628 (M47) into multi-scale hub-filament structures using the CO (2-1) line map. All molecular clouds as potential hubs were classified into three categories: leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C. For leaf-HFs-C, the density contrast between the hub and the surrounding diffuse gas is not significant. Both leaf-HFs-A and leaf-HFs-B feature distinct central hubs, but leaf-HFs-A exhibit the best-defined hub-filament morphology. Leaf-HFs-A also have the highest density contrast, the largest mass and the lowest virial ratio. Leaf-HFs-C are not necessarily low-density structures, but their density is more uniformly distributed. Currently, leaf-HFs-C lack a clear gravitational collapse process that would produce a pronounced density contrast. We found a clear correlation between density contrast and the virial ratio, with the density contrast decreasing as the virial ratio increases. The density contrast (C) effectively gauges the degree of gravitational collapse and the depth of the gravitational potential well that shape the hub-filament morphology. In terms of understanding the development and evolutionary stage of molecular clouds, the density contrast is more indicative than density itself. The study of the formation and evolution of three types of structures mentioned above is important to understand what are the main mechanisms that regulate star formation in galaxies. Molecular gas in galaxies ranges from diffuse gas to regions of localized clumping, followed by a further increase in density due to gravitational collapse and accretion, which eventually results in star formation. The physical factors that influence the development of molecular cloud's density contrast are the same factors that affect star formation in galaxies. As shown in J. W. Zhou et al. (2024a), as long as C is different, other physical quantities will also differ significantly. C can be used to represent the star formation capability of molecular clouds. The correlation between C and various physical quantities can reflect the impact of those physical processes on star formation in molecular clouds. In this work, we are not concerned with the formation of molecular clouds, but rather with the physical processes that create variations in density contrast among molecular clouds in NGC 628. Particularly, we focus on the role of different types of tides. The work presented in this section is reproduced from the publication by J. W. Zhou & Dib (2025).

3.4.1 Density contrast distribution

For all clouds in NGC 628, the median value of the density contrast C is ≈ 1.63. Based on the distance to the galaxy center ($R_{\rm G}$), we divided the galaxy into different rings. As shown in Fig.3.9(a), only rings with a radius of ≤ 4 kpc are complete. In Table.3.1, for $R_{\rm G} \leq 4$ kpc, the number of high-density contrast clouds rises notably as $R_{\rm G}$ increases. In order to extend the comparison to a wider spatial range within the galaxy, we calculate the proportion of clouds with $C > 1.63$ and $C < 1.63$ within each ring. The distributions of high- and low-density contrast clouds are relatively uniform within each ring, meaning that they are not biased toward any specific part of the ring. Despite the incompleteness of the outer rings,

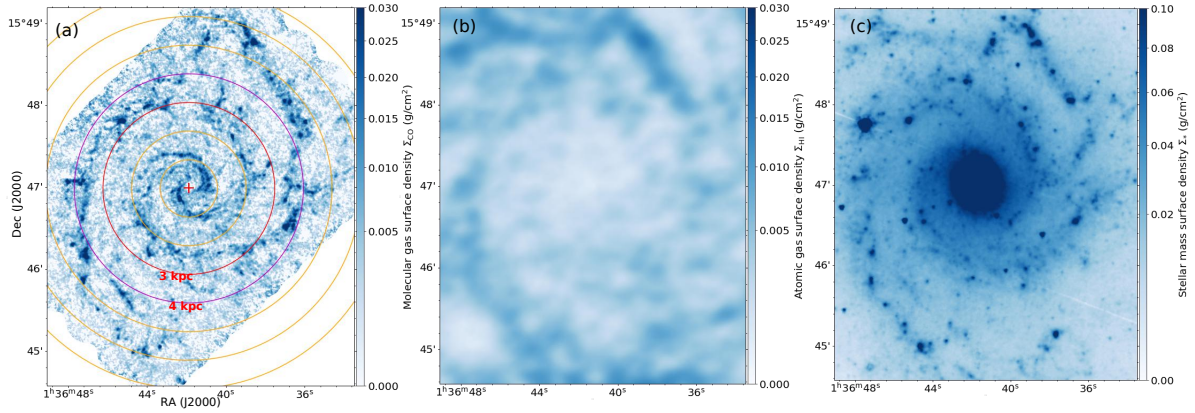


Figure 3.9: Surface density maps of CO (2-1), HI and stellar mass.

Table 3.1: The ratio of high-density contrast clouds to low-density contrast clouds at different distances from the galactic center.

distance (kpc)	$C > 1.63$	$C < 1.63$	proportion
$R_G < 1$	14	39	0.358
$1 < R_G < 2$	45	63	0.714
$2 < R_G < 3$	68	77	0.883
$3 < R_G < 4$	94	68	1.382
$4 < R_G < 5$	88	75	1.173
$5 < R_G < 6$	57	42	1.357
$6 < R_G < 7$	20	18	1.111

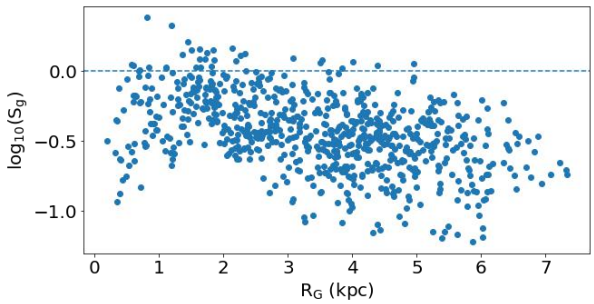


Figure 3.10: Correlation between the shear parameter (S_g) and the distance to the galaxy center (R_G).

calculating the ratio of high- to low-density contrast clouds remains feasible. In Table.3.1, $R_G < 3$ kpc has significantly lower proportion of clouds with $C > 1.63$, consistent with the fact that star formation is suppressed in the center of the galaxy (Kruijssen et al., 2014). When $R_G > 3$ kpc, the proportion is relatively stable. When $R_G < 4$ kpc, there is a clear trend where the proportion of clouds with $C > 1.63$ increases significantly with increasing R_G. Next, we will explore in detail the reasons behind the dependence of the cloud's density contrast on R_G.

3.4.2 Shear and tidal fields of the galaxy

As discussed in Hunter et al. (1998), Dib et al. (2012) and Thilliez et al. (2014), the shear parameter characterizes the effect of differential rotation on the stability of molecular clouds against gravitational

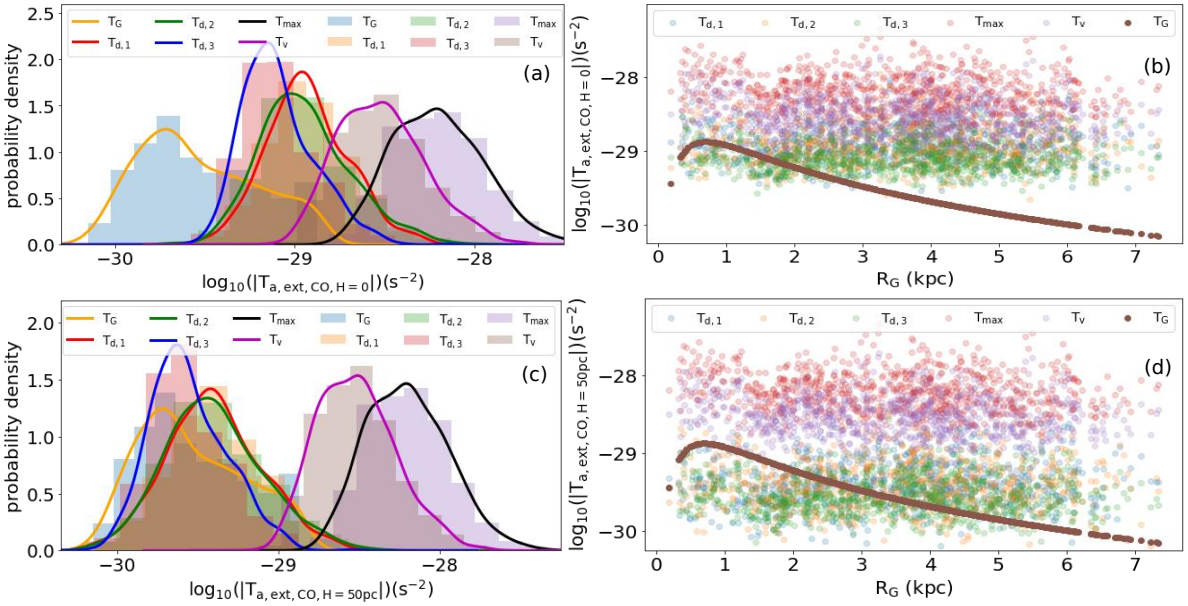


Figure 3.11: Comparison of the tidal strengths obtained from the tidal tensor analysis and the pixel-by-pixel computation based on the CO surface density map. (a) Distribution of the average tidal strength in molecular clouds, assuming the half-thickness of the galaxy $H = 0$; (b) Variation of the average tidal strength of molecular clouds with the distance of the molecular clouds to the galactic center. Panels (c) and (d) are same as panels (a) and (b), but for $H = 50$ pc.

collapse, assessing the time available for perturbations to collapse in the presence of local rotational shear caused by the galaxy's global rotation. The shear parameter is defined as

$$S_g = \frac{\Sigma_{\rm cri}}{\Sigma}, \tag{3.7}$$

where Σ is the local gas surface density. $\Sigma_{\rm cri}$ is the critical surface density, which represents the minimal surface density needed to resist the shear. Shear is effective at tearing apart the gas structure if $S_g > 1$ and ineffective if $S_g < 1$. In Fig.3.10, only a few structures have $S_g >1$, so the effect of shear can be neglected. This is consistent with the results of Dib et al. (2012) and Thilliez et al. (2014) for molecular clouds in the Milky Way and the Large Magellanic Cloud. Moreover, there is a clear trend where S_g decreases as R_G increases, which is inconsistent with the trend in Table.3.1, where the proportion of high- and low-density contrast clouds initially increases with R_G and then keeps relatively stable.

In a galactic context, the overall gas and stellar distribution can contribute to the confinement or disruption of gas structures. For a cloud of radius R at a distance R_G from the galactic center, the edge of the cloud from the galactic center is at a distance of R_G+R. The tidal acceleration per unit distance (called "tidal strength" in the later) defined in Stark & Blitz (1978) is,

$$T_G = \frac{V^2}{R_G^2} - \frac{\partial}{\partial R_G}\left(\frac{V^2}{R_G}\right). \tag{3.8}$$

Fig.3.11 shows that the tidal strength T_G is a strong function of R_G, which is also inconsistent with the trend in Table.3.1.

3.4.3 Tidal effects of neighboring material

Using the same method as in J. W. Zhou et al. (2024f), we directly derived the tidal field according to the surface density distribution obtained from CO (2−1) emission. Calculating the gravitational potential field $\phi(x, y)$ based on the surface density map $\Sigma(x, y)$ has been discussed in Gong & Ostriker (2011) and He et al. (2023). The tidal tensor of a structure includes both self-gravity and external tides, that is, $\mathbf{T} = \mathbf{T}_{\rm ext} + \mathbf{T}_{\rm int}$. The decomposition of the tidal tensor is discussed in detail in J. W. Zhou et al. (2024f). For three extreme cases, the external tides can be estimated by λ_1, λ_2, and $(\lambda_2 - \lambda_1)/2$, and they are called $T_{\rm d,1}$, $T_{\rm d,2}$, and $T_{\rm d,3}$, respectively. λ_1 and λ_2 are the eigenvalues of the tidal tensor.

For the surface density map, the half-thickness H is 0. Taking a characteristic scale height of the molecular gas in the galaxy as 50 pc (Bronfman et al., 1988; Stark & Y. Lee, 2005; Cox, 2005; Ellsworth-Bowers et al., 2013; Jeffreson et al., 2022), which is also close to the median value of the effective radii of all identified clouds (≈55 pc), then we have $H = 50$ pc. We calculated the tidal tensors for both $H = 0$ and $H = 50$ pc. In the case of 3D, for a cloud, the energy ($E_{\rm a,int,H}$) derived from $\mathbf{T}_{\rm int}$ should be comparable with its self-gravity energy ($E_{\rm g}$). As shown in J. W. Zhou & Dib (2025), $E_{\rm a,int,H=0}$ is comparable with $E_{\rm g}$. However, $E_{\rm a,int,H=50pc}$ is systematically smaller than $E_{\rm g}$. Thus, the computed 2D tidal field strength ($H = 0$) can roughly reflect the real 3D tidal field strength of the galaxy.

We further verify the above tidal tensor analysis through an original methodology, as done in J. W. Zhou et al. (2024f). The cumulative tidal strength at a pixel signifies the aggregate deformation resulting from external gravity at that point. To quantify the collective tidal strength within a structure induced by external gravity of the neighboring material, we adopt the scalar superposition. The surface density map of the galaxy was divided into two distinct regions using the cloud's mask obtained in J. W. Zhou et al. (2024a), i.e. the cloud itself and all other material outside of it. To compute the external tidal effects exerted by the surrounding material on a point within the cloud, we use equation.3.1 to directly calculate the tidal strength at that point pixel by pixel. The average tidal strength across all pixels within the cloud subsequently provides an estimate of the average tidal strength experienced by the cloud. Our computation is limited to a 2D plane. Actually, there is a characteristic scale height of the molecular gas in the galaxy. In equation.3.1, the tidal strength is very sensitive to the distance. If the external material and the calculated point in a cloud do not lie on the same plane, the 2D computation results in an underestimation of the distance between them, consequently leading to an overestimation of the tidal strength (the maximum computation, $T_{\rm max}$). To compare with the scalar superposition, we also calculate a vector superposition of tides ($T_{\rm v}$), enabling direct cancellation of tidal interactions.

3.4.4 Baryonic matter

Although we focus solely on molecular clouds as sites of star formation in the galaxy, it's inadequate to only consider molecular gas when accounting for external tidal effects on the clouds. Strictly speaking, we should include all baryonic matter. As a rough approximation, apart from molecular gas surface density ($\Sigma_{\rm CO}$), we also consider atomic gas surface density ($\Sigma_{\rm HI}$) and stellar mass surface density (Σ_*).

Using the recipes in Leroy et al. (2021b) and Sun et al. (2022), assuming optically thin 21 cm emission, the 21 cm line integrated intensity $I_{\rm HI}$ can be converted to atomic gas surface density $\Sigma_{\rm HI}$. Near-IR imaging data can be used to trace the old stellar mass distribution. Using the *Spitzer* IRAC 3.6 μm image, the stellar continuum intensity at 3.6 μm can be converted to the stellar mass surface density. As shown in Fig.3.9, the stellar mass surface density is significantly higher than the molecular and atomic gas surface density. We combined $\Sigma_{\rm CO}$, $\Sigma_{\rm HI}$ and Σ_* to create a total surface density map to approximate the baryonic matter distribution in the galaxy, which is dominated by Σ_*.

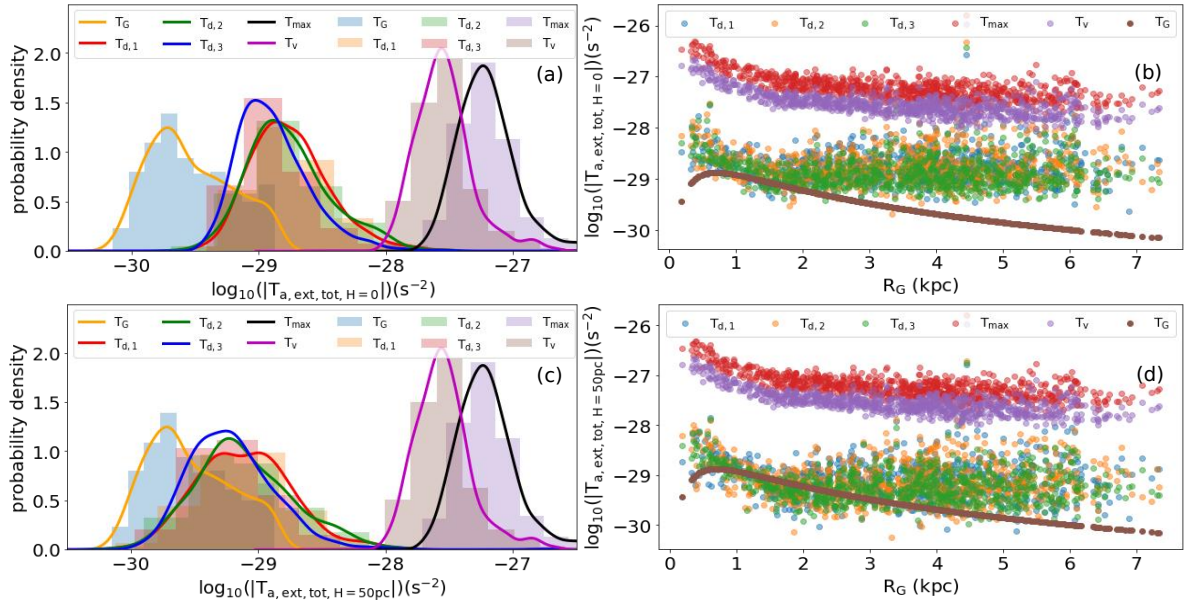


Figure 3.12: Same as Fig.3.11, but all tidal strengths derived from the total surface density map, rather than only the CO surface density map in Fig.3.11.

3.4.5 Comparison of different tidal calculations

For the tidal strengths calculated according to the CO surface density map, Fig.3.11(a) shows that the components $T_{d,1}$, $T_{d,2}$, and $T_{d,3}$ derived from the tidal tensor analysis are comparable. As expected, the values of T_{max} are significantly larger than others. The tidal strength T_G derived from the rotation curve is significantly smaller than others at large R_G, when $H = 0$. In Fig.3.11(b), we plot the variation of the tidal strengths with R_G. Except for T_G, the other quantities do not vary with the distance to the galaxy center. Because the calculations of the other quantities are based on the CO surface density map. In Fig.3.9(a), the galactic center does not have significantly higher CO surface density.

For the tidal strengths calculated based on the total surface density map, although the components $T_{d,1,tot}$, $T_{d,2,tot}$, and $T_{d,3,tot}$ derived from the tidal tensor analysis are comparable, they are significantly smaller than the tidal strengths obtained from pixel-by-pixel computation ($T_{v,tot}$ and $T_{max,tot}$). In Fig.3.12(b), all calculated tidal strengths change with R_G, and the scatterings are also smaller than those in Fig.3.11(b). Especially, in Fig.3.12(b), when $R_G < 4$ kpc, $T_{v,tot}$ and $T_{max,tot}$ decrease as R_G increases, and then remain relatively constant when $R_G > 4$ kpc. This aligns well with the dependence of the proportion of high- and low-density contrast clouds on R_G, as presented in Table.3.1. Therefore, the tidal effects of neighboring material have a significant impact on the development of molecular cloud's density contrast.

In the two cases mentioned above, $H = 50$ pc generates lower tidal strengths than $H = 0$, but is otherwise identical to $H = 0$ in all other respects. $T_{v,tot}$ and $T_{max,tot}$ are calculated in ways that are quite different, but their trends of variation with respect to R_G are consistent. The only difference is in the magnitude of the tidal strengths.

The pixel-by-pixel computation can reflect the variation in tidal strength with R_G more precisely than the tidal tensor analysis, as it has a smaller scatter in tidal strength, as shown in Fig.3.12. However, to accurately estimate the tidal strength using the pixel-by-pixel computation, we need to know the specific 3D spatial distribution of the matter in the galaxy, which is challenging.

Above, we conducted an extensive discussion aimed at summarizing an effective tidal analysis method

for future research. To explain the variation in the density contrast of molecular clouds with $R_{\rm G}$ in NGC 628, it is necessary to integrate a total (baryonic matter) surface density map that includes molecular gas, atomic gas, and stellar components. This is because the stellar mass surface density is significantly higher than the molecular gas surface density, indicating stronger tidal effects. The pixel-by-pixel computation can more precisely reflect the variation in tidal strength with $R_{\rm G}$ than the tidal tensor analysis but cannot provide a correct magnitude estimate of the tidal strength due to the lack of knowledge about the specific 3D spatial distribution of the matter in the galaxy. This issue can be resolved using the tidal tensor analysis. Therefore, by combining these two methods, we may achieve both appropriate variation and magnitude estimates of the tidal strength across the galaxy.

3.4.6 Tidal effects

In Fig.3.12, T_G calculated by the rotation curve consistently decreases with increasing $R_{\rm G}$. $T_{\rm v,tot}$ and $T_{\rm max,tot}$ represent local tidal strengths from neighboring material and are only sensitive to local material distribution. As shown in Fig.3.9, the surface density is significantly higher in galactic center. Therefore, $T_{\rm v,tot}$ and $T_{\rm max,tot}$ only decrease with increasing $R_{\rm G}$ in the central region of the galaxy. And they remains relatively constant in the outer regions of the galaxy, as the surface density distribution in the outer regions is relatively uniform. Therefore, one important reason for the low star formation rate in the galactic center is that the tidal effects from neighboring material acting on molecular clouds are strong, restricting the gravitational collapse of the clouds. As a result, the clouds exhibit low density contrasts.

Molecular clouds do not evolve in isolation. Their gravitational collapse is modulated by various external physical processes. Although the measured velocity gradients in J.-W. Zhou & T. Davis (2024) and J. W. Zhou et al. (2024a) support the gravitational collapse of gas structures on galaxy-cloud scales, the collapse is significantly slower than a pure free-fall gravitational collapse. In the case of free-fall, the velocity gradient (∇v) and the scale (R) satisfy $\nabla v \propto R^{-1.5}$. However, for NGC 628, the fitting in J. W. Zhou et al. (2024a) was $\nabla v \propto R^{-0.9}$. For NGC 4321 and NGC 5236, J.-W. Zhou & T. Davis (2024) obtained a similar slope, i.e. $\nabla v \propto R^{-0.8}$. As discussed in J.-W. Zhou & T. Davis (2024), the slowing down of a pure free-fall gravitational collapse can be caused by multiple physical processes. The results of this work suggest that the tidal effects of neighboring material can resist the gravitational collapse of molecular clouds.

3.5 Summary

3.5.1 On clump scale

Local dense gas structures were directly identified based on the 2D integrated intensity map of $H^{13}CO^+$ J=1-0 emission, followed by extraction of average spectra to fit the velocity dispersion. The slope of the $\sigma - N * R$ relation for leaf structures is ≈0.2, significantly smaller than the 0.5 expected for pure free-fall gravitational collapse, suggesting a deceleration of such collapse.

The effects of shear and tides from the protocluster on the surrounding dense gas structures were examined. Results indicate that shear and tides from the protocluster alone are not the dominant factors. However, a gas structure is also subject to tides from other nearby structures. To account for all tidal interactions, the tidal field was derived based on the density distribution, and the average strength of external tides on each structure was used to measure the total tidal effect. For comparison, an original pixel-by-pixel computation was also applied to estimate the average tidal strength, yielding consistent results across methods.

When tidal interactions are considered, the slope of the $\sigma - N * R$ relation approaches 0.5, and the correlation improves, indicating that deformation by external tides can effectively slow down the free-fall gravitational collapse of gas structures. While external tides tend to disrupt a structure, the accompanying external pressure resists this effect. The balance between external tidal forces and pressure acts to suppress free-fall collapse, leading to what may be termed "tide-regulated gravitational collapse".

3.5.2 On cloud scale

The 3D density distribution derived from the 3D dust map is essential for calculating the tidal field, while velocity information is only available from the projected CO (1–0) line emission. To reconcile these datasets, the 3D density cube was first projected onto the l–b plane, and the dendrogram algorithm was applied to identify molecular clouds. A good spatial correspondence was found between the projected CO and dust emissions.

For 90 non-overlapping clouds, systematic velocity, velocity dispersion, distance, and thickness were fitted based on average “spectra” extracted from the l–b–v and l–b–d cubes. These identified clouds were then used to assess tidal effects. Two independent approaches—tidal tensor analysis and pixel-by-pixel computation—were employed to estimate tidal strength, yielding comparable outcomes. The equivalent velocity dispersion derived from the average tidal strength closely matches the observed velocity dispersion, indicating that external tidal forces significantly contribute to shaping the internal kinematics of molecular clouds.

Extended gas structures in the interstellar medium (ISM) are particularly sensitive to the tidal forces exerted by surrounding matter. Incorporating these effects into star formation models is critical. Since star formation primarily occurs within dense clumps embedded in molecular clouds, it is important to consider these clumps in the context of their large-scale environment. Their interactions with the surrounding matter are non-negligible and may strongly influence their evolution and star-forming potential.

3.5.3 On galaxy-cloud scale

The physical factors influencing the density contrast of molecular clouds are the same ones that impact star formation across the galaxy. The proportion of high- and low-density contrast clouds initially increases with galactocentric radius (R_G) and then remains relatively stable. Spiral arms, bubbles, and magnetic fields do not account for the observed variations in density contrast among molecular clouds. Additionally, galactic shear and tidal strength—estimated from the rotation curve—consistently decrease with increasing R_G, a trend that does not match the radial variation in the density contrast proportions.

To investigate the role of neighboring material, tidal effects on each cloud were analyzed using two independent approaches: tidal tensor analysis and pixel-by-pixel computation. Although all calculations were performed in two dimensions, the 2D tidal field strength computed with a galactic half-thickness of $H = 0$ in the tidal tensor analysis can reasonably approximate the actual 3D tidal field strength of the galaxy. In pixel-by-pixel calculations, both scalar and vector superpositions of tidal forces were considered. Material located beyond four times a cloud’s effective radius was found to have negligible tidal influence. Since star formation occurs in molecular clouds, it is insufficient to consider only molecular gas when assessing external tidal effects. A total surface density map, combining molecular gas, atomic gas, and stellar mass surface densities, was used to trace the baryonic matter distribution in the galaxy.

For all cases, adopting a galactic half-thickness of $H = 50$ pc yields lower tidal strengths than $H = 0$, although the overall trends remain unchanged. Tidal strengths calculated from the CO (2–1) surface

density map show little variation with R_G, due to the lack of a prominent concentration of CO in the galactic center. In contrast, tidal strengths based on the total surface density map vary with R_G. Among the two methods, pixel-by-pixel computation produces results with lower scatter and more accurately reflects the radial trend. Specifically, for $R_G < 4$ kpc, tidal strength decreases with radius and then levels off beyond 4 kpc—a pattern that mirrors the distribution of density contrast among molecular clouds. These findings indicate that the tidal influence of nearby material plays a critical role in shaping cloud density contrast, which in turn regulates star formation activity.

CHAPTER 4

Feedback

The content presented in this chapter are reproduced from the publications by J. W. Zhou et al. (2024e) (Sec.4.2) and J. W. Zhou et al. (2024g) (Sec.4.1). We will focus on the feedback during the formation and evolution of star clusters.

4.1 On cloud scale

In J. W. Zhou et al. (2023), we find in the G333 molecular cloud complex that the larger-scale inflow is driven by the larger-scale structure, indicating hierarchical structure in the G333 complex and gas inflow from large to small scales. The large-scale gas inflow is driven by gravity, implying that the molecular clouds in the G333 complex may be in a state of global gravitational collapse. However, the broken morphology of some very infrared(IR)-bright structures in the G333 complex also indicates that feedback is disrupting star-forming regions. Here, we address the question of how dense molecular gas structures survive and maintain the gravitational collapse state in a strong feedback environment. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024g).

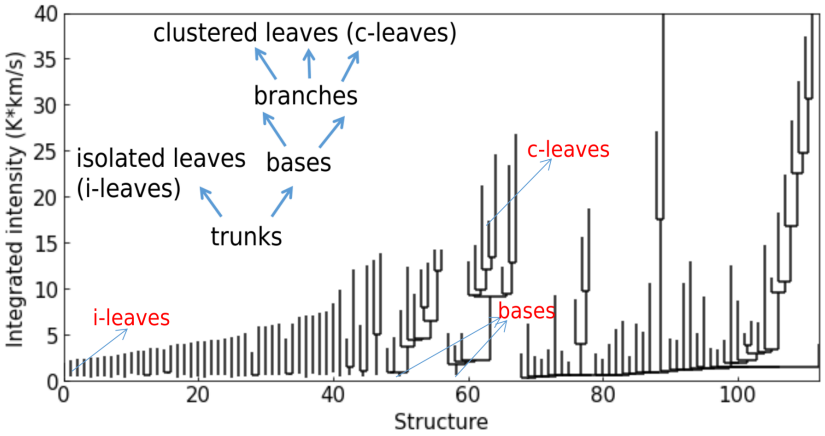


Figure 4.1: Hierarchical structures identified by the dendrogram algorithm. A segment of the dendrogram tree is used to illustrate the structure types output by the dendrogram algorithm.

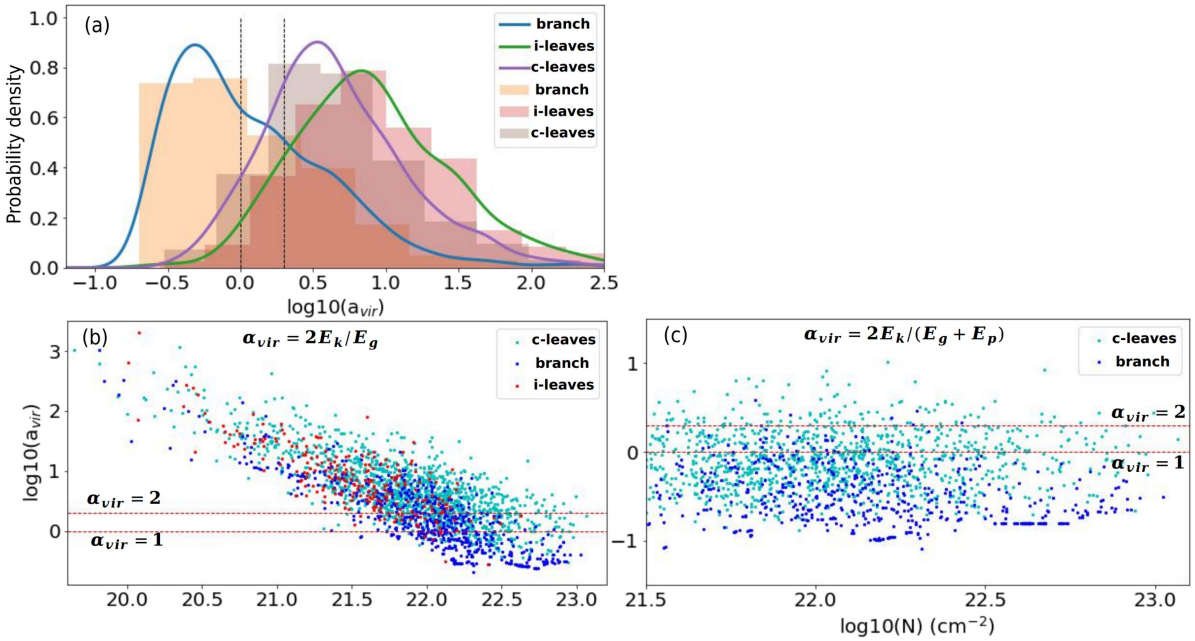


Figure 4.2: Virial parameters of dense gas structures. (a) Virial parameters of all type1 and type2 structures. Dashed lines mark the positions $\alpha_{\rm vir} = 1$ and $\alpha_{\rm vir} = 2$. The probability density is estimated using the KDE method. (b) Correlation between virial parameter $\alpha_{\rm vir} = 2E_{\rm k}/E_{\rm g}$ and the average column density of the structures. (c) Correlation between virial parameter $\alpha_{\rm vir} = 2E_{\rm k}/(E_{\rm g} + E_{\rm p})$ and average column density above the threshold of $\approx 3.2 \times 10^{21}$ cm^{-2}.

4.1.1 Virial parameter

For the G333 molecular cloud complex, we directly identified hierarchical structures based on the integrated intensity map of ^{13}CO (3−2) emission by the dendrogram algorithm, and then extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. As described in E. W. Rosolowsky et al. (2008), the dendrogram algorithm decomposes intensity data into hierarchical structures called leaves, branches, and trunks. The relationship between those structures is shown in Fig.4.1. Trunks are the largest continuous structures at the bottom of hierarchical structures ("bases"), but by definition they can also be isolated leaves ("i-leaves") without any parent structure. Thus, there are two kinds of "trunks", called "bases" and "i-leaves" in this work. Clustered leaves ("c-leaves") are defined as small-scale, bright structures at the top of the tree that do not decompose into further substructures; they are the smallest structures inside "branches". Branches are the relatively large-scale structures in the tree, and can be broken down into substructures. Between "bases" and "c-leaves", all hierarchical substructures are "branches", and therefore branches can span a wide range of scales.

To investigate the energy balance within the extracted structures, we determine the gravitational potential energy ($E_{\rm g}$) and the kinetic energy ($E_{\rm k}$) to compute the virial parameter ($\alpha_{\rm vir}$) (McKee, 1989; Bertoldi & McKee, 1992). Nonmagnetized cores with $\alpha_{\rm vir} < 2$, $\alpha_{\rm vir} \approx 1$, and $\alpha_{\rm vir} < 1$ are considered to be gravitationally bound, in hydrostatic equilibrium, and gravitationally unstable, respectively (Bertoldi & McKee, 1992; Kauffmann et al., 2013). Those with $\alpha_{\rm vir} > 2$ could be gravitationally unbound, and are therefore either pressure-confined, or in the process of dispersal.

Fig.4.2(a) shows the distribution of virial parameters for all identified structures. We can see that more than half of leaf structures are gravitationally unbound and only a small fraction are in gravitational collapse. However, in J. W. Zhou et al. (2023), we argue that molecular clouds in the G333 complex are in a state of global gravitational collapse, because the ubiquitous density and velocity fluctuations

towards hubs imply the widespread presence of local gravitational collapse. Our previous work provides a more comprehensive approach to study the gas kinematics in the clouds. The dense structures in the clouds are connected to the surrounding environment through filaments, and the gravitational state of the structures can be reflected by the velocity gradients along the filaments, which indicate the converging motions toward gravitational centers (hubs). In Fig.4.3, except for the low-column density structures, there is an obvious correlation between velocity dispersion and column density in most of the structures, which indicates a gravitational origin of velocity dispersion. Therefore, most of the structures must be gravitationally bound, or even in a state of gravitational collapse.

In the ISM, each of the structures is embedded in a larger-scale structure and one can therefore assume that they are confined by various external pressures (E. R. Keto & Myers, 1986; C. J. Lada et al., 2008; Field et al., 2011; Leroy et al., 2015; Kirk et al., 2017; H.-R. V. Chen et al., 2019; S. Li et al., 2020). To reconcile the evidence of gravitational collapse presented by J. W. Zhou et al. (2023) with the apparent small fraction of gravitationally unstable gas structures based on the classical virial analysis, we consider the additional effect of external pressure from the ambient cloud for each decomposed structure. After considering the external pressure energy, $E_{\rm p}$, the new estimation of the virial parameter is

$$\alpha_{\rm vir} = 2E_{\rm k}/(\left|E_{\rm g}\right| + \left|E_{\rm p}\right|). \tag{4.1}$$

For c-leaves and branches, the proportion of the structures above the density threshold ($\approx 3.2 \times 10^{21}$ cm^{-2}) is 82.5%. The proportions of $\alpha_{\rm vir} = 2E_{\rm k}/E_{\rm g} < 2$ and $\alpha_{\rm vir} = 2E_{\rm k}/E_{\rm g} \geq 2$ are 45.3% and 54.7%, respectively. After considering the external pressure ($\alpha_{\rm vir} = 2E_{\rm k}/(E_{\rm g} + E_{\rm p})$), their proportions become 93% and 7%. Now the majority of the structures are gravitationally bound, and susceptible to gravitational collapse, as shown in Fig. 4.2(c). However, the peripheral structures are at low column densities and are less bound by external pressure, and are therefore likely to be dispersed by feedback.

Here, we do not consider the HI halo of molecular clouds and also ignore the external pressure exerted by HII regions. The latter should also be important due to the strong feedback in the G333 complex. However, the energy injected into the clouds from HII regions might also destroy the clouds, and therefore we only consider the external pressure from the ambient cloud in binding the structures. The rough estimate in this section shows the important role of external pressure in confining the observed gas structures.

4.1.2 Scaling relation

The physical states of the structures can be reflected by the scaling relations. Fig.4.3(a) shows the velocity dispersion–scale ($\sigma - R$) relations of i-leaves, c-leaves, and branch structures. It appears that only branch structures show a clear correlation between velocity dispersion and scale. i-leaves roughly inherit the trend in the $\sigma - R$ relation of branch structures extending from large to small scales, and they can barely be linked behind branch structures in Fig.4.3(a), although their velocity dispersion shows no significant correlation with scale, which is similar to c-leaves. Moreover, c-leaves deviate more significantly from the Larson relation. The red dashed line, which is fitted to branch and i-leaves structures, has a gradient of 0.33±0.01.

Fig.4.3(b) shows the velocity dispersion–column density ($\sigma - N$) relation. For c-leaves, there is a moderate correlation between velocity dispersion and column density: the Pearson coefficient is ≈0.45. Interestingly, we can see different behaviors of high-column-density and low-column-density structures. For high-column-density structures, the velocity dispersion and column density show a clear correlation, while low-column-density structures do not. In recent simulations (Weis et al., 2022; Ganguly et al., 2024b), dense structures roughly follow the Heyer relation, and less dense structures show no trend with

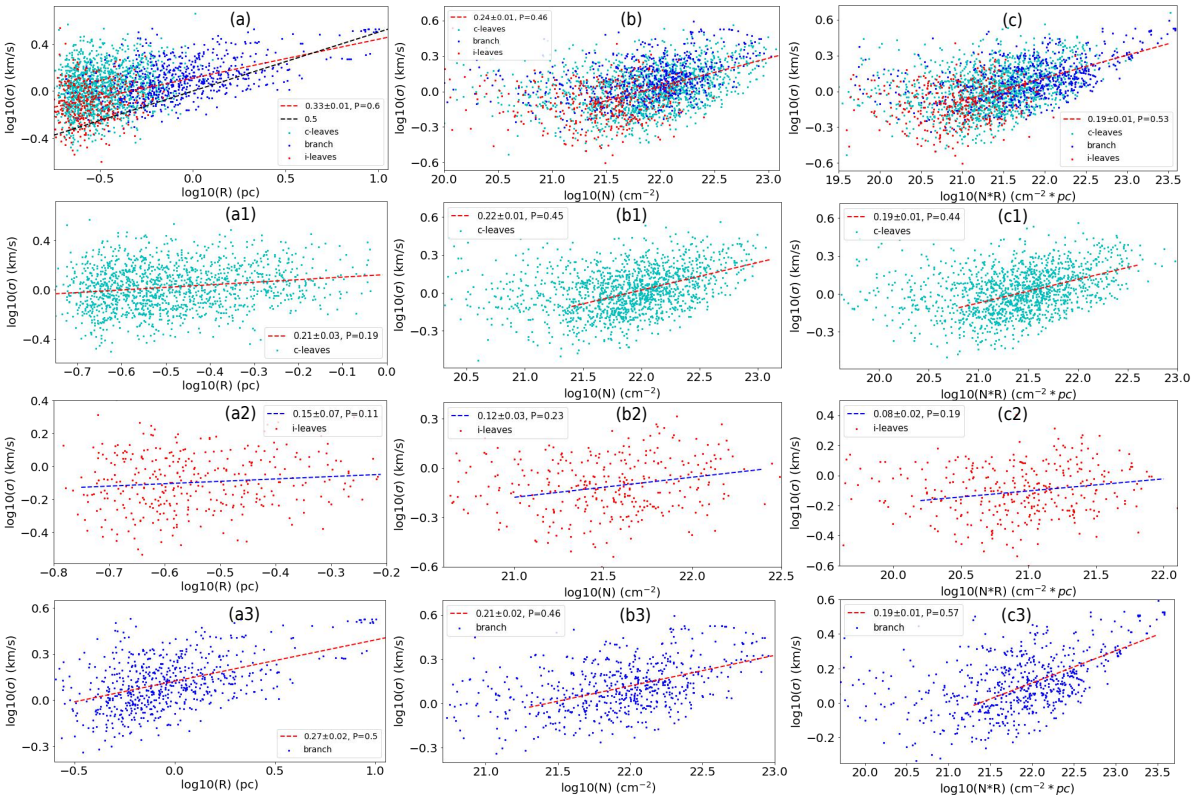


Figure 4.3: Scaling relations of leaf and branch structures. (a) $\sigma - R$; (b) $\sigma - N$; (c) $\sigma - N * R$. σ, R and N are the velocity dispersion, effective radius, and column density of each structure, respectively. Here, "P" represents the Pearson coefficient.

the column density, and populate a low-density tail in the Heyer relation, as shown in Fig.4.3(c). From Fig.4.3(a) and (b), we conclude that the velocity dispersion of branch structures correlates with both scale and column density, and that the velocity dispersion of c-leaves is only sensitive to column density, while the velocity dispersion of i-leaves shows no significant dependence on either scale or column density.

4.1.3 Feedback from HII regions

In the study of the G305 molecular cloud complex, Mazumdar et al. (2021b) argued that the 8 μm emission can be a good indicator of feedback strength. We calculated the average 8 μm surface brightness over each structure to measure the strength of feedback on each structure. In Fig. 4.4, the 8 μm surface brightness shows a strong positive correlation with column density for both c-leaves and branch structures, which might be an indication for triggering in the G333 complex. However, branch structures show a more obvious correlation between 8 μm surface brightness and velocity dispersion than c-leaves, which is consistent with the results of Mazumdar et al. (2021a), implying that feedback has a greater impact on large-scale structures. The small-scale structures are embedded in large-scale structures, and are therefore less affected by feedback. For large-scale structures, as shown in J. W. Zhou et al. (2024g), the more evolved regions are fragmenting into several pieces, potentially torn apart by the expanding HII regions. These results may explain why the velocity dispersion of branch structures shows a clear correlation with scale, while that of leaf structures does not, as shown in Fig.4.3. However, as shown in J. W. Zhou et al. (2024g), after filtering low-column-density and gravitationally unbound structures, the velocity dispersion of c-leaves appears to show a better correlation with scale than that in Fig.4.3, and

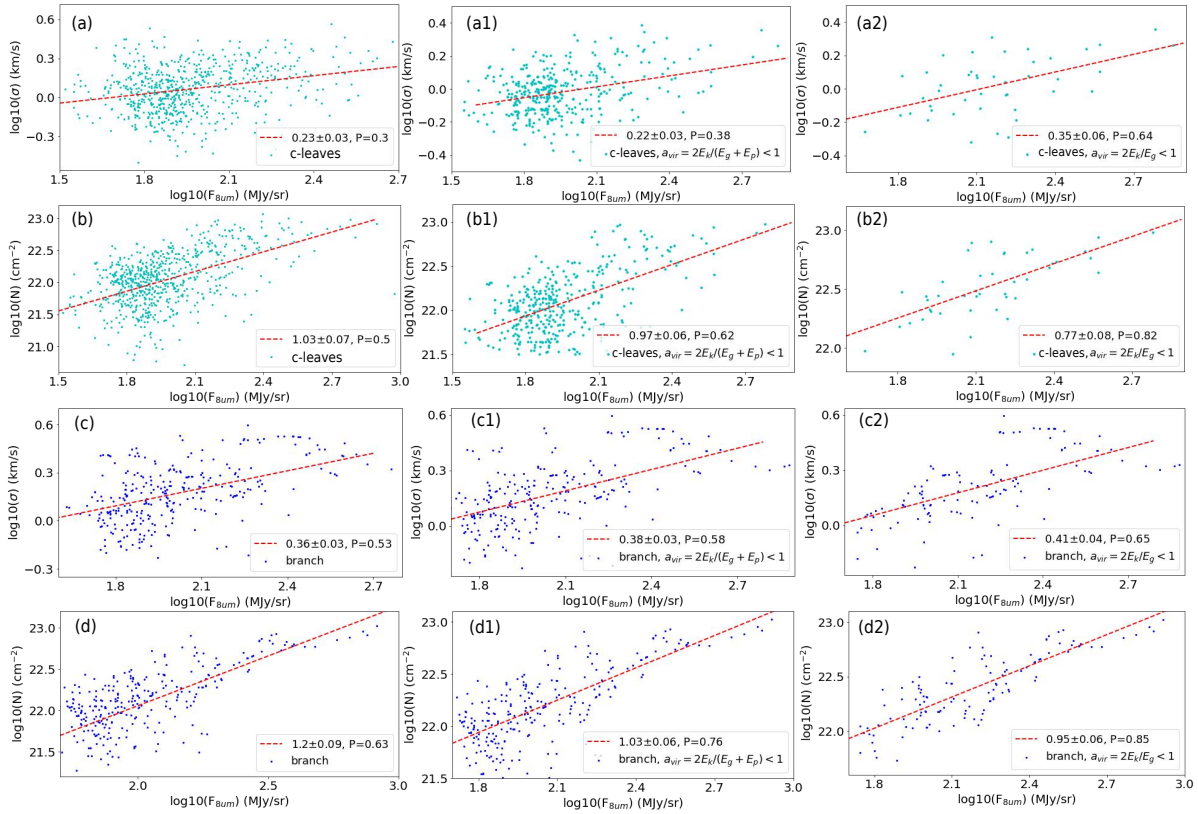


Figure 4.4: Correlation between the 8 μm surface brightness and velocity dispersion and column density of the structures. First column: All type1 c-leaves and branch structures. Second column: Type1 c-leaves and branch structures satisfying the condition $\alpha_{\rm vir} = 2E_{\rm k}/(E_{\rm g}+E_{\rm p}) < 1$. Third column: Type1 c-leaves and branch structures satisfying the condition $\alpha_{\rm vir} = 2E_{\rm k}/E_{\rm g} < 1$.

the correlation between the 8 μm surface brightness and velocity dispersion of c-leaves is also improved in Fig.4.4, which means that leaf structure is also affected by feedback. Here, we should remember that there is considerable overlap in the scale of leaf and branch structures.

Our analysis is based on structural identification, when we say that feedback increases the density and velocity dispersion of the structures, provided that these structures can exist stably. Structures with $\alpha_{\rm vir} = 2E_{\rm k}/(E_{\rm g}+E_{\rm p}) < 1$ and $\alpha_{\rm vir} = 2E_{\rm k}/E_{\rm g} < 1$ can be more tenacious in feedback than other structures, and thus exhibit better scaling relations, as presented in J. W. Zhou et al. (2024g). Dale et al. (2014) used simulations to examine the effects of photoionization and momentum-driven winds from O-stars on molecular clouds, and found that feedback is highly destructive to clouds of lower mass and density, but has little effect on more massive and denser clouds.

4.1.4 The Heyer relation in feedback

As discussed in J. W. Zhou et al. (2024g), self-gravitating structures can better fit the Heyer relation. Considering that global collapse may lag behind local collapse within the cloud (Heitsch et al., 2008), structures collapsing under self-gravity can be relatively independent of (or "decoupled from") the surrounding environment. Therefore, the explanations of the Heyer relation in M. Heyer et al. (2009) and Ballesteros-Paredes et al. (2011b) can hold. Contrarily, for non-self-gravitating structures, the exchange of energy with the surrounding environment will break the conversion between $\rm E_g$ and $\rm E_k$, thus breaking the Heyer relation.

Sun et al. (2018) measured cloud-scale molecular gas properties in 15 nearby galaxies, and observed an excess in the velocity dispersion σ at low surface density Σ relative to the expected relation for self-gravity-dominated gas. This behavior leads to a shallower $\sigma - \Sigma$ relation in several galaxies, clearly deviating from the $\sigma - \Sigma^{0.5}$ relation extrapolated from the high-surface-density regime. One of the explanations for the deviation is that gas structures in the low-surface-density regime may be more susceptible to external pressure originating from the ambient medium or motions due to the galaxy potential, which is similar to our previous explanation.

4.1.5 Cloud disruption and collapse under feedback

Peretto et al. (2023) performed an analysis of 27 infrared-dark clouds (IRDCs) embedded within 24 molecular clouds. These authors found that the clumps are decoupling from their surrounding cloud and concluded that the observations are best explained by a universal global collapse of dense clumps embedded within stable molecular clouds, thus discovering direct evidence of a transition regime in the dynamical properties of the gas within individual molecular clouds. As discussed in Heitsch et al. (2008) and Ballesteros-Paredes et al. (2011b), the decoupling may be due to the global collapse of the molecular cloud lagging behind the local collapse of dense clumps within it. Invoking the notion of the "funnel" structure in PPV space (J. W. Zhou et al., 2023), a similar statement is that relatively small-scale hub–filament structures will have a more recognizable "funnel" morphology than large-scale ones due to their strong local gravitational field. Substructure G332.83-0.55 in the G333 complex is a vivid example of the decoupling highlighted in Fig.1 of J. W. Zhou et al. (2023) and J. W. Zhou et al. (2024f); it is collapsing into a hub–filament structure and is separating from its surrounding environment.

One implication of the work of Peretto et al. (2023) is that star formation is likely to be mostly confined to parsec-scale collapsing clumps, which is also consistent with the results of J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023). In our previous works, for both ALMA and LAsMA data, most of the fitted velocity gradients concentrate on scales of ≈ 1 pc, a scale that is considered to be the characteristic scale of massive clumps (Urquhart et al., 2018). Velocity gradients measured around 1 pc show that the most frequent velocity gradient is $\approx 1.6\,\mathrm{km\,s^{-1}\,pc^{-1}}$. Assuming free-fall, we estimate that the kinematic mass corresponding to 1 pc is $\approx 1190\,\mathrm{M_\odot}$, which is also comparable with the typical mass of clumps in the ATLASGAL survey (Urquhart et al., 2018). Therefore, parsec-scale clumps are probably gravity-dominated collapsing objects.

The results in J.-W. Zhou et al. (2022), J. W. Zhou et al. (2023) and Peretto et al. (2023) show that the physical properties of parsec-scale clumps in two very different physical environments (infrared dark and infrared bright) are comparable. Therefore, feedback in infrared-bright star-forming regions, such as the G333 complex, will not significantly change the physical properties of parsec-scale clumps, which is also consistent with the finding from surveys that most Galactic parsec-scale massive clumps seem to be gravitationally bound no matter how evolved they are (T. Liu et al., 2016b; Urquhart et al., 2018; I. Evans N. J. et al., 2021). Although the clumps are exposed to feedback and part of their velocity dispersion is due to feedback, they remain self-gravitating sufficiently to continue their collapse, even after the lower-density material has been disrupted. Watkins et al. (2019b) found that stellar feedback from O stars does not have a significant impact on the dynamical properties of the dense gas that has already been assembled, but does clearly modify the structure of the larger-scale clouds. The broken morphology of some very infrared-bright structures in the G333 complex also indicates that the feedback is disrupting the molecular clouds.

The effects of feedback in star-forming regions can redistribute, disperse, and enhance preexisting gas structures, and create new structures (B. G. Elmegreen & C. J. Lada, 1977; Dale et al., 2007; H.-T. Lee & W. P. Chen, 2007; Nagakura et al., 2009; Krumholz et al., 2014; Fukui et al., 2021). According

to the physical picture described in J. W. Zhou et al. (2023), the hub–filament structures at different scales may be organized into a hierarchical system, extending up to the largest scales probed, through the coupling of gravitational centers at different scales. Large-scale velocity gradients always involve many intensity peaks, and the larger-scale inflow is driven by the larger-scale structure, implying that the clustering of local small-scale gravitational structures can act as a gravitational center on larger scale. Considering the hierarchical hub–filament structures and the coupling of local gravitational centers in molecular clouds, and feedback does not significantly impact the physical properties of dense gas structures, therefore although the feedback disrupting the molecular clouds may break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and changes in gravitational centers, but gravitational collapse is always ongoing.

4.2 On clump scale

The ATOMS survey has observed 146 active star-forming regions at ALMA Band 3. The data can be used to systematically investigate the spatial distribution of various dense gas tracers in a large sample of Galactic massive clumps. The dense molecular gas structures inside the molecular cloud serve as star-forming sites, their physical states directly determine the star formation capability of the cloud under feedback. A more systematic study is needed to evaluate how massive protostars influence the dense gas distribution and the star formation efficiency within their natal clumps. Using ALMA data, we zoomed into the dense gas structures close to or even connecting with protoclusters. The effects of feedback were then inferred through the physical properties of the surrounding dense gas structures. The sample of the ATOMS survey includes 146 active star-forming regions. In J.-W. Zhou et al. (2022), we find that hub-filament systems are very common within massive protoclusters, and that stellar feedback from HII regions gradually destroys the hub-filament systems as protoclusters evolve. The sources at different evolutionary stages are dominated by different feedback mechanisms. By comparing the physical properties of dense gas structures in the clumps at different evolutionary stages, we will be able to infer the relative importance of various feedback mechanisms. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024e).

4.2.1 Evolutionary stages

Generally, the infrared properties and the presence of $H_{40\alpha}$ emission can be used to assess the evolutionary stages of the sources (Chambers et al., 2009; Sanhueza et al., 2012). As discussed in J.-W. Zhou et al. (2022), stellar feedback from HII regions will gradually destroy the hub-filament systems as protoclusters evolve. Thus, the morphology of the hub-filament structure can also characterize the evolutionary stages. In Fig.2 and Fig.3 of J.-W. Zhou et al. (2022), for hub-filament systems, the overlaid contours of the 3 mm continuum should be located at the center of filamentary network with the strongest $H^{13}CO^+$ J=1-0 emission, where is the center of the gravitational potential well and the potential site for high-mass star formation. In the late stage of evolution, the 3 mm continuum emission is almost completely separated from $H^{13}CO^+$ J=1-0 emission, indicating that the relatively evolved HII regions have dispersed their natal gas. Thus, the separation between 3mm continuum emission and $H^{13}CO^+$ (1-0) emission can also be used for the classification. Finally, we divided ATOMS sources into three evolutionary stages, i.e. early, middle and late. Three examples are shown in Fig.4.5.

For each source, we directly identified hierarchical (sub-)structures based on the 2D integrated intensity map of $H^{13}CO^+$ (1−0) emission, and then extracted the average spectrum of each structure to investigate

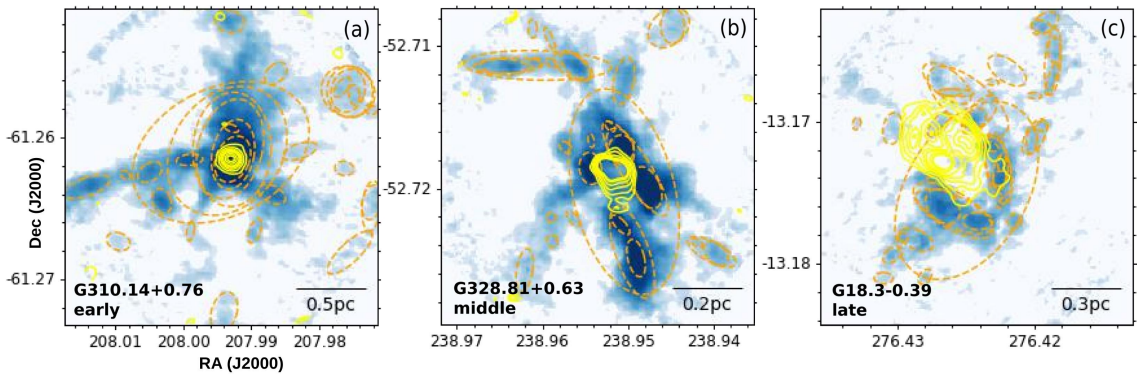


Figure 4.5: Moment 0 maps of $H^{13}CO^+$ J=1-0 for three ATOMS sources at three evolutionary stages. Yellow contours show 3 mm continuum emission with its minimum contour level of 5σ. Orange eclipses represent the retained dendrogram structures.

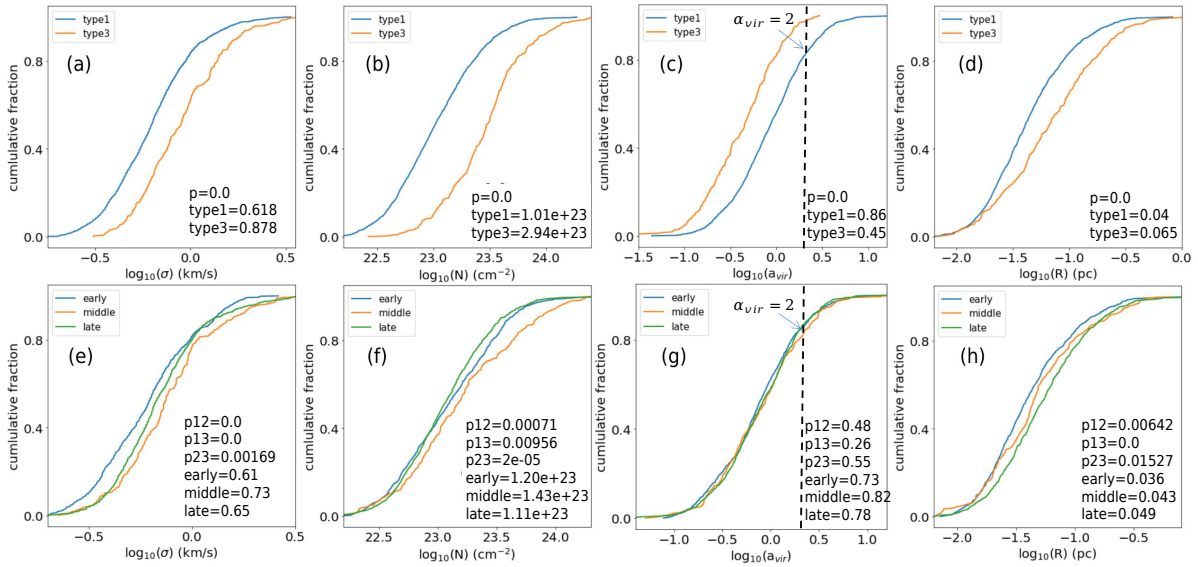


Figure 4.6: Physical properties of different structure types at different evolutionary stages. The first row is the comparison between type1 and type3 structures. The second row is the comparison between early, middle and late sources. The dashed lines in panels (c) and (g) mark the positions of α_{vir}=2. "p" is the p value of the Kolmogorov-Smirnov (KS) test. “p12,” “p13,” and “p23” are the p values of KS tests between the structures inside early and middle sources, early and late sources, and middle and late sources, respectively. The median values also show in the legends.

its velocity components and gas kinematics. According to the line profiles, all the averaged spectra of 2420 structures were divided into three categories: type1 (single velocity component, 2038), type2 (separated velocity components, 42), and type3 (blended velocity components, 340). We also refer to Fig.6 of J. W. Zhou et al. (2024g) for examples of typical line profiles. In the subsequent analysis, we treat type3 structures as independent structures like type1 structures, but with more complex gas motions.

4.2.2 Statistics of physical quantities

From the first line of Fig.4.6, at different evolutionary stages, type3 structures show higher velocity dispersion, greater column density, and lower virial ratios than type 1 structures, indicating that type3 structures are mainly concentrated in the hub regions, and thus present multiple velocity components. In

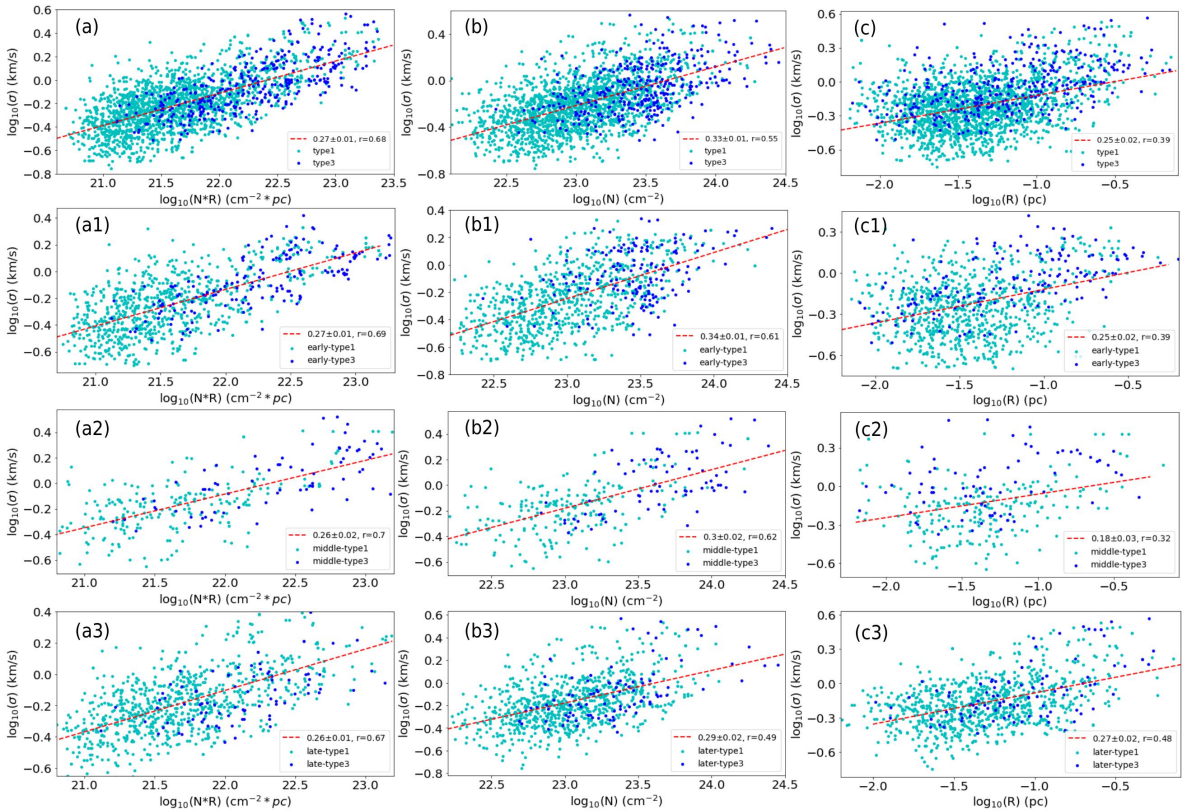


Figure 4.7: Scaling relations of type1 and type3 structures. (a) $\sigma - N * R$, (b) $\sigma - N$, and (c) $\sigma - R$. σ, R and N are the velocity dispersion, effective radius, and column density of each structure, respectively. "r" represents the Pearson correlation coefficient. The values in the legends are the exponents of the power laws. The second, third, and fourth rows are the same as the first row, but for early, middle and late sources, respectively.

J.-W. Zhou et al. (2022), from the grid maps of $H^{13}CO$ (1−0) emission, the hub regions always show multiple velocity components that should be attributed to the complex gas motions, not the superposition of uncorrelated foreground or background velocity components.

From the second line of Fig.4.6, although late sources have powerful HII regions, their inner gas structures do not have significantly larger velocity dispersions and virial ratios, compared to those in early and middle sources. In Fig.4.6(e), the embedded dense structures inside middle sources even have slightly larger velocity dispersions than those inside late sources. These statistical results are unusual. We need to reconsider the effects of feedback from expanding HII regions.

4.2.3 Scaling relations

The scaling relations between different quantities contain information on the physical conditions of the structures. The scaling relations between the velocity dispersion, σ, effective radius, R, and column density, N, shown in Fig.4.7 are similar to those in J. W. Zhou et al. (2024g). $\sigma - N * R$ always has a stronger correlation compared to $\sigma - N$ and $\sigma - R$. However, the slope of the $\sigma - N * R$ relation is ≈0.27, which may indicate the slowing down of a pure free-fall gravitational collapse. There are significant correlations between velocity dispersion and column density, which may imply that the velocity dispersion partly originates from the gravitational collapse, something which is also revealed by the measured velocity gradients in the next section. $\sigma - R$ has a poor correlation, that is, a significant deviation from the Larson-relation (R. B. Larson, 1981). Additionally, it is noteworthy that both type1

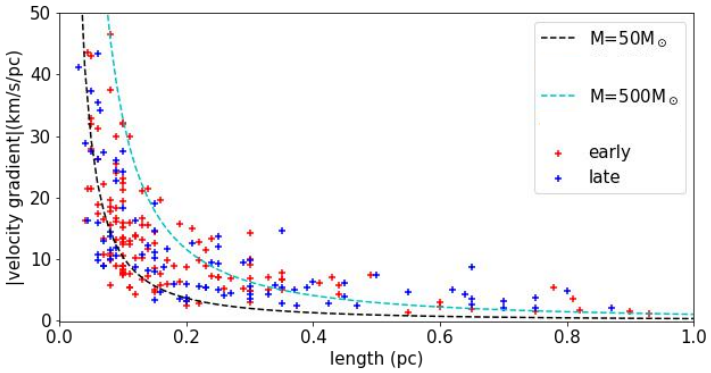


Figure 4.8: Velocity gradient versus the length over which the gradient has been fitted. Blue and red crosses represent the velocity gradients fitted in early and late sources, respectively. The dashed lines show free-fall velocity gradients for comparison. For the free-fall model, black and cyan lines denote masses of $50M_{\odot}$ and $500M_{\odot}$, respectively.

and type3 structures present similar scaling relations, which further confirms that type3 structures are independent like type1 structures. The structures generated by random overlap should not show good scaling relations.

4.2.4 Velocity gradient

In J.-W. Zhou et al. (2022), we used the integrated intensity map of $H^{13}CO^{+}$ J=1-0 to identify filaments in ATOMS sources, and extracted the velocity and intensity along the filament skeletons from the moment 1 and moment 0 maps. Clear velocity and density fluctuations are seen along the filaments, allowing us to fit velocity gradients around the intensity peaks. The filaments like the streamlines, which can trace how gas inflows converge to local dense structures. Local dense structures as gravitational centers will accrete the surrounding diffuse gases and then form local hub-filament structures. A hub-filament structure can be translated into a funnel structure in the PPV space, as described in J. W. Zhou et al. (2023). As illustrated in Fig.9 of J. W. Zhou et al. (2023), the gradient of the funnel profile can reflect the strength of the gravitational field. In J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2023), the filaments will pass through multiple local hub-filament structures presented as velocity and density fluctuations. Therefore, the local velocity gradients along the filaments can reveal the strength of local gravitational fields. Then, by comparing local velocity gradients, we can assess the differences in the gravitational states of the embedded dense gas structures across different feedback environments.

For all ATOMS sources, the identified filaments, as well as the extracted velocity and density along the filaments, have been shown in J.-W. Zhou et al. (2022). We first estimated two overall velocity gradients between the velocity peaks and valleys at the two sides of the center of the gravitational potential well and ignored the local velocity fluctuations, then we also derived local velocity gradients around local intensity peaks of the $H^{13}CO^{+}$ emission (see Fig. 6 in J.-W. Zhou et al. 2022). In Fig.4.6(g), the virial ratios of the embedded dense gas structures in late and early sources are comparable. In Fig.4.8, the measured velocity gradients of the structures in early and late sources are also comparable, once again demonstrating that feedback rarely affects the physical properties of the embedded dense gas structures.

4.2.5 Relative importance of feedback processes

In the analysis of the gas kinematics in early sources, we mainly focused on the outflows ($v_{\rm turb,out}$) and inflows ($v_{\rm in}$). When investigating the gas kinematics in late sources, we primarily considered the ionized

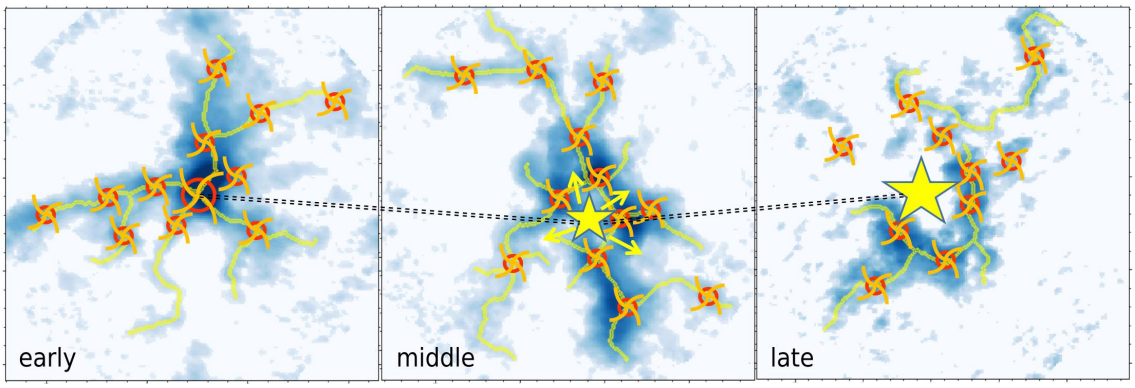


Figure 4.9: Schematic diagram of the scenario argued in this work. Evolution of a cloud with multi-scale hub-filament structures with increasing effects of feedback. The yellow lines are the filament skeletons identified in J.-W. Zhou et al. (2022). Considering the competitive accretion scenario, the precursor of the protocluster is located in the central position of the hub. With the evolution of the protocluster, feedback from HII regions destroys the hub by mechanical force, i.e., the expanding HII regions push away and reshape the hub region, accompanied by the redistribution of small-scale dense structures that may be local hub-filament structures. We note that small-scale hub-filament structures were not directly identified in this work due to the limitation of the resolution.

gas pressure ($v_{\rm HII}$). Moreover, the radiation pressure ($v_{\rm rad}$) exists in all sources at different evolutionary stages. But late sources must have stronger radiation pressure than early sources, because they embed luminous ionizing (young) stellar objects. We note that inflow is not a feedback mechanism; but as a line-broadening mechanism, it is comparable to other feedback mechanisms. The velocity dispersion generated by the four physical processes is quantitatively estimated in J. W. Zhou et al. (2024e).

$v_{\rm HII}$ is significantly larger than other velocities, consistent with the conclusion in Lopez et al. (2014) that the ionized gas feedback dominates over the other considered feedback mechanisms. In this work, we did not intend to study which feedback mechanism is the strongest, and thus we ignored some feedback mechanisms, such as dust-reprocessed radiation and stellar wind, which may also play important roles (Pabst et al., 2019; Olivier et al., 2021; Henshaw et al., 2022). Here, we note that the estimated $v_{\rm HII}$ is obviously larger than the observed line widths in $H^{13}CO^+$ J=1-0. This is not surprising because the energy is deposited into the ionized gas and photodissociation regions. As expected, late sources present significantly stronger radiation pressure than early sources. For early sources, the velocity dispersion caused by the outflows is significantly smaller than that from the infalls and the radiation pressure.

As discussed in Sec.4.2.4, even under strong feedback, the embedded dense structures in late sources can still retain gravitational collapse. Feedback does not seem to affect the internal dynamics of the embedded dense gas structures. As illustrated in Fig.4.9, feedback mainly redistributes local dense gas structures. The results are consistent with the conclusions in Sec.4.1.

4.2.6 The role of feedback

The significantly larger $v_{\rm HII}$ will naturally make us think that late sources should have the strongest feedback effects. However, as presented above, the velocity dispersions and virial ratios of the structures inside early and late sources are comparable. As shown in Sec.4.2.4, the gas structures in late sources are also in gravitational collapse like those in early sources. Gravitational collapse may also contribute significantly to the velocity dispersion of the gas structures in late sources. Thus, the contribution of expanding HII regions in late sources to the velocity dispersion of gas structures decreases further. The clouds exist as gas reservoirs that contain dense gas structures as local gravitational centers. Feedback

from HII regions destroys the clouds by mechanical force; in other words, the expanding HII regions push away and reshape the loose cloud complex (McLeod et al., 2015; Beuther et al., 2022; Bonne et al., 2023). But the processes do not significantly affect the physical properties of the embedded dense gas structures in the clouds, suggesting that HII regions may not effectively inject momentum into the embedded dense gas structures as star-forming sites. However, the triggering does not necessarily result from momentum being transferred inside the gas structures but by the gas structures being compressed, and so becoming gravitationally unstable.

Although the feedback from HII regions is very strong, it is acting from the outside on the embedded dense gas structures. By comparison, the outflows and inflows can go deep inside of the embedded dense gas structures, and thus the effect of them on the physical properties of those gas structures is not necessarily weaker than the feedback from HII regions. We should not only focus on the strength of a feedback mechanism itself, but also assess the extent to which it can affect the embedded dense gas structures. In Fig.4.6(e), the structures in middle sources display slightly larger velocity dispersions than those in late sources. Notably, middle-stage sources possess larger-scale hub-filament structures than their late-stage counterparts. The complex gas kinematics in these hub regions within middle-stage sources may contribute to the higher velocity dispersion observed. Furthermore, early sources also exhibit large-scale hub-filament structures, but their gas structures present smaller velocity dispersions compared to those in middle sources. This suggests that feedback from UC-HII regions in middle sources plays a significant role. At least, we can conclude that feedback from protocluster-driven HII regions may not be the most powerful factor affecting the physical properties of local dense gas structures in the clumps. It is important to note that the above discussion treats each clump in isolation, without considering the potential effects of the surrounding gas environment on the physical properties of the internal structures of the clumps, due to the limited FOV of ALMA observations. Given that the protocluster-driven HII regions may still be bound by the abundant surrounding gas, possible quasi-static expanding processes also indicate the limited effect of feedback from HII regions on the embedded dense gas structures. As shown in Kuhn et al. (2019), star clusters and associations that are still embedded in molecular clouds are less likely to be expanding than those that are partially or fully revealed.

As illustrated in Fig.4.9, hierarchical hub-filament structures of star-forming regions mean that star-forming regions exist as loose gravitational complexes. Feedback will reshape the overall structures of star-forming regions. This process is accompanied by a redistribution of the clouds' internal structures, while the gravitational coupling between the internal structures also reorganizes them into new gravitationally governed configurations around new gravitational centers. Generally, feedback will disperse the clouds. To some extent, the reorganization is a resistance to the dispersing process, although the resistance will in the end fail. A negative effect of the feedback is that it destroys the gas reservoir inhabited by dense gas structures, inhibiting the further growth of dense structures by accreting the surrounding gas, and thus suppressing the formation of massive stars. On the other hand, if feedback causes dense gas structures to become more dispersed, and they remain within a gas-rich environment, this looser distribution may enhance their material accumulation efficiency by reducing competitive accretion. As a result, it could ultimately lead to the formation of more massive stars.

4.3 Summary

4.3.1 On cloud scale

The kinematics and dynamics of gas structures under feedback in the G333 complex were investigated. Dense gas structures were identified using the dendrogram algorithm based on the integrated intensity map of ^{13}CO (3−2), yielding 3608 structures. The averaged spectra of these structures were extracted

and fitted individually. Physical quantities for each structure were then derived from the corresponding line profiles.

Column density across the observed field was determined using ATLASGAL+Planck 870 μm data, Hi-GAL data, and multiple transitions of ^{13}CO (J=1–0, 2–1, and 3–2). Pixel-by-pixel comparisons of column density estimates revealed that values derived from ATLASGAL+Planck and Hi-GAL are consistent, while molecular line emission yielded significantly lower column densities than continuum-derived estimates. The peak column density ratios between ^{13}CO transitions were found to be $N_{2-1}/N_{1-0} \approx 0.5$, $N_{3-2}/N_{1-0} \approx 0.3$, and $N_{3-2}/N_{2-1} \approx 0.5$. These values are consistent with the predictions from the NLTE molecular radiative transfer code RADEX for typical gas densities of $\approx 4.2 \times 10^3$ cm^{-3}. A correction factor of 0.3 was therefore applied to the column density derived from ^{13}CO $J = 3 - 2$ to obtain values more representative of total column density.

Classical virial analysis indicates that many structures are unbound; however, this assessment does not fully reflect their physical state. When external pressure from the ambient cloud is taken into account, nearly all structures with column densities exceeding the threshold of $\approx 3.2 \times 10^{21}$ cm^{-2} appear gravitationally bound or even undergoing gravitational collapse. A positive correlation between velocity dispersion and column density further supports the gravitational origin of the velocity dispersion.

The average 8 μm surface brightness was used as a proxy for feedback strength and shows a strong positive correlation with column density. The correlation between 8 μm brightness and velocity dispersion is more pronounced in branch structures, suggesting that feedback has a more significant impact on larger-scale structures. This indicates that while both gravitational collapse and feedback influence the velocity dispersion of large-scale structures, gravitational collapse remains the dominant factor for small-scale structures. The contribution of feedback at small scales warrants further investigation.

For both leaf and branch structures, the $\sigma - N * R$ relation consistently shows a stronger correlation than either $\sigma - N$ or $\sigma - R$ relations. Scaling relations become more pronounced and display steeper slopes when limited to self-gravitating structures, aligning more closely with the Heyer relation. However, in the G333 complex, where feedback is strong, only a minority of structures are in a state of self-gravitational collapse.

Despite the disruptive effects of feedback on molecular clouds, the substructures of the original complex can reorganize into new gravitationally bound configurations around new gravitational centers. This evolutionary process involves both structural disruption and regeneration, along with shifts in gravitational centers, while gravitational collapse continues uninterrupted.

4.3.2 On clump scale

A total of 64 ATOMS sources at different evolutionary stages (early, middle and late) were selected to investigate the kinematics and dynamics of gas structures under feedback. Dense gas structures were directly identified based on the 2D integrated intensity map of H^{13}CO$^+$ (1–0) emission, followed by extraction of the average spectrum of each structure to investigate their velocity components and gas kinematics. Differences in the physical properties of various structure types at different evolutionary stages were examined. Type3 structures with blended velocity components exhibit larger velocity dispersion, higher column density, and lower virial ratios than type1 structures with a single velocity component, which can be attributed to the tendency of type3 structures to reside in hub regions. Overall, the velocity dispersion and virial ratios of structures within early and late sources are comparable. Structures within middle sources even display slightly larger velocity dispersions than those within late sources. These unexpected results raise questions about the effectiveness of feedback from HII regions in late sources.

Significant correlations between velocity dispersion and column density suggest that the velocity

dispersion may partly originate from gravitational collapse, which is also supported by the measured velocity gradients. Dense gas structures in late sources are undergoing gravitational collapse, similar to those in early sources. While large-scale hub-filament structures are absent in late sources, the embedded dense gas structures display kinematic properties comparable to those in early and middle sources, possibly explained by the presence of multi-scale hub-filament structures within the clouds. Variations in the gravitational states of dense gas structures across different feedback environments were explored through a comparison of local velocity gradients. Comparable velocity gradients in structures within early and late sources indicate that feedback has a limited effect on the kinematic properties of local dense gas structures.

The velocity dispersions generated by outflows, inflows, ionized gas pressure, and radiation pressure were quantitatively estimated, revealing that ionized gas feedback is the most significant among these mechanisms. Late sources exhibit substantially stronger radiation pressure than early sources. In early sources, outflow-driven velocity dispersion is significantly smaller than that caused by infall or radiation pressure. However, despite the dominance of HII region feedback, it does not strongly impact the physical properties of embedded dense gas structures, implying that HII regions may not efficiently inject momentum into the star-forming clumps.

This clump-core scale analysis aligns with findings from the G333 molecular cloud complex study on cloud-clump scale in J. W. Zhou et al. (2024g), supporting the idea that although feedback from cloud to core scales disrupts the original cloud complex, the substructures may be reorganized into a new gravitationally governed configuration around emerging gravitational centers. This process involves both structural destruction and regeneration, accompanied by shifts in gravitational centers, while gravitational collapse continues throughout.

CHAPTER 5

Formation and evolution of star clusters

The content presented in this chapter are reproduced from the publications by J. W. Zhou et al. (2024d) (Sec.5.2), J. W. Zhou et al. (2024b) (Sec.5.3), J.-w. Zhou et al. (2024b) (Sec.5.4), J.-W. Zhou et al. (2024) (Sec.5.6), J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025) (Sec.5.7), and three submitted papers (Sec.5.8, Sec.5.9 and Sec.5.10). Starting from massive star and cluster formation in clumps, these studies aim to track the evolutionary pathways from clumps to embedded clusters and ultimately to open clusters.

5.1 Data and sample

5.1.1 Embedded clusters

In J.-w. Zhou et al. (2024b), we identified a sample of embedded clusters based on the surface density maps of stars for 17 MSFRs with distances less than 3.6 kpc from the the Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX) project (Feigelson et al., 2013). To this, we also added the embedded cluster catalogs of E. A. Lada et al. (1991), Carpenter et al. (1993), C. J. Lada & E. A. Lada (2003), Carpenter (2000), Kumar et al. (2006).

5.1.2 Clumps

In Urquhart et al. (2022), 5007 ATLASGAL clumps have been classified into four evolutionary stages, i.e. quiescent, protostellar, young stellar objects (YSOs) and HII regions. In this work, we focus on the clumps at the most evolved stage, called HII-clumps. The physical parameters of 1246 HII-clumps have been calculated and listed in Table.4 of Urquhart et al. (2022). Clump sizes are determined based on the pixel count within the Full Width at Half Maximum (FWHM) contour, i.e. above 50 percent of the peak of the ATLASGAL dust continuum emission. The clump's mass, denoted as $M_{\rm cl}$ and measured at FWHM, is determined by integrating the flux density at 870 μm within the 50 percent contour, as outlined in Urquhart et al. (2022). The bolometric luminosity ($L_{\rm bol,obs}$) is derived from the spectral energy distributions (SEDs) of the clumps, which are constructed using continuum data spanning from mid-infrared to sub-millimeter wavelengths (Konig et al., 2017; Urquhart et al., 2022). There are 1226 HII-clumps with mass and bolometric luminosity estimates. Then they were matched with radio sources in the catalogs of the CORNISH-North (Purcell et al., 2013) and the CORNISH-South (Irabor et al., 2023) surveys. The separation between the central coordinates of a radio source and a HII-clump was required to be smaller than the radius of the clump. Generally, one HII-clump matches with one radio source. If one HII-clump includes more than one radio source, we only consider the radio source with the strongest radio emission, because we are only interested in the most massive star in a HII-clump.

Finally, the numbers of radio sources matched with HII-clumps are 244 and 148 in the CORNISH-South and the CORNISH-North survey catalogs, respectively.

5.1.3 Open clusters

Hunt & Reffert (2023) conducted a blind, all-sky search for open clusters using 729 million sources from Gaia DR3 down to magnitude G≈20, creating a homogeneous catalog of 7167 clusters. By measuring cluster masses and Jacobi radii, the members in the star cluster catalog are classified into two categories, i,e, bound open clusters (OCs) and unbound moving groups (MGs). In Hunt & Reffert (2024), the catalog includes 5647 OCs and 1309 MGs, where 3530 OCs and 539 MGs are of high-quality with median color-magnitude diagram (CMD) classifications greater than 0.5 and signal-to-noise ratio (S/N) greater than 5σ.

To calculate the photometric masses of the clusters, Hunt & Reffert (2024) first derived the photometric masses of the member stars in each cluster using the PARSEC isochrones (Bressan et al., 2012). They then corrected for selection effects and applied a correction for unresolved binary stars. Subsequently, mass functions were fitted and integrated to determine the total cluster mass.

5.2 Embedded clusters in ATLASGAL clumps

HII-clumps represent the final stage of embedded cluster formation in the clumps, which are critical to build the bridge between the stellar kinematics and abundances community and the star-formation community, because protoclusters or embedded clusters in HII-clumps are the first stage of gas-free star clusters. Currently, it is difficult to construct complete star catalogs for early-stage and distant star clusters from observations (Feigelson et al., 2013; Romine et al., 2016; Kuhn et al., 2017). If there is a continuous transition from protoclusters to embedded clusters (J.-w. Zhou et al., 2024b), and assuming the stellar IMF is a universal optimal distribution function rather than a probability density (Kroupa et al., 2013; Yan et al., 2023)[1], we can directly infer the complete census of stars in HII-clumps. In this work, we will investigate the physical properties of embedded clusters in HII-clumps and estimate the final star formation efficiency of the clumps. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024d).

5.2.1 Synthetic embedded clusters

The HII-clump mass ($M_{\rm cl}$) and its embedded star cluster mass ($M_{\rm ecl}$) satisfy $M_{\rm ecl}$ = SFE * $M_{\rm cl}$, where SFE is the final star formation efficiency of the clump. In this work, we take SFE = 0.33 (C. J. Lada & E. A. Lada, 2003; Megeath et al., 2016; J. W. Zhou et al., 2024b). For all HII-clumps, the mass range is $\log_{10}(M_{\rm cl}/M_{\odot}) \in [1, 5]$, and we take the values uniformly at a step size of 0.1 in this range. Then according to the corresponding $M_{\rm ecl}$, a population of stars is synthesized using the GalIMF code (Schulz et al., 2015; Yan et al., 2017).

In Yan et al. (2023), assuming that the pre-main sequence stars have bolometric luminosities comparable to those of zero age main sequence (ZAMS) stars, the bolometric luminosity of each optimally sampled star in an embedded star cluster was calculated by the empirical mass–luminosity relations summarised in equation.(12) of Yan et al. (2023). Then, the luminosity of all stars were added together to get the luminosity of the embedded cluster. Yan et al. (2023) found that the optimal sampling model qualitatively agrees with the observation in Urquhart et al. (2022). However, the observational data

[1] See however Marks et al. (2012) and Dib (2023b) for how the slope of the IMF in the intermediate-to high-mass end can vary with the environment.

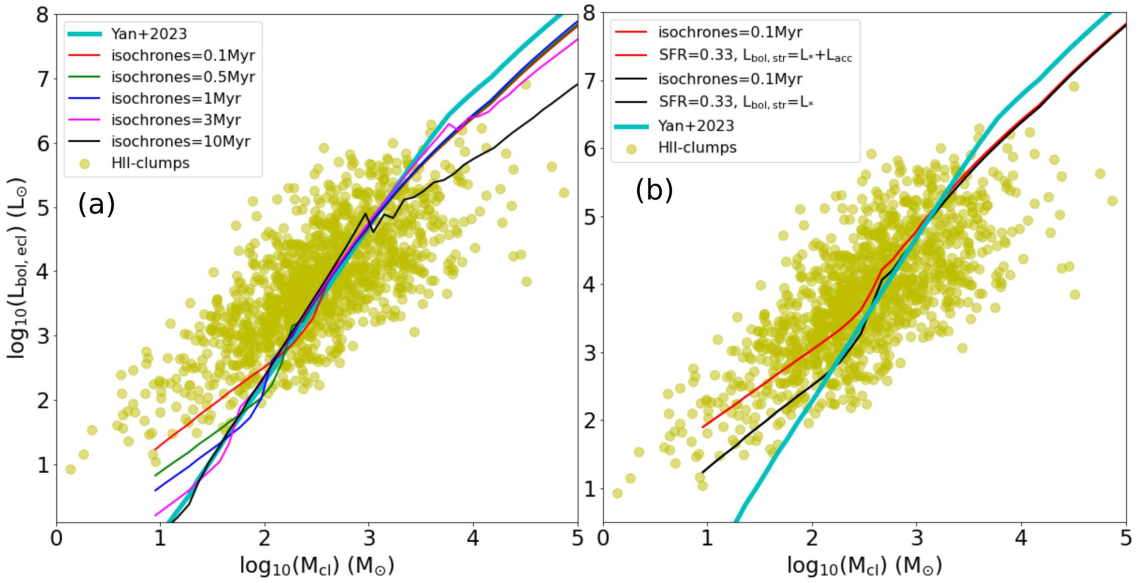


Figure 5.1: Fitting the observed HII-clumps in Urquhart et al. (2022) using the optimal sampling model. (a) Assuming $L_{\rm bol,str}$=L_*, L_* were estimated by different isochrones; (b) Assuming $L_{\rm bol,str}$=L_*+$L_{\rm acc}$, L_* were estimated by the 0.1 Myr isochrone.

systematically become more luminous than the optimal sampling model for low-mass HII-clumps, as shown in Fig.5.1(a). Low-mass embedded clusters in low-mass HII-clumps are predominantly populated by low-mass stars. As these low-mass stars become observable within HII-clumps, they are still in the process of contracting onto the main sequence with significantly larger luminosities. Consequently, it is essential to incorporate their pre-main sequence evolutionary tracks into the optimal sampling model. To achieve this goal, we use pre-main sequence isochrones to estimate the bolometric luminosity for each star rather than using the empirical mass–luminosity relation. For a synthetic embedded cluster, the mass of each star is known. Using the isochrones, the bolometric luminosity of a star can be estimated from its mass.

We use the isochrones from the MESA Isochrones and Stellar Tracks (MIST) project (Choi et al., 2016). As shown in Fig.5.1(a), using the 0.1 Myr isochrone appears to best fit the observed data points. Although it is unlikely that all the stars in a HII-clump are that young.

5.2.2 Accretion luminosity

In Fig.5.1(a), after using the isochrones, the fitting to the data is better than that in Yan et al. (2023) who only considered the zero age main sequence luminosities of all stars in the embedded cluster. However, the models are still systematically lower than the observations for low-mass HII-clumps, even using the 0.1 Myr isochrone. For HII-clumps, feedback from the HII region may cause part of the clump mass to be dispersed. Therefore, the observed data points will shift to the left, thus deviating from the models. In order to check this possibility, we compared the mass-size relations of the ATLASGAL clumps at four different evolutionary stages, and found that they are quite comparable, suggesting that the mass dispersal of HII clumps is not significant.

Given that HII-clumps have ongoing star formation, the accretion luminosity $L_{\rm acc}$ from protostars may contribute to the total bolometric luminosity, and it is not explicitly included in the pre-main sequence evolutionary tracks. For each accreting protostar in a clump, its total bolometric luminosity ($L_{\rm bol,str}$) is the sum of the accretion luminosity ($L_{\rm acc}$) and the luminosity of the protostar (L_*). As shown in Fig.4 of

Hosokawa & Omukai (2009), when M_*<10 $M_\odot$, $L_{\rm acc}$ is much larger than L_*. L_* can be calculated using the isochrones. For the accretion, Davies et al. (2011) compared different accretion models, and found that the prescription of accretion in McKee & Tan (2003b) excellently fits the observations. Thus, we compute $L_{\rm acc}$ using equations (41) and (73) in McKee & Tan (2003b). As shown in Fig.5.1(b), after adding the accretion luminosity for each star, the model can fit the observed data.

5.2.3 Age distribution

In the above analysis, we only considered the 0.1 Myr isochrone. Actually, many stars in HII-clumps should be older than 0.1 Myr. There is certainly an age spread in the young star and protostar populations in embedded clusters of HII-clumps. Thus, we should choose the appropriate isochrones based on the ages of the stars. Moreover, not all stars in HII-clumps are in an accretion phase, which also depends on their ages. In order to address these issues, we use the method presented in Dib et al. (2025) to create an age distribution of the stellar population in each synthetic cluster to properly distribute the isochrone to each star. Then, we sorted out the fraction of stellar objects that are still protostars to properly add their accretion luminosities. We examine scenarios where the star formation history (SFH) remains constant and others where it varies over time.

The typical duration of the young protostellar stage is ≈0.5 Myr (≈ 0.1 Myr for the Class 0 phase and ≈ 0.44 Myr for the Class I phase) and that of the Class II phases is ≈ 2 Myr (I. Evans N. J. et al., 2009; Hsieh & Lai, 2013; Pokhrel et al., 2020; Megeath et al., 2022; Dib et al., 2025). In this work, only the protostars with ages less than 0.5 Myr are used to calculate the accretion luminosity.

For the case of a flat SFH, nearly 90% stars are excluded. Thus, the case with a flat SFH can not account for difference in luminosity. In the cases with time-dependent SFHs, as long as 20% of the stars within the embedded cluster are still accreting, the contribution of accretion luminosity will be significant for low-mass clumps. Currently, we have assigned corresponding isochrones based on the age of each star, rather than using the 0.1 Myr isochrone for all stars in Fig.5.1(b). As shown in J. W. Zhou et al. (2024d), the fitting lines remain comparable to Fig.5.1(b).

5.2.4 Critical clump mass

In Fig.4 of Hosokawa & Omukai (2009), when M_*>10 $M_\odot$, L_* begins to exceed $L_{\rm acc}$. Thus, for high-mass stars, the contribution from L_* to $L_{\rm bol,str}$ is more important than $L_{\rm acc}$. For the synthetic embedded clusters, when $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.2$, stars with masses larger than 10 $M_\odot$ begin to appear. In Fig.5.1(b), we can see the black line indeed begins to approach the red line around $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.2$. But now, the number of low-mass stars is still much larger than high-mass stars in the synthetic embedded cluster, the total accretion luminosity of low-mass stars remains dominant. When $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.8$, the luminosity of high-mass stars becomes dominant and $L_{\rm acc}$ can be ignored. Therefore, there is a mass limit of $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.8$, when a HII-clump mass is less than this limit, one must consider the contribution of the accretion luminosity from protostars to the observed bolometric luminosity ($L_{\rm bol,obs}$) of the HII-clump.

5.3 Star formation efficiency of ATLASGAL clumps

For ATLASGAL clumps with HII regions, in J. W. Zhou et al. (2024d), we synthesized their embedded clusters under the assumption that the stellar initial mass function follows a universal optimal distribution function rather than a probability density (Kroupa et al., 2013; Yan et al., 2023). We utilized the 0.1 Myr isochrone to estimate the bolometric luminosity of individual stars within an embedded cluster,

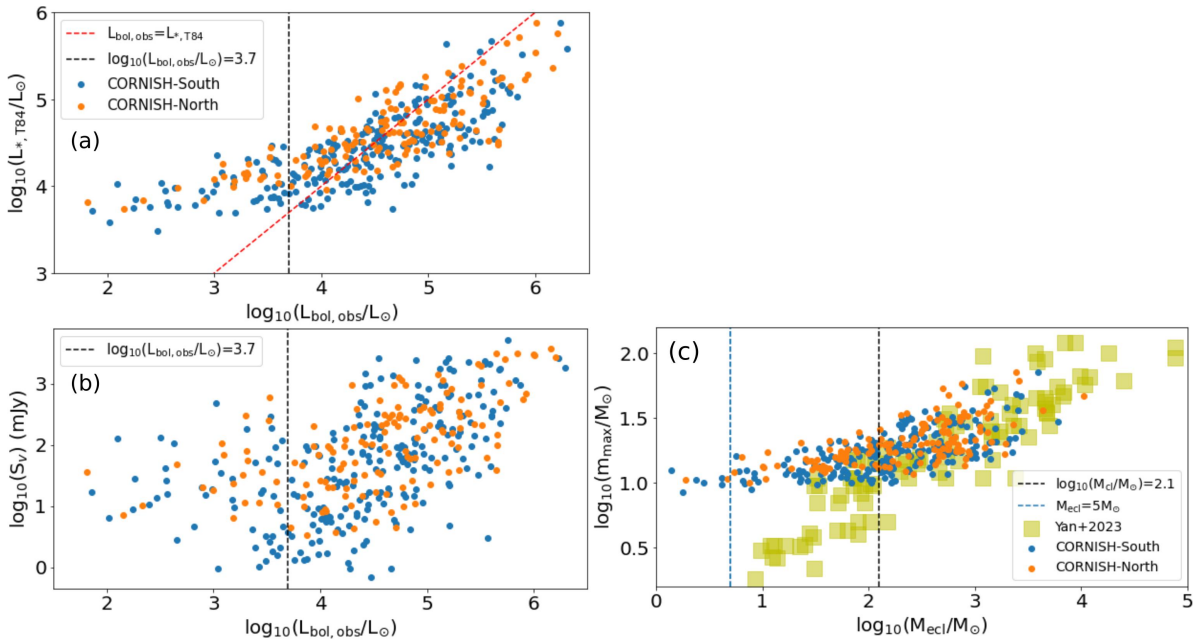


Figure 5.2: (a) The correlation between $L_{*,\mathrm{T84}}$ and $L_{\mathrm{bol,obs}}$; (b) The correlation between S_ν and $L_{\mathrm{bol,obs}}$; (c) The $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation. Yellow squares represent the star clusters from the table of Yan et al. (2023). The vertical dashed lines mark the turning points, see text for the details.

augmenting this with the accretion luminosity of each star in the cluster. The cumulative bolometric luminosity of the synthetic embedded clusters aligns closely with the observed bolometric luminosity of HII-clumps, validating the effectiveness of the method. As a follow up study based on J. W. Zhou et al. (2024d), in this work, we focus on the ionizing sources in HII-clumps. They may represent the most massive stars in the embedded clusters. We explore their physical properties by comparing the observations with numerical simulations. The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2024b).

5.3.1 Accretion

We calculated the physical parameters of the matched radio sources in Sec.5.1 following the formalism of Mezger & Henderson (1967) and Rubin (1968) using the equations collected in Sánchez-Monge et al. (2013). In this work, we determined the stellar luminosity $L_{*,\mathrm{T84}}$ corresponding to the calculated N_{Ly} (Lyman continuum flux) by interpolating the stellar luminosity given as a function of the Lyman continuum flux in Table.1 of R. I. Thompson (1984).

The stellar luminosity $L_{*,\mathrm{T84}}$ of the brightest star should be smaller than the bolometric luminosity $L_{\mathrm{bol,obs}}$ of the corresponding HII-clump. However, in Fig. 5.2(a), for low-luminosity HII-clumps, $L_{*,\mathrm{T84}}$ is significantly larger than $L_{\mathrm{bol,obs}}$, which is abnormal. In Fig. 5.2(b), there is a turning point around $\log_{10}(L_{\mathrm{bol,obs}}/L_\odot) \approx 3.7$. For HII-clumps with luminosity lower than this value, there is no clear correlation between $L_{\mathrm{bol,obs}}$ and radio flux density S_ν. The turning point $\log_{10}(L_{\mathrm{bol,obs}}/L_\odot) \approx 3.7$ corresponds to the HII-clump mass $\log_{10}(M_{\mathrm{cl}}/M_\odot) \approx 2.55$ in J. W. Zhou et al. (2024d). Using the 0.5 Myr isochrone, we estimated the stellar mass $m_{*,\mathrm{iso}}$ according to the stellar luminosity $L_{*,\mathrm{T84}}$. The ionizing sources represent the most massive stars in HII-clumps, i.e. $m_{*,\mathrm{iso}} \approx m_{\mathrm{max}}$. Considering the $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation (Weidner & Kroupa, 2006; Weidner et al., 2013a; Yan et al., 2023), a correlation between the embedded star cluster mass M_{ecl} and the mass of the most massive star m_{max}, there should be a correlation between $m_{*,\mathrm{iso}}$ and the mass of the corresponding embedded cluster. The clump mass

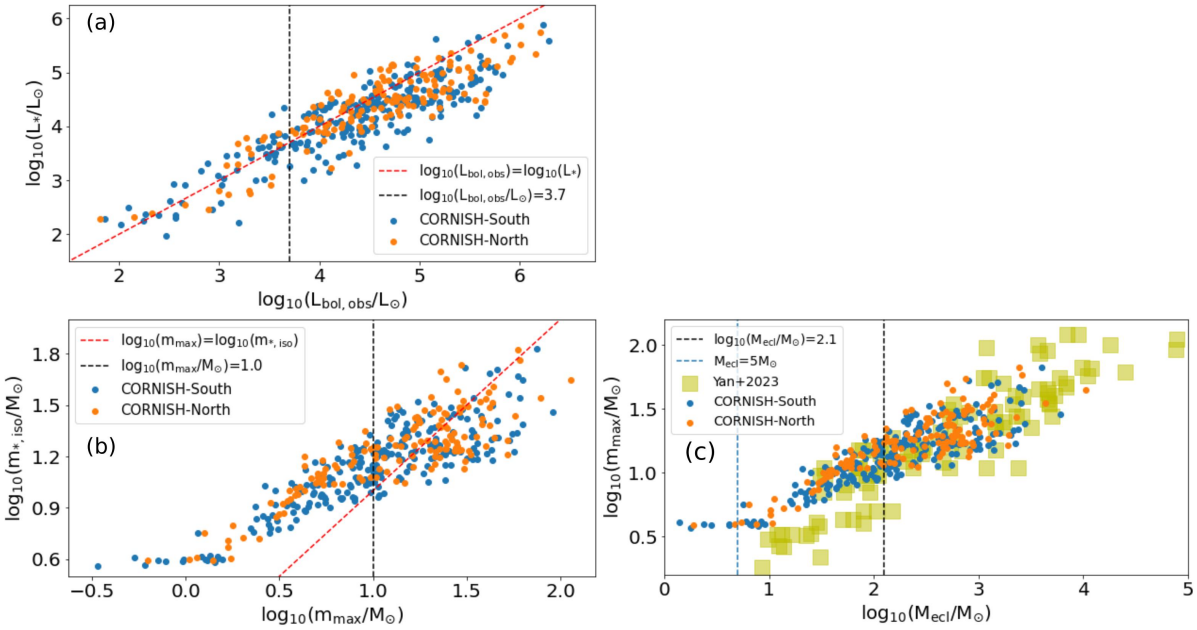


Figure 5.3: (a) The correlation between L_* and $L_{\rm bol,obs}$; (b) Comparison between $m_{\rm max}$ from the synthetic clusters and $m_{*,\rm iso}$ from the observation; (c) The $m_{\rm max} - M_{\rm ecl}$ relation. Yellow squares represent the star clusters from the table of Yan et al. (2023).

$M_{\rm cl}$ and its embedded cluster mass $M_{\rm ecl}$ satisfy $M_{\rm ecl}$ = SFE $\times M_{\rm cl}$, where SFE is the final star formation efficiency of the clump, and which we assumed to be SFE=0.33 (C. J. Lada & E. A. Lada, 2003; Megeath et al., 2016; J. W. Zhou et al., 2024d). For 100 embedded clusters with age < 5 Myr, the table in Yan et al. (2023) gives the values of their $m_{\rm max}$ and $M_{\rm ecl}$. This sample fits the $m_{\rm max} - M_{\rm ecl}$ relation well. In Fig. 5.2(c), we compared our sample with the sample from Yan et al. (2023). When $\log_{10}(M_{\rm ecl}/M_\odot) >$ 2.1, the $m_{*,\rm iso} - M_{\rm ecl}$ relation follows the $m_{\rm max} - M_{\rm ecl}$ relation. Assuming SFE=0.33, from the turning point $\log_{10}(M_{\rm ecl}/M_\odot) \approx 2.1$, we have $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.55$, consistent with the findings above.

The turning point $\log_{10}(\rm M_{cl}/M_\odot) \approx 2.55$ is used to judge the relative importance of accretion luminosity in J. W. Zhou et al. (2024d). Therefore, the abnormal tail in Fig. 5.2 on the left of the vertical dashed lines means that if the accretion luminosity is dominant, the formalism of Mezger & Henderson (1967) and Rubin (1968) cannot be used, because the main Lyman continuum flux or radio emission are not predominantly from the stars. The high radio flux density of low-luminosity and low-mass HII-clumps in Fig. 5.2(b) indicate that accretion is an effective way to produce radio emission. Therefore, the Lyman continuum flux $N_{\rm Ly}$ derived from radio flux density S_ν includes a significant contribution from the accretion process. The previous studies of HII regions in Lumsden et al. (2013), Urquhart et al. (2013) and Sánchez-Monge et al. (2013) have noted that there is an excess of Lyman photons over what would be expected from the ionizing stars compared to the bolometric luminosities. The inclusion of the luminosity contribution from accretion may provide an explanation.

We obtained the stellar luminosity $L_{*,\rm T84}$ from the Lyman continuum flux $N_{\rm Ly}$. Actually, $L_{*,\rm T84}$ is not only the stellar luminosity L_*, but the sum of L_* and the accretion luminosity $L_{\rm acc}$, i.e. $L_{*,\rm T84}$=L_*+$L_{\rm acc}$. Using the method described in J. W. Zhou et al. (2024d), we calculated the accretion luminosity $L_{\rm acc}$ of the most massive stars of the synthetic embedded clusters, and subtracted $L_{\rm acc}$ from the corresponding $L_{*,\rm T84}$. Then we repeated the above analysis using L_* rather than $L_{*,\rm T84}$. As shown in Fig. 5.3(a), the abnormal tail disappears.

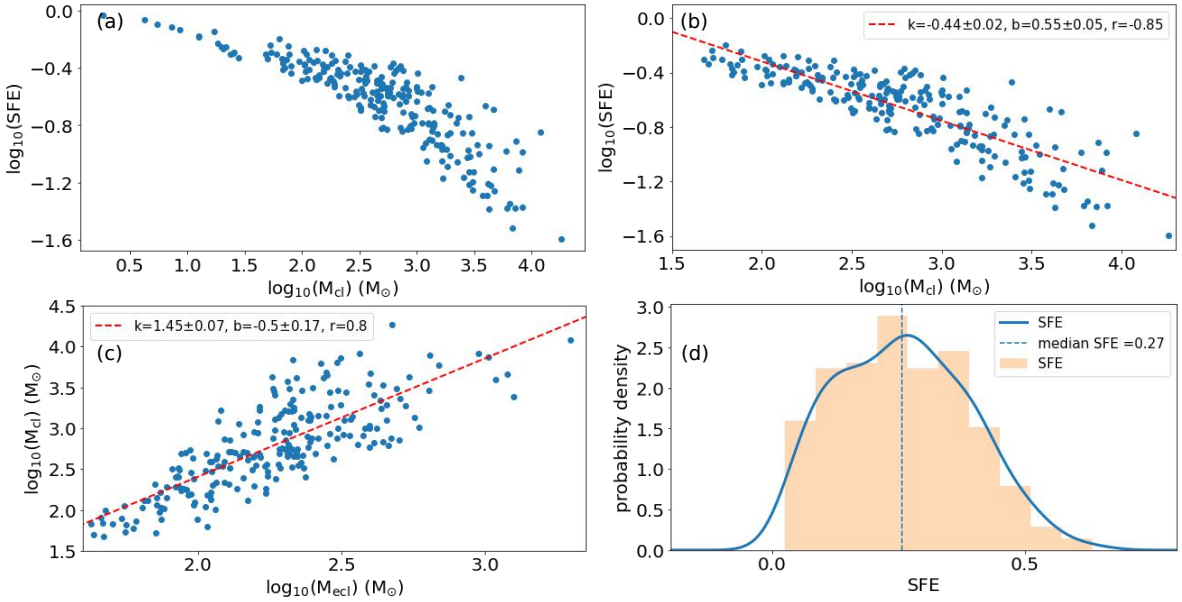


Figure 5.4: Star formation efficiency derived from the observations. (a) The correlation between the star formation efficiency and the clump mass for all the data points presented in Fig.5.3(b); (b) Eliminated the abnormal tail in panel (a); (c) The correlation between the embedded cluster mass and the corresponding HII-clump mass; (d) The distribution of the SFEs. Red dashed line is the linear fitting to the data points. "k" and "b" are the slope and the intercept, "r" is the Pearson correlation coefficient.

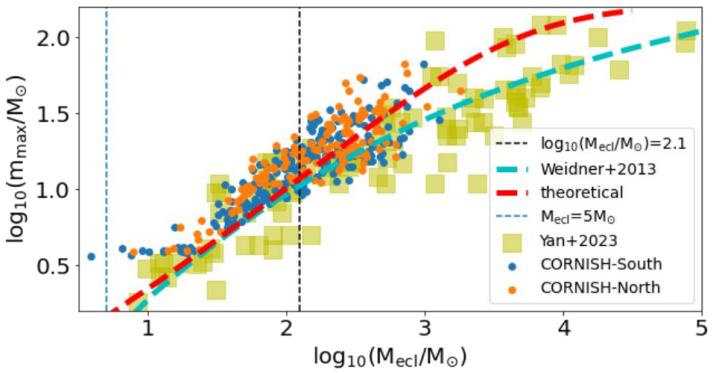


Figure 5.5: Same as Fig.5.3(c), but after modifying the SFE using equation.5.2. Cyan dashed line is the $m_{\rm max} - M_{\rm ecl}$ relation fitted in Weidner et al. (2013a) based on the observation. Red dashed line is the $m_{\rm max} - M_{\rm ecl}$ relation derived from the IMF theory, i.e. equation.5.1.

5.3.2 Star formation efficiency

In Fig. 5.3(c), for low-mass clusters, there is a mismatch between our sample and the sample in Yan et al. (2023). As discussed above, we may be overestimating the stellar mass. Another possibility is that low-mass clumps have higher SFEs (> 0.33). Assuming all measured maximum stellar masses follow the $m_{\rm max} - M_{\rm ecl}$ relation, we can calculate the mass of the embedded cluster and then the star formation efficiency for each clump. The third-order polynomial fit to the observed $m_{\rm max} - M_{\rm ecl}$ relation presented in Weidner et al. (2013a) has large errors. As an alternative, we derived a theoretical $m_{\rm max} - M_{\rm ecl}$ relation directly from the initial mass function (IMF) in J. W. Zhou et al. (2024b). The embedded cluster mass,

$M_{\rm ecl}$, is given by

$$M_{\rm ecl} \approx \frac{5.37}{\frac{0.77}{m_{\rm max}^{1.3}} - 0.001} - \frac{3.33}{m_{\rm max}^{0.3}\left(\frac{0.77}{m_{\rm max}^{1.3}} - 0.001\right)}, \tag{5.1}$$

where we use the $m_{\rm max}$ values displayed in Fig .5.3(c) to derive the corresponding value of $M_{\rm ecl}$ for each embedded cluster.

The abnormal tail in Fig. 5.3(b) is visible in Fig. 5.4(a) again, presented as the SFEs close to 1. We excluded this tail in the subsequent analysis by requiring the clump mass larger than $10^{1.5}$ $M_\odot$. Fig. 5.4(b) displays a strong correlation between the SFE and the clump mass. The SFE decreases with increasing clump mass, with an upper limit of ≈0.65, This is also visible in Fig.5 of Wells et al. (2022). We fit the relation between $M_{\rm cl}$ and the SFE with a power law of the form SFE $\propto M_{\rm cl}^k$ and find a value of $k = -0.44 \pm 0.02$. This anti-correlation between the mass of the clumps and their SFEs has been predicted theoretically by Dib et al. (2013). For the simple case of uniform density clumps, Dib et al. (2013) predicted the clump mass and the SFE to follow SFE $\propto M_{\rm cl}^{-0.6}$ and to be the result of an increasingly stronger effect of stellar feedback in clumps with higher masses which leads to a faster gas expulsion and limits the value of the SFE. Interestingly, the median value of the SFEs is ≈0.3, consistent with our previous work (J. W. Zhou et al., 2024d). An SFE ≈ 0.33 is also consistent with values obtained from hydrodynamic simulations (Machida & Matsumoto, 2012; Bate et al., 2014) and observations of embedded clusters in the solar neighborhood (C. J. Lada & E. A. Lada, 2003; Megeath et al., 2016).

In order to assess the uncertainty in the observations, now we independently study the clump's SFE using theoretical models. In Fig. 5.4(b), the upper limit of the SFE is ≈0.65. In Fig.4 of Wells et al. (2022), a SFE larger than 0.5 is necessary to fit the low-mass HII-clumps. Therefore, we consider the SFEs that fall in the range [0.05, 0.65] with a step of 0.05. Then each HII-clump has 13 alternative SFEs. We consider both the stellar luminosity and accretion luminosity of each star in the embedded cluster of a HII-clump, and also apply the pre-main sequence evolutionary track for each star. There is certainly an age spread of the stellar populations in embedded clusters within HII-clumps. Therefore, selecting the appropriate isochrones according to the stars' ages is essential. Additionally, not all stars in HII-clumps are in the accretion phase, as this too is age-dependent. We should identify the fraction of stellar objects that remain protostars to properly add their accretion luminosities. In order to address these issues, we apply the method described in J. W. Zhou et al. (2024d) to generate an age distribution for the stellar population in each synthetic cluster. The age distribution of stars within a HII-clump depends on the star formation history (SFH) of the clump. As discussed in J. W. Zhou et al. (2024d), compared to a constant SFH, burst-like and time-dependent SFHs can better fit the observational data. Thus, we only consider the time-dependent SFHs in this work. As described in J. W. Zhou et al. (2024d), there are 12 different SFHs with different age peaks ($t_{\rm p}$) and age standard deviations ($\sigma_{\rm t}$), i.e. $t_{\rm p} = 0.1, 0.25, 0.5, 1$ Myr and $\sigma_{\rm t} = 0.25, 0.5, 1$ Myr. Finally, each clump has 156 different combinations of SFEs and SFHs, resulting in 156 distinct bolometric luminosities. The combination of a SFE and a SFH that produces the bolometric luminosity of the synthetic embedded cluster closest to the observed bolometric luminosity of the corresponding HII-clump is considered as the optimal one. Overall, the SFEs of HII-clumps derived from the observation and the model are comparable. From the model, we fit the relations

$$\log_{10}({\rm SFE}) = (-0.37 \pm 0.01) \times \log_{10}(M_{\rm cl}) + (0.42 \pm 0.04), \tag{5.2}$$

and

$$\log_{10}(M_{\rm cl}) = (1.02 \pm 0.02) \times \log_{10}(M_{\rm ecl}) + (0.52 \pm 0.05). \tag{5.3}$$

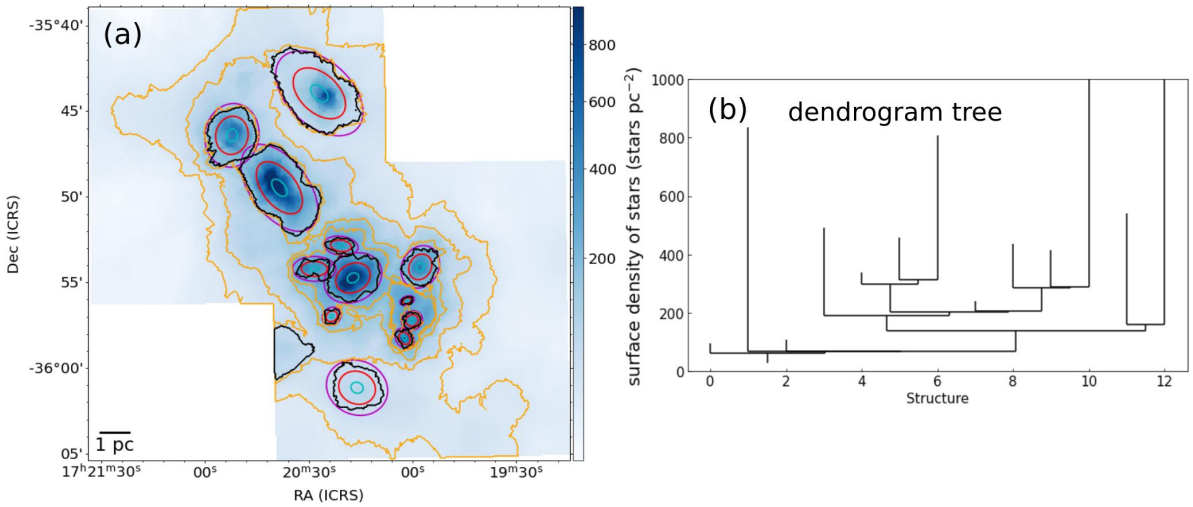


Figure 5.6: Hierarchical structures of the clusters in NGC 6334. (a) The background is the surface density of stars. The black and orange contours show the masks of leaf and branch clusters, respectively. The magenta, red, and cyan cycles are the approximated ellipses with effective radii $R_{\rm eff}$, $R_{\rm eff,1}$, and $R_{\rm eff,2}$, respectively. (b) The dendrogram tree shows the hierarchical cluster structures.

Then we use equation. 5.2 to calculate the SFE or the mass of the embedded cluster for each HII-clump. As shown in Fig. 5.5, after the modification of the SFE, the mismatch in Fig. 5.3(c) indeed disappears.

5.4 Isolated evolution from clumps to embedded clusters

Observations of MSFRs at various stages of their star-forming lifespan unveil diverse stellar cluster structures. Notably, a striking feature of the spatial distributions of young stars is their organization into clusters and subclusters. In addition to the main clusters within these regions, smaller stellar groups can be observed outside the main clusters, or even as subclusters within the main clusters (C. J. Lada & E. A. Lada, 2003; Portegies Zwart et al., 2010b; Krumholz et al., 2019). Stars form from hierarchically collapsing molecular clouds, which often results in clustered star formation that mirrors the intricate structure of the original molecular cloud. In a similar way to what J. W. Zhou et al. (2024e) and J. W. Zhou et al. (2024g) reported for molecular clouds and clumps, we used the dendrogram algorithm to decompose the surface density distributions of stars into hierarchical structures. This revealed the multi-scale structures of star clusters. The work presented in this section is reproduced from the publication by J.-w. Zhou et al. (2024b).

5.4.1 Cluster mass

Fig.5.6 displays an example that shows the hierarchical structures of the clusters in NGC 6334, decomposed by the dendrogram algorithm. The *astrodendro* package can output the mask of each structure, which can be directly used to calculate the physical quantities of the identified structures, such as the mass and the effective radius. The algorithm can also approximate the morphology of each structure as an ellipse. However, as shown in Fig.5.6, the long and short axes output by the dendrogram, a and b, define an ellipse that significantly underestimates the size of the identified structure, but multiplying by a factor of 2.5 is appropriate to cover the structure (J. W. Zhou et al., 2024e; J. W. Zhou et al., 2024g). The effective physical radius of an ellipse is then $R_{\rm eff} = \sqrt{2.5a \times 2.5b} * d$, with d being the distance to each star-forming region.

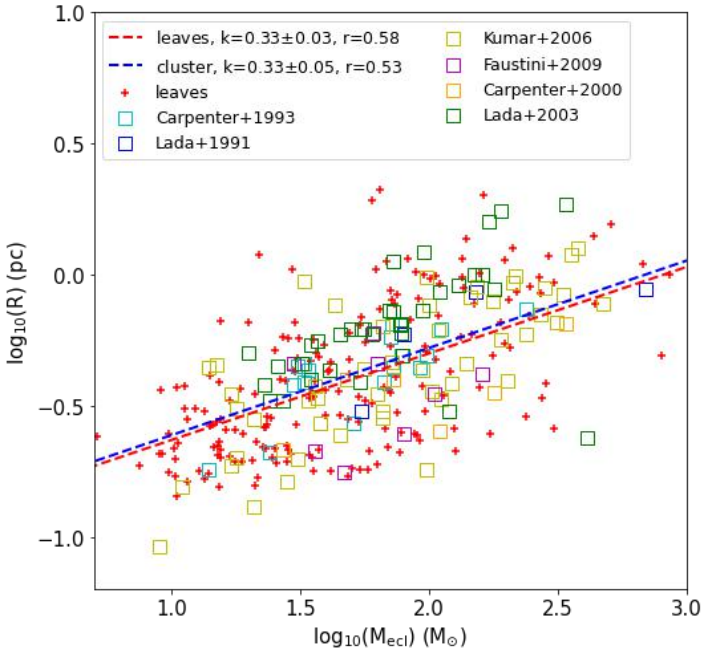


Figure 5.7: Mass-size relations of leaf clusters (red plus) identified in this work fit by the dashed red line, and embedded clusters (colored squares) in the catalogs of E. A. Lada et al. (1991), Carpenter et al. (1993), C. J. Lada & E. A. Lada (2003), Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009) fit by the dashed black line. The slope for all is ≈0.33.

5.4.2 Self-similar mass-size relations

Fig.5.7 shows that after the observed radii are unified to the half-mass radii (J.-W. Zhou et al., 2024), the embedded clusters in the catalogs of E. A. Lada et al. (1991), Carpenter et al. (1993), C. J. Lada & E. A. Lada (2003), Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009) are quite comparable with the leaf clusters identified in this work. Their mass-size relations have almost the same slope, ≈0.33. As marked in Fig.5.6(a) by colored ellipses, we calculated the physical parameters of leaf clusters by taking the effective radius of the ellipse as $R_{\rm eff}$ =$\sqrt{2.5a \times 2.5b} * d$, $R_{\rm eff,1}$ =$\sqrt{1.5a \times 1.5b} * d$, and $R_{\rm eff,2}$ =$\sqrt{0.5a \times 0.5b} * d$. As shown in Fig.5.8(a), for these three cases, the data points mainly move down to the lower left side along the line of the mass-size relation with a constant slope of ≈0.3. Therefore, the slope of the mass-size relation is independent of the cluster radius and mass measurements, implying self-similar cluster structures. Interestingly, when the three different radii and mass measurements are mixed together, a steeper mass-size relation is obtained, similar to the mass-size relation found by Pfalzner et al. (2016).

Branch clusters are the complexes of leaf clusters. As shown in Fig 5.8(b), they follow a similar trend as those shown in Fig 5.8(a). When leaf and branch clusters were mixed together, we again obtained the steeper mass-size relation. The data points in Table.1 of Kuhn et al. (2015) also follow this trend, although they were identified and measured with completely different methods. The radii and masses of the subclusters in Kuhn et al. (2015) are quite comparable with the middle-scale branch clusters in this work, and they have similar slopes for the mass-size relations in Fig.5.8(b). In conclusion, different measurements for embedded clusters at different scales and ages give comparable mass-size relations.

Here and in Kuhn et al. (2015), a unified measurement of the samples results in a shallower slope for the mass-size relation. In contrast, when different measurements at different scales are mixed, the mass-size relation is always steeper. Although the mixing may not make sense, it is unavoidable when

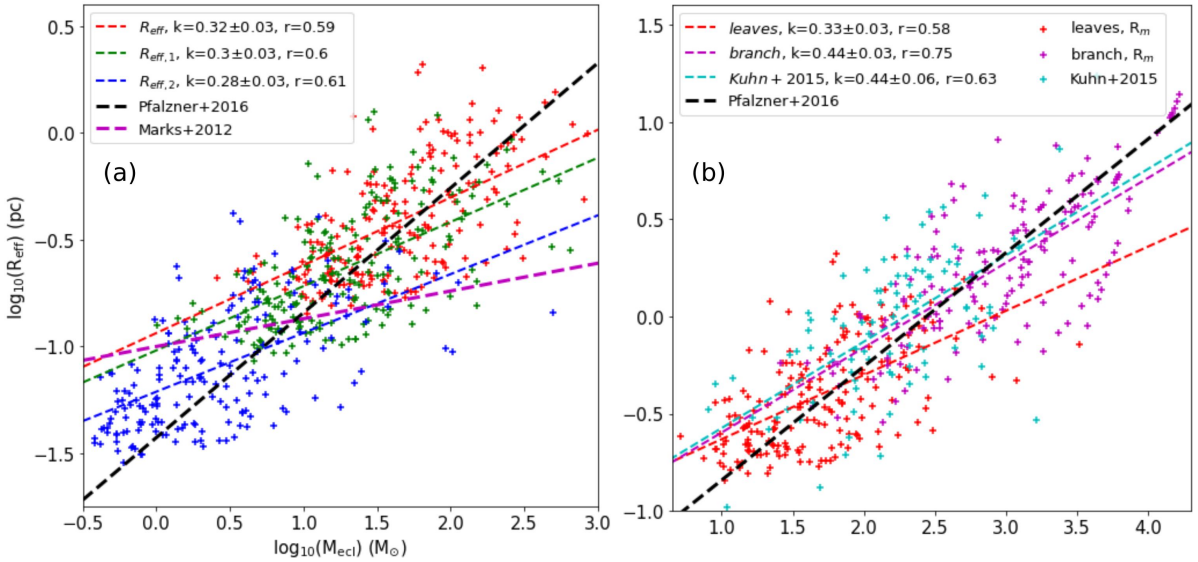


Figure 5.8: Mass-size relation of star clusters measured at different scales. (a) Red, green, and blue pluses represent leaf clusters with effective radii $R_{\rm eff}$, $R_{\rm eff,1}$, and $R_{\rm eff,2}$, respectively. (b) Magenta, cyan, and red pluses represent branch clusters, the subclusters in Table.1 of Kuhn et al. (2015), and leaf clusters, respectively. k and r are the slope of the mass-size relation and the Pearson correlation coefficient, respectively.

different catalogs from different observations are collected together. Thus, in J.-W. Zhou et al. (2024), we unified the observed radii from different observations to the half-mass radii using the templates obtained from direct N-body simulations. In this work, the unified measurement of the samples from the MYStIX project and the unified identification of the clusters effectively avoid this issue.

The exponent $k \approx 0.33$ in Fig. 5.7 and Fig. 5.8(a) indicates a uniform distribution of the mass through the scale if the structures were to be roundish. This is consistent with the suggestion made in J.-W. Zhou et al. (2024), namely that all subclusters initially have a Plummer profile. The Plummer radius serves as a measure of the core radius, which is the region where the density remains relatively constant. For a Plummer profile with the Plummer radius $r_{\rm pl}$, the mass within the radius r is given by

$$M(r) = M_{\rm ecl} \frac{\left(\frac{r}{r_{\rm pl}}\right)^3}{\left[1 + \left(\frac{r}{r_{\rm pl}}\right)^2\right]^{\frac{3}{2}}}. \tag{5.4}$$

Thus, different radii in different observations can be unified to the half-mass radii. This operation does not change the slope of the mass-size relation, which only shifts the line of the mass-size relation, as shown in Fig.5.8(a).

5.4.3 Isolated evolution

It remains controversial whether massive clusters are formed by the merger of clusters with lower masses. If there is a considerable merger between embedded clusters, we expect the separations between embedded clusters to be significantly smaller than their precursors, that is, clumps. In order to test this hypothesis, we used the clumps from the ATLASGAL survey, which provides the most well-characterized currently available sample of high-mass star-forming clumps (Urquhart et al., 2022). Seven of the 17 MSFRs in Kuhn et al. (2015) are covered in the ATLASGAL survey, which are the Lagoon nebula,

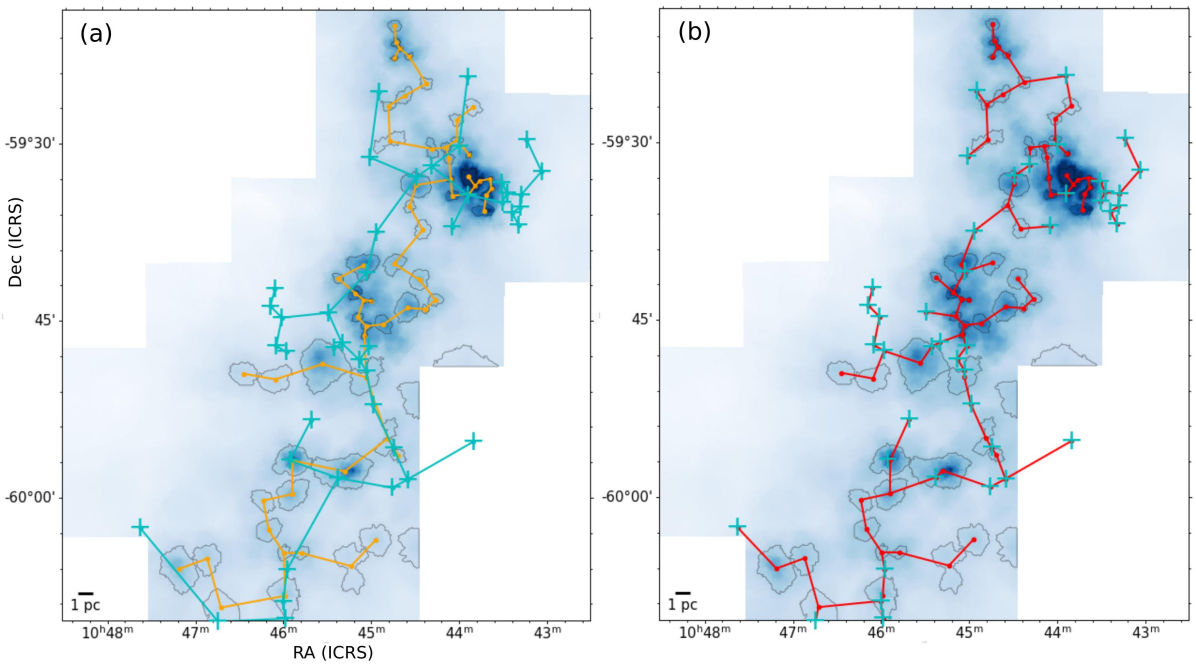


Figure 5.9: The Carina Nebula serves as an example illustrating the use of the minimum spanning tree (MST) method. The cyan plus marks the positions of ATLASGAL clumps. The black contours show the masks of leaf clusters. (a) Orange and cyan lines show the minimum spanning trees between clusters and between clumps, respectively. (b) MST between clusters and clumps.

NGC 6334, NGC 6357, the Eagle nebula, M 17, the Carina nebula, and the Trifid nebula. In Fig.5.9, the minimum spanning tree (MST) method (Cartwright & Whitworth, 2004; Dib et al., 2018; Dib & Henning, 2019) was used to measure the separation between each two neighboring objects (between clusters, between clumps, and between clusters and clumps). Interestingly, the peak values of the separation distributions in three cases are comparable. This indicates the following: (1) The evolution from clump to embedded cluster is fixed at the initial position of the clump. There is no significant positional change between the embedded cluster and its precursor; (2) There is no significant merger between embedded clusters for the samples in the MYStIX project; (3) Feedback from embedded clusters does not redistribute the surrounding clumps, consistent with the conclusions in J. W. Zhou et al. (2024e) and J. W. Zhou et al. (2024g), which show that feedback from embedded clusters does not significantly change the physical properties of the surrounding dense gas structures.

In short, the evolution from a clump to an embedded cluster proceeds in isolation and locally, with no significant effect on the surrounding objects. The well-established $m_{\rm max} - M_{\rm ecl}$ relation (R. B. Larson, 2003; Bonnell et al., 2004; Weidner & Kroupa, 2006; Weidner et al., 2013a; Stephens et al., 2017; Yan et al., 2023; J. W. Zhou et al., 2024b) also does not support the hypothesis that mergers are common during the formation of embedded clusters. Using Gaia DR2 data, Kuhn et al. (2019) investigated the kinematics of some subclusters identified in Kuhn et al. (2014). They found no evidence that these groups are merging. We note that the imbalanced populations of star clusters and clumps in terms of numbers may impact the statistics of the spacing. More samples are necessary for further validation.

5.5 N-body simulations

The work presented in this section is reproduced from the publications by J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025).

5.5.1 Individual cluster simulation

We computed three clusters with stellar masses of 300 $M_\odot$, 1000 $M_\odot$, and 3000 $M_\odot$ in the first 100 Myr. The parameters for the simulation are summarized from previous works (Kroupa et al., 2001; Baumgardt & Kroupa, 2007; Banerjee et al., 2012; Banerjee & Kroupa, 2013; Banerjee & Kroupa, 2014; Banerjee & Kroupa, 2015; Oh et al., 2015; Oh & Kroupa, 2016; Banerjee & Kroupa, 2017; Brinkmann et al., 2017; Oh & Kroupa, 2018; L. Wang et al., 2019; Pavlik et al., 2019; Dinnbier et al., 2022; J.-W. Zhou et al., 2024; J.-w. Zhou et al., 2024a). The previous simulations have already demonstrated the effectiveness and rationality of the parameter settings (see below). The influence of different parameter settings on simulation results and the discussion of the multidimensional parameter space can also be found in the works cited above. The initial density profile of all clusters is the Plummer profile (Aarseth et al., 1974; D. Heggie & Hut, 2003; Kroupa, 2008), an appropriate choice since the molecular clouds' filaments in which stars form have been found to have Plummer-like cross sections (Malinen et al., 2012; P. J. André et al., 2022), and open star clusters can also be described by the Plummer model (Röser et al., 2011; Röser & Schilbach, 2019). Moreover, such a specific initial profile does not significantly affect the overall expansion rate of a cluster, as discussed in Banerjee & Kroupa (2017), which is primarily governed by the total stellar mass loss and the dynamical interactions occurring within the inner part of the cluster. The half-mass radius, r_h, of the cluster is given by the $r_\mathrm{h} - M_\mathrm{ecl}$ relation (Marks & Kroupa, 2012):

$$\frac{r_\mathrm{h}}{\mathrm{pc}} = 0.10^{+0.07}_{-0.04} \times \left(\frac{M_\mathrm{ecl}}{M_\odot}\right)^{0.13\pm0.04} . \tag{5.5}$$

All clusters are fully mass segregated (S=1), with no fractalization, and in a state of virial equilibrium (Q=0.5). S and Q are the degree of mass segregation and the virial ratio of the cluster, respectively. More details can be found in Küpper et al. (2011) and the user manual for the `McLuster` code. The initial segregated state is detected for young clusters and star-forming clumps/clouds (Littlefair et al., 2003; L. Chen et al., 2007; Portegies Zwart et al., 2010b; Kirk & Myers, 2011; Getman et al., 2014; Lane et al., 2016; Alfaro & Román-Zúñiga, 2018; Plunkett et al., 2018; Pavlik et al., 2019; Nony et al., 2021; Y. Zhang et al., 2022; F. Xu et al., 2024), but the degree of mass segregation is not clear. In simulations of the very young massive clusters R136 and NGC 3603 with gas expulsion by Banerjee & Kroupa (2013), mass segregation does not influence the results. In J.-W. Zhou et al. (2024), we compared S=1 (fully mass segregated) and S=0.5 (partly mass segregated) and found similar results. We also discussed settings with and without fractalization in J.-W. Zhou et al. (2024); the results of the two are also consistent. The initial mass functions of the clusters are chosen to be canonical (Kroupa, 2001), with the most massive star following the $m_\mathrm{max} - M_\mathrm{ecl}$ relation of Weidner et al. (2013a):

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \tag{5.6}$$

where $y = \log_{10}(m_\mathrm{max}/M_\odot)$, $x = \log_{10}(M_\mathrm{ecl}/M_\odot)$, a_0 = -0.66, a_1 = 1.08, a_2 = -0.150, and a_3 = 0.0084. We assumed the clusters to be at solar metallicity (i.e., Z = 0.02; von Steiger & Zurbuchen, 2016). The clusters travel along circular orbits within the Galaxy, positioned at a galactocentric distance of 8.5 kpc, and are moving at a speed of 220 km s^{-1}.

The initial binary setup follows the method described in L. Wang et al. (2019). All stars are initially in binaries, that is to say, f_b=1, where f_b is the primordial binary fraction. Kroupa (1995b) and Kroupa (1995d) propose that stars with masses below a few $M_\odot$ are initially formed with a universal binary distribution function and that star clusters start with a 100% binary fraction. Inverse dynamical population synthesis was employed to derive the initial distributions of binary periods and mass ratios. Belloni et al., 2017 introduces an updated model of pre-main-sequence eigenevolution, originally developed by

Kroupa, 1995d, to account for the observed correlations between the mass ratio, period, and eccentricity in short-period systems. For low-mass binaries, we adopted the formalism developed in Kroupa (1995b) and Kroupa (1995d) and Belloni et al. (2017) to characterize the period, mass ratio, and eccentricity distributions. For high-mass binaries (OB stars with masses > 5 $M_\odot$), we utilized the Sana et al. (2012) distribution, which is derived from O stars in OCs. The distribution functions of the period, mass ratio, and eccentricity are presented in Oh et al., 2015 and Belloni et al., 2017.

Accurately modeling gas removal from embedded clusters is challenging due to the complexity of radiation hydrodynamical processes, which involve uncertain and intricate physical mechanisms. To simplify the approach, we simulated the key dynamical effects of gas expulsion by applying a diluting, spherically symmetric external gravitational potential to a model cluster, following the method presented in Kroupa et al. (2001) and Banerjee & Kroupa (2013). This analytical approach is partially validated by Geyer & Burkert (2001), who conducted comparison simulations using the smoothed particle hydrodynamics method to treat the gas. The hydrodynamics+N-body simulations in Farias et al. (2024) also confirm that the exponential decay model presented in equation.5.7 generally provides a good description of gas removal driven by radiation and wind feedback. Specifically, we used the potential of the spherically symmetric, time-varying mass distribution:

$$\begin{aligned} M_g(t) &= M_g(0) && t \leq \tau_d, \\ M_g(t) &= M_g(0)\exp\left(-\tfrac{(t-\tau_d)}{\tau_g}\right) && t > \tau_d. \end{aligned} \tag{5.7}$$

Here, $M_g(t)$ is the total mass of the gas; it is spatially distributed with the Plummer density distribution (Kroupa, 2008) and starts depleting after a delay of τ_d, and is totally depleted within a timescale of τ_g. The Plummer radius of the gas distribution is kept time-invariant (Kroupa et al., 2001). This assumption is an approximate model of the effective gas reduction within the cluster in the situation that gas is blown out while new gas is also accreting into the cluster along filaments such that the gas mass ends up being reduced with time but the radius of the gas distribution remains roughly constant. As discussed in Urquhart et al. (2022), the mass and radius distributions of the ATLASGAL clumps at different evolutionary stages are quite comparable. We used an average velocity of the outflowing gas of $v_g \approx 10$ km s^{-1}, which is the typical sound speed in an HII region. This gives $\tau_g = r_h(0)/v_g$, where $r_h(0)$ is the initial half-mass radius of the cluster. As for the delay time, we take the representative value of $\tau_d \approx 0.6$ Myr (Kroupa et al., 2001), this being about the lifetime of the ultracompact HII region. As shown in Banerjee & Kroupa (2013), varying the delay time, τ_d, primarily results in a temporal shift in the cluster's rapid expansion phase, without significantly impacting its subsequent evolution for times greater than τ_d. Protoclusters typically form in hub-filament systems (Motte et al., 2018b; Vázquez-Semadeni et al., 2019; Kumar et al., 2020; J.-W. Zhou et al., 2022), which are located in hub regions. Compared to the surrounding filamentary gas structures, the hub region, as the center of gravitational collapse, is usually more regular, as shown in J.-W. Zhou et al. (2022) and J. W. Zhou et al. (2024f). Thus, modeling a spherically symmetric mass distribution is appropriate.

In this work, we assumed a SFE ≈ 0.33 as a benchmark (i.e., $M_g(0) = 2M_{\rm ecl}(0)$). This value has been widely used in the simulations cited above and has been proven effective in reproducing the observational properties of stellar clusters. Such a SFE is also consistent with the value obtained from hydrodynamical calculations that include self-regulation (Machida & Matsumoto, 2012; Bate et al., 2014) and as well with observations of embedded systems in the solar neighborhood (C. J. Lada & E. A. Lada, 2003; Megeath et al., 2016). In J.-w. Zhou et al. (2024b), by comparing the mass functions of the ATLASGAL clumps and the identified embedded clusters, we found that a SFE of ≈ 0.33 is typical for the ATLASGAL clumps. In J. W. Zhou et al. (2024d), assuming SFE = 0.33, it was shown that the total bolometric

luminosity of synthetic embedded clusters can fit the observed bolometric luminosity of ATLASGAL clumps with HII regions. In J. W. Zhou et al. (2024b), we directly calculated the SFE of ATLASGAL clumps with HII regions and found a median value of ≈0.3.

5.5.2 Slow and moderate gas expulsions

Embedded clusters form in clumps. More massive clumps can produce more massive clusters, leading to stronger feedback and a higher gas expulsion velocity (Dib et al., 2013). There should be correlations between the feedback strength, the clump (or cluster) mass, and the gas expulsion velocity (v_g). And low-mass clusters should have a slower gas expulsion process compared with high-mass clusters. As shown in Pang et al. (2021), low-mass clusters tend to agree with the simulations of slow gas expulsion. Except for the feedback strength, the SFE determines the total amount of the remaining gas, which also influences the timescale of gas expulsion. The gas expulsion process is driven by feedback, and the effectiveness of the feedback will depend on the geometric shape of the gas shell surrounding the embedded cluster (Wünsch et al., 2010; Rahner et al., 2017). Therefore, a complex or nonspherical gas distribution would also change the timescale of gas expulsion. In short, these parameter uncertainties can ultimately be incorporated into the timescale of gas expulsion. Thus, apart from the fast gas expulsion with the gas decay time, τ_g, described above, we also simulated slow and moderate gas expulsions. For the slow gas expulsion, the gas decay time was set to $10\tau_g$. The moderate gas expulsion is between the fast and slow gas expulsions. Considering the large parameter space between τ_g and $10\tau_g$, we simulated two kinds of moderate gas expulsions: $2\tau_g$ ("moderate2") and $5\tau_g$ ("moderate1").

5.5.3 Star formation efficiency

The total amount of the residual gas not only affects the gas expulsion timescale, but also significantly influences the strength of the gas potential. The strength of the external gas potential may have a considerable impact on the evolution of embedded star clusters. The total amount of residual gas is determined by the SFE of the clump. In the simulations described above, we assumed a SFE = 0.33 for all star clusters. In J. W. Zhou et al. (2024b), we calculated the SFE of ATLASGAL clumps with HII regions and found a strong anticorrelation between the SFE and the clump mass. We can directly derive the residual gas mass using equation.6 in J. W. Zhou et al. (2024b), which encodes the variation in SFE with clump mass.

5.5.4 Procedure

The `McLuster` program (Küpper et al., 2011) was used to set the initial configurations. The dynamical evolution of the model clusters was computed using the state-of-the-art `PeTar` code (L. Wang et al., 2020). `PeTar` employs well-tested analytical single and binary stellar evolution recipes (SSE/BSE) (Hurley et al., 2000; Hurley et al., 2002; Giacobbo et al., 2018; Tanikawa et al., 2020; Banerjee et al., 2020).

5.5.5 Simulation results analysis

In Fig.5.10, as expected, the mass loss of the clusters strongly depends on the gas expulsion modes. In the scenario of fast gas expulsion, gas is removed almost instantaneously, which strongly impacts the clusters. In contrast, with slow gas expulsion, the gas disperses over a timescale significantly longer than the cluster's crossing (dynamical) time, enabling the cluster to adapt to the changing gravitational potential and expand while maintaining near dynamical equilibrium, as discussed in Baumgardt &

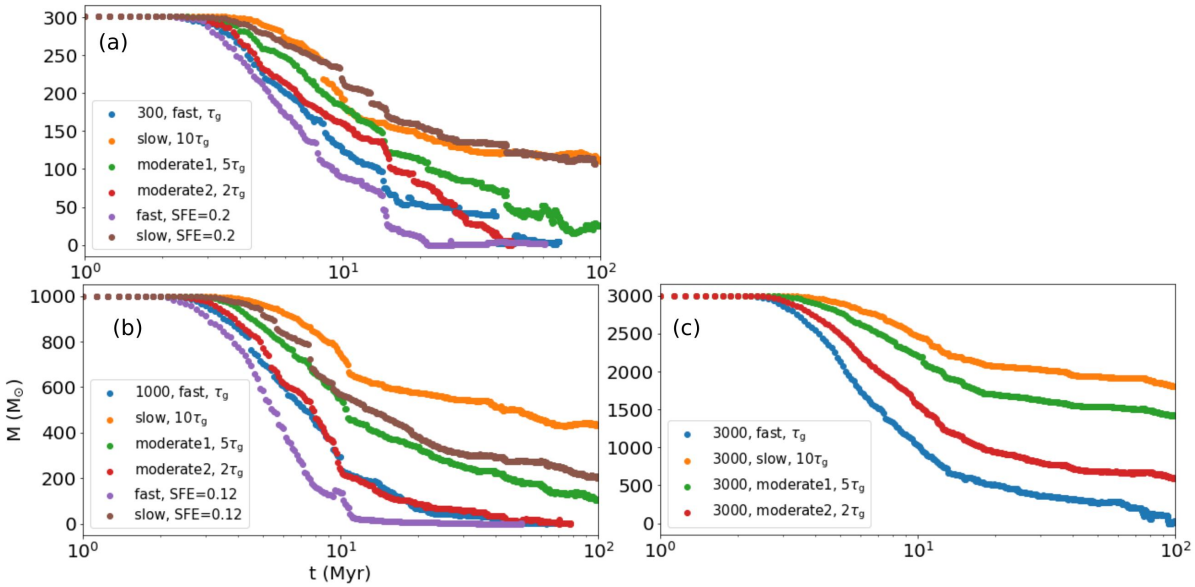


Figure 5.10: Evolution of the cluster mass within the tidal radius over time under different gas expulsion modes and with different SFEs.

Kroupa (2007). This quasi-static (or adiabatic) evolution results in substantially less stellar mass loss, as shown in Fig.5.10. Therefore, it best fits the mass distribution of the observed OCs. However, it fails to match the radius distribution, as shown in Fig.5.11(d1). For the fast gas expulsion, the clusters cannot survive more than 100 Myr. Similarly, in Kroupa (2001) and Brinkmann et al. (2017), the massive cluster containing many O stars most likely survives as an OC. Low-mass embedded clusters or groups dissolve quickly, in part due to the loss of their residual gas. A lower SFE means more residual gas, leading to more intense expansion and mass loss of the cluster after gas expulsion, as shown in Fig.5.10(b). A lower SFE is equivalent to a shorter gas expulsion timescale.

For high-mass clusters, such as 3000 $M_\odot$, τ_g and $2\tau_g$ will give significantly different results. Thus, the simulations are very sensitive to the gas decay time. We need to carefully constrain this parameter from observations in future work. Since $\tau_g = r_h(0)/v_g$, with $r_h(0)$ being constrained by equation.5.5, we can focus on the measurement of the gas expulsion velocity (v_g) in observations, especially the correlation between v_g and the clump (or cluster) mass.

In Fig.5.11, fast gas expulsion is required to explain the radius distribution of the observed OCs, but to account for the mass distribution, slow gas expulsion is necessary. This contradiction makes both fast and slow gas expulsions inappropriate. To reconcile this contradiction, the gas expulsion should be somewhere between fast and slow. In terms of both mass and radius, "moderate2" can best fit the observed OCs, and the main contribution is from the high-mass cluster (3000 $M_\odot$). Therefore, to explain the mass and radius distributions of the observed OCs, initial embedded clusters with masses higher than 3000 $M_\odot$ are necessary. However, in Fig.5.19, the upper limit of the embedded cluster sample is less than 1000 $M_\odot$, and only a few clumps have a mass higher than 3000 $M_\odot$. For an OC, its progenitor should have a significantly higher mass. Thus, the currently observed clumps cannot be the "direct" precursors of the currently observed OCs.

5.5.6 Cluster coalescence simulation

For the 17 MSFRs investigated in the MYStIX project, nine of them are subcluster complexes. As shown in Kuhn et al. (2014), some of these nine MSFRs exhibit similar morphologies. Finally, we selected

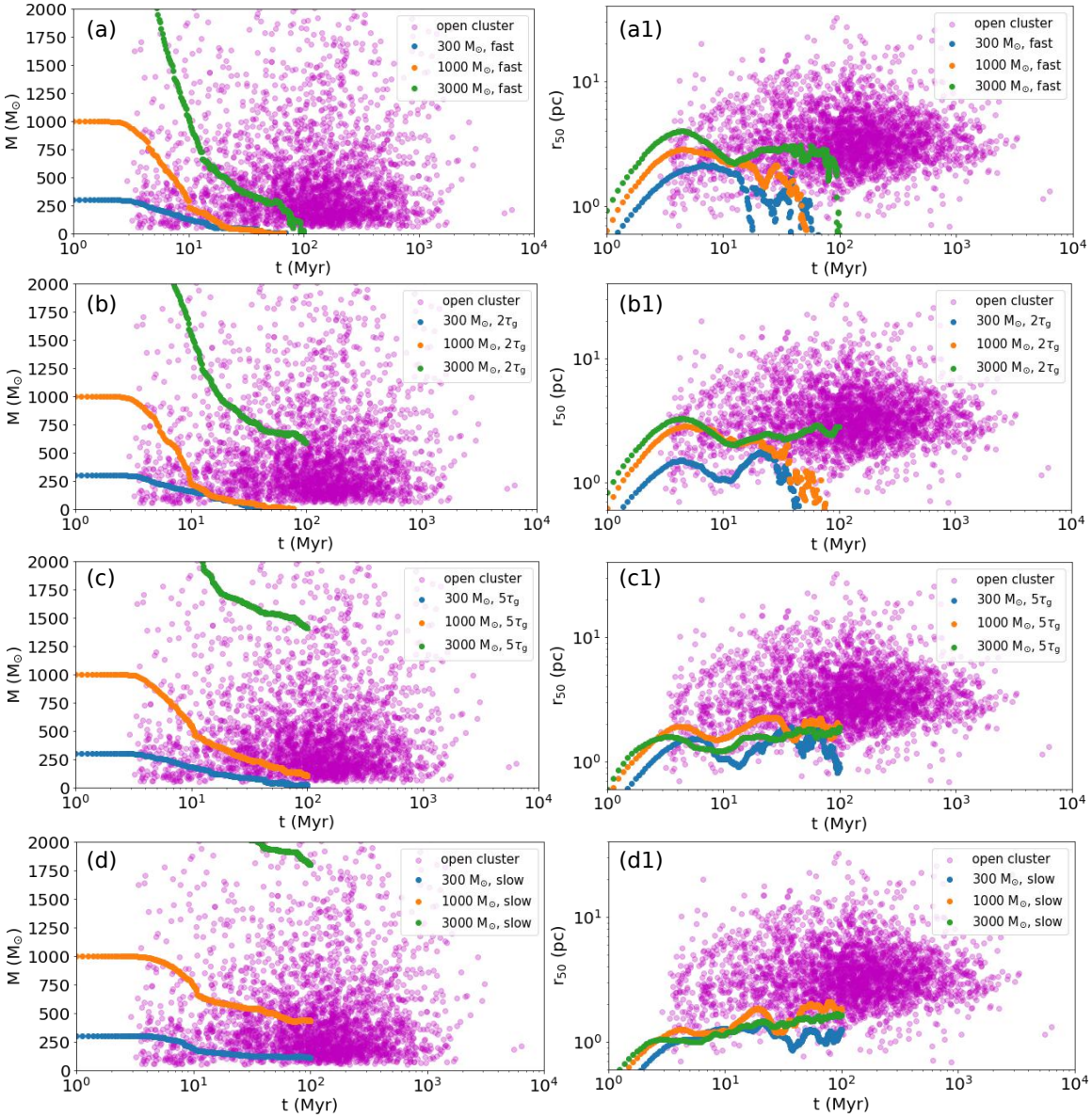


Figure 5.11: Fitting the physical parameters of OCs in Hunt & Reffert (2024) using individual embedded cluster simulations.

Table 5.1: "Terminated time" is the time when the simulation is terminated. The cluster may or may not have dissolved by this time. We consider two types of gas expulsion modes, i.e. "fast" (τ_g) and "moderate" ($5\tau_g$). τ_g is the gas depletion time in equation.5.7. "yes" or "no" indicate whether spatial separations or relative velocities between embedded clusters are accounted for in the simulations. Both "NGC1893-fast-vd-d0.5" and "NGC1893-fast-vd" incorporate spatial separations between embedded clusters, but the separation of the former is half of the latter.

Cases	Gas expulsion	Spatial separation	Relative velocity	Terminated time (Myr)
NGC1893-fast	fast	no	no	100
NGC1893-fast-vd	fast	yes	yes	100
NGC1893-fast-vd-v0	fast	yes	no	100
NGC1893-fast-vd-d0.5	fast	yes, * 0.5	yes	100
NGC1893-moderate-vd	moderate	yes	yes	100
NGC6334-fast	fast	no	no	614.25
NGC6334-fast-vd	fast	yes	yes	463.75
NGC6334-fast-vd-v0	fast	yes	no	100
NGC6334-moderate-vd	moderate	yes	yes	259.25
Carina-fast	fast	no	no	556.5
Carina-fast-vd	fast	yes	yes	324.5

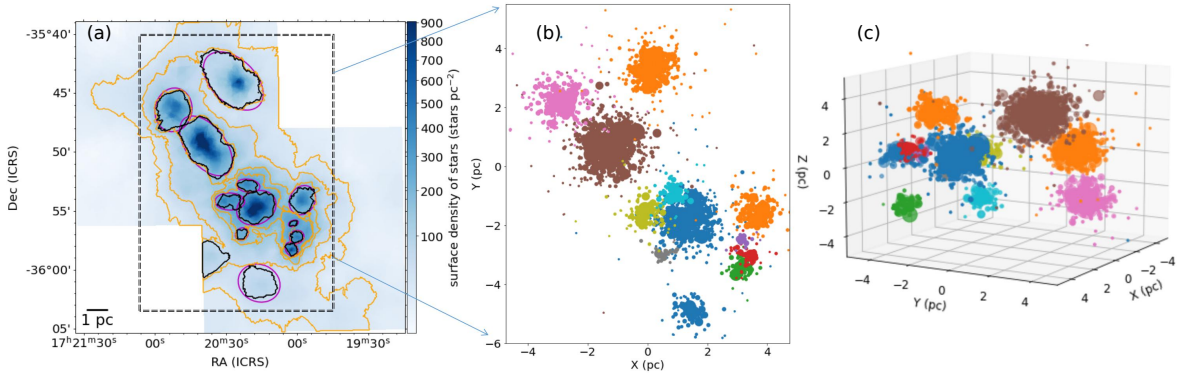


Figure 5.12: The initial conditions of the simulations drawn from the observations of NGC 6334. (a) Background is the surface density of stars. Black contours show the masks of embedded clusters identified by the dendrogram algorithm in J.-w. Zhou et al. (2024b). Magenta ellipses are the approximate ellipses that encompass the embedded clusters; (b) The collection of all simulated embedded clusters which are used to create an initial configuration of the embedded cluster complex similar to the observation; (c) Considering spatial separations between embedded clusters based on panel (b). In panels (b) and (c), the size of the dot scales with the stellar mass using an arbitrary scaling.

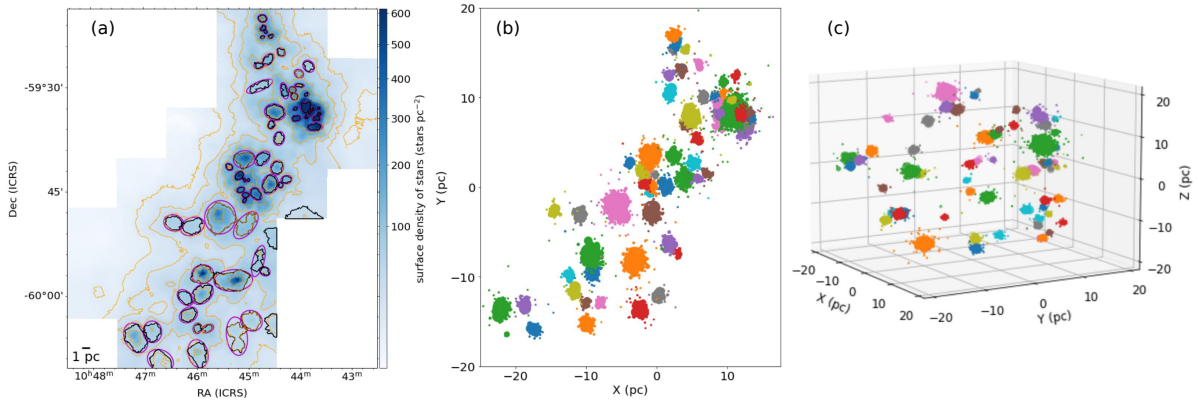


Figure 5.13: Same as Fig. 5.12, but for Carina.

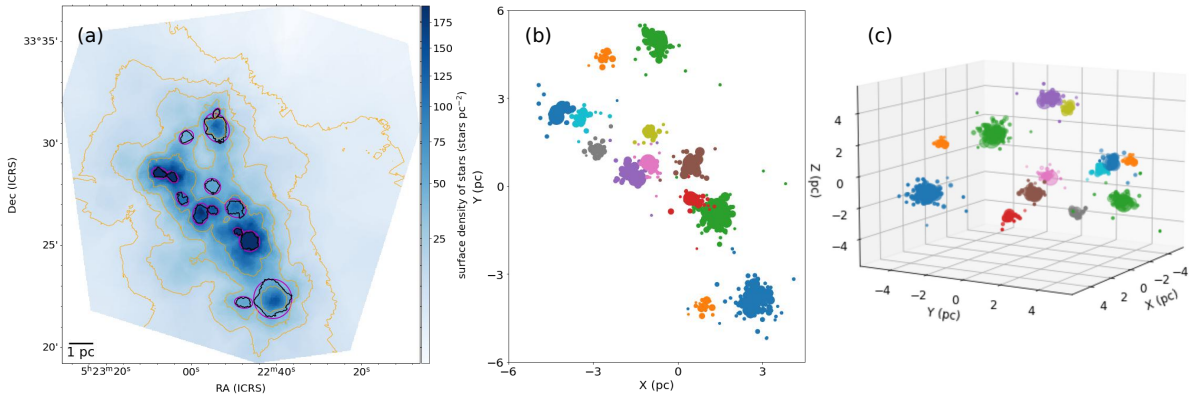


Figure 5.14: Same as Fig. 5.12, but for NGC 1893.

NGC 1893, NGC 6334 and the Carina nebula as representatives. In these three MSFRs, the spatial distribution of their subclusters varies, particularly with significant differences in the total mass of the subclusters, which helps to fit the observed physical parameters of open clusters. We have identified the subclusters in these three MSFRs using the dendrogram algorithm according to the surface density distributions of stars in J.-w. Zhou et al. (2024b), and calculated their physical parameters. Considering that the mass-radius relation of embedded clusters can be well fitted by the ≈1 Myr expanding line in J.-W. Zhou et al. (2024), we simulated the evolution of the subclusters in each MSFR for the first 1 Myr using the recipe described in Sec.5.5.1. The initial masses of the subclusters strictly follow the observed values, and each subcluster is simulated individually. Then, all subclusters were collected together to create an initial configuration similar to the observation, as shown in Fig.5.12. The separations between subclusters in Fig.5.12(b) strictly adhere to the observed values.

Currently, we assume all embedded clusters are initially at rest and located on the same plane. A more realistic scenario is that clusters have relative velocities and line-of-sight spatial separations. To take these two factors into account, we approximate the molecular cloud containing the clusters as an ellipsoid. To derive the lengths of the three axes of the ellipsoid (semi-major axis, a, intermediate axis, b, semi-minor axis, c), we fit the coordinates of the clusters located at the periphery to determine a and b. Given that the real cloud shape should be more sheet-like instead of spherical (Shetty & Ostriker, 2006; Inutsuka et al., 2015; Arzoumanian et al., 2018; Kohno et al., 2021; Arzoumanian et al., 2022; Rezaei Kh. & Kainulainen, 2022; J.-W. Zhou et al., 2023; J. W. Zhou et al., 2023; S. D. Clarke et al., 2023; Ganguly et al., 2023), we assume $c = b$. Then we randomly distribute clusters along the Z-axis

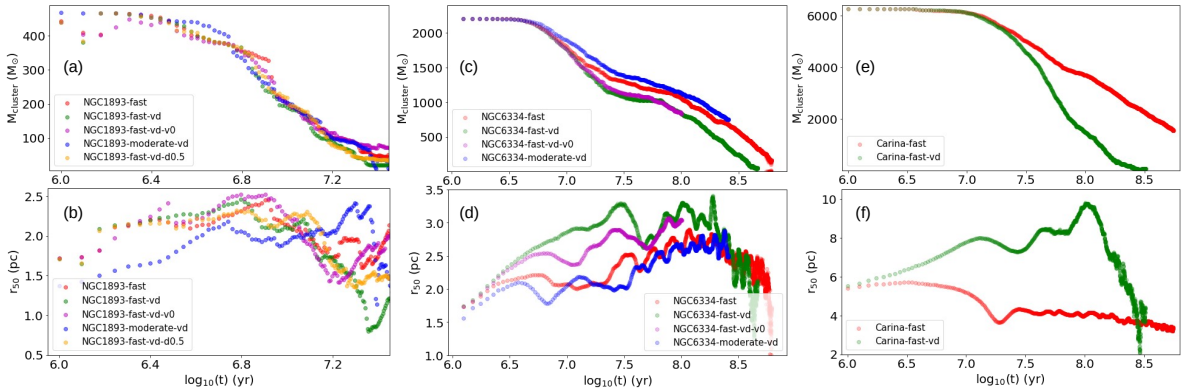


Figure 5.15: Mass and radius evolution of the embedded cluster complexes over time in different cases listed in Table.5.1. r_{50} is the radius containing 50% of the members within the tidal radius of the cluster.

(line-of-sight). As shown in Fig.5.12, Fig.5.13 and Fig.5.14, the spatial distribution of clusters is nearly spherical. Therefore, we have overestimated the spacing between clusters.

For molecular clouds in the catalog of Miville-Deschênes et al. (2017) across the Galactic plane, the most probable cloud size is ≈30 pc, and the most probable value of the velocity dispersion is ≈1.95 km s^{-1} on the scale of clouds. In our case, the sizes of the MSFRs are ≈5 pc for NGC 6334 and NGC 1893, and ≈20 pc for the Carina MSFR. As a conservative estimate, we take the velocity dispersion of the original molecular clouds in the three MSFRs as 2 km s^{-1} and assume that the clusters inherit the velocity dispersion of the clouds. The system's center has a velocity of 0 km s^{-1}, and the velocity of the outermost cluster is 2 km s^{-1}. The velocities of other clusters are distributed according to the Larson relation, i.e. $v \propto L^{0.5}$, where L is the distance to the system's center.

5.5.7 Simulation results analysis

In the simulations, before 1 Myr, the evolution of each cluster is approximately independent. After further expansion, they start to interact upon contact, leading to subsequent mutual influence in their evolution. The spatial distribution, relative velocities, mass distribution, and gas expulsion modes of embedded clusters will affect the dynamics of their coalescence as well as the stability of the coalesced products. We simulated all possible cases listed in Table.5.1 for NGC 1893. The embedded clusters in NGC 1893 have low mass and are few in number (13). They will quickly disperse after expansion regardless of the initial conditions, as shown in Fig. 5.15(a). There may not be a real coalescence between them, but rather the formation of a loose association.

For NGC 6334, we simulated four cases, i.e. "fast", "fast-vd", "fast-vd-v0" and "moderate-vd". Overall, they produce similar results, but there are also differences. The line-of-sight spatial separations and relative velocities between embedded clusters ("fast-vd") lead to a larger radius and faster mass loss of the coalesced cluster, but this effect can be counteracted by a slower gas expulsion of the embedded clusters ("moderate-vd"). Therefore, there is a degeneracy between the line-of-sight spatial separations and the relative velocities between embedded clusters and the gas expulsion modes of embedded clusters. Thus, "moderate-vd" and "fast" give comparable simulation results, as shown in Fig.5.15(c) and (d). Comparing "fast-vd" and "fast-vd-v0", we can see that the line-of-sight spatial separations are more important than the relative velocities in determining the mass loss and radius of the coalesced cluster.

NGC 6334 and NGC 1893 both have a small number of embedded clusters, but NGC 6334 contains more massive embedded clusters. As shown in J.-w. Zhou et al. (2025), massive embedded clusters provide a more stable gravitational core, significantly aiding the coalescence and enabling the resulting

cluster to remain stable within the Galactic tidal field. Strong gravitational forces between the embedded cluster complexes are essential to counteract the system's expansion and the Galactic tidal influence, ensuring the coalescence process continues. This strong local gravitational field is predominantly generated by the massive embedded clusters.

Most of the embedded clusters in the Carina complex are low-mass (<100 $M_\odot$), but they are numerous and densely distributed. Additionally, the Carina also contains a few massive embedded clusters (>300 $M_\odot$). Therefore, even though the embedded clusters in Carina were initially given large line-of-sight spatial separations (the maximum ≈40 pc), they are still able to coalesce after expansion, which is significantly different from the case of NGC 1893. However, this coalescence complex is not stable enough and gradually dissipates in the Galactic tidal field, as shown in J.-w. Zhou et al. (2025). The large line-of-sight spatial separations lead to insufficient coalescence between embedded clusters, preventing the formation of a stable core. In contrast, since there is no line-of-sight spatial separation in the case of "Carina-fast", the embedded clusters are able to sufficiently coalesce, forming a stable cluster core that can persist for over 300 Myr, as shown in J.-w. Zhou et al. (2025).

5.6 Expanding very young star clusters

Increasing observational evidence suggests that early expansion plays a fundamental role in the dynamical evolution of young star clusters. *Gaia* DR2 and DR3 data have opened a new window onto the internal kinematics of young star clusters and large star-forming complexes. Several case studies based on *Gaia* DR2 and DR3 data and multi-object spectroscopy have revealed the expansion of very young clusters in star-forming regions (Wright et al., 2019; Cantat-Gaudin et al., 2019; Kuhn et al., 2020; Lim et al., 2020; Swiggum et al., 2021; Lim et al., 2022; Mužić et al., 2022; Das et al., 2023; Flaccomio et al., 2023; Jadhav et al., 2024). Statistically, in Kuhn et al. (2019), for 28 clusters and associations with ages of about $1-5$ Myr, proper motions from *Gaia* DR2 reveal that at least 75% of these systems are expanding. Della Croce et al. (2023) presented a comprehensive kinematic analysis of almost all known young ($t < 300$ Myr) Galactic clusters based on the improved astrometric quality of the *Gaia* DR3 data, finding that a remarkable fraction (up to 80%) of clusters younger than ≈ 30 Myr are currently experiencing significant expansion. Wright et al. (2024) presented a large-scale 3D kinematic study of ≈ 2000 spectroscopically confirmed young stars ($t < 20$ Myr) in 18 star clusters and OB associations (groups) from the combination of *Gaia* astrometry and *Gaia*-ESO Survey spectroscopy, showing that nearly all of the groups studied are in the process of expansion. Based on the universality of the early cluster expansion, in this work, we collected various cluster samples to characterize the early cluster expansion using the numerical simulation recipe presented in Sec.5.5, with our main goal being to study the evolution of the initial mass–radius relation of embedded clusters. The work presented in this section is reproduced from the publication by J.-W. Zhou et al. (2024).

5.6.1 Evolution of cluster radii in N-body simulations

The computed model parameters are listed in Table.5.2. We studied the variations of the half-mass radius $r_{\rm h}$ (the radius within which 50% of the mass in stars is found) and the full radius $r_{\rm f}$ (the radius within which 90% of the mass in stars is found) of the cluster in the first 4 Myr. The $r_{\rm h} - M_{\rm ecl}$ relation (equation.5.5) predicts the initial half-mass radius of the clusters at t = 0 Myr. In Fig.5.16, we fitted the $r_{\rm h} - M_{\rm ecl}$ relations at different time nodes, namely t = 1 Myr, 2 Myr, 3 Myr, and 4 Myr. In the subsequent analysis, these are referred to as the cluster's 1 Myr, 2 Myr, 3 Myr, and 4 Myr expanding lines. In the first 4 Myr, the cluster mass in stars does not change significantly, and we therefore assume stellar mass conservation, although the stars are allowed to evolve. To determine $r_{\rm h}(t)$ at a given mass of

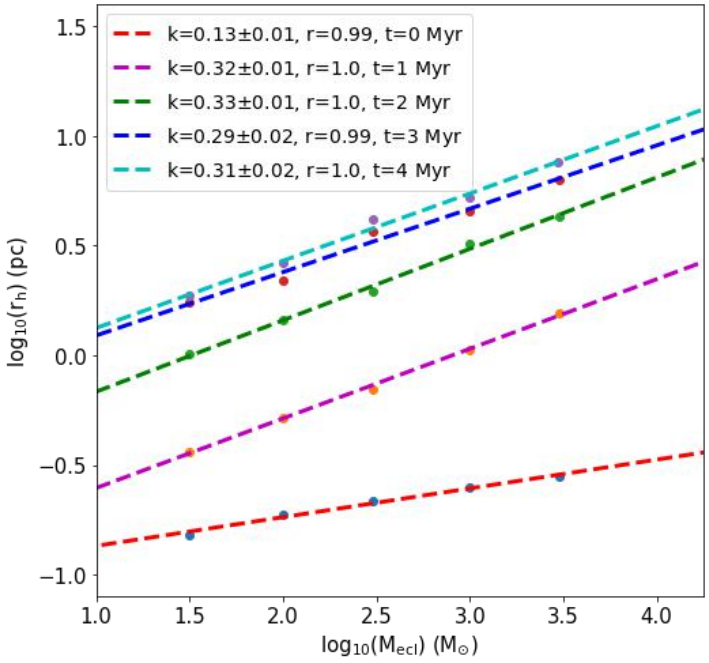


Figure 5.16: Fits to the expanding cluster in simulations by least-squares. The dots are from different time nodes; i.e., t = 0 Myr, 1 Myr, 2 Myr, 3 Myr, and 4 Myr. Here, "k" is the slope of the linear fit with the unit pc/$M_\odot$, and "r" represents the Pearson correlation coefficient.

Table 5.2: Computed model parameters.

Cluster	$M_{\rm ecl}(M_\odot)$	$r_{\rm h}(pc)$	$m_{\rm max}(M_\odot)$	$\tau_{\rm g}(Myr)$
1	30	0.16	5	0.012
2	100	0.18	10	0.014
3	300	0.21	17	0.016
4	1000	0.25	30	0.019
5	3000	0.28	43	0.022

the embedded cluster in stars, $M_{\rm ecl}$, the distance from the star to the cluster center is recorded, within which half of the cluster's mass is contained. At time t, the data $r_{\rm h}(M_{\rm ecl};t)$ are fitted by a linear relation using least-squares, as shown in Fig.5.16.

5.6.2 Mass–radius relation

Assuming all embedded clusters have the Plummer profile with the Plummer radius $r_{\rm pl}$, the full radius (0.9 mass fraction) and the half-mass radius (0.5 mass fraction) satisfy $r_{\rm f} \approx 3.707 r_{\rm pl}$ and $r_{\rm h} \approx 1.305 r_{\rm pl}$, and therefore $r_{\rm f} \approx 2.84 r_{\rm h}$ (Kroupa, 2008). Thus, for the cluster samples in this work, we can unify their radii to the half-mass radii. As shown in Fig.5.17, different samples can indeed merge together after the unification. We fitted the mass–radius relation for all samples, which gives a slope of ≈0.33 (dashed cyan line in Fig.5.17). If we fit the samples in E. A. Lada et al. (1991), Carpenter et al. (1993) and C. J. Lada & E. A. Lada (2003) and Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009) separately, the slopes are ≈0.43 (dashed black line in Fig.5.17) and ≈0.32 (dashed blue line in Fig.5.17), respectively. Only using the samples in Kuhn et al. (2015) gives a slope of ≈0.36.

The equation.4 of Pfalzner et al. (2016) has a steeper slope of ≈0.58 (dashed yellow line in Fig.5.17).

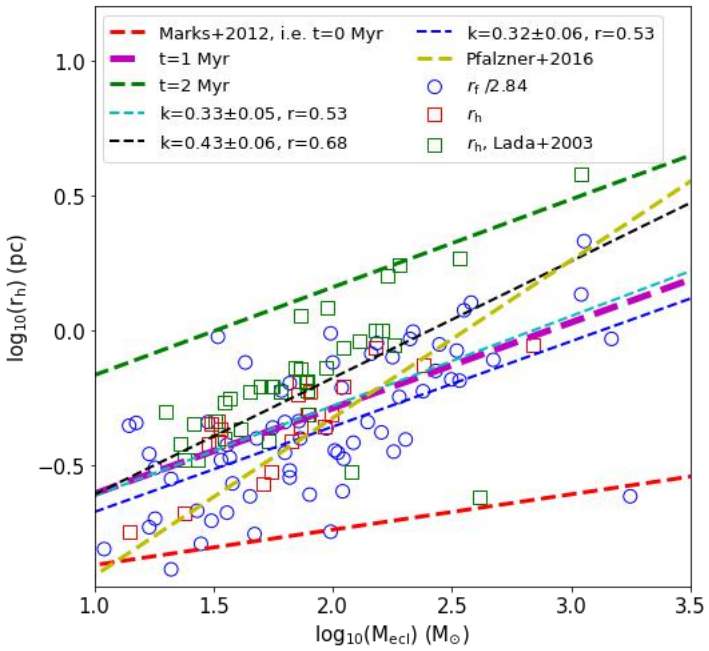


Figure 5.17: Fitting the mass-radius relation of embedded clusters in observations using N-body simulations. The radii of all samples were unified to the half-mass radius. The yellow dashed line is the mass–radius relation of Pfalzner et al. (2016). Dashed cyan, black, and blue lines represent the mass–radius relations of all samples, the samples in E. A. Lada et al. (1991), Carpenter et al. (1993) and C. J. Lada & E. A. Lada (2003) and the samples in Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009), respectively. Here, "k" is the slope of the fitting with the unit $pc/M_\odot$, and "r" represents the Pearson correlation coefficient.

Pfalzner et al. (2016) also found that all collected data in the K-band show a mass–radius relation with a similar slope. These authors therefore argued that clusters might "grow" from low-mass clusters to their final mass without changing their size, and their size would be determined by the extent of the clump. However, as shown in this work, the unique mass–radius relation of Pfalzner et al. (2016) actually reflects the expansion of embedded clusters. There is no self-similar mass–radius development during the evolution of embedded clusters, because the expanding lines in Fig.5.16 are not parallel at the beginning. The real initial mass–radius relation of embedded clusters (equation.5.5) is significantly different from the relation of Pfalzner et al. (2016). Any observable cluster is already in a state of expansion to varying degrees. Thus, the current mass–radius relation of clusters cannot reveal their initial formation conditions. In Pfalzner et al. (2016), there is no cluster containing a thousand stars and with a half mass radius of < 0.4 pc in their samples, because all the samples are expanding and thus detach from the initial compact configuration predicted by the initial mass–radius relation of embedded clusters (equation.5.5). As confirmed in the following section, even the embedded clusters in ATLASGAL clumps with HII regions are already in expansion.

In Pfalzner et al. (2016), all collected samples show a mass–radius relation with a similar slope, also shown in Fig.5.17, although different methods were used to determine the radius and mass in different observations. As explained in Pfalzner et al. (2016), assuming a unique mass–radius relation and a self-similar mass–radius development throughout, the different mass and radius measurements simply move the data points along the line of the mass–radius relation. In the discussion above, we have excluded this interpretation. Another simple explanation is that all clusters initially have the Plummer profile and then expand through the expulsion of residual gas, as modeled in Sec.5.5. Different radii of star clusters in different observations can be unified to the half-mass radii. This operation will not change the slope

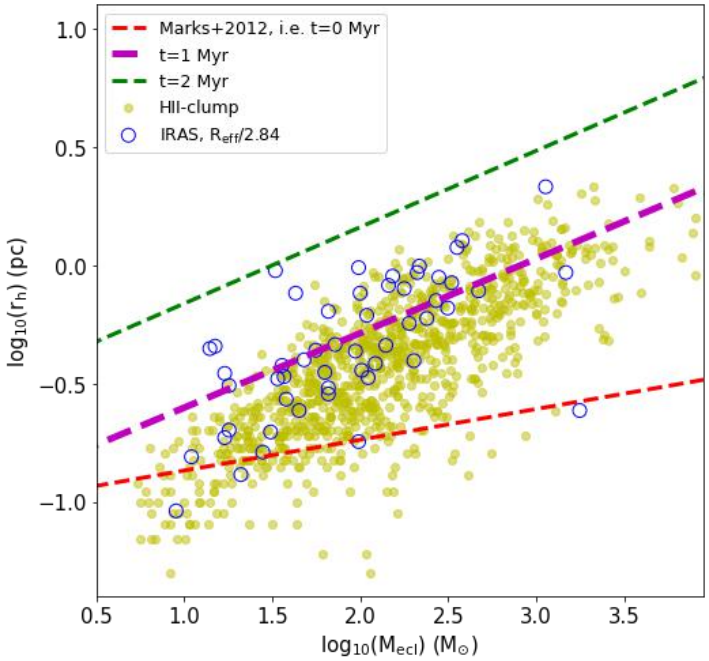


Figure 5.18: Fitting the masses and sizes of embedded clusters in HII clumps (yellow dots) using the expanding lines in Fig.5.16. Blue circles represent the clusters in Kumar et al. (2006).

of the mass–radius relation, which only shifts the line of the mass–radius relation, as shown in Pfalzner et al. (2016) and this work, also see the discussion in J.-w. Zhou et al. (2024b).

5.6.3 Expansion of embedded clusters in HII-clumps

The sample in Kumar et al. (2006) consists of the youngest stellar clusters associated with massive protostellar candidates (IRAS sources). As such, they are comparable to embedded clusters in HII-clumps. The radius of HII-clumps, $R_{\rm cl}$, in Urquhart et al. (2022) is determined from the number of pixels within the FWHM contour. As discussed in J.-W. Zhou et al. (2024), $R_{\rm cl}$ is close to the half-mass radii of embedded clusters in HII-clumps. A HII-clump mass $M_{\rm cl}$ and its embedded star cluster mass $M_{\rm ecl}$ satisfy $M_{\rm ecl} = {\rm SFE} * M_{\rm cl}$, where SFE is the final star formation efficiency of the clump. Here, we take SFE = 0.33 (C. J. Lada & E. A. Lada, 2003; Megeath et al., 2016; J. W. Zhou et al., 2024d; J. W. Zhou et al., 2024b). In Fig.5.18, the masses and radii of the embedded clusters in HII-clumps can be well fitted on the $r_{\rm h} - M_{\rm ecl}$ plane. Generally, the samples are between the 0 and 1 Myr expanding lines, which is consistent with their ages ($\leq$ 1 Myr). However, some samples are located under the initial 0 Myr line. This may be due to the fact that these clumps reach a higher gas densities during gas collapse than those implied by the $r_{\rm h} - M_{\rm ecl}$ relation for embedded clusters, or it could be that the measurement of their radii is biased or affected by resolution. As shown in Urquhart et al. (2022), evolutionary sequences still exist within HII-clumps. For the samples under the initial 0 Myr line, they may be so young that the inner cluster formation is not yet complete.

5.7 Origin of open clusters

The work presented in this section is reproduced from the publications by J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025).

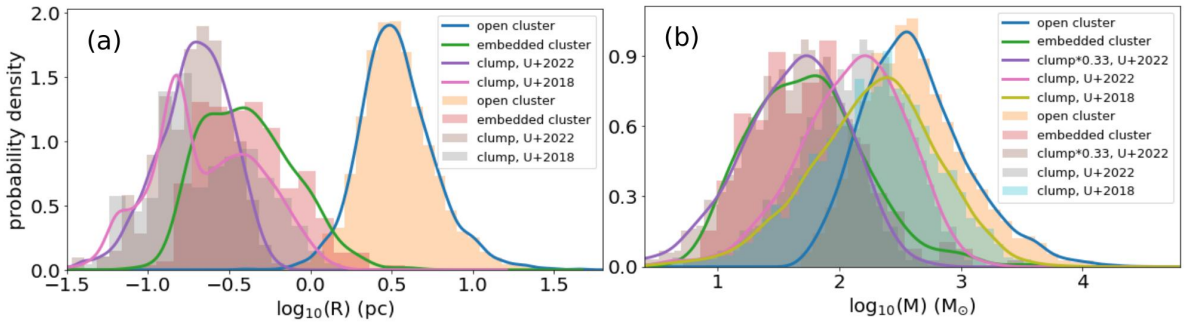


Figure 5.19: Physical properties of Galactic clumps, embedded clusters, and open clusters. (a) Radius distribution; (b) Mass distribution. The lines are from the histograms.

5.7.1 Mass and radius

The radii of ATLASGAL clumps with HII regions (HII clumps) were confirmed to be comparable with the half-mass radii of embedded clusters in the clumps (J.-W. Zhou et al., 2024). In Hunt & Reffert (2024), the radius of OCs was determined by including half of the members (r_{50}) that is close to the half-mass radius. For embedded clusters, their radii were unified to the half-mass radii, as done in J.-W. Zhou et al. (2024). Therefore, the radii of clumps, embedded clusters, and OCs in Sec.5.1 can be directly compared. Generally, the radii of OCs are significantly larger than those of embedded clusters and clumps. The radii of embedded clusters are larger than those of clumps, which may be due to the expansion of embedded clusters, as presented in J.-W. Zhou et al. (2024).

In Fig.5.19(b), the mass distribution decreases from OCs to clumps to embedded clusters. Interestingly, the mass distribution of clumps, when multiplied by 0.33, is comparable to that of embedded clusters, implying that 0.33 could be a typical star formation efficiency for clumps in the Milky Way (J. W. Zhou et al., 2024d; J. W. Zhou et al., 2024b). The mass distribution of the clumps is significantly smaller than that of the OCs. The mass of ATLASGAL clumps in Urquhart et al. (2018) is the full mass rather than the FWHM mass in Urquhart et al. (2022), which is still significantly lower than the mass of OCs.

These offsets raise questions about the origin of OCs with masses significantly higher than those of their progenitors (clumps or embedded clusters). From clumps to embedded clusters, the SFE is ≈0.3, and there is considerable mass loss during the evolution. Therefore, for a single clump to evolve into a single embedded cluster and then evolve into a single OC, with the current clump mass distribution in the Milky Way, this evolutionary sequence cannot produce the observed OCs' mass distribution, as verified by the direct N-body simulations in Sec.5.5.

If an OC originates from an embedded cluster, the initial mass of the embedded cluster should be much higher than that of the OC due to the considerable mass loss during the evolution. However, the opposite is seen in Fig.5.19(b), where the mass distribution of OCs is significantly larger than that of embedded clusters (by about one order of magnitude), suggesting that most OCs do not originate from single embedded clusters.

5.7.2 Separations of open clusters

We used the minimum spanning tree (MST) analysis to measure the inter-cluster spacing based on the 3D coordinates of OCs. The MST connects all clusters in a network with the least amount of total distance, without creating any loops. The length of the line connecting two clusters represents the separation between them. The typical inter-cluster separation of 3103 high-quality OCs is ≈80 pc. Considering all 5858 OCs and moving groups (< 3.6 kpc), the lower limit of the separation is ≈60 pc. However, the most

probable size for molecular clouds in the entire Milky Way disk is ≈30 pc (Miville-Deschênes et al., 2017). This implies that most molecular clouds may only form one OC. Each molecular cloud typically includes many embedded clusters. If each individual embedded cluster could evolve into an OC, the spacing between OCs should not be so large; it should at least be smaller than the size of the parental molecular cloud.

In J.-w. Zhou et al. (2024b), for the MSFRs in the MYStIX project, the separations between clusters, between clumps, and between clusters and clumps are quite comparable. The typical separation is ≈1 pc. Therefore, as the embedded clusters in each MSFR expand, they are able to undergo post-gas expulsion coalescence (J.-w. Zhou et al., 2025). Significant differences in separation for embedded clusters and OCs indicate that there is considerable coalescence between embedded clusters as they evolve into OCs.

5.7.3 Mergers at different stages

During different stages of cluster formation and evolution, various types of mergers can occur.

Formation of clumps (Stage1)

This stage should be dominated by the hydrodynamics of the gas. Clumps are locally dense gas structures in molecular cloud, which serve as local star-forming sites within molecular cloud.

Formation of embedded clusters (Stage2)

Once clumps form, their evolution and internal star formation proceed together until the formation of embedded clusters. Clumps are dynamically decoupled from their parent molecular clouds (J. W. Zhou et al., 2023; Peretto et al., 2023). As discussed in Urquhart et al. (2022), the differences observed for the physical parameters (such as radius, mass, density) of clumps at different evolutionary stages are in fact the result of observational biases, rather than the evolution of the embedded objects. This implies that the clumps remain stable throughout their evolutionary process and do not merge. It also suggests that the material lost through outflow and accretion onto stars is roughly balanced by the inflow of new material onto the clumps during star formation. As discussed in J.-w. Zhou et al. (2024b), the evolution from clump to embedded cluster proceeds in isolation and locally, and does not affect the surrounding objects significantly.

Evolution of embedded clusters (Stage3)

As the effects of feedback from the embedded clusters become increasingly stronger, they are responsible for gradually dispersing the gas of the clumps (Dale & Bonnell, 2008; Dale et al., 2012; Dib et al., 2011; Dib et al., 2013; J.-W. Zhou et al., 2022; Lewis et al., 2023). Due to the gas expulsion, the embedded clusters will expand and subsequently interact and coalesce. Recent observational progress has highlighted the crucial role of early expansion in the evolution of young star clusters. In Appendix.A of J.-w. Zhou et al. (2024a), the ATLASGAL+Planck 870 μm data (Csengeri et al., 2016) was used to trace the gas distribution around embedded or very young star clusters. On the maps, clusters and clumps are well separated. In Stage3, feedback from the newly formed embedded clusters will disperse the surrounding gas, then these clusters will be located within the cavities of HII regions. Finally, the clusters or their coalesced remnants will be gas-free. In this work, we only focus on the coalescence of the formed embedded clusters following their expansion, which occurs in Stage3.

The dynamics of both stars and gas have an important influence on the structure and evolution of embedded star clusters. A comprehensive simulation should combine both components. However,

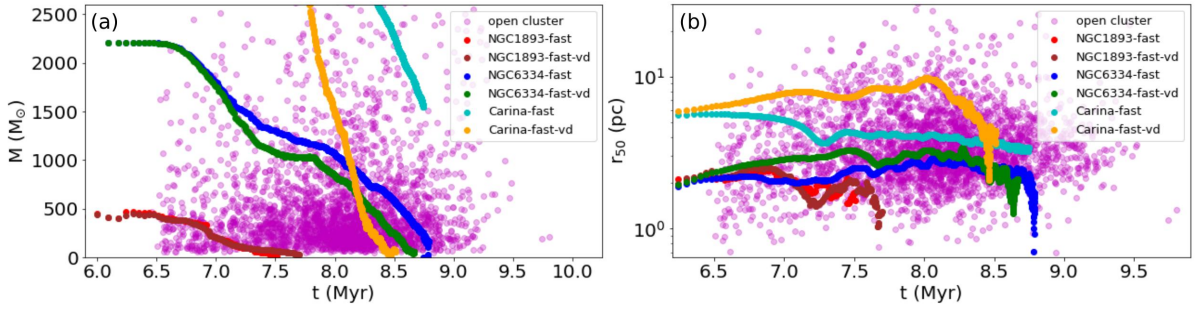


Figure 5.20: Fitting the physical parameters of open clusters in Hunt & Reffert (2024) using the coalescence simulations. r_{50} is the radius containing 50% of the members within the tidal radius of an open cluster. M is the mass of the open cluster.

considering the limitations of computational resources, we have to simplify the simulation by focusing only on the first aspect, namely stellar dynamics. The complexity of the simulation arises from the interaction between the cluster and the surrounding gas during the formation and evolution of the cluster. If the evolution of the cluster is relatively independent of the gas, as in Stage3, we can primarily focus on N-body dynamics.

5.7.4 Post-gas expulsion coalescence

The primary goal of this section is to evaluate whether the post-gas expulsion coalescence of embedded clusters can account for the observed physical parameters of open clusters. Therefore, we only consider two extreme cases, i.e. "fast" and "fast-vd" presented in J.-w. Zhou et al. (2025). The simulation results from other parameter choices should lie in between these two extremes.

As discussed in J.-w. Zhou et al. (2024a), the current mass distribution of clumps in the Milky Way suggests that the evolutionary process, where a single clump develops into an embedded cluster and eventually into an open cluster, cannot account for the observed open clusters with relatively old ages and large masses. However, each molecular cloud typically contains multiple clumps, which are the local sites of star formation within the molecular cloud and the precursors of embedded clusters. A composite of multiple low-mass embedded clusters could, through coalescence, cover the parameter space observed in open clusters, as shown in Fig.5.20, where the mass and radius of the clusters are displayed as a function of their age. In the post-gas expulsion coalescence scenario, massive embedded clusters are not necessary. This is consistent with the current mass distribution of clumps in the Milky Way. We note that, as presented in J.-w. Zhou et al. (2024a), the evolution of individual embedded clusters can also account for part of the observed parameter space of open clusters.

5.8 Evolution of the $m_{\rm max} - M_{\rm ecl}$ relation

In this study, we explore how the initial $m_{\rm max} - M_{\rm ecl}$ relation for embedded clusters evolves over time, with a particular focus on the similarities and differences between this relation and the $m_{\rm max} - M_{\rm cluster}$ relation for open clusters. Furthermore, we discuss the indications of the temporal changes in the initial $m_{\rm max} - M_{\rm ecl}$ relation to the dynamics of cluster evolution and the formation processes of open clusters.

5.8.1 N-body simulations

In J.-w. Zhou et al. (2024a), we computed the dynamical evolution of three clusters with the masses in stars of 300 $M_\odot$, 1000 $M_\odot$ and 3000 $M_\odot$ for the first 100 Myr under four different gas expulsion modes.

Table 5.3: N-body simulations modeling both the evolution of individual clusters and the coalescence of multiple clusters under different gas expulsion modes. More details can refer to J.-w. Zhou et al. (2024a) and J. W. Zhou et al. (2024c). "Time" is the time when the simulation is terminated. The cluster may or may not have dissolved by this time.

Cases	Type	Mass ($M_\odot$)	Time (Myr)	Dissolved
10000-fast	individual	10000	263	yes
10000-moderate1	individual	10000	100	no
10000-moderate2	individual	10000	270	no
10000-slow	individual	10000	359.25	no
3000-fast	individual	3000	100	yes
3000-moderate1	individual	3000	441	no
3000-moderate2	individual	3000	256.25	yes
3000-slow	individual	3000	950.25	yes
1000-fast	individual	1000	67.5	yes
1000-moderate1	individual	1000	145.5	yes
1000-moderate2	individual	1000	78	yes
1000-slow	individual	1000	533.5	yes
300-fast	individual	300	69	yes
300-moderate1	individual	300	100	yes
300-moderate2	individual	300	45.25	yes
300-slow	individual	300	346.5	yes
NGC1893-fast	coalescence	468	305.75	yes
NGC1893-fast-vd	coalescence	468	283.25	yes
NGC6334-fast	coalescence	2207	614.25	yes
NGC6334-fast-vd	coalescence	2207	463.75	yes
Carina-fast	coalescence	6271	556.5	no
Carina-fast-vd	coalescence	6271	324.5	yes

In this work, we continue the simulations from the previous work until the clusters dissolve, as shown in Table.5.3. We also simulate a 10000 $M_\odot$ cluster using the same recipe as in J.-w. Zhou et al. (2024a). In J.-w. Zhou et al. (2025), we simulated the post-gas expulsion coalescence of embedded clusters formed within the same parental molecular cloud that undergo a phase of expansion caused by the effects of gas expulsion. In this work, we only consider two extreme cases in the coalescence simulation, i.e. "fast" and "fast-vd".

5.8.2 Evolutionary patterns of the $m_{\rm max} - M_{\rm ecl}$ relation

In the simulations, an embedded cluster evolves from a point on the initial $m_{\rm max} - M_{\rm ecl}$ relation. During the evolution, both the mass of the most massive star in the cluster and the total mass of the cluster decrease over time. As illustrated in Fig.5.21, the possible evolutionary patterns are as follows: If the cluster's total mass decreases more rapidly, it will deviate above the initial $m_{\rm max} - M_{\rm ecl}$ relation (red arrow, pattern2); If they decrease at the same rate, the evolution will exactly follow the initial $m_{\rm max} - M_{\rm ecl}$ relation (orange arrow, pattern3); If the mass of the most massive star decreases more rapidly, it will deviate below the initial $m_{\rm max} - M_{\rm ecl}$ relation (green arrow, pattern4); The black (pattern1) and red (pattern5) dashed lines represent two extremes, where either the mass of the most massive star or the cluster's total mass does not change over time; The purple arrow (pattern6) represents an impossible evolutionary pattern for individual embedded clusters.

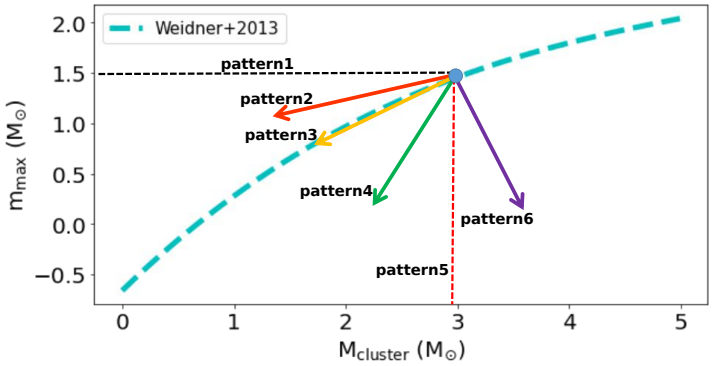


Figure 5.21: Evolutionary patterns of the $m_{\rm max} - M_{\rm ecl}$ relation over time.

5.8.3 The $m_{\rm max} - M_{\rm cluster}$ relation in observation

We picked out the most massive star in each observed open cluster. Fig.5.22 displays the $m_{\rm max} - M_{\rm cluster}$ relations of open clusters in different age ranges. Overall, as the cluster ages increase, the $m_{\rm max} - M_{\rm cluster}$ relation gradually deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation, and ultimately manifests as age stratification. The clusters are gradually loosing the most massive stars through evolution.

The dataset in Weidner et al. (2013a), used to fit the initial $m_{\rm max} - M_{\rm ecl}$ relation, requires clusters with age estimates of less than 5 Myr. However, some of these estimates come with significant uncertainties, suggesting that certain clusters could be as old as 9 Myr within the margin of error (Yan et al., 2023). In Fig.5.22(a), we focus on the open clusters younger than 5 Myr. The $m_{\rm max} - M_{\rm cluster}$ relation of these clusters roughly follows the initial $m_{\rm max} - M_{\rm ecl}$ relation.

As shown in Fig.5.22(h), in different age ranges, the older the open clusters are, the higher the upper mass limit of $M_{\rm cluster}$, similar to the impossible "pattern6" evolutionary pattern for individual embedded clusters in Fig.5.21. One might argue that this is a selection effect, where only massive clusters can survive for a long time. If that's the case, these old, massive open clusters must originate from the past, as there are no corresponding younger progenitors in Fig.5.22(h). Given the current mass distribution of clumps in the Milky Way, it is also impossible for such massive open clusters' precursors to form today (J.-w. Zhou et al., 2024a). However, as verified in J.-w. Zhou et al. (2025), the post-gas expulsion coalescence of a group of embedded clusters can form massive open clusters. In this scenario, massive embedded clusters are not necessarily required as precursors for the formation of massive open clusters.

5.8.4 Comparison of observation and simulation

In Fig.5.22, the $m_{\rm max} - M_{\rm cluster}$ relation of the observed high-quality OCs and MGs is mainly located between "pattern3" and "pattern6" (see Fig.5.21). For the simulations shown in Fig.5.23, the gas expulsion modes can determine the $m_{\rm max} - M_{\rm cluster}$ relation, because the mass loss of the clusters strongly depends on the gas expulsion modes. From fast to slow gas expulsions, the $m_{\rm max} - M_{\rm cluster}$ relation shifts from "pattern2" or "pattern1" to "pattern5". To some extent, the mass of a cluster and the gas expulsion mode are equivalent. Embedded clusters with different masses may achieve similar evolutionary trajectories on the $m_{\rm max} - M_{\rm cluster}$ diagram under different gas expulsion modes. For example, the evolutionary track on the $m_{\rm max} - M_{\rm cluster}$ diagram of a 1000 $M_\odot$ embedded cluster undergoing slow gas expulsion is similar to that of a 3000 $M_\odot$ embedded cluster with moderate gas expulsion. Comparing Fig.5.23 and Fig.5.25 with Fig.5.24 and Fig.5.26, although each subcluster in the coalescence simulation is in fast gas expulsion, the evolution of their coalesced product is similar to a cluster with slow gas expulsion in the individual cluster simulation.

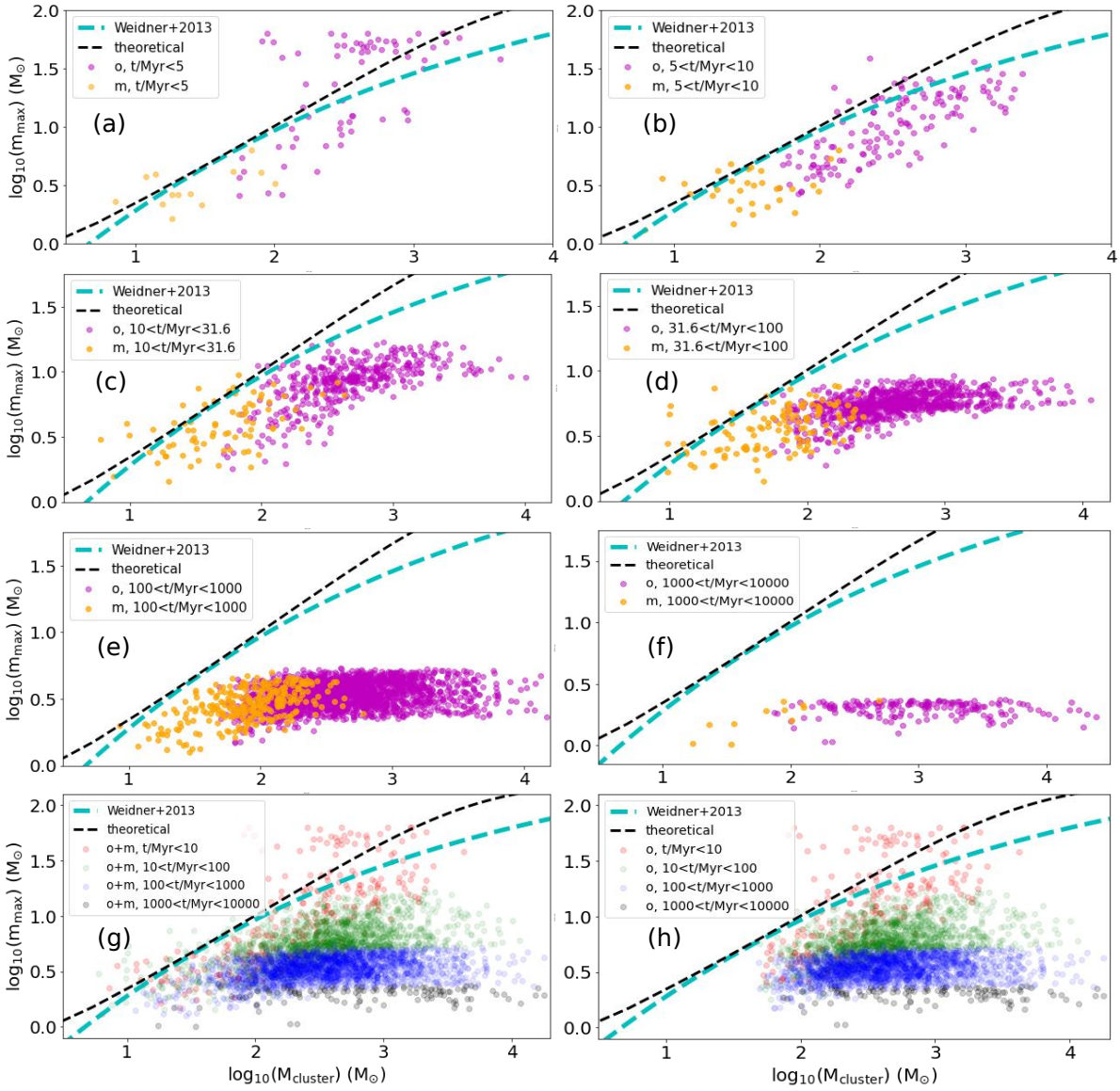


Figure 5.22: The $m_{\rm max} - M_{\rm cluster}$ relation of the observed open clusters evolves over time. The observational and theoretical $m_{\rm max} - M_{\rm ecl}$ relations correspond to the cyan and black dashed curves, respectively. "o" and "m" are the bound open clusters and unbound moving groups in the catalog of Hunt & Reffert (2024).

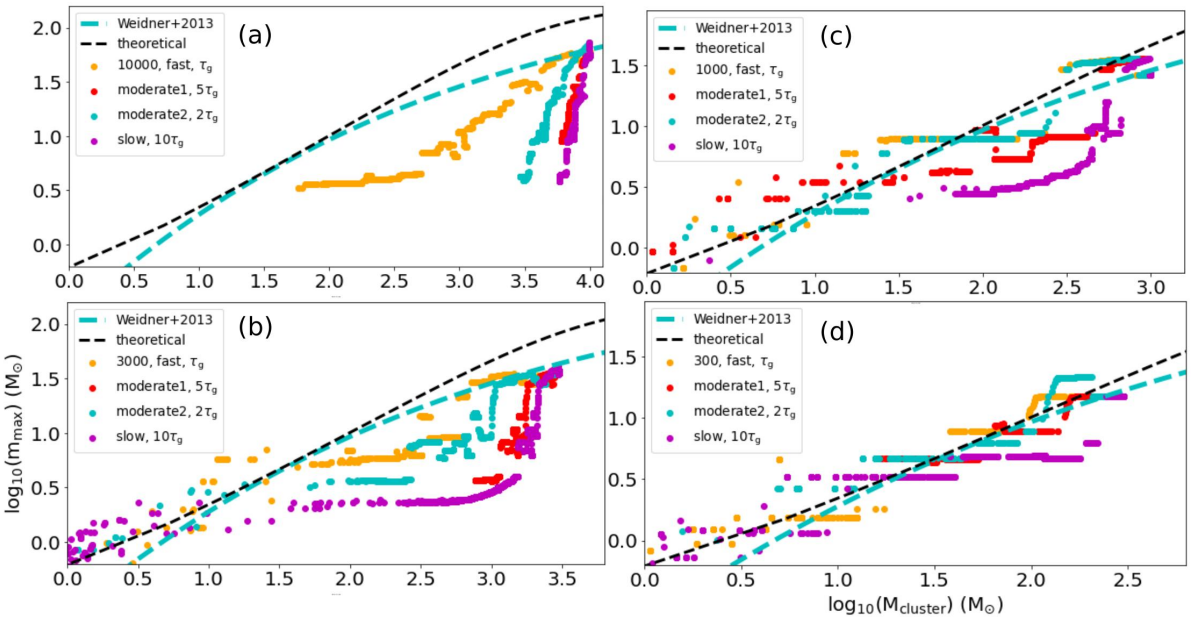


Figure 5.23: The initial $m_{max} - M_{ecl}$ relation evolves over time in individual cluster simulations. "fast", "moderate1", "moderate2" and "slow" are four different gas expulsion modes. More details can refer to J.-w. Zhou et al. (2024a).

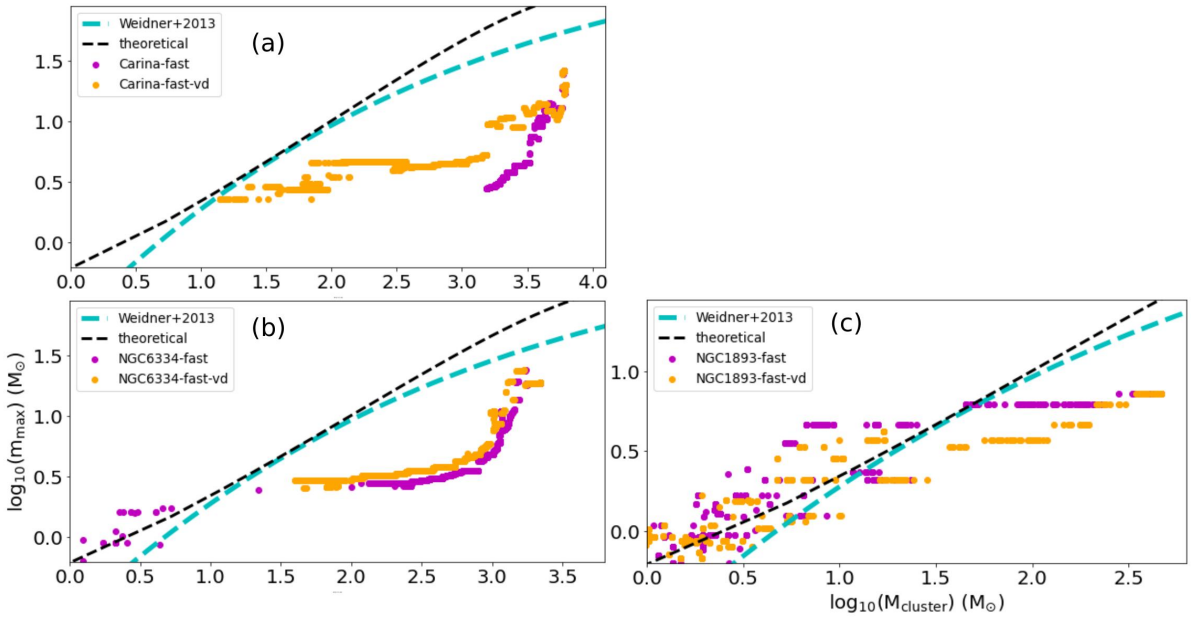


Figure 5.24: The initial $m_{max} - M_{ecl}$ relation evolves over time in cluster coalescence simulations. More details can refer to J.-w. Zhou et al. (2025).

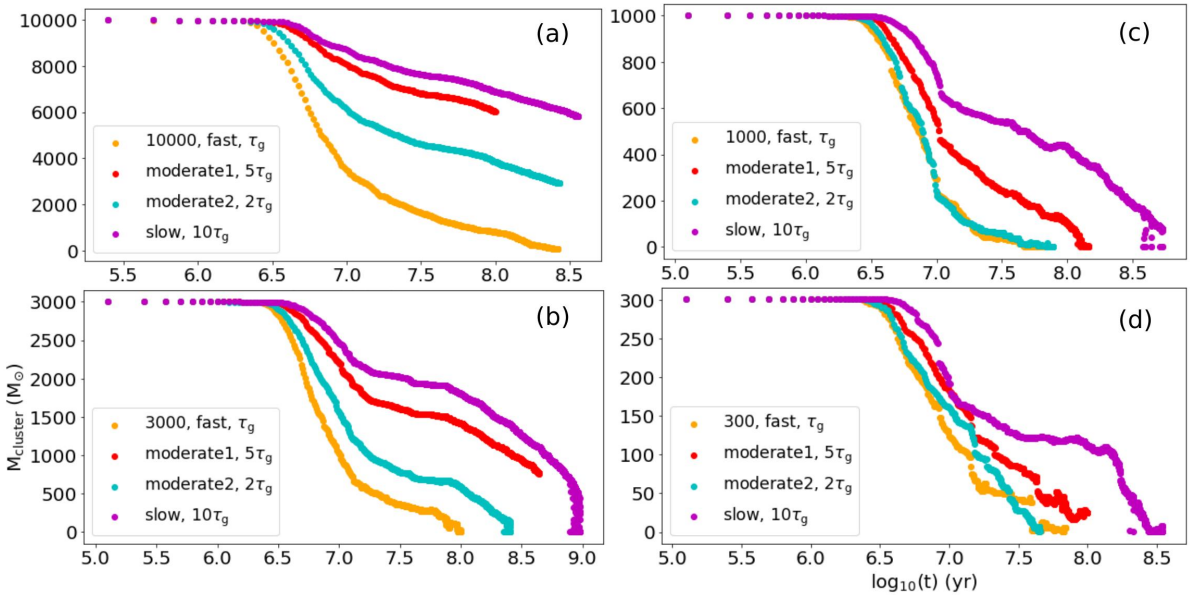


Figure 5.25: The mass of individual embedded cluster evolves over time in individual cluster simulations.

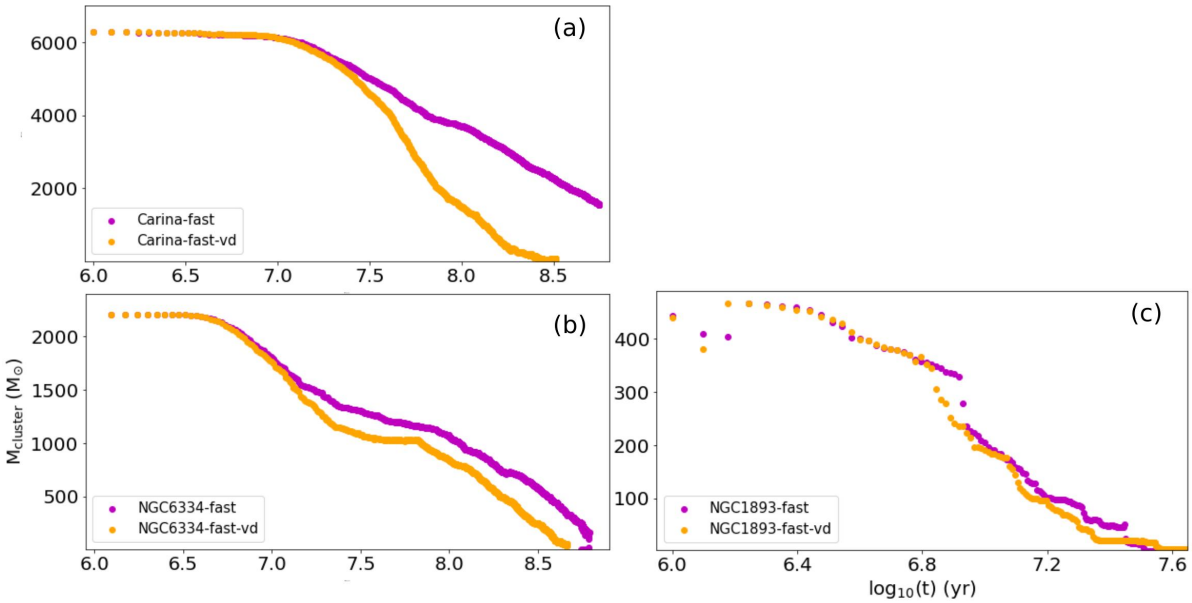


Figure 5.26: The mass of embedded cluster complex evolves over time in cluster coalescence simulations.

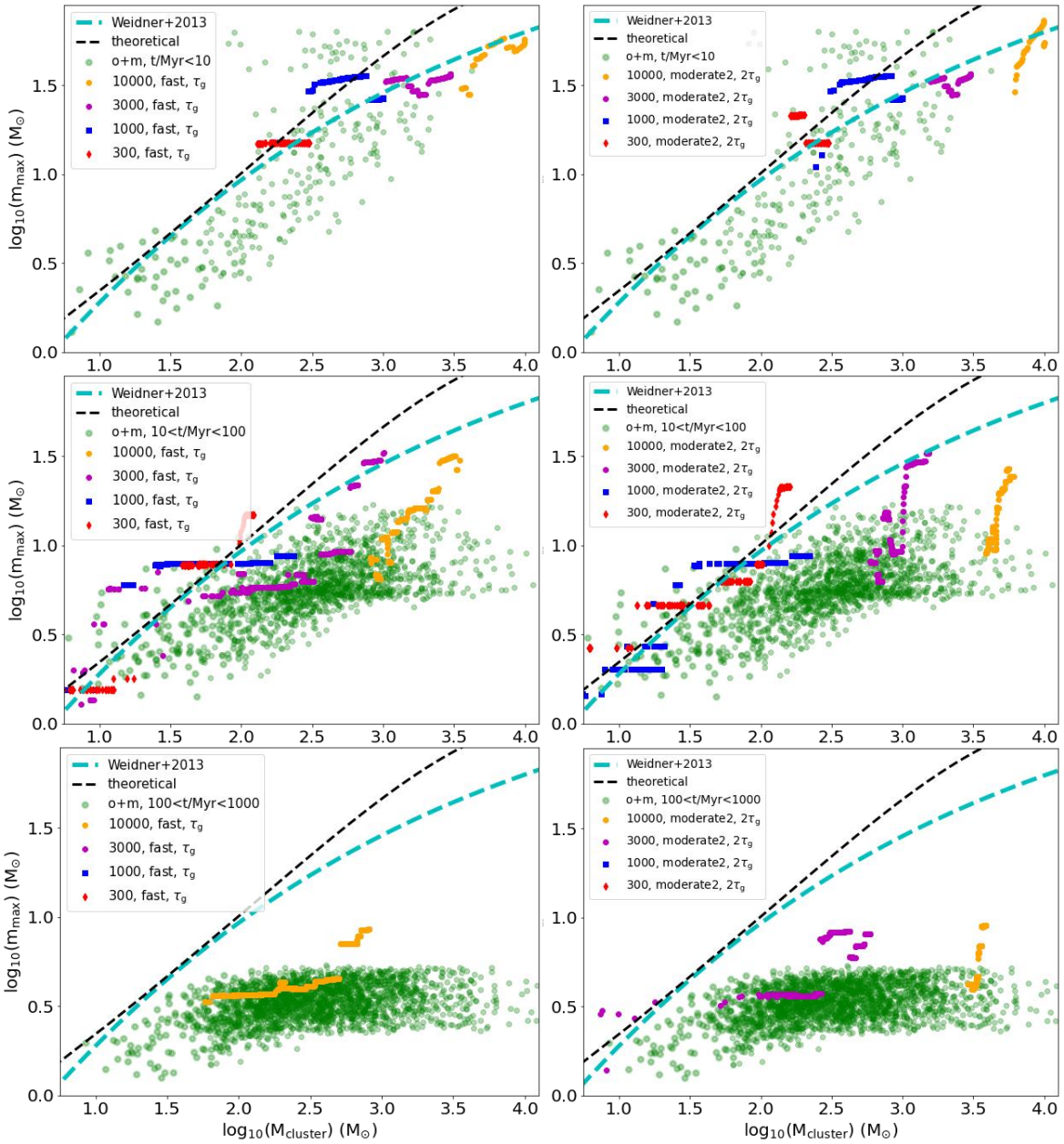


Figure 5.27: Fitting the observed $m_{\rm max} - M_{\rm cluster}$ relations of open clusters in different age ranges using the individual cluster simulations in the corresponding age ranges. The left and right panels represent "fast" and "moderate2" gas expulsion modes, respectively.

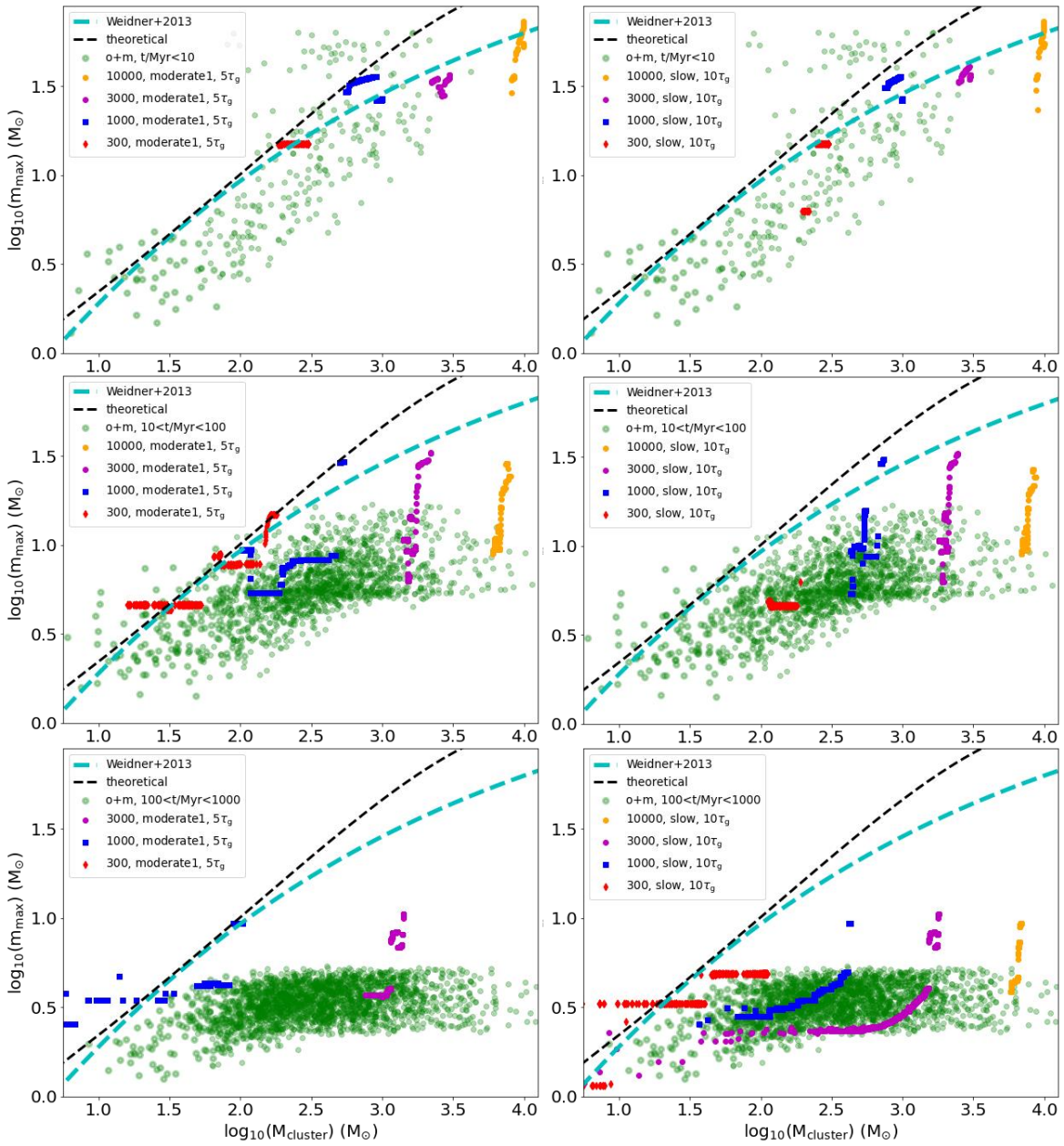


Figure 5.28: Same as Fig.5.27. The left and right panels represent "moderate1" and "slow" gas expulsion modes, respectively.

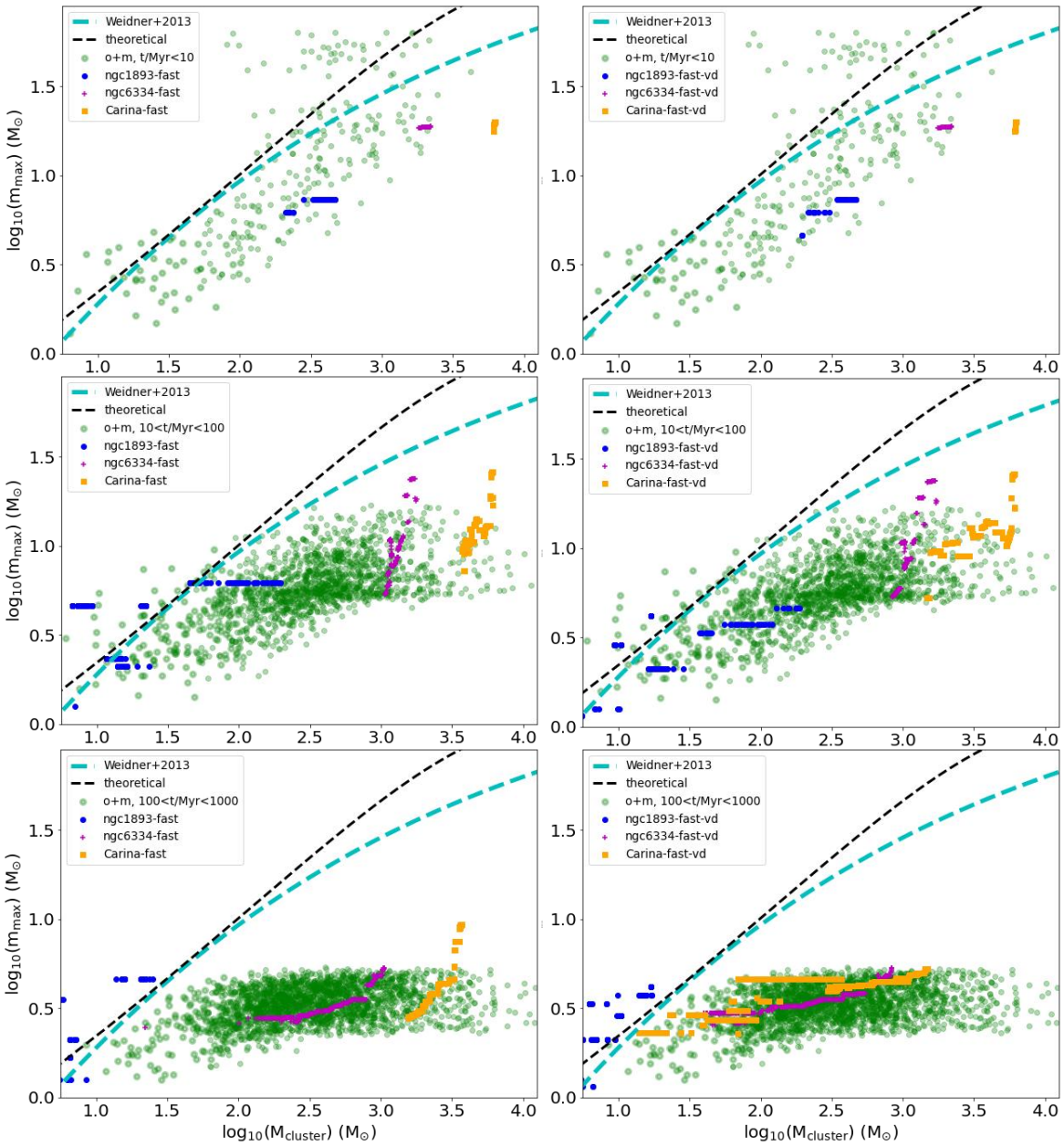


Figure 5.29: Fitting the observed $m_{\max} - M_{\text{cluster}}$ relations of open clusters in different age ranges using the cluster coalescence simulations in the corresponding age ranges.

For the fast gas expulsion, the clusters ($\leq$3000 $M_\odot$) cannot survive more than 100 Myr, as shown in Fig.5.25. And the $m_{\rm max} - M_{\rm ecl}$ relation evolves following "pattern3" or "pattern2" in Fig.5.23, which can only cover the upper boundary of the observed $m_{\rm max} - M_{\rm cluster}$ relation of open clusters in Fig.5.22. As shown in Fig.5.27 and Fig.5.28, a slow gas expulsion mode in individual cluster simulation is necessary to fit the observed $m_{\rm max} - M_{\rm cluster}$ relation of open clusters.

In Fig.5.27 and Fig.5.28, both observation and simulation show the mass of the most massive stars can be larger than the prediction of the initial $m_{\rm max} - M_{\rm ecl}$ relation. However, our simulations started from the initial $m_{\rm max} - M_{\rm ecl}$ relation of Weidner et al. (2013a). It means the merger of stars during the cluster evolution. Similar results were also presented in Oh & Kroupa (2018), where they found that more than half of the simulated clusters host a merger product as its most massive member within the first 5 Myr of cluster evolution.

5.8.5 Individual cluster simulation vs. cluster coalescence simulation

In Fig.5.22, open clusters with the age around 10 Myr have clearly detached from the initial $m_{\rm max} - M_{\rm ecl}$ relation in observation. This is consistent with the coalescence simulation shown in Fig.5.29. However, in the individual cluster simulation, the clusters with the age around 10 Myr still roughly follow the initial $m_{\rm max} - M_{\rm ecl}$ relation, as presented in Fig.5.27 and Fig.5.28, which is inconsistent with the observation. The youngest cluster in the catalog of Hunt & Reffert (2024) has the age of $\approx$2.9 Myr. In Fig.5.22, overall, when the age <10 Myr, $m_{\rm max}$ in open clusters are systematically lower than those predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation. This may not be due to the evolution of embedded clusters. In Fig.5.27 and Fig.5.28, even for clusters with ages around 10 Myr in individual cluster simulations, $m_{\rm max}$ still remains close to the initial $m_{\rm max} - M_{\rm ecl}$ relation, if they initially followed it as we set in the simulations. A possible explanation is that $m_{\rm max}$ in these clusters are smaller than those predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation from the beginning.

If open clusters are formed through the coalescence of embedded clusters, and each embedded cluster satisfies the initial $m_{\rm max} - M_{\rm ecl}$ relation, then $m_{\rm max}$ in the coalesced cluster will naturally fall below the value predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation, as shown in Fig.5.29.

In short, after comparing Fig.5.27, Fig.5.28 and Fig.5.29, we conclude that the coalescence simulation matches the observation better than the individual cluster simulation. The post-gas expulsion coalescence of embedded clusters may thus be a primary mechanism in open cluster formation.

5.9 Constraining the gas expulsion modes of embedded clusters

In N-body simulations, any gas expulsion mode (fast, moderate, or slow) can be assumed. However, what determines the specific mode a cluster follows? To address this, it is essential to integrate observations with simulations to identify which clusters correspond to each expansion mode and further investigate the underlying factors that drive a cluster into a particular expansion mode. This work aims to constrain the gas expulsion mode of initially embedded star clusters by inversely analyzing the expansion speed of open clusters (OCs). The clusters are required to be younger than 5 Myr, so that the initial dynamical state can be more easily traced (Weidner et al., 2013a; Yan et al., 2023; J.-W. Zhou et al., 2024).

5.9.1 Kolmogorov-Smirnov test

To compare with the observed star clusters, we transformed the simulated star clusters from the Galactic coordinate system in N-body simulation to the equatorial coordinate system using *astropy*. During this process, the star cluster's 3D velocities (v_x, v_y and v_z) were transformed into the proper motions and

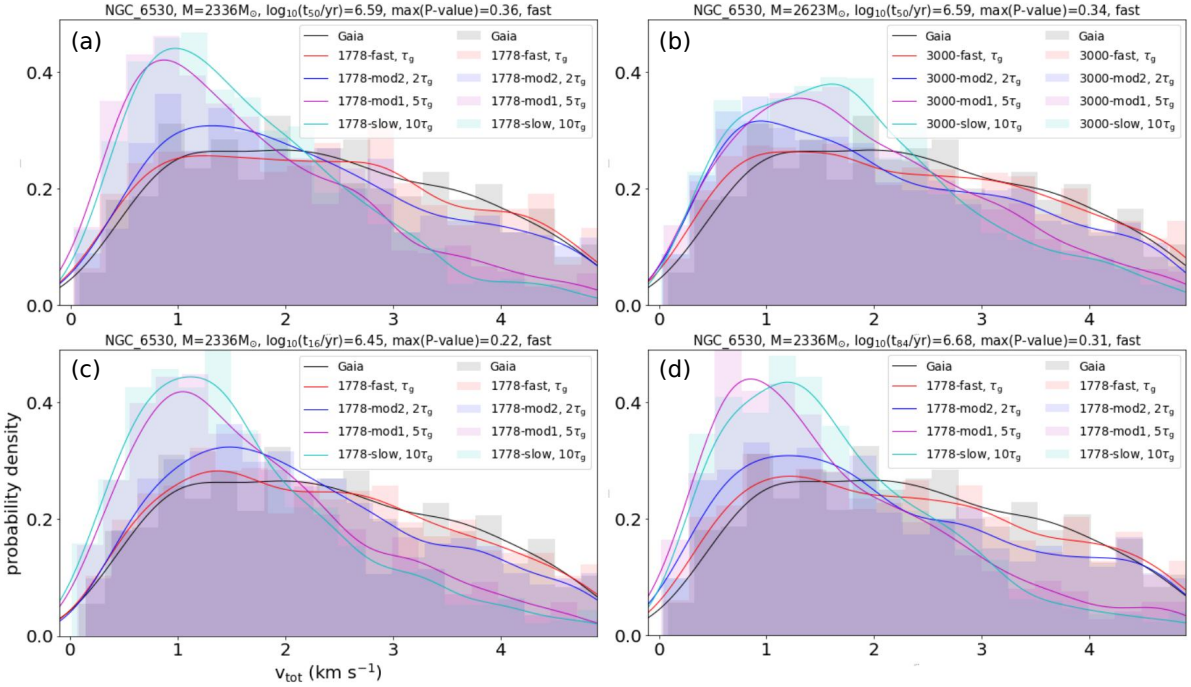


Figure 5.30: Five different velocity distributions of the cluster NGC 6530, namely the observed one and those under four different gas exclusion modes in the simulation. (a) The mass and age of the observed cluster are 2336 $M_{\odot}$ and $10^{6.59}$ yr, respectively. The average p-value of 1000 Kolmogorov-Smirnov (KS) tests using the Bootstrap method is 0.36, which strongly indicates that the cluster is in a fast gas expulsion mode; (b) Same as panel (a), but the cluster mass is 2623 $M_{\odot}$; (c) Same as panel (a), but the cluster age is $10^{6.45}$ yr; (d) Same as panel (a), but the cluster age is $10^{6.68}$ yr.

the radial velocity, which can directly compare with the observation. we also used the minimum and maximum stellar masses in the observed cluster to constrain the simulated cluster's members.

Each cluster has five different velocity distributions, namely the observed one and those under four different gas expulsion modes in the simulation. The velocity distribution under a gas expulsion mode most closely matches the observed velocity distribution suggests that the observed cluster is most likely to have that gas expulsion mode. Thus, we have a total of four sets of comparisons for each cluster. Fig.5.30 displays an example.

In observations, the velocities of stars in a cluster often have significant uncertainties. To account for this, we applied the Bootstrap method for resampling within each comparison and used the Kolmogorov-Smirnov (KS) test to evaluate whether they originate from the same distribution. The number of Bootstrap resampling iterations was set to 1000. Finally, the average p-value of all the KS tests was calculated. Generally, a p-value greater than 0.05 is considered as indicating that the two distributions are the same. Since there are only three possible gas expulsion modes, i.e. slow, moderate and fast. In this work, the one with the largest p-value among all four comparisons is considered the closest to the observed velocity distribution, regardless of whether the p-value is greater or less than 0.05. For all clusters, we found that the results are not significantly affected by errors in velocity, mass, or age of the clusters, as shown in Fig.5.30.

Then we compared the differences in physical properties among star clusters with different gas expulsion modes. More than half of the clusters are classified as "slow". To increase the sample size, we combined the "moderate" and "fast" categories and then compare them with the "slow" category. As presented in J.-w. Zhou et al. (2024a), star clusters undergoing rapid gas expulsion tend to have

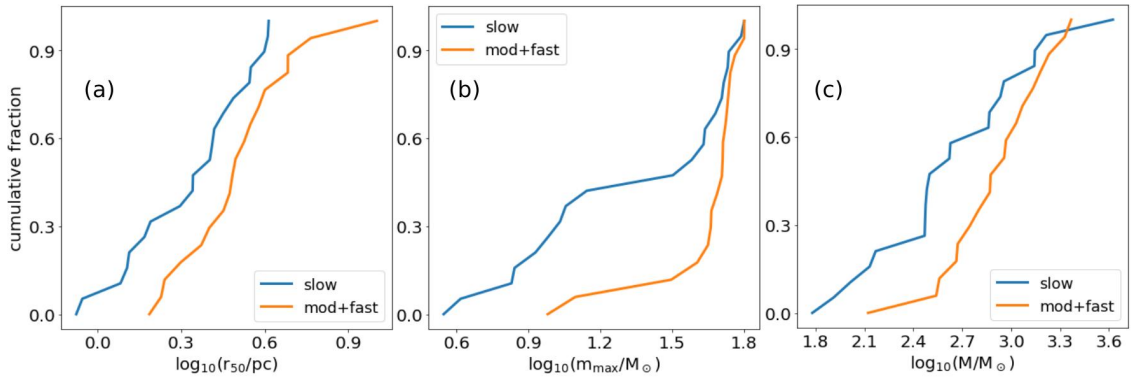


Figure 5.31: Physical properties among star clusters with different gas expulsion modes. (a) r_{50} is the radius containing 50% of the members within the tidal radius of a cluster; (b) The mass of the most massive star in a cluster; (c) The total mass of a cluster.

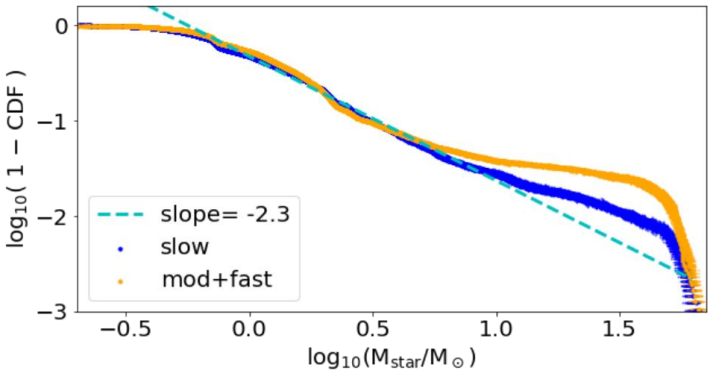


Figure 5.32: The complementary cumulative distribution function (CCDF) plot of the IMF for clusters with different gas expulsion modes. Here, we use Monte Carlo simulation to account for the uncertainty in stellar mass, randomly sampling within the mass error range and repeating the CCDF calculation 100 times.

larger radii than those of the same age experiencing slower gas expulsion. In Fig.5.31(a), star clusters classified as "slow" indeed exhibit smaller radii. This further validates the reliability of using the velocity distribution of stars within a cluster to constrain the initial gas expulsion mode of the cluster.

5.9.2 Possible causes of slow expansion

Mass distribution of stars

Star clusters are typically assumed to undergo an initial fast gas expulsion process in simulations, resulting in significant expansion (Kroupa et al., 2001; Bastian & Goodwin, 2006; Goodwin & Bastian, 2006; Banerjee & Kroupa, 2013; Banerjee & Kroupa, 2014; Banerjee & Kroupa, 2015). However, the above results indicate that many star clusters actually expand much more slowly, if they are indeed in expansion. The gas expulsion process is typically dominated by the feedback of massive stars within the cluster. In Fig.5.32 and Fig.5.31(b), star clusters classified as "slow" tend to contain relatively fewer massive stars, and their most massive stars ($m_{\rm max}$) also have lower masses. However, the IMF of these clusters is actually significantly top-heavy, regardless of which gas expulsion mode they are in. While the top-heavy could be an observational bias, since less-massive stars are harder to detect in very young clusters. But we can at least conclude that the slow expansion of clusters is not necessarily due to a lack of massive stars.

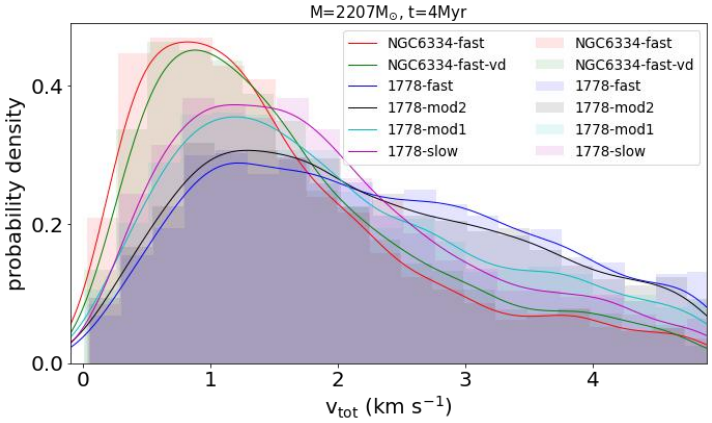


Figure 5.33: Stellar velocity distributions of a coalesced cluster (NGC 6334) and an individual cluster at the age of 4 Myr with masses of 2207 $M_\odot$ and 1778 $M_\odot$, respectively.

As shown in Sec.5.8, open clusters from the catalog of Hunt & Reffert (2024) with ages less than 5 Myr generally follow the initial $m_{\rm max} - M_{\rm ecl}$ relation of embedded clusters (Weidner et al., 2013a; Yan et al., 2023; J. W. Zhou et al., 2024b), where $M_{\rm ecl}$ represents the total stellar mass of embedded clusters. In clusters undergoing slow expansion, their most massive stars have lower masses. Consequently, in Fig.5.31(c), the total mass of slow expansion clusters is also smaller.

Gravitational potential of the surrounding material

Very young star clusters may still be associated with their birth environment (molecular clouds), and the gravitational potential of the surrounding material may influence their evolution and dynamical state. Lallement et al. (2022) provides a 3D dust extinction density map with a $6 \times 6 \times 0.8$ kpc^3 volume around the Sun. The pixel size of the 3D map is 10 pc. We placed all clusters onto this 3D map and compared the extinction density of the pixels they occupy, and found that there is no significant difference in extinction density for clusters with different expansion modes. This suggests that the clusters' slow expansion may not be caused by the gravitational potential of the surrounding material.

Star cluster formation

As discussed in J.-w. Zhou et al. (2024a), in the Milky Way, open clusters have significantly higher masses than embedded clusters, by about one order of magnitude. Given the current mass distribution of clumps in the Milky Way, the evolutionary sequence from a single clump evolving into an embedded cluster and subsequently into an open cluster cannot account for the observed parameter space of open clusters in the Milky Way. However, the coalescence of multiple low-mass embedded clusters in the same parental molecular cloud can explain the observed characteristics of open clusters (J.-w. Zhou et al., 2025). In Sec.5.8, the evolution of the initial $m_{\rm max} - M_{\rm ecl}$ relation over time also aligns better with the cluster coalescence simulation rather than the individual cluster simulation. As shown in J.-w. Zhou et al. (2025), embedded clusters that are in close proximity can coalesce within the first 5 Myr.

The coalescence of expanding star clusters may reduce the overall expansion speed due to dynamical friction, energy redistribution, and an increase in overall density. To investigate this possibility, we compared the cluster coalescence simulation with the individual cluster simulation presented in J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025). As shown in Fig.5.33, the coalesced cluster indeed exhibits significantly lower stellar velocity distribution, despite each subcluster undergoing fast gas expulsion

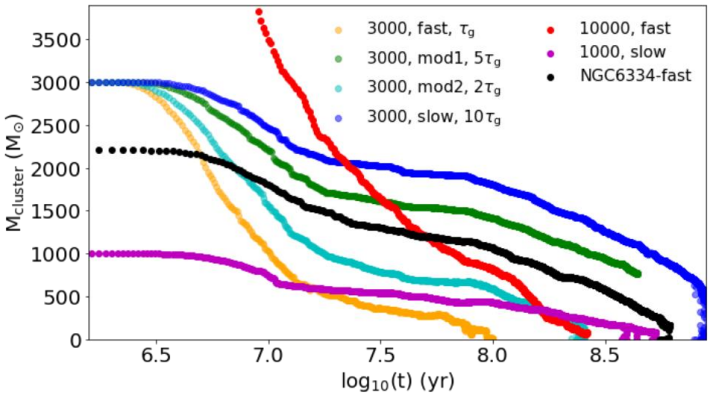


Figure 5.34: The evolution of cluster mass over time.

before coalescence. However, the exact change in expansion speed is influenced by various factors, including the mass ratio, relative velocity, collision angle, and initial expansion speed of subclusters. A detailed discussion on using the cluster coalescence model to explain the slow expansion of observed open clusters will be presented in future work.

5.10 Black hole formation in open clusters

Recent proper motion data from *Gaia*'s Data Release 3 (Gaia Collaboration et al., 2023) has led to the discovery of two dormant black holes (BHs, i.e. *Gaia* BH1 and *Gaia* BH2) in BH–star binary systems within the Galactic field (Chakrabarti et al., 2023; El-Badry et al., 2023b; El-Badry et al., 2023a; Tanikawa et al., 2023). As part of the validation process for the upcoming fourth *Gaia* data release, Gaia Collaboration et al. (2024) analyzed preliminary astrometric binary solutions and identified a black hole with a mass of ≈33 $M_\odot$ in a binary system, now referred to as *Gaia* BH3. Following the detection of *Gaia* BH1 and *Gaia* BH2, discussions and hypotheses emerged regarding the origins of these systems. The formation histories of *Gaia* BH1 and *Gaia* BH2 are challenging to reconcile with isolated binary (IB) evolution, as their orbital properties do not align with the expected outcomes of common envelope evolution (El-Badry et al., 2023a; El-Badry et al., 2023b). Their near-solar metallicities and inferred orbits within the Galactic plane suggest that they were likely formed in open clusters (Rastello et al., 2023; Di Carlo et al., 2024; Marın Pina et al., 2024; Tanikawa et al., 2024). Young and open star clusters host most massive stars, which are the precursors of compact objects. Consequently, the majority of black holes in the Milky Way likely spent their early years in these clusters, engaging in dynamic interactions. N-body simulations are powerful tools for studying star cluster dynamics and black hole formation within clusters. Physically, the initial conditions of N-body simulations are determined by the star cluster formation process. Given the significant progress in observational studies of massive star and star cluster formation, these observational results should be translated into the initial conditions for star cluster dynamical simulations. This work examines how initial conditions and the cluster formation process influence the evolutionary outcomes of star clusters.

5.10.1 Cluster evolution

The evolution of clusters under different simulation scenarios has been extensively analyzed in J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025) and Sec.5.8, and references therein. Here, we summarize some key conclusions. For individual cluster simulations, the mass loss of a cluster is strongly influenced by

the mode of gas expulsion: the faster the gas expulsion, the greater the mass loss. However, there exists a degeneracy between cluster mass and gas expulsion timescale. Clusters of varying initial masses can exhibit similar evolutionary outcomes. As shown in Fig.5.34, a 1000 $M_\odot$ cluster undergoing slow gas expulsion evolves in a manner similar to a 3000 $M_\odot$ cluster experiencing fast gas expulsion. Conversely, a 10000 $M_\odot$ cluster with fast gas expulsion exhibits similar evolutionary behavior to a 3000 $M_\odot$ cluster with slow gas expulsion.

In cluster coalescence simulations, factors such as the spatial distribution, relative velocities, mass distribution, and gas expulsion modes of subclusters influence both the dynamics of their coalescence process and the stability of the final coalesced product. As shown in Fig.5.34, despite each subcluster undergoing fast gas expulsion, the evolution of the coalesced cluster closely resembles that of a single cluster with moderate gas expulsion in individual cluster simulations. This suggests a degeneracy between the cluster mass, the initial gas expulsion mode, and the cluster formation mechanism (whether monolithic or through coalescence).

5.10.2 Black hole formation

The 300 $M_\odot$ cluster is unable to form a black hole at its center under any gas expulsion mode. A black hole can form at the center of the 1000 $M_\odot$ cluster only if the gas expulsion is slow. For the 3000 $M_\odot$ cluster, black holes can form at the center under moderate or slow gas expulsion modes. The 10000 $M_\odot$ cluster can form central black holes even with fast gas expulsion, whereas the 3000 $M_\odot$ cluster cannot under the same conditions. Moreover, despite each initial subcluster undergoing fast gas expulsion, the coalesced cluster with the mass of 2207 $M_\odot$ can still form a black hole at its center. These results reflect the aforementioned degeneracy again.

Variations in cluster mass, gas expulsion modes, and formation mechanisms ultimately influence the cluster's expansion speed, which in turn affects its mass loss rate and black hole formation rate. A star cluster's expansion speed reflects the overall distribution of stellar velocities and plays a crucial role in determining whether massive stars and their remnants remain bound to the cluster.

In Fig.5.35, we can see closely spaced black holes, their ability to form multiple systems (binary or triple) depends on the dynamical state of the star cluster. The initial gas expansion mode is crucial to the cluster's dynamical evolution and has a significant impact on its lifetime. Black holes in slowly expanding clusters experience a more stable dynamical environment and have substantially more time to evolve within the cluster. The number of black holes at the cluster centers of Fig.5.35(a) and (b) is 2 and 3, respectively. Fig.5.36 shows the evolution of the distance between each pair of black holes over time. Three black holes in Fig.5.35(b) eventually form a triple system, whereas two black holes in Fig.5.35(a) fail to form a binary system even until the cluster dissolves.

In Sec.5.9, we tried to constrain the gas expulsion mode of initially embedded star clusters by inversely analyzing the expansion speed of very young open clusters (< 5 Myr) using the *Gaia* DR3 data. After comparing with N-body simulations, we found that all gas expulsion modes (fast, moderate and slow) are possible. As discussed above, slowly expanding clusters are more likely to retain black holes and black hole-formed multiple systems, making them promising candidates for black hole searches with *Gaia* observations.

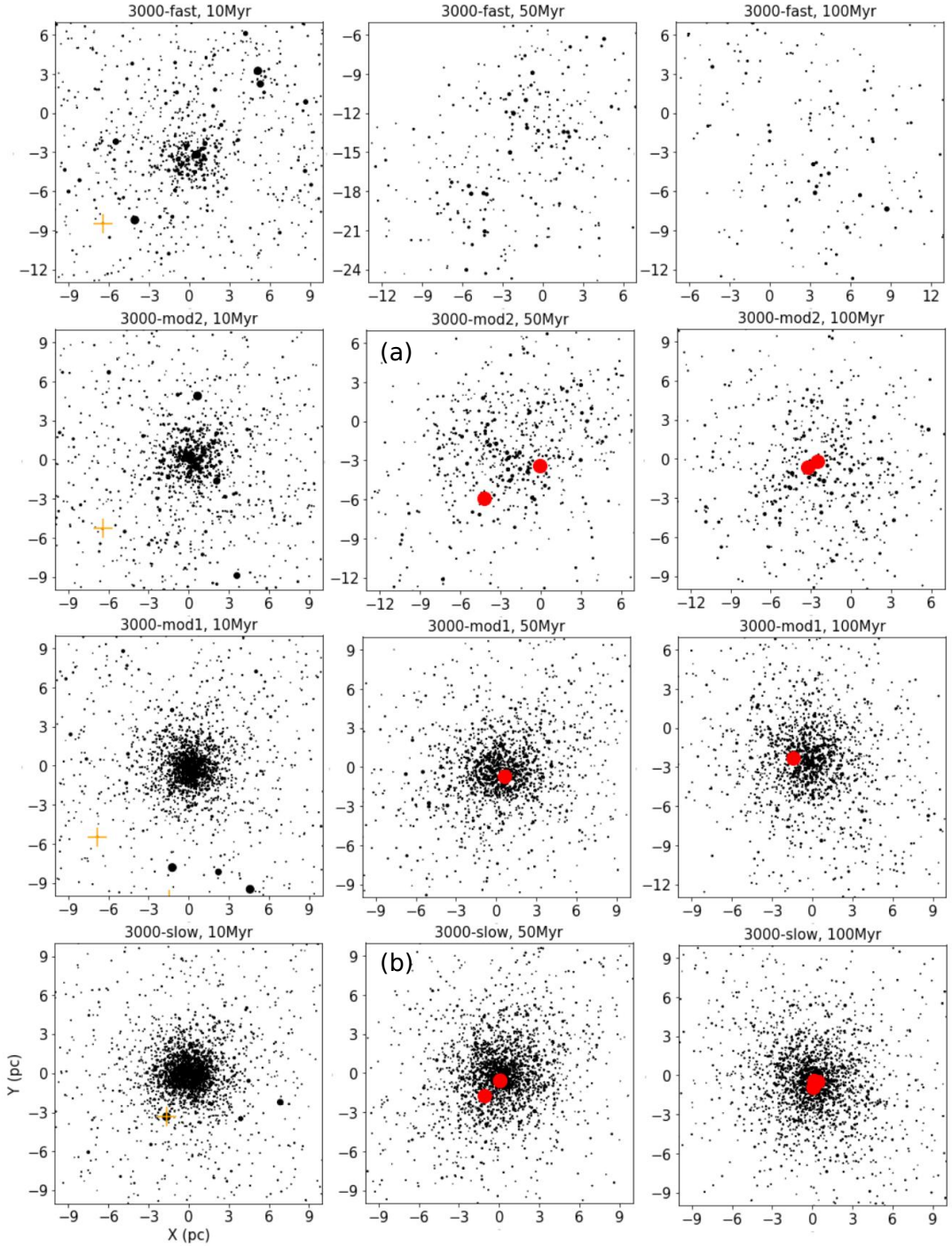


Figure 5.35: The evolution of the 3000 $M_{\odot}$ cluster under different gas expulsion modes. Snapshots projected on the XY plane (20 pc × 20 pc around the cluster center) at ages 10 Myr, 50 Myr, and 100 Myr are shown here. The size of the dots reflects the stellar mass. Orange pluses and red dots represent neutron stars and black holes, respectively. The central regions of the star clusters in panels (a) and (b) contain 2 and 3 black holes, respectively. In panel (b), two of the three black holes are too close to be distinguished on the map.

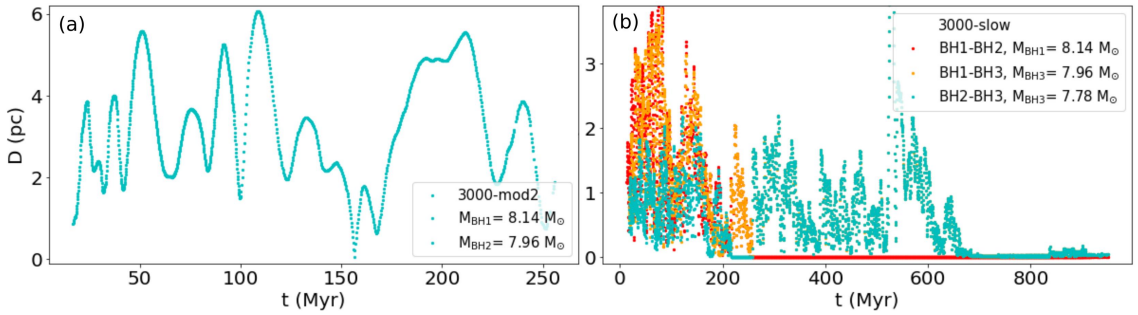


Figure 5.36: The temporal evolution of the distance between each pair of black holes.

5.11 Summary

5.11.1 Embedded clusters in ATLASGAL clumps

Using the optimal sampling model, the embedded clusters of ATLASGAL clumps with HII regions (HII-clumps) were synthesized. The 0.1 Myr isochrone was used to estimate the bolometric luminosity of each star in an embedded cluster, and the accretion luminosity of each star was also added. The total bolometric luminosity of the synthetic embedded clusters closely matches the observed bolometric luminosity of HII-clumps. To provide a more realistic scenario, age spreads in the young star and protostar populations of embedded clusters in HII-clumps were taken into account by modeling both constant and time-varying star formation histories (SFHs). Based on the age distribution of the stellar population, appropriate isochrones were assigned to each star, and the fraction of stellar objects still in the protostellar phase (Class 0 and Class I) was identified to properly include their accretion luminosities. Compared to a constant SFH, burst-like and time-dependent SFHs yield a better fit to the observational data. As long as 20% of the stars in the embedded cluster are still accreting, the contribution of accretion luminosity becomes significant to the total bolometric luminosity of low-mass HII-clumps with mass $\log_{10}(M_{\rm cl}/M_{\odot}) < 2.8$. Further investigation into variations in accretion rate, star formation efficiency (SFE), and the initial mass function (IMF), along with additional physical processes such as external heating from HII regions and flaring from pre-main sequence (PMS) stars and protostars, is required to fully explain the excess luminosity of low-mass HII-clumps.

5.11.2 Star formation efficiency of ATLASGAL clumps

A total of 1226 HII-clumps from the ATLASGAL survey with estimates of clump mass ($M_{\rm cl}$) and bolometric luminosity ($L_{\rm bol,obs}$) were matched with radio sources in the CORNISH-North and CORNISH-South survey catalogs. Among these, 392 HII-clumps were found to have corresponding radio sources. These clumps are predominantly located at the upper end of the $L_{\rm bol,obs}$–$M_{\rm cl}$ diagram and are more luminous than the remaining clumps, suggesting the presence of high-mass stars. The radio sources are excited by the most massive stars within the embedded clusters of these HII-clumps. The emission measure (EM), electron density ($n_{\rm e}$), and Lyman continuum flux ($N_{\rm Ly}$) were calculated for each radio source. Using the relationship between stellar parameters and Lyman continuum flux from Table.1 of R. I. Thompson (1984), the stellar luminosity ($L_{*,\rm T84}$) was derived based on the calculated $N_{\rm Ly}$.

A turning point was identified around $\log_{10}(L_{\rm bol,obs}/L_{\odot}) \approx 3.7$, corresponding to a clump mass of $\log_{10}(M_{\rm cl}/M_{\odot}) \approx 2.55$. Below this threshold, $L_{*,\rm T84}$ tends to significantly exceed $L_{\rm bol,obs}$, resulting in an abnormal tail in the $L_{\rm bol,obs}$–$L_{*,\rm T84}$ diagram. The stellar mass ($m_{*,\rm iso}$) was estimated using the 0.5 Myr isochrone based on $L_{*,\rm T84}$.

To explore the physical origin of this turning point, the optimal sampling algorithm was used to generate stellar populations for embedded clusters via the publicly available GalIMF code. The most massive star in each synthetic cluster, with mass $m_{\rm max}$, was identified and compared to the corresponding $m_{*,\rm iso}$. For $\log_{10}(m_{\rm max}/M_\odot) > 1$, a reasonable agreement between $m_{\rm max}$ and $m_{,\rm iso}$ was observed. The turning point at $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.55$ coincides with $\log_{10}(m_{\rm max}/M_\odot) \approx 1$, which is also adopted as a threshold for the dominance of accretion luminosity, as discussed in J. W. Zhou et al. (2024d). The abnormal tail suggests that, when accretion luminosity is dominant, stellar parameters cannot be reliably inferred from radio emission alone. To address this, the accretion luminosity ($L_{\rm acc}$) of the most massive star in each synthetic cluster was calculated and subtracted from $L_{*,\rm T84}$, yielding a corrected stellar luminosity $L = L_{*,\rm T84} - L_{\rm acc}$. However, this corrected luminosity may still be overestimated due to additional physical processes beyond accretion. For instance, magnetic flaring, common in low-mass stars, was considered. The ratio of optical to bolometric luminosity ($L_{\rm opt}/L_{\rm bol}$) in pre-main sequence stars decreases rapidly with stellar mass and remains below unity even in low-mass stars. Hence, luminosity from flaring activity can be considered negligible, though other contributing processes require further investigation.

In the analysis above, a constant star formation efficiency (SFE) of 0.33 was initially assumed. To explore the variation of SFE with clump mass, the mass of the most massive star in each HII-clump was used along with the theoretical $m_{\rm max}$–$M_{\rm ecl}$ relation to estimate the total embedded cluster mass and derive the SFE. A strong anti-correlation between SFE and clump mass was identified, with SFE decreasing with increasing clump mass. The median SFE value was approximately 0.3.

An independent estimation of SFE was also performed based on the model of J. W. Zhou et al. (2024d). SFEs within the range [0.05, 0.65] in steps of 0.05 were considered. For each HII-clump, 12 distinct star formation histories (SFHs) were applied to assign stellar ages. Both stellar and accretion luminosities were considered for each star using pre-main sequence evolutionary tracks. Each clump, therefore, yielded 156 combinations of SFEs and SFHs, resulting in 156 different bolometric luminosities. The optimal combination was selected by minimizing the difference with the observed bolometric luminosity of the clump. Overall, the SFEs derived observationally and through modeling showed good agreement.

5.11.3 Isolated evolution from clumps to embedded clusters

The dendrogram algorithm was employed to decompose the surface density distributions of stars from the MYStIX project into hierarchical structures. This algorithm reveals the multiscale structures of star clusters. Measurements of embedded clusters across different scales and evolutionary stages yield comparable mass–size relations. A straightforward explanation is that embedded clusters generally follow a Plummer profile. When a unified measurement is applied to the same sample, the resulting mass–size relation exhibits a shallower slope. In contrast, combining measurements taken at different scales tends to produce a steeper slope. This discrepancy is inevitable when assembling data from various catalogs and observational sources. Therefore, observed radii must be standardized prior to analysis.

The minimum spanning tree (MST) method was used to determine the separations between clusters and clumps within each massive star-forming region (MSFR). The comparable distances observed between cluster–cluster, clump–clump, and cluster–clump pairs suggest that the transformation from a clump to an embedded cluster occurs at the clump's original position, with no significant spatial displacement. No substantial mergers between embedded clusters are evident in the MYStIX sample. Additionally, feedback from embedded clusters appears not to significantly redistribute the surrounding clumps. Nevertheless, further validation requires analysis of a larger sample.

5.11.4 Expanding very young star clusters

The evolution of the mass-size relations during the first 4 Myr was traced using direct N-body simulations. Expanding lines at various time nodes were adopted as benchmarks to classify the observed radii of very young star clusters. The strong agreement between the simulated expanding lines and the observed data compellingly suggests that these young clusters are undergoing expansion. After unifying the radii of all samples to half-mass radii, mass–size relations were fitted for various categories, revealing consistency with the relation documented by Pfalzner et al. (2016). This indicates that the mass–size relation reported by Pfalzner et al. (2016) reflects the expansion of embedded clusters originating from the initial conditions outlined in Marks & Kroupa (2012). It is further proposed that even embedded clusters within ATLASGAL clumps associated with HII regions (HII-clumps) are already in an expansion phase.

All analyzed samples exhibit mass–size relations with similar slopes, potentially indicating that embedded clusters are initially formed with a phase-space density distribution resembling that of the Plummer model (e.g., Aarseth et al. 1974; D. Heggie & Hut 2003). This observation is noteworthy, as the Plummer phase-space density distribution function is the simplest analytical (two-parameter) solution to the time-independent collisionless Boltzmann equation (e.g., Kroupa 2008). Moreover, molecular filaments—where star formation is observed—often exhibit Plummer-like cross sections (P. J. André et al., 2022), and both open clusters (Röser et al., 2011; Röser & Schilbach, 2019) and globular clusters (Plummer, 1911) are well described by Plummer models.

5.11.5 Origin of open clusters

Samples of Galactic clumps, embedded clusters, and open clusters (OCs) were collected to compare their physical properties. Overall, the radii of OCs are significantly larger than those of embedded clusters and clumps. The radii of embedded clusters exceed those of clumps, likely due to the expansion of embedded clusters. The mass distribution of OCs spans a significantly broader range—by about one order of magnitude—than that of embedded clusters. Given the current mass distribution of clumps in the Milky Way, it is unlikely that a single clump evolving into a single embedded cluster, and subsequently into a single OC, could explain the population of observed OCs with relatively old ages and high masses. This conclusion is also supported by N-body simulations of individual embedded clusters.

A comparison of the separations between OCs and the typical sizes of molecular clouds in the Milky Way shows that most molecular clouds are only capable of forming one OC. However, each molecular cloud generally contains many embedded clusters, with typical separations of ≈1 pc (J.-w. Zhou et al., 2024b). After undergoing expansion, these embedded clusters may be able to coalesce.

Based on the spatial and mass distributions of embedded clusters in three massive star-forming regions (NGC 1893, NGC 6334, and the Carina Nebula) , simulations were conducted to evolve the embedded clusters in each region to ≈1 Myr. These clusters were then combined to reproduce the initial configuration of embedded cluster complexes similar to those observed in each MSFR. In the simulations, prior to 1 Myr, the evolution of each embedded cluster proceeds independently. As expansion continues, interaction begins upon contact, influencing their subsequent evolution.

The dynamics of the coalescence process, and the stability of the resulting cluster, are affected by the spatial distribution, relative velocities, mass distribution, and gas expulsion modes of the embedded clusters. Large line-of-sight separations inhibit effective coalescence and the formation of a stable core. Greater spatial separation and higher relative velocities contribute to larger final radii and accelerated mass loss in the coalesced cluster; however, these effects can be mitigated by slower gas expulsion rates.

The embedded clusters in NGC 1893 are low in mass and few in number, leading to rapid dispersion after expansion regardless of initial simulation conditions. Rather than true coalescence, a loose

association may result. Both NGC 6334 and NGC 1893 host a small number of embedded clusters, but NGC 6334 contains more massive ones, facilitating cluster coalescence. Thus, embedded cluster mass plays a critical role in enabling coalescence. Massive clusters offer a stable gravitational core, which aids the coalescence process and resists both expansion and Galactic tidal forces. A comparison between the Carina Nebula and NGC 1893 suggests that a larger number of embedded clusters can also promote coalescence.

Currently observed clumps in the Milky Way are not sufficiently massive to directly form massive embedded clusters. However, a population of lower-mass embedded clusters, through coalescence, could reproduce the parameter space of observed open clusters. In the post-gas expulsion coalescence scenario, massive open clusters need not originate from single, massive embedded clusters.

5.11.6 Evolution of the $m_{\rm max} - M_{\rm ecl}$ relation

A crucial factor in determining whether a star cluster should be modeled using stochastic or optimal sampling from the initial mass function (IMF) is the presence of a significant correlation between the mass of the most massive star ($m_{\rm max}$) and the total stellar mass of the embedded cluster ($M_{\rm ecl}$). Analysis of the Gaia DR3 open cluster catalog indicates that, as clusters evolve, the $m_{\rm max} - M_{\rm cluster}$ relation in open clusters gradually deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation observed in embedded clusters, ultimately displaying an age-dependent stratification.

Across age groups, older open clusters tend to exhibit a higher upper mass limit of $M_{\rm cluster}$. If open clusters originate from individual embedded clusters, the relatively old and massive open clusters must have formed in earlier epochs, as no younger counterparts with similar properties are observed. Alternatively, massive open clusters may form through the post-gas expulsion coalescence of multiple embedded clusters, thereby eliminating the necessity for massive embedded clusters as their direct progenitors. Additionally, considering the current mass distribution of clumps in the Milky Way, the formation of such massive open cluster precursors under present conditions appears unfeasible (J.-w. Zhou et al., 2024a).

The observed $m_{\rm max} - M_{\rm cluster}$ relation of open clusters was compared with results from N-body simulations of both individual cluster evolution and cluster coalescence. In the individual cluster simulations, gas expulsion modes significantly influence the $m_{\rm max} - M_{\rm cluster}$ relation, as mass loss is highly sensitive to the nature of gas expulsion. Under fast gas expulsion, clusters with initial masses $\leq$3000 M$\odot$ fail to survive beyond 100 Myr, and the resulting $m_{\rm max} - M_{\rm cluster}$ relation only aligns with the upper boundary of the observed relation for open clusters. To reproduce the observed $m_{\rm max} - M_{\rm cluster}$ relation, a slow gas expulsion mode is necessary.

For clusters younger than 10 Myr, $m_{\rm max}$ values in open clusters are systematically lower than those predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation. This discrepancy is unlikely to result from the natural evolution of embedded clusters. In simulations of individual clusters, even at an age of $\approx$10 Myr, $m_{\rm max}$ remains near the initial $m_{\rm max} - M_{\rm ecl}$ relation. One possible explanation is that $m_{\rm max}$ in these open clusters was already lower than predicted by the initial relation at birth, suggesting a coalescence scenario in which each individual embedded cluster initially follows the $m_{\rm max} - M_{\rm ecl}$ relation. Furthermore, observations indicate that open clusters around 10 Myr old have clearly diverged from the initial $m_{\rm max} - M_{\rm ecl}$ relation, whereas clusters of similar age in individual cluster simulations continue to roughly follow it—an inconsistency with observational data. Both of these discrepancies in the individual cluster simulations are effectively resolved in cluster coalescence simulations. This outcome supports the conclusion that post-gas expulsion coalescence of embedded clusters is likely a dominant formation mechanism for open clusters, consistent with the findings of J.-w. Zhou et al. (2024a) and J.-w. Zhou et al. (2025).

5.11.7 Constraining the gas expulsion modes of embedded clusters

Recent observational progress has highlighted the critical role of early expansion in the evolution of young star clusters. N-body simulations suggest that the gas expulsion mode of initially embedded star clusters plays a decisive role in shaping their dynamical state and subsequent evolution. In this study, the gas expulsion mode of these clusters is constrained by inversely analyzing the expansion speeds of very young star clusters (< 5 Myr) using Gaia DR3 data. A series of star clusters with varying masses were also simulated under different gas expulsion modes (fast, moderate, and slow) for comparison with observations. To evaluate the similarity between the velocity distributions of stars in observed and simulated clusters, the Kolmogorov–Smirnov test was applied, accounting for observational uncertainties in velocity, mass, and age. The results suggest that all gas expulsion modes are possible for the observed clusters. Contrary to common assumptions, gas expulsion timescales are not determined solely by the rate of gas removal driven by stellar feedback. Instead, they also reflect uncertainties arising from additional physical factors, such as star formation efficiency and the spatial distribution of residual gas. Given the current limitations of N-body simulations, a range of gas expulsion modes should be considered when comparing simulations with observations. Clusters undergoing slow gas expulsion tend to contain fewer massive stars, and their most massive members are also relatively smaller. Nevertheless, the IMF of all clusters remains top-heavy across different gas expulsion modes. This implies that the slow expansion observed in some clusters is not necessarily attributable to a lack of massive stars. Furthermore, the extinction density around clusters shows no significant variation across different gas expulsion modes, suggesting that gravitational interactions with surrounding material are unlikely to be the primary cause of slow expansion. The coalescence of expanding star clusters can result in a reduction of overall expansion speeds. Future work will investigate the potential of this mechanism to account for the slow expansion observed in young star clusters.

5.11.8 Black hole formation in open clusters

This study examines how the star cluster formation process influences the evolutionary outcomes of star clusters. The progenitors of black holes are massive stars. Whether a star cluster is modeled using stochastic or optimal sampling from the IMF directly determines the initial population of massive stars. A significant relationship between the mass of the most massive star ($m_{\rm max}$) and the total mass of the embedded cluster in stars ($M_{\rm ecl}$) supports the optimal sampling scenario. Recent observational advances have underscored the critical role of early expansion in the evolution of young star clusters. In N-body simulations, the gas expulsion mode (or timescale) of initially embedded clusters exerts a near-decisive influence on their dynamical state and subsequent evolution. This, in turn, strongly affects the probability of ejecting massive stars and compact objects, thereby influencing the black hole formation rate within clusters. Furthermore, cluster survival time varies under different gas expulsion modes, which determines how long the evolutionary products of massive stars remain in the cluster environment—potentially impacting the formation of multiple systems involving black holes.

Regarding the cluster formation mechanism, clusters that form through the coalescence of multiple subclusters exhibit significantly different dynamics and evolutionary outcomes compared to those formed as single entities. Given that the post-gas expulsion coalescence of embedded clusters within a common parental molecular cloud may be a dominant mechanism in the formation of open clusters, exploring black hole formation in various coalescence scenarios appears promising.

A cluster's evolution is closely tied to its formation process, rather than occurring independently. Variations in initial cluster mass, gas expulsion modes, and formation mechanisms ultimately affect the expansion speed, which in turn influences both the rate of mass loss and the likelihood of black

hole formation. The expansion speed reflects the overall stellar velocity distribution, which plays a critical role in determining whether massive stars and their remnants remain gravitationally bound to the cluster. These physical factors should therefore be incorporated into the initial conditions of N-body simulations. Achieving this requires a comprehensive physical framework that captures the full lifecycle of star cluster formation and evolution. Traditionally, the evolution of star clusters and compact object formation, as well as the formation of massive stars and star clusters, have been studied by separate research communities. However, a thorough understanding of black hole formation within open clusters demands integration of insights from both fields to reconstruct a unified physical picture.

CHAPTER 6

Mass functions in star-forming regions

The content presented in this chapter are reproduced from the publications by J. W. Zhou et al. (2025) (Sec.6.2), and two submitted works (Sec.6.3 and Sec.6.4). These studies aim to link star-forming regions and stellar populations using the IMF and integrated IMF theories, and explore the transition from core mass functions to the initial mass function.

6.1 Data

6.1.1 Molecular cloud, clump and core catalogs

For three high-mass star-forming regions, W43-MM1&MM2&MM3, we used the core catalogs in Motte et al. (2018a), Pouteau et al. (2022), Nony et al. (2023) and Louvet et al. (2024). For the same observation in the ALMA-IMF survey (Motte et al., 2022), the core catalogs of W43-MM1&MM2&MM3 are presented in Pouteau et al. (2022), Nony et al. (2023) and Louvet et al. (2024). For W43-MM1, there is an independent catalog derived from the observation in Motte et al. (2018a). For the same region, different catalogs have different selection criteria and identifications, varying in the number of cores and their mass distribution, which is reflected in the variations observed in the CMF, as shown in Fig.6.1.

Urquhart et al. (2022) provides an up-to-date version of the ATLASGAL clump catalogue. The individual ATLASGAL clumps were organized into complexes using friends-of-friends clustering analysis. This method associates clumps that are closely situated in terms of longitude, latitude, and velocity. The study of Miville-Deschênes et al. (2017) presents a catalog of 8107 molecular clouds identified by CO (1-0) emission that covers the entire Galactic plane. We matched the ATLASGAL clump complexes in Urquhart et al. (2022) with the molecular clouds in Miville-Deschênes et al. (2017) by minimizing the differences in their central coordinates and systemic velocities. We also used the CO (1-0) cube of Dame et al. (2001)[1].

6.1.2 Star cluster and stellar association catalogs

The catalogs for compact clusters and multi-scale associations across five galaxies (NGC 1433, NGC 1566, NGC 1559, NGC 3351 and NGC 3627) are being made available as the DR3 release of the PHANGS-HST survey. A detailed description of these catalogs can be found in J. C. Lee et al. (2022), Whitmore et al. (2021), Thilker et al. (2022) and K. L. Larson et al. (2023). In our analysis, we exclude

[1] https://lweb.cfa.harvard.edu/rtdc/CO/

Table 6.1: Resolutions of the datasets. NGC 1559 is not observed by JWST currently. "D" and "i" are the distances and the inclination angles of the galaxies, respectively.

	D (Mpc)	i (deg)	angular (") ALMA	linear (pc) ALMA	linear (pc) JWST
NGC 1433	18.6	28.6	1.10	99.1	58.6
NGC 1559	19.4	65.4	1.25	117.5	–
NGC 1566	17.7	29.5	1.25	107.6	55.7
NGC 3351	10.0	45.1	1.46	70.7	31.5
NGC 3627	11.3	57.3	1.63	89.2	35.6

clusters and associations whose estimated masses are less than their mass estimation errors, as their masses are not well determined.

6.1.3 CO (2-1) cube and 21 μm map

For the five galaxies mentioned above, we used the combined 12m+7m+TP PHANGS-ALMA CO (2−1) data cubes to investigate gas kinematics and dynamics, which have a spectral resolution of 2.5 km s^{-1}. We also used the JWST 21 μm (F2100W) maps from the PHANGS-JWST survey. The full width at half maximum (FWHM) of the original F2100W point spread function (PSF) is ≈0.65" (T. G. Williams et al., 2024). The angular and linear resolutions of these datasets are listed in Table.6.1. Overviews of the PHANGS-ALMA and PHANGS-JWST surveys' science goals, sample selection, observation strategy, and data products are described in Leroy et al. (2021b), Leroy et al. (2021a) and J. C. Lee et al. (2023). All the data are available on the PHANGS team website [2].

6.2 From CMF to IMF

In this work, we start from the canonical IMF to inversely synthesize the observed CMF of high-mass star formation regions, taking into account variations in multiplicity and core's star formation efficiency ($\epsilon_{\rm core}$). The work presented in this section is reproduced from the publication by J. W. Zhou et al. (2025).

6.2.1 Observed CMFs

For W43-MM2&MM3, the core catalogs in Nony et al. (2023) were directly selected from those in Pouteau et al. (2022), which only retains a portion of the original lists. The CMFs derived from the core catalogs in Pouteau et al. (2022) and Nony et al. (2023) have similar slopes, but the catalogs in Nony et al. (2023) with fewer cores shift systematically upward. For W43-MM1, three independent identifications in different observations give comparable CMFs. For W43-MM3, two independent identifications in the same observation give quite different CMFs. Therefore, numerous observational constraints affect the CMF, including the calculation of the core's physical parameters, observational resolution, identification methods, and the completeness of the core population. However, the shallow nature of the CMF slope in W43 seems fairly robust at least.

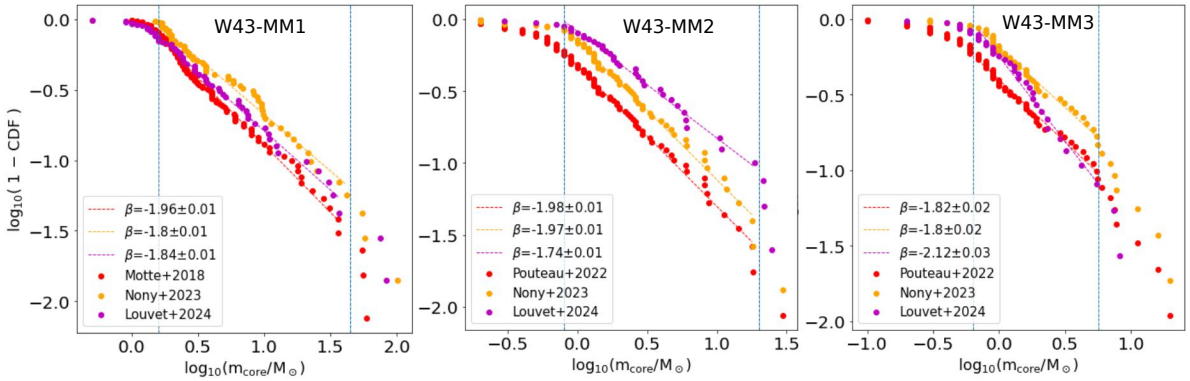


Figure 6.1: Complementary cumulative distribution function (CCDF) plot and the linear fitting of the observed CMF in the mass range marked by two vertical dashed lines. β is the slope of the corresponding dN/dM mass function derived from linear fitting of the CCDF. The lower mass of the adopted fitting range corresponds to the observational mass completeness level estimated by Louvet et al. (2024).

Table 6.2: Physical parameters of the corresponding ATLASGAL clumps. $M_{\rm fwhm}$ is the clump mass defined in Urquhart et al. (2022). $M_{\rm ecl}$ is the mass of the embedded cluster finally formed in a clump. $m_{\rm max}$ is the mass of the most massive star in an embedded cluster. $N_{\rm star}$ is the total number of stars in the cluster.

region	clump	$\log_{10}(M_{\rm fwhm})$ $(M_\odot)$	$M_{\rm ecl}$ $(M_\odot)$	$m_{\rm max}$ $(M_\odot)$	$N_{\rm star}$
W43-MM3	AGAL030.718-00.082	3.173	399	26	798
W43-MM2	AGAL030.703-00.067	3.211	435	27	870
W43-MM1	AGAL030.818-00.056	3.248	473	29	946

6.2.2 Clumps and embedded clusters

Clumps are localized high-density structures within a molecular cloud and serve as the primary sites of star formation within the cloud (Urquhart et al., 2018; Urquhart et al., 2022). Cores are assumed to be overdense regions with respect to their parental clump. Generally, a cloud includes many clumps, and a clump includes many cores. They appear as hierarchical, multiscale hub-filament structures (Kumar et al., 2020; J.-W. Zhou et al., 2022; J. W. Zhou et al., 2023; J. W. Zhou et al., 2024f; J.-W. Zhou & T. Davis, 2024). To trace the evolution from the CMF to the IMF, the observation should cover the entire clump, as was done in the ALMA-IMF survey. A clump eventually forms an embedded cluster with an IMF (Yan et al., 2017; J. W. Zhou et al., 2024d; J. W. Zhou et al., 2024b). This IMF corresponds to the initial CMF in the clump. Therefore, clumps should be regarded as the fundamental units for studying the origin of the IMF (Kroupa et al., 2024). W43 is a molecular cloud complex. The main star-forming regions in W43 are the clumps. The three subregions of W43 (W43-MM1&MM2&MM3) correspond to three independent clumps displayed in Table.6.2. They have independent star formation (J.-w. Zhou et al., 2024b), and thus different CMFs, as is shown in Fig.6.1.

The ATLASGAL clumps in Urquhart et al. (2022) corresponding to each region are listed in Table.6.2. J. W. Zhou et al. (2024d) and J. W. Zhou et al. (2024b) systematically studied the relationship between the physical parameters of the embedded clusters within ATLASGAL clumps with HII regions (HII-clumps) and the physical parameters of the clumps themselves. We derived a $M_{\rm cl} - M_{\rm ecl}$ relation in J. W. Zhou et al. (2024b). Therefore, according to the clump mass ($M_{\rm cl}$), we can directly estimate the embedded cluster mass in stars ($M_{\rm ecl}$) in the clump. Although this relation is derived from HII-clumps, considering

[2] https://sites.google.com/view/phangs/home

that the clumps in different evolutionary stages have similar mass distributions (Urquhart et al., 2022), it can also be used for younger clumps.

6.2.3 Initial stellar populations

Based on the recipe described in J.-w. Zhou et al. (2024a), the initial stellar population in each high-mass star-forming region was generated using the updated `McLuster` program (Küpper et al., 2011; L. Wang et al., 2019), according to the embedded cluster mass ($M_{\rm ecl}$). Assuming that the average mass of stars in a cluster is 0.5 $M_{\odot}$, the total number of stars in a cluster ($N_{\rm star}$) is $M_{\rm ecl}$/(0.5 $M_{\odot}$). The proportion of binaries in the initial stellar population ($f_{\rm b}$) is an adjustable parameter in the `McLuster` code. Since the primary and secondary mass of the binary were selected randomly from the canonical IMF, for each embedded cluster, we ran `McLuster` 100 times, and overlapped these 100 stellar populations to derive a synthesized core population. For binaries, we summed the masses of each pair corresponding to one core. For single stars, we retained the generated mass. Then the stellar mass was divided by $\epsilon_{\rm core}$ to construct the mass distribution of the cores.

6.2.4 Comparison of synthesized and observed CMFs

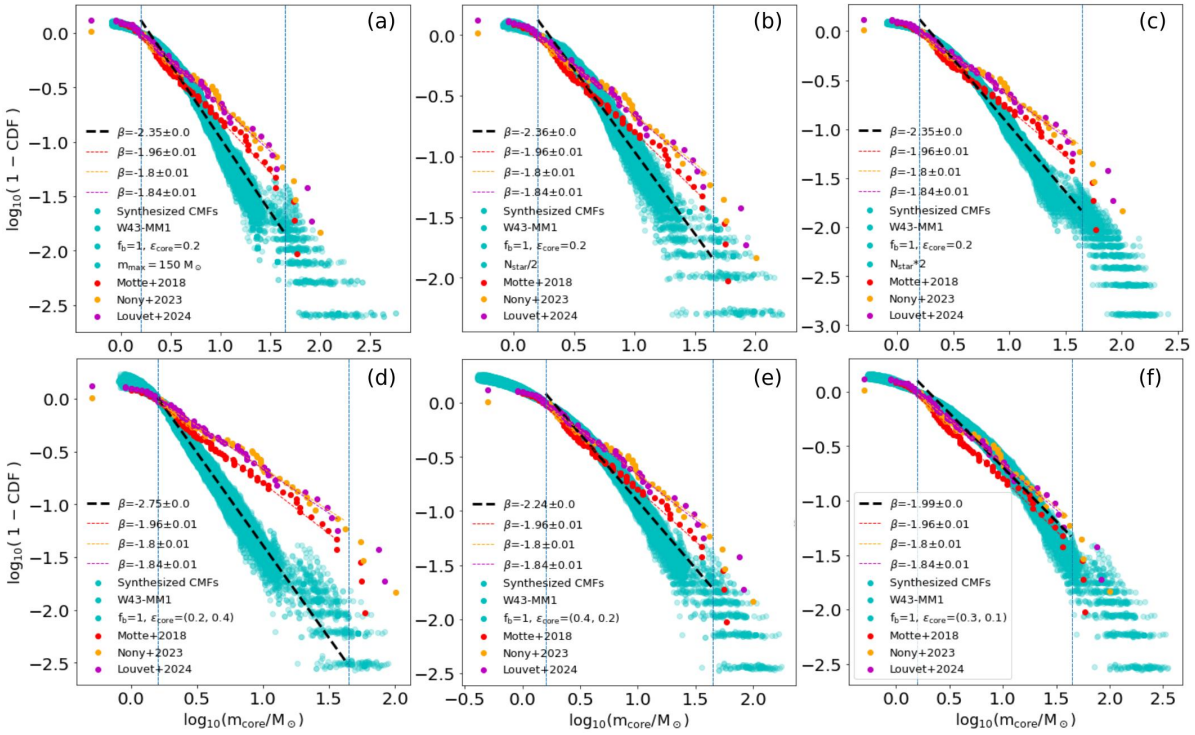


Figure 6.2: Complementary cumulative distribution function (CCDF) plot and the linear fitting of the synthesized CMFs (in cyan) for W43-MM1 by setting $f_{\rm b}$=1 and $\epsilon_{\rm core}$=0.2. (a)-(c) The cases of $m_{\rm max}$=150 $M_{\odot}$, half the original number of stars ($N_{\rm star}$/2) and twice the original number of stars ($N_{\rm star}$*2), respectively; (d) The case in which $\epsilon_{\rm core}$ decreases with decreasing core mass. $\epsilon_{\rm core}$ was randomly selected from the range (0.2, 0.4); (e) The case in which $\epsilon_{\rm core}$ increases with decreasing core mass. $\epsilon_{\rm core}$ was randomly selected from the range (0.2, 0.4); (f) Same as panel (e), but $\epsilon_{\rm core}$ is in the range (0.1, 0.3). All selected $\epsilon_{\rm core}$ in panels (d), (e), and (f) are sorted in descending or ascending order. All CCDFs (both observed and synthesized) were normalized to 1 at the mass completeness level marked by the left vertical dashed line.

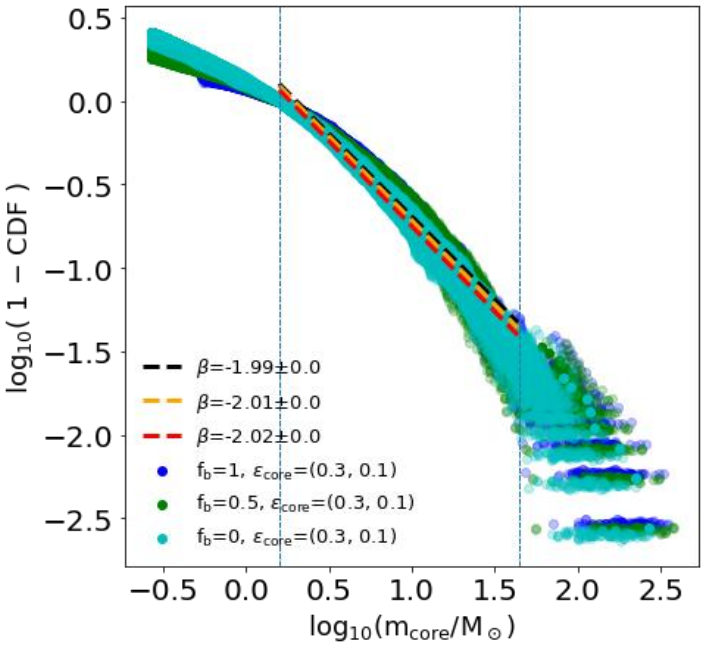


Figure 6.3: Comparison of the synthesized CMFs for W43-MM1 for f_b=1, 0.5 and 0. All synthesized CCDFs were normalized to 1 at the mass completeness level marked by the left vertical dashed line.

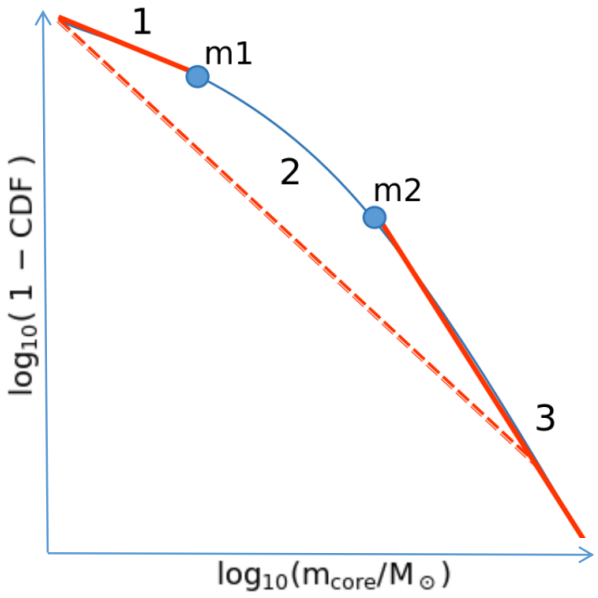


Figure 6.4: The synthesized CMF can be roughly divided into three segments, i.e., “1”, “2”, and “3”. Segment “2” is the transition part. “m1” and “m2” mark the mass range of the transition. Segment “3” is the real CMF of high-mass cores. The dashed red line is the observed CMF, for which observational biases prevent one from distinguishing the transition, resulting in a combined fit of three different segments with a shallower slope. Thus, the observed top-heavy CMF is an artifact.

For all three regions in Fig.6.1, the low-mass end of the CMFs does not exhibit the clear transition seen in the IMF, which may be caused by the uncertainty in core mass estimation. As is discussed in Louvet et al. (2024), the primary sources of error in the mass estimates stem from uncertainties in the opacity index and core temperature. To account for these uncertainties, they conducted a Monte Carlo simulation on 330 cores exceeding the 1.64 $M_{\odot}$ completeness threshold, allowing simultaneous variations in the opacity index, temperature, and source flux. Fig.6 of Louvet et al. (2024) presents the cumulative CMFs obtained from 10^3 simulation trials, where we can see a clear slope transition.

According to the observed CMF, we have fixed the mass range of the linear fitting for each clump in Fig.6.1. For the synthesized core population of each clump, we strictly follow the fixed mass range to do linear fitting (the dashed black segments in Fig.6.2 and Fig.6.3). A direct comparison between the synthesized and observed CMFs is only meaningful above the estimated mass completeness level of the observed sample. Therefore, for the purpose of direct comparison, all CCDFs (both observed and synthesized) were normalized to 1 at the mass completeness level marked by the left vertical dashed line in Fig.6.1.

In Fig.6.2(f), there is a noticeable transition region in the synthesized CMFs, where the slope changes from flat to steep; that is, segment "2" in Fig.6.4. As is illustrated by the dashed red line in Fig.6.4, if the fitting range lies precisely within this transition region, it encompasses both the flat and steep slopes, resulting in an overall flat slope. If an observed CMF truly exhibits such a slope transition but observational biases obscure it, leading to a combined fit with a shallower slope, then the apparent top-heavy CMF would be merely an artifact. To accurately fit the high-mass end of the CMF, it is essential to identify whether the critical masses for the transition – "m1" and "m2" in Fig.6.4 – exist. If so, this critical mass, m2, should be used as the starting point for fitting the slope of the high-mass end, which can then be compared to the Salpeter value. These issues appear in Fig.6 of Louvet et al. (2024), in which a combined fit results in a shallower CMF slope. When focusing only on the high-mass end of this figure, the CMF slope is actually comparable to the Salpeter value.

By setting $f_{\rm b} = 1$ and $\epsilon_{\rm core} = 0.2$, in Fig.6.2(a-c), we explored how the selected $m_{\rm max}$ and $N_{\rm star}$ (or $M_{\rm ecl}$) influence the synthesized CMFs, and found that the synthesized CMFs are not sensitive to these parameters. A larger $m_{\rm max}$ will increase the mass at the high-mass end, which is beyond the observed core mass range. For the same reason, we do not consider systems of a higher order than binaries (such as triples and quadruples) in this work. In Fig.6.3, the synthesized CMFs with $f_{\rm b} = 1$ do not provide better fits than those with $f_{\rm b} = 0.5$ or $f_{\rm b} = 0$. A high multiplicity fraction in the synthesis is not necessary.

In Fig.6.2(a-c), the synthesized CMFs do not align with the observed ones. The reason may be that, in each case, we apply a constant $\epsilon_{\rm core}$ for all the synthesized cores. A likely possibility is that cores of different masses have different $\epsilon_{\rm core}$. In Fig.6.2(d-e), we compare two scenarios in which $\epsilon_{\rm core}$ either increases or decreases with decreasing core mass, respectively. In each case, $\epsilon_{\rm core}$ was randomly selected from a range, then, all selected $\epsilon_{\rm core}$ were sorted in descending or ascending order. We found that in order to match the observed CMFs, $\epsilon_{\rm core}$ must increase with decreasing core mass. A possible explanation is that higher-mass cores contain more gas, leading to the formation of more massive stars, and stronger feedback ultimately disperses more gas. This is similar to the observational results at the clump scale (J. W. Zhou et al., 2024b), where the star formation efficiency is strongly correlated with the clump mass: higher-mass clumps exhibit lower SFEs. Due to observational resolution limitations, many of the so-called high-mass cores may actually be sub-clumps, which eventually form multiple systems, rather than just a single star. Furthermore, we should keep in mind that the cores in high-mass star formation regions can continue to accrete and accumulate mass during the process of star formation. These need to be considered when estimating $\epsilon_{\rm core}$.

To align the synthesized CMF with the observed CMF, we need to assume lower-mass cores that have higher $\epsilon_{\rm core}$, and then adjust the range of $\epsilon_{\rm core}$ accordingly. As shown in J. W. Zhou et al. (2025),

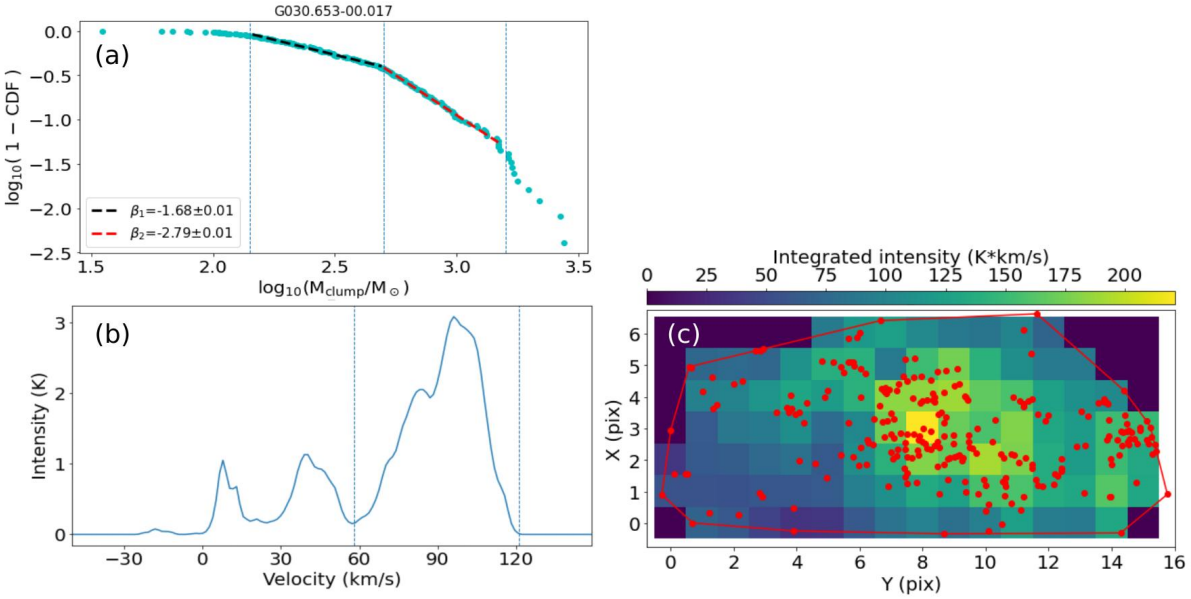


Figure 6.5: Clump complex G030.653-00.017 as an example. (a) Clump mass function. The complementary cumulative distribution function (CCDF) plot and the linear fitting of the observed CMF in the mass range marked by the vertical dashed lines. The two segments of the clump mass function correspond to three mass endpoints, i.e. M_1, M_2 and M_3; (b) Average spectrum of CO (1-0) emission for all molecular gas in the region marked in panel (c). Two vertical lines mark the velocity range of the molecular gas covering the clump complex; (c) The convex hull of G030.653-00.017, which is the smallest convex shape that encloses all given clumps. Red dots are the clumps.

this approach successfully aligns the synthesized CMFs of W43-MM1/MM2/MM3 with the observed ones. Since the currently observed CMF may be not robust, we do not discuss the selection of optimal parameters further, as this depends on the specific CMF. Ideally, the mass dependence of $\epsilon_{\rm core}$ would be directly determined from observations in future work.

In this work, we have assumed a canonical IMF and inversely reconstructed the observed CMFs to investigate potential factors contributing to a top-heavy CMF. The factors discussed in this work may or may not be valid, implying that the top-heavy CMF could be genuine. The variation in the IMF with environmental conditions has been extensively studied (J. C. Lee et al., 2009; Gunawardhana et al., 2011; Mor et al., 2019; Dib, 2023b; Kroupa et al., 2024; del Alcázar-Julià et al., 2025). Rather than asserting the universality of the IMF, this work emphasizes potential issues in the CMF fitting that warrant careful consideration.

6.3 Clump populations in molecular clouds

Since only molecular clouds within the Milky Way can be resolved at the level of their internal structures, they serve as a crucial reference for understanding the internal composition and star formation processes of molecular clouds beyond our galaxy. In this study, we explore the connection between molecular clouds and their internal clumps in the Milky Way, which may serve as a template for dissecting molecular clouds in nearby galaxies. Especially, we focus on the clump mass function of individual molecular clouds, which allows us to generate the population of clumps within a molecular cloud based on its mass. This approach facilitates the study of stellar population formation and evolution within molecular clouds.

Table 6.3: Physical parameters of the clump mass function for the 31 largest clump complexes in the Milky Way.

Name	N	l (deg)	b (deg)	D (kpc)	Q	C_1	C_2	β_1	β_2	M_1 ($M_\odot$)	M_2 ($M_\odot$)	M_3 ($M_\odot$)	$M_{\rm clump,tot}$ ($M_\odot$)	$M_{\rm clump,min}$ ($M_\odot$)	$M_{\rm clump,max}$ ($M_\odot$)	$M_{\rm cloud,as}$ ($M_\odot$)	$M_{\rm cloud,re}$ ($M_\odot$)
G030.653-00.017	250	30.653	-0.017	4.90	0.64	4.71	7.88	1.68	2.79	2.15	2.70	3.20	5.10	1.55	3.47	6.06	6.73
G333.121-00.353	245	333.121	-0.353	3.57	0.60	4.23	7.22	1.89	3.09	2.05	2.45	3.05	4.82	1.49	3.22	6.24	6.43
G024.322+00.187	61	24.322	0.187	7.60	0.66	4.09	6.83	1.29	1.89	2.25	2.75	3.30	4.89	2.09	3.82	6.01	6.01
G336.996-00.002	107	336.996	-0.002	4.70	0.66	4.18	6.90	1.60	2.23	2.25	2.95	3.60	4.99	1.56	4.09	5.76	5.98
G012.960-00.168	85	12.96	-0.168	2.60	0.63	3.41	6.09	1.53	3.08	1.70	2.25	2.80	4.24	1.67	2.85	5.82	5.93
G331.389-00.106	87	331.389	-0.106	3.90	0.64	3.66	6.34	1.91	3.38	2.20	2.65	3.25	4.50	1.59	3.23	5.46	5.92
G018.936-00.350	58	18.936	-0.35	4.00	0.68	3.46	6.13	1.41	3.14	1.80	2.40	2.95	4.30	1.71	3.22	5.75	5.91
G305.453+00.065	102	305.453	0.065	4.00	0.63	3.91	6.58	1.91	3.58	2.35	2.80	3.15	4.76	1.88	3.45	5.83	5.90
G023.072-00.349	41	23.072	-0.349	5.00	0.52	3.57	6.24	1.70	2.61	2.15	2.75	3.30	4.42	1.89	3.40	5.91	5.89
G049.261-00.318	93	49.261	-0.318	5.20	0.59	4.14	6.77	1.34	2.51	2.10	2.85	3.70	5.01	2.08	3.90	5.09	5.83
G006.183-00.358	78	6.183	-0.358	3.00	0.61	3.46	6.03	1.26	2.40	1.60	2.20	2.75	4.37	1.51	3.20	5.34	5.72
G332.124-00.070	67	332.124	-0.07	3.10	0.47	3.32	5.84	1.48	2.94	1.80	2.30	2.95	4.24	1.44	2.99	5.28	5.64
G340.248-00.273	61	340.248	-0.273	3.56	0.72	3.30	5.75	1.23	2.71	1.85	2.20	2.80	4.26	1.85	2.98	5.43	5.52
G010.147-00.237	65	10.147	-0.237	2.90	0.71	3.16	5.61	1.41	2.11	1.65	2.10	2.55	4.13	1.47	2.82	4.98	5.51
G345.479+00.954	51	345.479	0.954	2.38	0.63	3.18	5.62	1.61	2.64	1.95	2.40	2.70	4.15	1.48	3.12	5.35	5.49
G331.094-00.414	53	331.094	-0.414	3.98	0.52	3.42	5.85	1.34	2.22	1.85	2.40	3.10	4.39	1.86	3.24	4.98	5.49
G337.690-00.053	38	337.688	-0.053	3.57	0.76	3.03	5.46	1.38	3.75	1.75	2.35	2.70	4.00	1.45	3.07	5.22	5.48
G014.477-00.149	40	14.477	-0.149	3.08	0.54	3.13	5.54	1.37	2.54	1.65	2.40	2.80	4.11	1.67	3.14	5.34	5.44
G327.006-00.161	48	327.006	-0.161	3.46	0.65	3.13	5.49	1.37	2.04	1.90	2.25	2.75	4.12	1.70	2.87	5.47	5.37
G343.620-00.165	45	343.62	-0.165	2.60	0.57	2.83	5.18	1.58	3.07	1.60	2.10	2.50	3.83	1.39	2.71	5.52	5.33
G008.581-00.329	38	8.581	-0.329	4.45	0.52	3.29	5.56	1.25	2.60	1.80	2.45	3.10	4.32	1.51	3.60	5.23	5.20
G018.179-00.303	41	18.179	-0.303	3.41	0.86	2.93	5.18	1.48	2.68	1.70	2.25	2.80	3.96	1.43	2.97	4.82	5.16
G328.166-00.540	41	328.166	-0.54	2.73	0.69	3.02	5.26	1.45	2.20	1.85	2.25	2.68	4.06	1.89	2.87	5.20	5.14
G335.294-00.261	36	335.294	-0.261	2.92	0.76	3.00	5.24	1.43	2.99	1.80	2.25	2.65	4.04	1.84	3.19	5.10	5.13
G351.206+00.703	159	351.206	0.703	1.30	0.52	3.10	5.32	2.03	2.91	1.40	2.30	2.70	4.14	0.72	3.02	4.84	5.11
G016.996+00.719	62	16.996	0.719	1.86	0.58	2.77	4.93	1.77	2.87	1.45	2.10	2.45	3.83	1.28	2.81	4.96	4.99
G015.079-00.625	55	15.079	-0.625	2.00	0.70	2.96	5.09	1.76	2.44	1.75	2.10	2.80	4.02	1.33	3.33	5.02	4.94
G353.118+00.702	136	353.118	0.702	1.35	0.72	2.81	4.88	1.49	2.91	1.20	1.70	2.15	3.87	0.84	2.24	4.80	4.84
G326.616+00.648	67	326.616	0.648	1.82	0.64	2.76	4.81	1.35	2.50	1.20	1.80	2.53	3.83	1.13	2.61	4.72	4.80
G344.102-00.632	38	344.102	-0.632	2.21	0.72	2.94	4.98	1.34	2.23	1.75	2.25	2.70	4.01	1.76	2.76	4.78	4.78
G014.227-00.571	66	14.227	-0.571	1.50	0.61	2.77	4.79	1.15	2.18	1.20	1.75	2.16	3.84	1.20	2.52	4.19	4.75

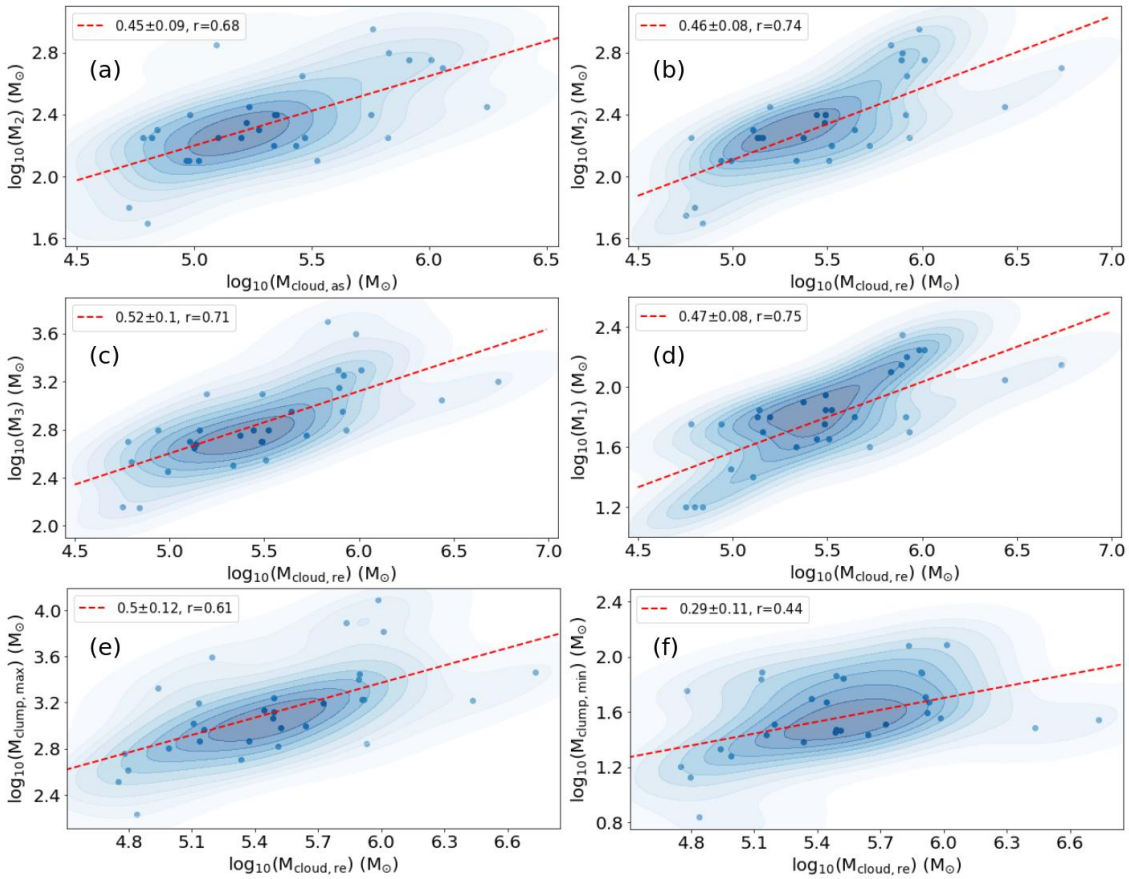


Figure 6.6: Correlations between the broken masses of the clump mass function (M_1, M_2 and M_3), the smallest and largest clump masses in the clump complex ($M_{\rm clump,min}$ and $M_{\rm clump,max}$) and the cloud mass ($M_{\rm cloud,as}$ or $M_{\rm cloud,re}$). r is the Pearson correlation coefficient.

6.3.1 Clump mass function in observation

In order to fit the clump mass function of each clump complex in Urquhart et al. (2022), it is necessary to ensure a sufficient number of clumps. For all clump complexes, we first sorted them by the number of clumps and then performed fitting sequentially. We found that when the number of clumps in a clump complex is less than 35, the clump mass function can not exhibit a clear shape. Therefore, we only considered clump complexes with more than 35 clumps. Typically, the power-law part of the clump mass function for each clump complex is divided into two segments (flat+steep), as shown in Fig.6.5(a). We fitted the two segments with different mass ranges separately, and obtained the slopes β_1 and β_2, respectively. As marked in Fig.6.5(a), the two mass ranges correspond to three mass endpoints, denoted as M_1, M_2 and M_3.

The mass of the associated molecular clouds ($M_{\rm cloud,as}$) in the catalog of Miville-Deschênes et al., 2017 has a clear correlation with M_2 in Fig.6.6(a). However, the molecular cloud associated with a clump complex may not fully encompass the entire complex. In other words, it may not accurately reflect the gaseous environment in which the clump complex resides. Therefore, based on the spatial distribution and coordinates of the clumps, we computed the convex hull of each clump complex, which is the smallest convex shape that encloses all given clumps, as shown in Fig.6.5(c). Then we used the same convex shape of each clump complex to extract the average spectrum according to the CO cube and fixed the velocity range of the CO emission covering the clump complex to calculate the total molecular

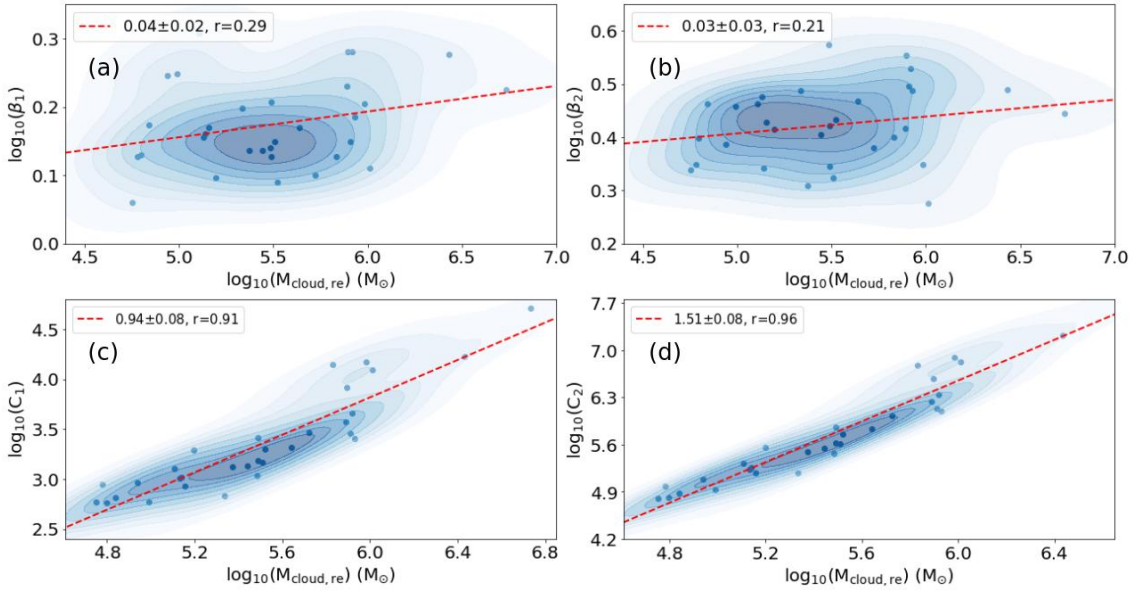


Figure 6.7: Correlations between the slopes (β_1 and β_2) and the normalization constants (C_1 and C_2) of the clump mass function and the cloud mass. r is the Pearson correlation coefficient.

gas mass ($M_{\rm cloud,re}$), following the method presented in Miville-Deschênes et al., 2017. Comparing Fig.6.6 (a) and (b), $M_{\rm cloud,re}$ exhibits a slightly better correlation with M_2 than $M_{\rm cloud,as}$. Moreover, M_1, M_3, $M_{\rm clump,min}$ and $M_{\rm clump,max}$ also have clear correlations with $M_{\rm cloud,re}$. $M_{\rm clump,min}$ and $M_{\rm clump,max}$ are the smallest and largest clump masses in the clump complex. Both β_1 and β_2 have weak correlations with $M_{\rm cloud,re}$. For all clouds, the median values of β_1 and β_2 are 1.45 and 2.64.

6.3.2 Analytical model of the clump mass function

As discussed in Urquhart et al. (2022), the mass distribution of the ATLASGAL clumps at different evolutionary stages are quite comparable. Thus, we assume that the clump mass remains constant during the evolution process. Since all clumps within a single molecular cloud follow the two-segment distribution presented above, an analytical expression of the clump mass function for a single molecular cloud is,

$$\xi_{\rm clump}(M_{\rm clump}, M_{\rm cloud}) = \begin{cases} C_1 * M_{\rm clump}^{-\beta_1}, \\ M_{\rm clump,min} \leq M_{\rm clump} \leq M_2; \\ C_2 * M_{\rm clump}^{-\beta_2}, \\ M_2 \leq M_{\rm clump} \leq M_{\rm clump,max}, \end{cases} \tag{6.1}$$

where C_1 and C_2 are the normalization constants. According to the fitted correlations in Fig.6.6 and Fig.6.7,

$$M_{\rm clump,min} = 10^{-0.02\pm0.59} * M_{\rm cloud}^{0.29\pm0.11}, \tag{6.2}$$

$$M_{\rm clump,max} = 10^{0.35\pm0.66} * M_{\rm cloud}^{0.50\pm0.12}, \tag{6.3}$$

$$M_2 = 10^{-0.21\pm0.43} * M_{\rm cloud}^{0.46\pm0.08}, \tag{6.4}$$

$$\beta_1 = 10^{-0.03\pm0.13} * M_{\rm cloud}^{0.04\pm0.02}, \tag{6.5}$$

$$\beta_2 = 10^{0.25\pm0.15} * M_{\rm cloud}^{0.03\pm0.03}. \tag{6.6}$$

The total mass of the clumps within a cloud is

$$M_{\rm clump,tot} = \int_{M_{\rm clump,min}}^{M_{\rm clump,max}} M_{\rm clump}\, \xi_{\rm clump}(M_{\rm clump}, M_{\rm cloud})\, {\rm d}M_{\rm clump}. \tag{6.7}$$

The continuity condition of the clump mass function is

$$C_1 * M_2^{-\beta_1} = C_2 * M_2^{-\beta_2}. \tag{6.8}$$

Solving these equations, we derived a correlation between $M_{\rm clump,tot}$ and $M_{\rm cloud}$. After substituting the values of $M_{\rm clump,tot}$ and $M_{\rm cloud,re}$ in Table.6.3, we obtained the values of C_1 and C_2 for each cloud. In Fig.6.7, C_1 and C_2 have strong correlations with $M_{\rm cloud,re}$,

$$C_1 = 10^{-1.81\pm0.44} * M_{\rm cloud}^{0.94\pm0.08}, \tag{6.9}$$

$$C_2 = 10^{-2.51\pm0.44} * M_{\rm cloud}^{1.51\pm0.08}. \tag{6.10}$$

Fig.6.8 shows a strong correlation between $M_{\rm clump,tot}$ and $M_{\rm cloud,re}$. The linear fitting gives

$$M_{\rm clump,tot} = 10^{0.7\pm0.41} * M_{\rm cloud}^{0.65\pm0.07}. \tag{6.11}$$

To test the validity of the model presented above, we compared the theoretical and observed $M_{\rm clump,tot} - M_{\rm cloud,re}$ relations, and found that the model can roughly follow the observation, despite the significant uncertainty in the model parameters. Alternatively, in addition to the analytical formalism, we provided the specific parameters of the clump mass function for each molecular cloud in Table.6.3, allowing for the selection of an appropriate set of parameters based on the cloud mass.

6.3.3 Spatial distribution of clumps

Next, we further investigated the spatial distribution of clump complexes within molecular clouds. We first utilized the Q parameter to examine the spatial structure of the cluster complex. The Q parameter defined in Cartwright & Whitworth (2004) is used to quantify the spatial distribution of stars in a cluster. The Q value quantifies the degree of subclustering and the overall radial density gradient within a specific region. Q $\geq$ 0.8 indicates a centrally concentrated spatial distribution, while Q < 0.8 suggests the presence of subclustering, with smaller values representing a higher degree of subclustering.

By analogy between "clumps and clump complexes" with "stars and star clusters", the above analysis can also be applied to study the spatial structure of clump complexes. As shown in Table.6.3, almost all clump complexes have the Q parameters less than 0.8, with an average value $\approx$0.64, which implies the widespread presence of substructures within clump complexes.

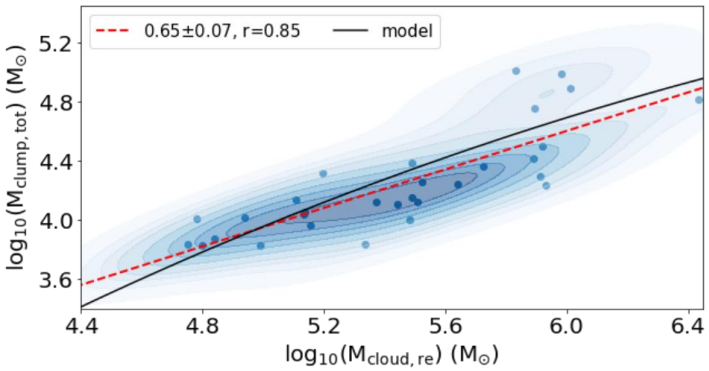


Figure 6.8: Correlation between the total mass of the clumps within a cloud ($M_{\rm clump,tot}$) and the cloud mass. r is the Pearson correlation coefficient. Red dashed line and black line represent the linear fitting and the model, respectively.

6.3.4 Mass segregation

We further investigated the mass segregation of clumps within the clump complex. Allison et al. (2009) proposed the mass segregation ratio (Λ) as a metric for detecting and quantifying mass segregation in star clusters. All clump complexes can be divided into three categories: (1) Clear mass segregation (type1). For almost all 35 most massive members, $\Lambda > 1$; (2) Vague (type2). For most of the 35 most massive members, Λ is around 1; (3) No clear mass segregation (type3). For a significant proportion of the 35 most massive members, $\Lambda < 1$.

Clump complexes with clear mass segregation are associated with more massive molecular clouds. As verified in J. W. Zhou et al. (2024a), more massive molecular clouds are more evolved. Therefore, as a molecular cloud evolves, clumps within it gradually become segregated. A molecular cloud is a gravitational network formed by multi-scale hub-filament structures, with clumps acting as the nodes of this network, see Sec.2.6 for more details. The emergence of segregation in this gravitational network over time may be a natural process under gravity. As presented above, more massive molecular clouds can form more massive clumps, indicating that the mass growth of molecular clouds and their internal clumps occurs simultaneously, which aligns with the picture of hierarchical, multi-scale hub-filament structures and gas inflows from large to small scales (Motte et al., 2018b; Vázquez-Semadeni et al., 2019; Kumar et al., 2020; Henshaw et al., 2020; J.-W. Zhou et al., 2022; J. W. Zhou et al., 2023; J. W. Zhou et al., 2024f; J.-W. Zhou & T. Davis, 2024). In the case of mass segregation, more massive clumps are located closer to the center of the clump complex, which may be due to gravitational mass settling or the fact that centrally located clumps can more efficiently accrete gas and thus accumulate mass more effectively. The latter is similar to a clump-scale "competitive accretion" scenario (Bonnell et al., 1997; Bonnell et al., 2001), which originally proposed to explain the formation of massive stars in the clump (core-scale).

6.4 The ICIMF theory

With high-resolution observations across various wavelengths, we are now able to break down star formation into individual molecular clouds in nearby galaxies (Leroy et al., 2021b; Leroy et al., 2021a; J. C. Lee et al., 2022; Emsellem et al., 2022; J. C. Lee et al., 2023). High-resolution CO imaging data from ALMA observations enable systematic studies of molecular clouds across nearby galaxies. The JWST and HST observations can reveal embedded clusters, young star clusters, and stellar associations

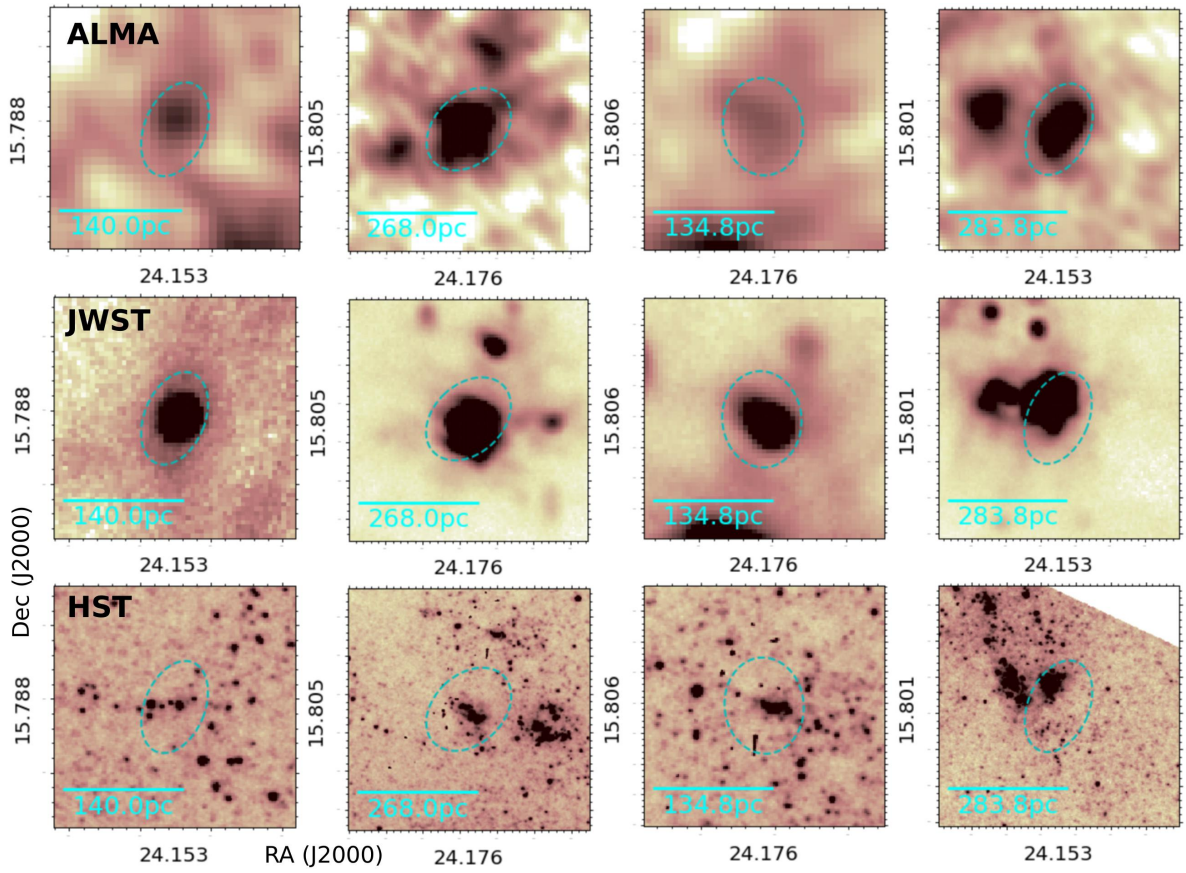


Figure 6.9: Some examples from J. W. Zhou et al. (2024a) are used to demonstrate the spatial distribution of molecular clouds in ALMA observation (the first row) and stellar populations in JWST (the second row) and HST (the third row) observations. Dashed cyan ellipses represent the effective ellipses of the molecular clouds.

within molecular clouds (Gouliermis et al., 2017; Whitmore et al., 2021; Thilker et al., 2022; K. L. Larson et al., 2023; Levy et al., 2024). The combination of these high-resolution observations provides an unprecedented look at the connection between individual clouds and the stellar populations (Turner et al., 2022). Some examples of multi-band observations are shown in Fig.6.9. In this work, we first look for correlations between the physical parameters of molecular clouds and their internal stellar populations, and then use the integrated cloud-wide initial mass function (ICIMF) theory to interpret the observed correlations.

6.4.1 Correlations between molecular clouds and their internal stellar populations

For the five galaxies (NGC 1433, NGC 1566, NGC 1559, NGC 3351 and NGC 3627) considered in this work, we conducted a direct identification of hierarchical structures based on the 2D intensity maps. For the identified molecular structures, we extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. Large-scale velocity gradients from the galaxy's rotation contribute to the non-thermal velocity dispersion. To address this, before extracting the average spectra, we removed the bulk motion due to galaxy rotation by creating gas dynamical models using the KinMS package (T. A. Davis et al., 2013).

Since the beam sizes of all five galaxies are more than 64 pc, we only focus on the associations identified on the scale level of 64 pc. If the distances between the center coordinates of the associations and a molecular cloud are less than the effective radius of the cloud, we distribute these associations to

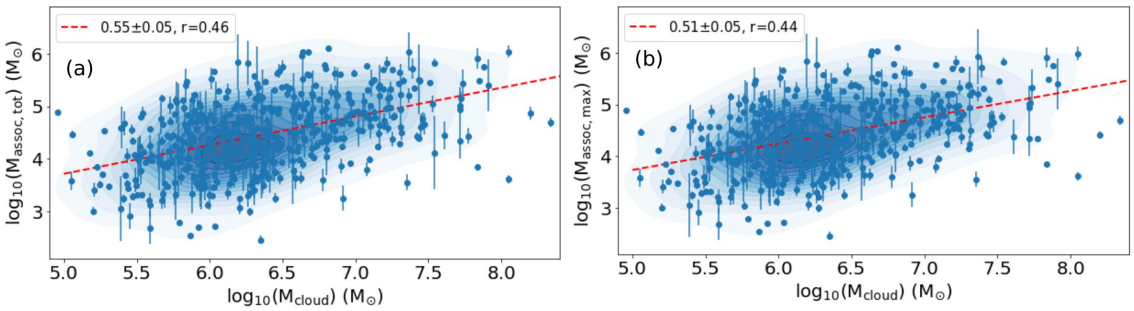


Figure 6.10: (a) The correlation between the mass of the cloud ($M_{\rm cloud}$) and the total mass of the associations in the cloud ($M_{\rm assoc,tot}$), i.e. the $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relation; (b) The correlation between $M_{\rm cloud}$ and the mass of the most massive association ($M_{\rm assoc,max}$) in the cloud, i.e. the $M_{\rm assoc,max}$-$M_{\rm cloud}$ relation. r is the Pearson correlation coefficient.

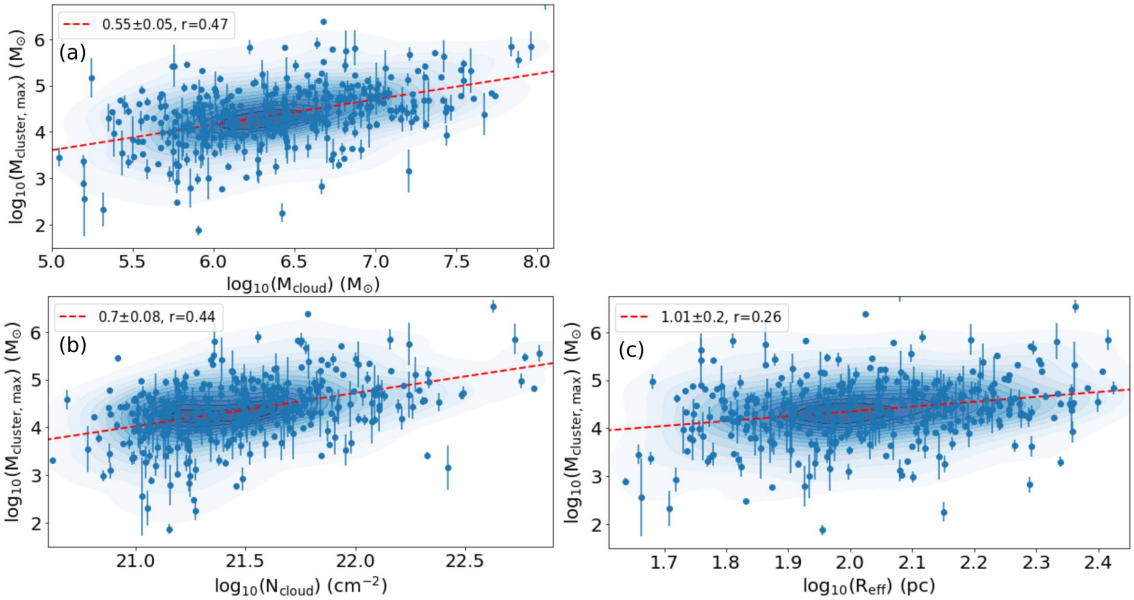


Figure 6.11: (a) The correlation between the mass of the cloud and the mass of the most massive cluster ($M_{\rm cluster,max}$) in the cloud, i.e. the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation; (b) The correlation between the column density of the cloud and $M_{\rm cluster,max}$, i.e. the $M_{\rm cluster,max}$-$N_{\rm cloud}$ relation; (c) The correlation between the effective radius of the cloud and $M_{\rm cluster,max}$. r is the Pearson correlation coefficient.

the cloud. If the associations from the V-band selection and the NUV-band selection overlap, we select the one with a larger area which includes more stars. Actually, the association from the V-band selection always has a larger area, as described in K. L. Larson et al. (2023). Finally, we added the masses of all independent associations in each cloud. Combining the data of all five galaxies together, Fig.6.10(a) shows a clear correlation between the mass of the cloud and the total mass of the associations in the cloud, i.e. the $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relation. For each cloud, we also singled out the most massive association with the mass $M_{\rm assoc,max}$. As displayed in Fig.6.10(b), the $M_{\rm assoc,max}$-$M_{\rm cloud}$ and $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relations are very similar, and this is because one cloud typically corresponds to one association identified at the 64 pc level.

Using the same criterion, if the distances between the center coordinates of the clusters and a molecular cloud are less than the effective radius of the cloud, we distribute these clusters to the cloud. Generally, one cloud will contain multiple star clusters. We first focus on the most massive one. Fig.6.11(a) displays

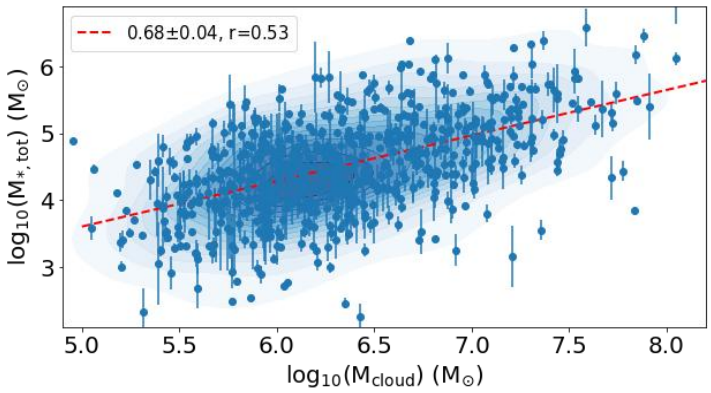


Figure 6.12: The correlation between the mass of the cloud and the mass of the total stellar populations ($M_{*,\mathrm{tot}}$) in the cloud, i.e. the $M_{*,\mathrm{tot}}$-M_{cloud} relation. r is the Pearson correlation coefficient.

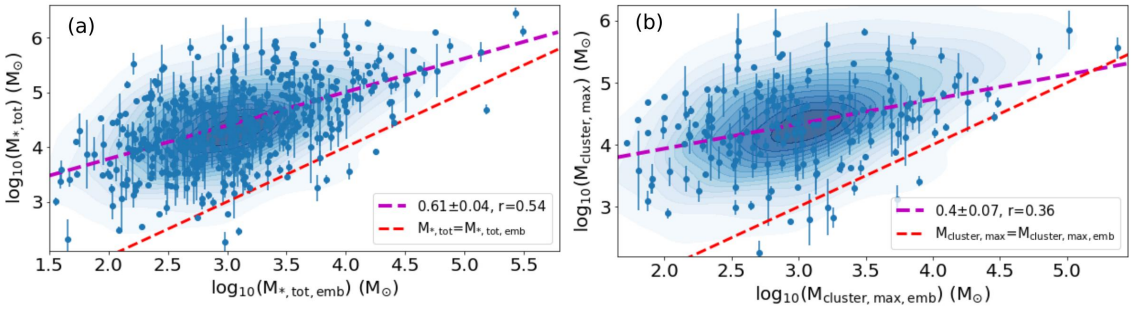


Figure 6.13: (a) Comparison between $M_{*,\mathrm{tot,emb}}$ and $M_{*,\mathrm{tot}}$; (b) Comparison between $M_{\mathrm{cluster,max,emb}}$ and $M_{\mathrm{cluster,max}}$. r is the Pearson correlation coefficient.

a clear correlation between the mass of the cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\mathrm{cluster,max}}$-M_{cloud} relation. Moreover, there is also a clear correlation between the column density of the cloud (N_{cloud}) and $M_{\mathrm{cluster,max}}$, i.e. the $M_{\mathrm{cluster,max}}$-N_{cloud} relation. However, there is only a weak correlation between the size of the cloud (R_{eff}) and $M_{\mathrm{cluster,max}}$.

A molecular cloud may contain only star clusters (case1), only stellar associations (case2), or both star clusters and stellar associations (case3). For case3, the clusters may be inside the associations or independent of them. To estimate the total mass of the stellar populations within the cloud, we need to sum the masses of both the independent clusters and the associations. To avoid redundant mass calculations, the clusters that fall within the associations should be excluded. Similarly, if the distances between the center coordinates of the clusters and an association are less than the effective radius of the association, we distribute these clusters to the association and do not count them. Fig.6.12 displays a clear correlation between the mass of the cloud and the mass of the total stellar populations in the cloud, i.e. the $M_{*,\mathrm{tot}}$-M_{cloud} relation.

Since star clusters and associations being associated with molecular clouds inherently indicates their youth, we did not impose age restrictions on the samples in the above analysis, as age estimates are subject to significant uncertainties.

6.4.2 Embedded stellar populations

The embedded stellar populations in molecular clouds may not be observed in the PHANGS-HST survey, which is more sensitive to the exposed stellar populations. As done in J. W. Zhou et al. (2024a), the JWST 21 μm emission was used to estimate the total mass of the embedded stellar populations in a cloud.

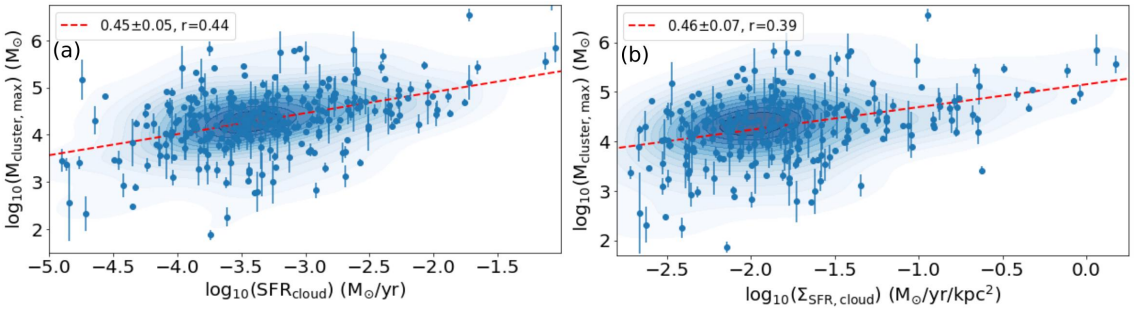


Figure 6.14: (a) The correlation between the star formation rate (SFR) of the cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\rm cluster,max}$-$\rm SFR_{cloud}$ relation; (b) The correlation between the SFR surface density of the cloud and $M_{\rm cluster,max}$, i.e. the $M_{\rm cluster,max}$-$\Sigma_{\rm SFR,cloud}$ relation. r is the Pearson correlation coefficient.

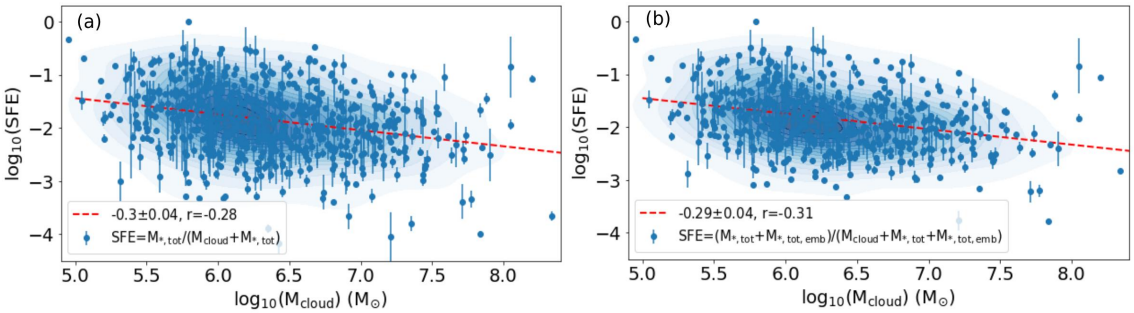


Figure 6.15: The correlation between the star formation efficiency (SFE) of the cloud and the mass of the cloud. (a) Embedded stellar populations were not considered in the calculation of SFE; (b) Embedded stellar populations were considered. r is the Pearson correlation coefficient.

Using the prescriptions described in Leroy et al. (2021b), the JWST 21 μm data can be used to calculate the local star formation rate (SFR) surface density. For NGC 628, J. Kim et al. (2023) defined the duration of the embedded phase of star formation as the time during which CO and 21 μm emissions are found to be overlapping, i.e. $t_{\rm fb,21\mu m}$, and they obtained $t_{\rm fb,21\mu m} \approx 5.1$ Myr. We assume this timescale is also applicable to the five galaxies in this work. Following the same method as in J. W. Zhou et al. (2024a), the spatial separation between 21 μm and CO emission peaks is used to estimate the typical age ($t_{\rm age}$) of the embedded stellar populations in the cloud. Then, the total mass of the embedded stellar populations is

$$M_{*,\rm tot,emb} = \Sigma_{\rm SFR} \times A \times t_{\rm age}, \tag{6.12}$$

where A is the cloud's area. Here, we consider all 21 μm emission enclosed in the cloud. The same method is also applied to calculate the mass of bright 21 μm sources ($M_{\rm cluster,emb}$), which represent the embedded clusters. As shown in Fig.6.13, both $M_{*,\rm tot,emb}$ and $M_{\rm cluster,max,emb}$ are systematically smaller than the corresponding $M_{*,\rm tot}$ and $M_{\rm cluster,max}$ observed in the PHANGS-HST survey. This suggests a short duration of the embedded phase. J. Kim et al. (2021) found that the embedded phase of star formation in six nearby galaxies lasts for 2−7 Myr and constitutes 17−47 percent of the cloud lifetime. Fig.6.14(a) and (b) show clear correlations between the SFR and $\Sigma_{\rm SFR}$ of a cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\rm cluster,max}$-$\rm SFR_{cloud}$ relation and the $M_{\rm cluster,max}$-$\Sigma_{\rm SFR,cloud}$ relation. The star formation efficiency (SFE) of each cloud is defined as

$$\rm SFE \approx \frac{\mathit{M}_{*,tot}}{\mathit{M}_{cloud} + \mathit{M}_{*,tot}}. \tag{6.13}$$

Table 6.4: All the correlations are fitted with simple linear functions in the form of $y = k * x + b$. r is the Pearson correlation coefficient.

y	x	k	Δk	b	Δb	r
$\log_{10}(M_{\rm assoc,tot}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.546	0.047	0.983	0.304	0.464
$\log_{10}(M_{\rm assoc,max}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.510	0.047	1.177	0.302	0.442
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.548	0.054	0.864	0.350	0.469
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(N_{\rm cloud}/{\rm cm}^{-2})$	0.695	0.075	-10.573	1.632	0.435
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(R_{\rm eff}/{\rm pc})$	1.009	0.201	2.329	0.410	0.255
$\log_{10}(M_{*,\rm tot}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.681	0.042	0.194	0.269	0.534
$\log_{10}(M_{*,\rm tot}/M_\odot)$	$\log_{10}(M_{*,\rm tot,emb}/M_\odot)$	0.612	0.041	2.558	0.131	0.544
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(M_{\rm cluster,max,emb}/M_\odot)$	0.397	0.074	3.141	0.230	0.362
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}[{\rm SFR_{cloud}}/(M_\odot {\rm yr}^{-1})]$	0.446	0.054	5.796	0.183	0.442
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}[\Sigma_{\rm SFR,cloud}/(M_\odot {\rm yr}^{-1}{\rm kpc}^{-2})]$	0.459	0.065	5.153	0.124	0.388
$\log_{10}$(SFE)	$\log_{10}(M_{\rm cloud}/M_\odot)$	-0.293	0.039	0.013	0.250	-0.306

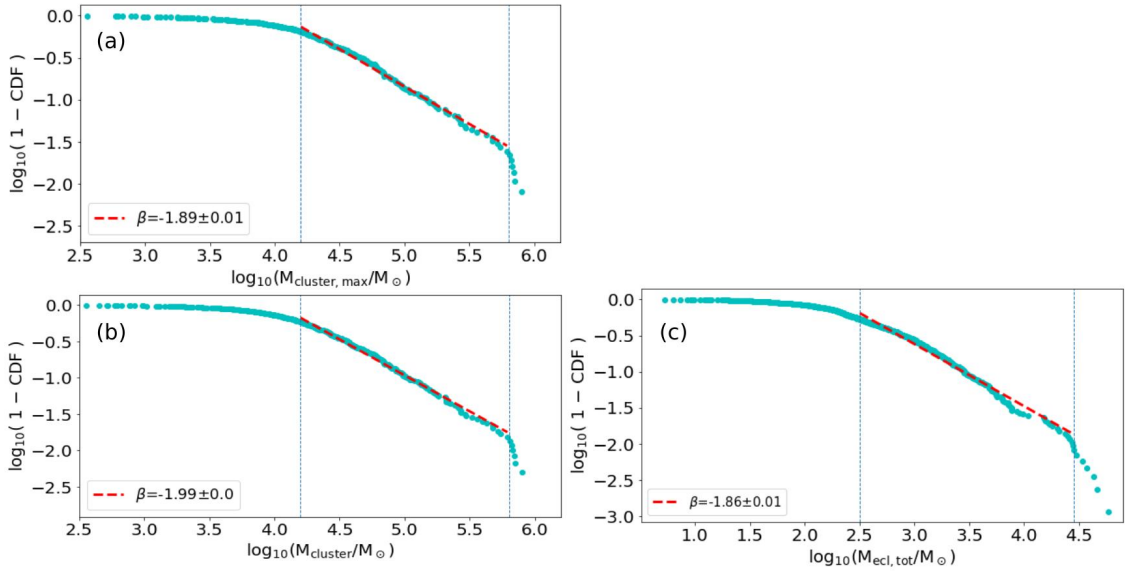


Figure 6.16: The complementary cumulative distribution function (CCDF) plot and the linear fitting of the cluster mass function. The vertical dashed lines mark the mass range of the linear fitting. (a) The mass function of the most massive clusters associated with molecular clouds in Fig.6.11(a); (b) The mass function of all clusters associated with molecular clouds; (c) The mass function of $M_{\rm ecl,tot}$, which represents the total mass of all embedded clusters in a cloud in the Milky Way.

The median value of SFEs in Fig.6.15 is ≈1.4%, which is comparable with the values listed in J. Kim et al. (2022) and Chevance et al. (2023).

6.4.3 The ICIMF theory

Table.6.4 summarizes all the fitted correlations between the physical parameters of molecular clouds and their internal stellar populations. In this section, we will verify that the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation can be understood theoretically on the basis of the integrated cloud-wide IMF (ICIMF) theory. In the ICIMF theory, the IMFs assembled within a molecular cloud combine to form the overall IMF of the cloud. Strictly speaking, this process involves first generating a series of embedded star clusters from the embedded cluster mass function (ECMF), then producing a population of stars for each embedded

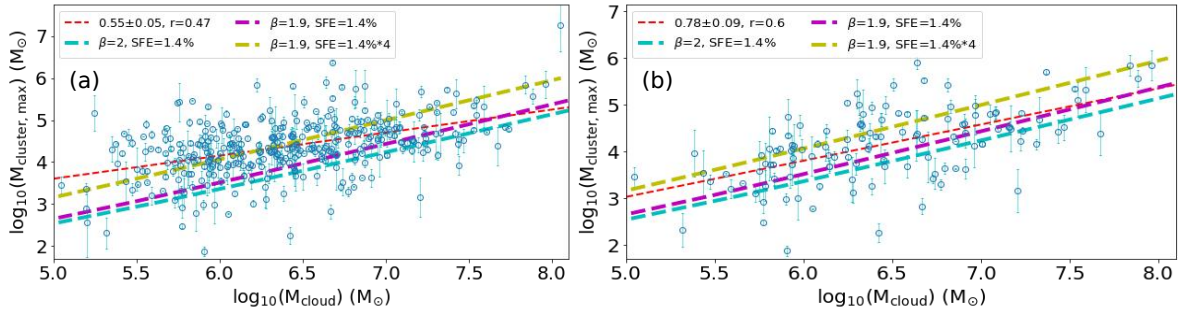


Figure 6.17: The $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation fitted by the model based on the ICIMF theory. r is the Pearson correlation coefficient. β is the slope of the cluster mass function. SFE is the star formation efficiency. (a) All clusters selected in Sec.6.4.1; (b) Only clusters younger than 10 Myr.

cluster from the IMF, and finally summing the stars from all clusters to construct the cloud-wide IMF. Thus, the ICIMF theory consists of two layers (hierarchical). In this work, we focus on the embedded cluster mass function in molecular clouds, which represents one layer of the ICIMF theory.

We start by assuming all the stars in a cloud form in a population of embedded clusters with the mass function

$$\xi_{\rm ecl}(M_{\rm ecl}) = k_{\rm ecl} M_{\rm ecl}^{-\beta} (M_{\rm ecl,min} \leqslant M_{\rm ecl} \leqslant M_{\rm ecl,max}), \tag{6.14}$$

where $k_{\rm ecl}$ is a normalization constant and $M_{\rm ecl}$ is the stellar mass of the embedded cluster. The mass of the smallest stellar group observed in Kuhn et al., 2015 is $\approx 5 M_\odot$. Thus, we take $M_{\rm ecl,min} = 5 M_\odot$, which is also consistent with the value adopted in Yan et al., 2017. The embedded cluster population mass conservation gives

$$M_{*,\rm tot} = \int_{M_{\rm ecl,min}}^{M_{\rm ecl,max}} M_{\rm ecl}\, \xi_{\rm ecl}(M_{\rm ecl})\, {\rm d}M_{\rm ecl}, \tag{6.15}$$

where $M_{*,\rm tot}$ is the total stellar mass formed in a cloud. $M_{*,\rm tot}$ and the cloud mass are linked by the star formation efficiency, i.e.

$$M_{*,\rm tot} \approx {\rm SFE} * M_{\rm cloud}. \tag{6.16}$$

To determine the stellar masses within a stellar system based on a given IMF, one needs to apply a specific sampling method. Here, we adopt the optimal sampling. The details of this method and its validity are available in Kroupa et al., 2013; Schulz et al., 2015; Kroupa, 2015; Yan et al., 2017; Yan et al., 2023. The optimal sampling normalization condition is

$$1 = \int_{M_{\rm ecl,max}}^{M_{\rm ecl,up}} \xi_{\rm ecl}(M_{\rm ecl})\, {\rm d}M_{\rm ecl}. \tag{6.17}$$

We take the upper integration limit (the hypothetical physical upper bound of embedded cluster mass) as $M_{\rm ecl,up} \approx 10^9$ M$_\odot$, following Yan et al., 2017. This normalization form is based on the assumption that there exists exactly one object within the mass range from $M_{\rm ecl,max}$ to $M_{\rm ecl,up}$ (Schulz et al., 2015). By solving these equations, we can build the relations between $M_{\rm ecl,max}$ ($\approx M_{\rm cluster,max}$), $M_{*,\rm tot}$ and $M_{\rm cloud}$.

For the slope β of the cluster mass function, generally, the value is ≈2 (C. J. Lada & E. A. Lada, 2003; Krumholz et al., 2019; Mok et al., 2020; Wainer et al., 2022). For massive and compact clusters, such as Super Star Clusters (SSCs), the slope can be ≈1.8 (Cuevas-Otahola et al., 2023). Using the most massive clusters associated with molecular clouds presented in Fig.6.11(a), we fitted the cluster mass function in Fig.6.16(a). It indeed provides a slightly flatter slope, with $\beta \approx 1.9$. For all clusters associated with molecular clouds, the slope is ≈2. However, as shown in Fig.6.16, the most massive

clusters and all clusters have comparable mass distributions, rather than the most massive clusters having a systematically larger mass. Since we will use the model to fit the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation, taking the slope fitted in Fig.6.16(a) is more appropriate. As shown in Fig.6.17, when fixing SFE=1.4%, a smaller β can provide a better fit. To some extent, a smaller β and a larger SFE are equivalent. For low-mass molecular clouds, larger SFE will yield a better fit. The fact that low-mass molecular clouds have higher SFEs can be observed in Fig.6.15. When the age of the cluster is less than 10 Myr, the model also provides a better fit due to the stronger connection between young and embedded star clusters.

On the galaxy scale, Weidner et al. (2013b) suggests the embedded cluster mass function to be flatter (top heavy) for galaxies with a high SFR, with a dependence that is given by

$$\beta = -0.106 \log_{10} \mathrm{SFR} + 2, \tag{6.18}$$

where the SFR is in unit of $[\mathrm{M}_\odot/\mathrm{yr}]$. In Weidner et al. (2004) and Weidner & Kroupa (2005), the observed most massive embedded star cluster in a galaxy is determined through the total galaxy-wide SFR

$$\frac{M_{\rm ecl,max}}{M_\odot} \propto \left(\frac{\mathrm{SFR}}{\mathrm{M}_\odot/\mathrm{yr}}\right)^{0.75}. \tag{6.19}$$

Fig.6.14(a) shows a clear correlation between $M_{\rm cluster,max}$ and $\mathrm{SFR}_{\rm cloud}$, although the slope is shallower, ≈0.5. Given that the estimation of the SFR in this work is different from that in Weidner et al. (2013b), the comparison is not straightforward. But similar to equation.6.18, a β-SFR relation may also exist on the cloud scale. Surveying adequately complete cluster populations in molecular clouds of nearby galaxies to fit the cloud-scale β remains challenging, but it is feasible for molecular clouds in the Milky Way. In order to better fit the observational data in Fig.6.17, we need to understand how β varies with the properties of molecular clouds.

Both SFE and β vary with the physical properties of molecular clouds. Therefore, using the same parameter settings for all molecular clouds can only result in a rough fit. Moreover, in this work, we use an unified CO−to−H_2 conversion factor (i.e. $\alpha_{\rm CO}^{2-1} \approx 6.7\ \mathrm{M}_\odot\mathrm{pc}^{-2}(\mathrm{K\ km\ s}^{-1})^{-1}$) for all five galaxies. However, $\alpha_{\rm CO}^{2-1}$ may vary between galaxies, and even within different regions of the same galaxy. Therefore, it is important to correct $\alpha_{\rm CO}^{2-1}$ for each molecular cloud to accurately estimate the mass of the cloud. This will be left for future work.

6.4.4 Cluster complex rather than single cluster

The typical number of clusters associated with the clouds is only ≈2. For open clusters in the Milky Way, after comparing the separation of open clusters and the typical size of molecular clouds in J.-w. Zhou et al. (2024a), we found that most molecular clouds may only form one open cluster. The pc-scale clumps within molecular clouds in the Milky Way are the basic units of star cluster formation. One molecular cloud can produce multiple embedded clusters in the Milky Way, which can coalesce into more massive clusters. The N-body simulations and the Gaia DR3 observations suggest that the post-gas expulsion coalescence of embedded clusters may be a dominant mechanism in the formation of open clusters in the Milky Way, see Chapter.5 for more details. As shown in J.-w. Zhou et al. (2024b), the typical separation of embedded clusters in a cloud can be ≈1 pc. Due to the limited observational resolution, the observed clusters in nearby galaxies may be cluster complexes rather than single clusters. To further explore this possibility, we need to refer to molecular clouds within the Milky Way, as molecular clouds within the Milky Way can have their internal structures resolved. We matched the ATLASGAL clump complex catalog of Urquhart et al. (2022) and the molecular cloud catalog of Miville-Deschênes et al. (2017) by minimizing the differences in their central coordinates and systemic velocities. In Fig.6.18(a), there is a

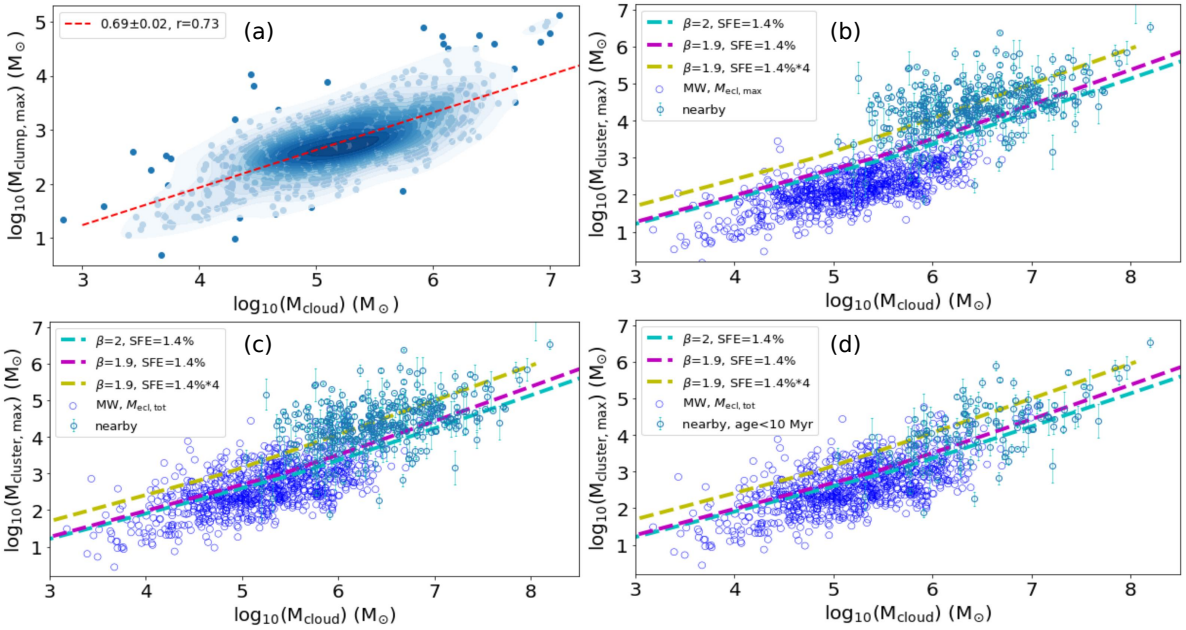


Figure 6.18: (a) Correlation between the mass of the most massive clump ($M_{\rm clump,max}$) in a cloud and the mass of the cloud in the Milky Way; (b) Comparison of the $M_{\rm ecl,max} - M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max} - M_{\rm cloud}$ relation in nearby galaxies. $M_{\rm ecl,max}$ is the mass of the most massive embedded cluster formed in the most massive clump. (c) Comparison of the $M_{\rm ecl,tot} - M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max} - M_{\rm cloud}$ relation in nearby galaxies. $M_{\rm ecl,tot}$ is the total mass of all embedded clusters in a cloud; (d) Same as panel (c), but $M_{\rm cluster,max}$ represents the clusters younger than 10 Myr. r is the Pearson correlation coefficient. β is the slope of the cluster mass function.

strong correlation between the mass of the most massive clump ($M_{\rm clump,max}$) in a cloud and the mass of the cloud. In J. W. Zhou et al. (2024b), we derived a relation between the clump mass and the mass of the embedded cluster formed in the clump. Naturally, one would think that the most massive embedded cluster ($M_{\rm ecl,max}$) formed in the most massive clump will be the most massive exposed cluster in the cloud. However, when we directly compare the $M_{\rm ecl,max} - M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max} - M_{\rm cloud}$ relation in nearby galaxies in Fig.6.18(b), the former is systematically lower than the latter. Given that the coalescence of embedded clusters could be a key mechanism in the formation of open clusters in the Milky Way, we need to add the mass of neighboring embedded clusters in a cloud to approximate the mass of the final coalesced product ($M_{\rm ecl,tot}$). As a simplification, we sum the masses of all embedded clusters within a cloud. In Fig.6.18(c), the $M_{\rm ecl,tot} - M_{\rm cloud}$ and $M_{\rm cluster,max} - M_{\rm cloud}$ relations can link together. The mass function of $M_{\rm ecl,tot}$ is also similar to that of $M_{\rm cluster,max}$, as shown in Fig.6.16(c). In actual observations, as long as the objects are enclosed within the resolution limit, they are thought to be a single cluster. For $M_{\rm cluster,max}$ in Sec.6.4.1, they could be the products of coalescence or currently in the process of coalescing. The substructures within them may be inferred based on observations in the Milky Way. We should be cautious with the very massive and young star clusters in nearby galaxies, as they may not actually be real clusters. This fact should be taken into account in the study of stellar population synthesis as well as star cluster formation and evolution in nearby galaxies.

Although the overall star formation rates between galaxies will be different, according to the ICIMF theory or the clear correlations between $M_{\rm cloud}$ and $M_{\rm cluster,max}$ (or $M_{\rm ecl,max}$ and $M_{\rm clump,max}$), star formation in the galaxy is localized to molecular cloud. The connection in Fig.6.18(d) means that star formation in the Milky Way and nearby galaxies follow similar patterns on the cloud scale. Currently, only molecular clouds within the Milky Way can be studied in detail. Combined with the ICIMF theory,

we may infer the internal details of molecular clouds in nearby galaxies using the template from the Milky Way.

6.5 Summary

6.5.1 From CMF to IMF

Many studies have revealed that the core mass function (CMF) in high-mass star-forming regions is top-heavy. This work begins with the canonical initial mass function and inversely synthesizes the observed CMFs of high-mass star formation regions, accounting for variations in multiplicity and mass conversion efficiency from core to star ($\epsilon_{\rm core}$). To reproduce the observed CMFs, cores of different masses must exhibit varying $\epsilon_{\rm core}$, with lower-mass cores requiring higher conversion efficiencies. In contrast, the multiplicity fraction does not significantly affect the synthesized CMFs. Accurately fitting the high-mass end of the CMF depends critically on whether a slope transition occurs between the low-mass and high-mass ends. If such a slope transition exists but is concealed by observational biases—leading to a combined fit with a shallower slope—this may result in an artificially top-heavy CMF.

6.5.2 Clump populations in molecular clouds

The physical properties of the 31 largest clump complexes and their associated molecular clouds in the Milky Way were investigated. The physical parameters of the clump mass function for a given clump complex show strong correlations with the mass of its associated molecular cloud. An analytical model was constructed for the clump mass function of a cloud, in which the model parameters depend solely on the cloud mass. For molecular clouds in nearby galaxies, this model enables the generation of internal clump populations based on cloud mass. Almost all clump complexes exhibit Q parameters below 0.8, with an average value of ≈0.64, indicating the widespread presence of substructures within these complexes. Clump complexes displaying clear mass segregation are associated with more massive, and typically more evolved, molecular clouds. This mass segregation may arise from gravitational mass settling or from the enhanced gas accretion efficiency of centrally located clumps, enabling them to accumulate mass more effectively.

6.5.3 The ICIMF theory

The molecular clouds in five galaxies (NGC 1433, NGC 1566, NGC 1559, NGC 3351, and NGC 3627) were identified based on the integrated intensity maps of CO (2−1) emission from ALMA observations. The catalogs for compact clusters and multi-scale associations across these galaxies were matched with the identified clouds. JWST 21 μm data was used to estimate the star formation rate (SFR) surface density of the clouds and calculate the masses of the embedded stellar populations in the clouds, revealing that the embedded stellar population mass is systematically smaller than that of the exposed stellar population. This suggests a short duration of the embedded phase. The mass of the most massive cluster ($M_{\rm cluster,max}$) in a cloud was found to be positively proportional to the mass ($M_{\rm cloud}$), the column density, the SFR, and the SFR surface density of the cloud.

The mass of the ionized gas ($M_{\rm I}$) in each cloud is negligible relative to $M_{\rm cloud}$. Based on the masses of the total stellar populations and their corresponding clouds, a typical value of the cloud-scale SFE, ≈1.4%, was obtained. There is a trend of decreasing SFE as the cloud mass increases. More massive clouds lead to larger stellar populations and more massive star clusters, resulting in stronger feedback. This feedback may suppress star formation in massive clouds, leading to a relatively lower SFE. This self-regulated star formation also supports the choice of optimal sampling.

An integrated cloud-wide IMF (ICIMF) model was constructed to fit the observed $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation. When fixing SFE=1.4%, there is only one free parameter β (the slope of the cluster mass function) in the solutions. For low-mass molecular clouds, a higher SFE yields a better fit. When the cluster's age is under 10 Myr, the model also provides a better fit due to the stronger connection between young and embedded star clusters. The ICIMF theory offers a quantitative framework for understanding the correlations between molecular clouds and their internal stellar populations.

For molecular clouds in the Milky Way, a strong correlation exists between the mass of the most massive clump ($M_{\rm clump,max}$) in a cloud and the mass of the cloud. For the most massive embedded cluster ($M_{\rm ecl,max}$) formed in the most massive clump, direct comparison of $M_{\rm ecl,max} - M_{\rm cloud}$ and $M_{\rm cluster,max} - M_{\rm cloud}$ relations reveals that the former is systematically lower than the latter. Given that the coalescence of embedded clusters could be a key mechanism in the formation of open clusters in the Milky Way, the mass of neighboring embedded clusters in a cloud should be added to approximate the mass of the final coalesced product ($M_{\rm ecl,tot}$). The $M_{\rm ecl,tot} - M_{\rm cloud}$ and $M_{\rm cluster,max} - M_{\rm cloud}$ relations can then be linked together. The mass function of $M_{\rm ecl,tot}$ is also similar to that of $M_{\rm cluster,max}$.

This connection suggests that star formation in the Milky Way and nearby galaxies follows similar patterns on the cloud scale. Research on molecular clouds within the Milky Way can serve as a template for understanding molecular clouds and their internal stellar populations in nearby galaxies. The most massive young star clusters in the clouds of nearby galaxies should be cluster complexes rather than single clusters. The substructures within them need to be inferred based on observations in the Milky Way.

It is worth noting that, under the ICIMF theory, the individual IMFs formed within a molecular cloud collectively shape the cloud's total IMF. This hierarchical process begins with the generation of embedded star clusters according to the embedded cluster mass function (ECMF). Each cluster then forms its own stellar population based on the IMF, and the combined stellar content of all clusters yields the final, cloud-wide IMF. Therefore, the ICIMF framework comprises two distinct layers. In this study, we concentrate on the ECMF within molecular clouds, which constitutes one of the two layers.

CHAPTER 7

Conclusion and outlook

In terms of spatial scales, this thesis constructs a framework for hierarchical hub-filament structures, spanning from dense cores (≈1000 AU) to clumps (≈1 pc), molecular clouds (≈10–100 pc), and extending to galaxy-cloud scales (≈1000 pc). Within this hierarchy, a dense core serves as the hub within a clump, a clump functions as the hub within a molecular cloud, and a molecular cloud acts as the hub within a galaxy. In terms of temporal evolution, this thesis outlines the evolutionary pathways from clumps to embedded clusters and ultimately to open clusters. In a hub-filament system, the hub is the primary site of star formation. A clump evolves into an embedded cluster, defining an initial mass function (IMF). A cloud consists of multiple clumps, leading to an integrated cloud-wide IMF (ICIMF). Likewise, a galaxy contains numerous clouds, giving rise to an integrated galaxy-wide IMF (IGIMF). These hierarchical hub-filament structures (star-forming regions) correspond to hierarchically integrated IMFs [1]. Extending this self-similar physical framework allows us to establish connections between star formation in the Milky Way and nearby galaxies, linking star-forming regions with stellar populations across various scales.

Tidal and gravitational fields are intrinsically interconnected. The hierarchical gravitational coupling of gas structures results in extensive tidal interactions, which can hinder gravitational collapse and regulate star formation. During the formation and evolution of star clusters, although feedback from cloud to core scales will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and changes in gravitational centers, but gravitational collapse is always ongoing.

The thesis could serve as a framework for incorporating physics and objects that span various timescales and spatial scales. Phenomena in the universe should be understood within a system rather than considered in isolation. According to this framework, we can establish connections between various elements based on causality, as well as temporal and spatial sequences, which can also be traced through multi-wavelength observations. However, this framework still requires further development and refinement, a goal that entails a substantial amount of work.

[1] Clumps, which are self-gravitating structures within molecular clouds, evolve more rapidly than the clouds themselves. By the time large-scale molecular clouds condense into dense, localized regions within galaxies, star formation has often already taken place within these regions. The molecular clouds we observe are typically composites, containing both exposed and embedded stellar populations, as well as clumps. Nearly all of those stellar populations—whether visible or still embedded—originate from clumps. Thus, a cloud consists of multiple clumps, leading to an integrated cloud-wide IMF (ICIMF). These clumps may either be currently existing or have existed in the past. Past clumps are now observed as exposed and embedded stellar populations. This highlights the nested spatial structure of star formation, despite asynchronous formation times. Likewise, a galaxy contains numerous clouds, giving rise to an integrated galaxy-wide IMF (IGIMF). As with clumps, clouds can be either present-day structures or ones that existed in the past.

Based on the picture of multi-scale, hierarchical, and self-similar hub-filament structures from dense core to clump to molecular cloud to galaxy scales, as an extension in the spatial scale, we can continue generalizing this hierarchical picture, such as a galaxy as the hub within a galaxy cluster. Actually, galaxy cluster and cosmic web also present hub-filament structures. The reason should be the formation and evolution of structures at different scales in the universe are dominant by gravity. We need a more quantitative description of this hierarchical network. In a hub-filament system, the hub is the main star-forming site. We may continue generalizing the integrated IMFs to larger scales (such as galaxy cluster), and study star formation histories according to the variations of the integrated IMFs. In the Milky Way and nearby galaxies, the mass of the most massive star within a star cluster depends on the cluster's total mass, while the mass of the most massive star cluster within a molecular cloud depends on the cloud's total mass. By analogy, the mass of the most massive galaxy within a galaxy cluster should also depend on the total mass of the galaxy cluster. It would certainly be interesting to test it through observations.

An ATLASGAL clump with an HII region is comprised of an embedded cluster and a gas envelope (and gas in the cluster). When we study the evolution of the embedded cluster, only considering stellar dynamics by N-body simulation is not enough, as done in my current works. Because the gas environment also makes effects. Cluster formation in molecular clouds occurs in several stages: the formation of clumps in diffuse gas, the evolution of these clumps, the formation of embedded clusters within them, and the subsequent evolution of the embedded clusters, accompanied by clump destruction due to feedback. ATLASGAL clumps at various stages encompass all of these physical processes (Urquhart et al., 2022). The relative contributions of magnetohydrodynamics (MHD) and N-body dynamics vary throughout these different stages. For the coalescence of embedded clusters in massive star-forming regions, the coalescence process should account for both stellar dynamics and interactions with the surrounding gas, as well as the ongoing star formation.

Understanding the detailed feedback processes within clumps is crucial for constructing a complete evolutionary track from clumps to embedded clusters to open clusters, which is also important for determining the initial conditions of N-body simulations. This is especially true for the early gas expulsion phase in the evolution of embedded clusters, as the later dynamical state of the clusters is highly sensitive to this process. Gas expulsion is driven by feedback, and the feedback strength varies with the mass of the embedded clusters, resulting in different timescales for gas expulsion. In simulations, we can model clumps of different masses and study how various feedback processes affect the final SFE within these clumps. If we can obtain results from the models that are similar to observations, we might be able to identify the specific feedback processes occurring within the clumps or determine which feedback mechanisms are dominant. Extensive observational data, particularly high-resolution ALMA observations, is available for ATLASGAL clumps and can be utilized for comparison and as input parameters in modeling efforts.

Given the significant progress in observational studies of massive star and star cluster formation, these observational results should be translated into the initial conditions for star cluster dynamical simulations. It is essential to consider more physics of star cluster formation, as the evolution of a cluster is intricately linked to, and heavily dependent on, its formation process. The post-gas expulsion coalescence of embedded clusters formed in the same parental molecular cloud may be a dominant mechanism for the formation of open clusters. A star cluster forms through the coalescence of several smaller clusters, as opposed to being born as a single entity, will have completely different cluster dynamics, and its evolutionary outcome and end products will also be different. The evolution of star clusters and the formation of compact objects, as well as the formation of massive stars and star clusters, are typically studied by two separate research communities. However, to fully grasp black hole formation within open clusters, it is necessary to integrate knowledge from both fields, in order to reconstruct a complete

physical picture.

For embedded clusters within clumps, many of which still contain accreting cores or a mix of PMS stars, protostars, and prestellar cores, high-resolution observations are essential for quantifying the accretion rates of protostars at different evolutionary stages. To explain the origin of the IMF, high-resolution observations and simulations are necessary for investigating core multiplicity and examining how the mass conversion efficiency from core to star changes with core mass. Furthermore, the canonical IMF is segmented, with low-mass and high-mass stars following distinct power laws. This discrepancy may stem from differing formation processes of low-mass and high-mass stars. Is there a significant difference in the mass function between low-mass and high-mass cores? To explain the origin of the $m_{\rm max} - M_{\rm ecl}$ relation, we need more observations to investigate the relationship between cores and clumps as precursors of stars and clusters. Do their formation and evolution occur in parallel?

Bibliography

Aarseth, S. J. (1974), *Dynamical evolution of simulated star clusters. I. Isolated models.*, A&A **35** 237 (cit. on p. 5).

Aarseth, S. J., Henon, M. & Wielen, R. (1974), *A Comparison of Numerical Methods for the Study of Star Cluster Dynamics*, A&A **37** 183 (cit. on pp. 94, 126).

Alfaro, E. J. & Román-Zúñiga, C. G. (2018), *Primordial mass and density segregation in a young molecular cloud*, MNRAS **478** L110, arXiv: 1804.06961 [astro-ph.GA] (cit. on p. 94).

Allison, R. J. et al. (2009), *Using the minimum spanning tree to trace mass segregation*, Monthly Notices of the Royal Astronomical Society **395** 1449, ISSN: 0035-8711, eprint: https://academic.oup.com/mnras/article-pdf/395/3/1449/18235817/mnras0395-1449.pdf, URL: https://doi.org/10.1111/j.1365-2966.2009.14508.x (cit. on p. 141).

Alves, J., Lombardi, M. & Lada, C. J. (2007), *The mass function of dense molecular cores and the origin of the IMF*, A&A **462** L17, arXiv: astro-ph/0612126 [astro-ph] (cit. on pp. 7, 8).

André, P. et al. (2014), "From Filamentary Networks to Dense Cores in Molecular Clouds: Toward a New Paradigm for Star Formation", *Protostars and Planets VI*, ed. by H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning 27, arXiv: 1312.6232 [astro-ph.GA] (cit. on pp. 7, 12).

André, P. J., Palmeirim, P. & Arzoumanian, D. (2022), *The typical width of Herschel filaments*, A&A **667**, L1 L1, arXiv: 2210.04736 [astro-ph.GA] (cit. on pp. 94, 126).

André, P. et al. (2016), *Characterizing filaments in regions of high-mass star formation: High-resolution submilimeter imaging of the massive star-forming complex NGC 6334 with ArTéMiS*, A&A **592**, A54 A54, arXiv: 1605.07434 [astro-ph.GA] (cit. on p. 13).

Arzoumanian, D., Shimajiri, Y., Inutsuka, S.-i., Inoue, T. & Tachihara, K. (2018), *Molecular filament formation and filament-cloud interaction: Hints from Nobeyama 45 m telescope observations*, PASJ **70**, 96 96, arXiv: 1807.08968 [astro-ph.GA] (cit. on p. 100).

Arzoumanian, D. et al. (2022), *Velocity structure of the 50 pc long NGC 6334 filamentary cloud. Hints of multiple compressions and their impact on the cloud properties*, A&A **660**, A56 A56, arXiv: 2201.04267 [astro-ph.GA] (cit. on p. 100).

El-Badry, K. et al. (2023a), *A red giant orbiting a black hole*, MNRAS **521** 4323, arXiv: 2302.07880 [astro-ph.SR] (cit. on p. 121).

El-Badry, K. et al. (2023b), *A Sun-like star orbiting a black hole*, MNRAS **518** 1057, arXiv: 2209.06833 [astro-ph.SR] (cit. on p. 121).

Ballesteros-Paredes, J. (2006), *Six myths on the virial theorem for interstellar clouds*, MNRAS **372** 443, arXiv: astro-ph/0606071 [astro-ph] (cit. on p. 57).

Ballesteros-Paredes, J., Gómez, G. C., Loinard, L., Torres, R. M. & Pichardo, B. (2009a), *Tidal forces as a regulator of star formation in Taurus*, MNRAS **395** L81, arXiv: 0903.0603 [astro-ph.GA] (cit. on pp. 29, 43).

Ballesteros-Paredes, J., Gómez, G. C., Pichardo, B. & Vázquez-Semadeni, E. (2009b), *On the gravitational content of molecular clouds and their cores*, MNRAS **393** 1563, arXiv: 0811.3234 [astro-ph] (cit. on pp. 29, 43).

Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F. & Zamora-Avilés, M. A. (2011a), *Gravity or turbulence? Velocity dispersion-size relation*, MNRAS **411** 65, arXiv: 1009.1583 [astro-ph.GA] (cit. on pp. 2, 4).

Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F. & Zamora-Avilés, M. A. (2011b), *Gravity or turbulence? Velocity dispersion-size relation*, MNRAS **411** 65, arXiv: 1009.1583 [astro-ph.GA] (cit. on pp. 4, 25, 72, 73).

Ballesteros-Paredes, J., Vázquez-Semadeni, E., Palau, A. & Klessen, R. S. (2018), *Gravity or turbulence? - IV. Collapsing cores in out-of-virial disguise*, MNRAS **479** 2112, arXiv: 1710.07384 [astro-ph.GA] (cit. on pp. 4, 25).

Ballesteros-Paredes, J. et al. (2020), *From Diffuse Gas to Dense Molecular Cloud Cores*, Space Sci. Rev. **216**, 76 76, arXiv: 2006.01326 [astro-ph.GA] (cit. on p. 4).

Banerjee, S. et al. (2020), *BSE versus StarTrack: Implementations of new wind, remnant-formation, and natal-kick schemes in NBODY7 and their astrophysical consequences*, A&A **639**, A41 A41, arXiv: 1902.07718 [astro-ph.SR] (cit. on p. 96).

Banerjee, S. & Kroupa, P. (2013), *Did the Infant R136 and NGC 3603 Clusters Undergo Residual Gas Expulsion?*, ApJ **764**, 29 29, arXiv: 1301.3491 [astro-ph.GA] (cit. on pp. 6, 94, 95, 119).

Banerjee, S. & Kroupa, P. (2014), *A Perfect Starburst Cluster made in One Go: The NGC 3603 Young Cluster*, ApJ **787**, 158 158, arXiv: 1403.4601 [astro-ph.GA] (cit. on pp. 6, 94, 119).

Banerjee, S. & Kroupa, P. (2015), *The formation of NGC 3603 young starburst cluster: 'prompt' hierarchical assembly or monolithic starburst?*, MNRAS **447** 728, arXiv: 1412.1473 [astro-ph.GA] (cit. on pp. 6, 94, 119).

Banerjee, S. & Kroupa, P. (2017), *How can young massive clusters reach their present-day sizes?*, A&A **597**, A28 A28, arXiv: 1510.04293 [astro-ph.GA] (cit. on pp. 6, 94).

Banerjee, S., Kroupa, P. & Oh, S. (2012), *Runaway Massive Stars from R136: VFTS 682 is Very Likely a "Slow Runaway"*, ApJ **746**, 15 15, arXiv: 1111.0291 [astro-ph.GA] (cit. on p. 94).

Bastian, N. & Goodwin, S. P. (2006), *Evidence for the strong effect of gas removal on the internal dynamics of young stellar clusters*, MNRAS **369** L9, arXiv: astro-ph/0602465 [astro-ph] (cit. on pp. 5, 119).

Bate, M. R. (1998), *Collapse of a Molecular Cloud Core to Stellar Densities: The First Three-dimensional Calculations*, ApJ **508** L95, arXiv: astro-ph/9810397 [astro-ph] (cit. on p. 5).

Bate, M. R., Tricco, T. S. & Price, D. J. (2014), *Collapse of a molecular cloud core to stellar densities: stellar-core and outflow formation in radiation magnetohydrodynamic simulations*, MNRAS **437** 77, arXiv: 1310.1092 [astro-ph.SR] (cit. on pp. 89, 95).

Baumgardt, H. & Kroupa, P. (2007), *A comprehensive set of simulations studying the influence of gas expulsion on star cluster evolution*, MNRAS **380** 1589, arXiv: 0707.1944 [astro-ph] (cit. on pp. 94, 96).

Belloni, D., Askar, A., Giersz, M., Kroupa, P. & Rocha-Pinto, H. J. (2017), *On the initial binary population for star cluster simulations*, MNRAS **471** 2812, arXiv: 1707.04271 [astro-ph.GA] (cit. on pp. 94, 95).

Benjamin, R. A. et al. (2003), *GLIMPSE. I. An SIRTF Legacy Project to Map the Inner Galaxy*, PASP **115** 953, arXiv: astro-ph/0306274 [astro-ph] (cit. on p. 11).

Bertoldi, F. & McKee, C. F. (1992), *Pressure-confined Clumps in Magnetized Molecular Clouds*, ApJ **395** 140 (cit. on pp. 2, 69).

Beuther, H., Kuiper, R. & Tafalla, M. (2025), *Star formation from low to high mass: A comparative view*, arXiv e-prints, arXiv:2501.16866 arXiv:2501.16866, arXiv: 2501.16866 [astro-ph.GA] (cit. on p. 1).

Beuther, H. et al. (2022), *FEEDBACK from the NGC 7538 H II region*, A&A **659**, A77 A77, arXiv: 2201.05532 [astro-ph.GA] (cit. on p. 79).
Bonne, L. et al. (2023), *The SOFIA FEEDBACK [CII] Legacy Survey: Rapid molecular cloud dispersal in RCW 79*, A&A **679**, L5 L5, arXiv: 2310.01657 [astro-ph.GA] (cit. on p. 79).
Bonnell, I. A., Bate, M. R., Clarke, C. J. & Pringle, J. E. (1997), *Accretion and the stellar mass spectrum in small clusters*, MNRAS **285** 201 (cit. on pp. 2, 4, 141).
Bonnell, I. A., Bate, M. R., Clarke, C. J. & Pringle, J. E. (2001), *Competitive accretion in embedded stellar clusters*, MNRAS **323** 785, arXiv: astro-ph/0102074 [astro-ph] (cit. on pp. 2, 4, 141).
Bonnell, I. A., Bate, M. R. & Vine, S. G. (2003), *The hierarchical formation of a stellar cluster*, MNRAS **343** 413, arXiv: astro-ph/0305082 [astro-ph] (cit. on p. 6).
Bonnell, I. A., Vine, S. G. & Bate, M. R. (2004), *Massive star formation: nurture, not nature*, MNRAS **349** 735, arXiv: astro-ph/0401059 [astro-ph] (cit. on pp. 6, 93).
Bressan, A. et al. (2012), *PARSEC: stellar tracks and isochrones with the PAdova and TRieste Stellar Evolution Code*, MNRAS **427** 127, arXiv: 1208.4498 [astro-ph.SR] (cit. on p. 83).
Brinkmann, N., Banerjee, S., Motwani, B. & Kroupa, P. (2017), *The bound fraction of young star clusters*, A&A **600**, A49 A49, arXiv: 1611.05871 [astro-ph.GA] (cit. on pp. 5, 94, 97).
Bronfman, L., Cohen, R. S., Alvarez, H., May, J. & Thaddeus, P. (1988), *A CO Survey of the Southern Milky Way: The Mean Radial Distribution of Molecular Clouds within the Solar Circle*, ApJ **324** 248 (cit. on p. 63).
Bronfman, L., Nyman, L. .-. & May, J. (1996), *A CS(2-1) survey of IRAS point sources with color characteristics of ultra-compact HII regions.*, **115** 81 (cit. on p. 11).
Camacho, V. et al. (2016), *Energy Budget of Forming Clumps in Numerical Simulations of Collapsing Clouds*, ApJ **833**, 113 113, arXiv: 1609.09112 [astro-ph.GA] (cit. on p. 4).
Cantat-Gaudin, T. et al. (2019), *Expanding associations in the Vela-Puppis region. 3D structure and kinematics of the young population*, A&A **626**, A17 A17, arXiv: 1812.08114 [astro-ph.GA] (cit. on pp. 5, 102).
Cantat-Gaudin, T. et al. (2024), *Uniting Gaia and APOGEE to unveil the cosmic chemistry of the Milky Way disc*, A&A **683**, A128 A128, arXiv: 2401.05023 [astro-ph.GA] (cit. on p. 5).
Carpenter, J. M. (2000), *2MASS Observations of the Perseus, Orion A, Orion B, and Monoceros R2 Molecular Clouds*, AJ **120** 3139, arXiv: astro-ph/0009118 [astro-ph] (cit. on pp. 82, 91, 103, 104).
Carpenter, J. M., Snell, R. L., Schloerb, F. P. & Skrutskie, M. F. (1993), *Embedded Star Clusters Associated with Luminous IRAS Point Sources*, ApJ **407** 657 (cit. on pp. 82, 91, 103, 104).
Cartwright, A. & Whitworth, A. P. (2004), *The statistical analysis of star clusters*, MNRAS **348** 589, arXiv: astro-ph/0403474 [astro-ph] (cit. on pp. 93, 140).
Chakrabarti, S. et al. (2023), *A Noninteracting Galactic Black Hole Candidate in a Binary System with a Main-sequence Star*, AJ **166**, 6 6, arXiv: 2210.05003 [astro-ph.GA] (cit. on p. 121).
Chambers, E. T., Jackson, J. M., Rathborne, J. M. & Simon, R. (2009), *Star Formation Activity of Cores within Infrared Dark Clouds*, ApJS **181** 360 (cit. on p. 74).
Chen, B. .-. et al. (2019), *Three-dimensional interstellar dust reddening maps of the Galactic plane*, MNRAS **483** 4277, arXiv: 1807.02241 [astro-ph.GA] (cit. on p. 57).
Chen, H.-R. V. et al. (2019), *Filamentary Accretion Flows in the Infrared Dark Cloud G14.225-0.506 Revealed by ALMA*, ApJ **875**, 24 24, arXiv: 1903.04376 [astro-ph.GA] (cit. on pp. 2, 70).
Chen, L., de Grijs, R. & Zhao, J. L. (2007), *Mass Segregation in Very Young Open Clusters: A Case Study of NGC 2244 and NGC 6530*, AJ **134** 1368, arXiv: 0706.2723 [astro-ph] (cit. on p. 94).
Chen, M. C.-Y. et al. (2020), *Velocity-coherent Filaments in NGC 1333: Evidence for Accretion Flow?*, ApJ **891**, 84 84, arXiv: 2002.11736 [astro-ph.GA] (cit. on p. 2).

Chen, X. et al. (2020), *New maser species tracing spiral-arm accretion flows in a high-mass young stellar object*, Nature Astronomy **4** 1170 (cit. on p. 15).

Chen, X. et al. (2013), *SMA Observations of Class 0 Protostars: A High Angular Resolution Survey of Protostellar Binary Systems*, ApJ **768**, 110 110, arXiv: 1304.0436 [astro-ph.SR] (cit. on p. 8).

Cheng, Y. et al. (2018), *The Core Mass Function in the Massive Protocluster G286.21+0.17 Revealed by ALMA*, ApJ **853**, 160 160, arXiv: 1706.06584 [astro-ph.GA] (cit. on p. 7).

Chevance, M. et al. (2023), "The Life and Times of Giant Molecular Clouds", *Protostars and Planets VII*, ed. by S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida & M. Tamura, vol. 534, Astronomical Society of the Pacific Conference Series 1, arXiv: 2203.09570 [astro-ph.GA] (cit. on pp. 6, 42, 146).

Chevance, M. et al. (2022), *Pre-supernova feedback mechanisms drive the destruction of molecular clouds in nearby star-forming disc galaxies*, MNRAS **509** 272, arXiv: 2010.13788 [astro-ph.GA] (cit. on p. 7).

Choi, J. et al. (2016), *Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models*, ApJ **823**, 102 102, arXiv: 1604.08592 [astro-ph.SR] (cit. on p. 84).

Clarke, S. D., Whitworth, A. P. & Hubber, D. A. (2016), *Perturbation growth in accreting filaments*, MNRAS **458** 319, arXiv: 1602.07651 [astro-ph.GA] (cit. on p. 2).

Clarke, S. D. et al. (2023), *A Herschel study of G214.5-1.8: a young, cold, and quiescent giant molecular filament on the shell of a H I superbubble*, MNRAS **519** 3098, arXiv: 2211.02422 [astro-ph.GA] (cit. on p. 100).

Cox, D. P. (2005), *The Three-Phase Interstellar Medium Revisited*, ARA&A **43** 337 (cit. on p. 63).

Crutcher, R. M. (2012), *Magnetic Fields in Molecular Clouds*, ARA&A **50** 29 (cit. on pp. 30, 43).

Crutcher, R. M., Wandelt, B., Heiles, C., Falgarone, E. & Troland, T. H. (2010), *Magnetic Fields in Interstellar Clouds from Zeeman Observations: Inference of Total Field Strengths by Bayesian Analysis*, ApJ **725** 466 (cit. on pp. 30, 43).

Csengeri, T. et al. (2016), *The ATLASGAL survey: distribution of cold dust in the Galactic plane. Combination with Planck data*, A&A **585**, A104 A104, arXiv: 1511.07145 [astro-ph.SR] (cit. on pp. 12, 107).

Cuevas-Otahola, B., Mayya, Y. D., Arriaga-Hernández, J., Puerari, I. & Bruzual, G. (2023), *The cluster initial mass function of the M82 disc super star clusters*, MNRAS **525** 5262, arXiv: 2309.00719 [astro-ph.GA] (cit. on p. 147).

Dale, J. E. & Bonnell, I. A. (2008), *The effect of stellar winds on the formation of a protocluster*, MNRAS **391** 2, arXiv: 0808.1510 [astro-ph] (cit. on p. 107).

Dale, J. E., Bonnell, I. A. & Whitworth, A. P. (2007), *Ionization-induced star formation - I. The collect-and-collapse model*, MNRAS **375** 1291, arXiv: astro-ph/0612128 [astro-ph] (cit. on p. 73).

Dale, J. E., Ercolano, B. & Bonnell, I. A. (2012), *Ionizing feedback from massive stars in massive clusters - II. Disruption of bound clusters by photoionization*, MNRAS **424** 377, arXiv: 1205.0360 [astro-ph.GA] (cit. on pp. 3, 107).

Dale, J. E., Ngoumou, J., Ercolano, B. & Bonnell, I. A. (2014), *Before the first supernova: combined effects of H II regions and winds on molecular clouds*, MNRAS **442** 694, arXiv: 1404.6102 [astro-ph.GA] (cit. on pp. 3, 72).

Dale, J. E. (2017), *The effect of the virial state of molecular clouds on the influence of feedback from massive stars*, MNRAS **467** 1067, arXiv: 1701.03653 [astro-ph.GA] (cit. on p. 5).

Dame, T. M., Hartmann, D. & Thaddeus, P. (2001), *The Milky Way in Molecular Clouds: A New Complete CO Survey*, ApJ **547** 792, arXiv: astro-ph/0009217 [astro-ph] (cit. on pp. 52, 130).

Das, S. R., Gupta, S., Prakash, P., Samal, M. & Jose, J. (2023), *Membership Analysis and 3D Kinematics of the Star-forming Complex around Trumpler 37 Using Gaia-DR3*, ApJ **948**, 7 7, arXiv: 2302.14347 [astro-ph.GA] (cit. on pp. 5, 102).

Davies, B. et al. (2011), *The Red MSX Source survey: critical tests of accretion models for the formation of massive stars*, MNRAS **416** 972, arXiv: 1105.3984 [astro-ph.GA] (cit. on p. 85).

Davis, T. A. et al. (2013), *The ATLAS^{3D} Project - XIV. The extent and kinematics of the molecular gas in early-type galaxies*, MNRAS **429** 534, arXiv: 1211.1011 [astro-ph.CO] (cit. on pp. 27, 36, 142).

Deharveng, L., Zavagno, A. & Caplan, J. (2005), *Triggered massive-star formation on the borders of Galactic H II regions. I. A search for "collect and collapse" candidates*, A&A **433** 565, arXiv: astro-ph/0412602 [astro-ph] (cit. on p. 3).

del Alcázar-Julià, M., Figueras, F., Robin, A. C., Bienaymé, O. & Anders, F. (2025), *Joint inference of the Milky Way star formation history and IMF from Gaia all-sky $G < 13$ data*, arXiv e-prints, arXiv:2501.17236 arXiv:2501.17236, arXiv: 2501.17236 [astro-ph.GA] (cit. on p. 136).

Della Croce, A., Dalessandro, E., Livernois, A. & Vesperini, E. (2024), *Young, wild, and free: The early expansion of star clusters*, A&A **683**, A10 A10, arXiv: 2312.02263 [astro-ph.GA] (cit. on p. 5).

Della Croce, A. et al. (2023), *Ongoing hierarchical massive cluster assembly: The LISCA II structure in the Perseus complex*, A&A **674**, A93 A93, arXiv: 2303.15501 [astro-ph.GA] (cit. on p. 102).

Dewangan, L. K. et al. (2020), *New Insights into the H II Region G18.88-0.49: Hub-Filament System and Accreting Filaments*, ApJ **903**, 13 13, arXiv: 2009.05249 [astro-ph.GA] (cit. on p. 2).

Di Carlo, U. N., Agrawal, P., Rodriguez, C. L. & Breivik, K. (2024), *Young Star Clusters Dominate the Production of Detached Black Hole–Star Binaries*, ApJ **965**, 22 22, arXiv: 2306.13121 [astro-ph.GA] (cit. on p. 121).

Di Francesco, J. et al. (2020), *Herschel Gould Belt Survey Observations of Dense Cores in the Cepheus Flare Clouds*, ApJ **904**, 172 172, arXiv: 2010.09894 [astro-ph.GA] (cit. on p. 7).

Dib, S. (2022), *The galaxy-wide stellar initial mass function in the presence of cluster-to-cluster IMF variations*, A&A **666**, A113 A113, arXiv: 2204.09064 [astro-ph.GA] (cit. on p. 7).

Dib, S. (2023a), *The evolution of the internal structure of massive star-forming regions in the Milky Way as revealed by ALMA*, MNRAS **524** 1625, arXiv: 2209.12905 [astro-ph.GA] (cit. on p. 5).

Dib, S. (2023b), *Variation of the High-mass Slope of the Stellar Initial Mass Function: Theory Meets Observations*, ApJ **959**, 88 88, arXiv: 2309.10842 [astro-ph.GA] (cit. on pp. 83, 136).

Dib, S., Brandenburg, A., Kim, J., Gopinathan, M. & André, P. (2008), *Core Mass Function: The Role of Gravity*, ApJ **678** L105, arXiv: 0801.2257 [astro-ph] (cit. on p. 5).

Dib, S., Gutkin, J., Brandner, W. & Basu, S. (2013), *Feedback-regulated star formation - II. Dual constraints on the SFE and the age spread of stars in massive clusters*, MNRAS **436** 3727, arXiv: 1306.6338 [astro-ph.GA] (cit. on pp. 3, 89, 96, 107).

Dib, S., Helou, G., Moore, T. J. T., Urquhart, J. S. & Dariush, A. (2012), *The Lesser Role of Shear in Galactic Star Formation: Insight from the Galactic Ring Survey*, ApJ **758**, 125 125, arXiv: 1202.4699 [astro-ph.GA] (cit. on pp. 1, 61, 62).

Dib, S. & Henning, T. (2019), *Star formation activity and the spatial distribution and mass segregation of dense cores in the early phases of star formation*, A&A **629**, A135 A135, arXiv: 1808.07057 [astro-ph.GA] (cit. on p. 93).

Dib, S., Kim, J., Vázquez-Semadeni, E., Burkert, A. & Shadmehri, M. (2007), *The Virial Balance of Clumps and Cores in Molecular Clouds*, ApJ **661** 262, arXiv: astro-ph/0607362 [astro-ph] (cit. on pp. 43, 57).

Dib, S., Piau, L., Mohanty, S. & Braine, J. (2011), *Star formation efficiency as a function of metallicity: from star clusters to galaxies*, MNRAS **415** 3439, arXiv: 1102.3839 [astro-ph.GA] (cit. on p. 107).

Dib, S., Schmeja, S. & Parker, R. J. (2018), *Structure and mass segregation in Galactic stellar clusters*, MNRAS **473** 849, arXiv: 1707.00744 [astro-ph.GA] (cit. on p. 93).

Dib, S. et al. (2010), *The IMF of stellar clusters: effects of accretion and feedback*, MNRAS **405** 401, arXiv: 0908.4522 [astro-ph.GA] (cit. on p. 3).

Dib, S. et al. (2025), *Assessing the accuracy of star formation rate measurements by direct star count in molecular clouds*, A&A **693**, A51 A51, arXiv: 2405.00095 [astro-ph.GA] (cit. on p. 85).

Dinnbier, F., Kroupa, P. & Anderson, R. I. (2022), *Do the majority of stars form as gravitationally unbound?*, A&A **660**, A61 A61, arXiv: 2201.06582 [astro-ph.GA] (cit. on pp. 1, 7, 94).

Dobbs, C. L. & Pringle, J. E. (2013), *The exciting lives of giant molecular clouds*, MNRAS **432** 653, arXiv: 1303.4995 [astro-ph.GA] (cit. on p. 29).

Edenhofer, G. et al. (2024), *A parsec-scale Galactic 3D dust map out to 1.25 kpc from the Sun*, A&A **685**, A82 A82, arXiv: 2308.01295 [astro-ph.GA] (cit. on pp. 52, 57).

Elia, D. et al. (2017), *The Hi-GAL compact source catalogue - I. The physical properties of the clumps in the inner Galaxy (-71.0° < ℓ < 67.0°)*, MNRAS **471** 100, arXiv: 1706.01046 [astro-ph.GA] (cit. on p. 2).

Elia, D. et al. (2021), *The Hi-GAL compact source catalogue - II. The 360° catalogue of clump physical properties*, MNRAS **504** 2742, arXiv: 2104.04807 [astro-ph.GA] (cit. on p. 1).

Ellsworth-Bowers, T. P. et al. (2013), *The Bolocam Galactic Plane Survey. VIII. A Mid-infrared Kinematic Distance Discrimination Method*, ApJ **770**, 39 39, arXiv: 1304.4597 [astro-ph.GA] (cit. on p. 63).

Elmegreen, B. G. & Lada, C. J. (1977), *Sequential formation of subgroups in OB associations.*, ApJ **214** 725 (cit. on p. 73).

Elmegreen, B. G. (1989), *A Pressure and Metallicity Dependence for Molecular Cloud Correlations and the Calibration of Mass*, ApJ **338** 178 (cit. on pp. 43, 57).

Elmegreen, B. G. (2000), *Modeling a High-Mass Turn-Down in the Stellar Initial Mass Function*, ApJ **539** 342, arXiv: astro-ph/0005455 [astro-ph] (cit. on p. 6).

Elmegreen, B. G., Elmegreen, D. M. & Efremov, Y. N. (2018), *Regularly Spaced Infrared Peaks in the Dusty Spirals of Messier 100*, ApJ **863**, 59 59, arXiv: 1806.08957 [astro-ph.GA] (cit. on p. 2).

Elmegreen, D. M. et al. (2014), *Hierarchical Star Formation in Nearby LEGUS Galaxies*, ApJ **787**, L15 L15, arXiv: 1404.6001 [astro-ph.GA] (cit. on p. 6).

Emsellem, E. et al. (2022), *The PHANGS-MUSE survey. Probing the chemo-dynamical evolution of disc galaxies*, A&A **659**, A191 A191, arXiv: 2110.03708 [astro-ph.GA] (cit. on pp. 12, 141).

Enoch, M. L. et al. (2008), *The Mass Distribution and Lifetime of Prestellar Cores in Perseus, Serpens, and Ophiuchus*, ApJ **684** 1240, arXiv: 0805.1075 [astro-ph] (cit. on p. 7).

Evans Neal J., I., Heyer, M., Miville-Deschênes, M.-A., Nguyen-Luong, Q. & Merello, M. (2021), *Which Molecular Cloud Structures Are Bound?*, ApJ **920**, 126 126, arXiv: 2107.05750 [astro-ph.GA] (cit. on pp. 18, 73).

Evans Neal J., I. et al. (2009), *The Spitzer c2d Legacy Results: Star-Formation Rates and Efficiencies; Evolution and Lifetimes*, ApJS **181** 321, arXiv: 0811.1059 [astro-ph] (cit. on p. 85).

Farias, J. P., Offner, S. S. R., Grudić, M. Y., Guszejnov, D. & Rosen, A. L. (2024), *Stellar populations in STARFORGE: the origin and evolution of star clusters and associations*, MNRAS **527** 6732, arXiv: 2309.11415 [astro-ph.GA] (cit. on p. 95).

Faúndez, S. et al. (2004), *SIMBA survey of southern high-mass star forming regions. I. Physical parameters of the 1.2 mm/IRAS sources*, A&A **426** 97 (cit. on p. 11).

Faustini, F., Molinari, S., Testi, L. & Brand, J. (2009), *Properties of stellar clusters around high-mass young stars*, A&A **503** 801, arXiv: 0904.3342 [astro-ph.GA] (cit. on pp. 91, 103, 104).

Feigelson, E. D. et al. (2013), *Overview of the Massive Young Star-Forming Complex Study in Infrared and X-Ray (MYStIX) Project*, ApJS **209**, 26 26, arXiv: 1309.4483 [astro-ph.SR] (cit. on pp. 82, 83).

Field, G. B., Blackman, E. G. & Keto, E. R. (2011), *Does external pressure explain recent results for molecular clouds?*, MNRAS **416** 710, arXiv: 1106.3017 [astro-ph.GA] (cit. on p. 70).

Flaccomio, E. et al. (2023), *Spatial and dynamical structure of the NGC 2264 star-forming region*, A&A **670**, A37 A37, arXiv: 2211.09000 [astro-ph.GA] (cit. on pp. 5, 102).

Fukui, Y., Habe, A., Inoue, T., Enokiya, R. & Tachihara, K. (2021), *Cloud-cloud collisions and triggered star formation*, PASJ **73** S1, arXiv: 2009.05077 [astro-ph.GA] (cit. on p. 73).

Gaia Collaboration et al. (2023), *Gaia Data Release 3. Summary of the content and survey properties*, A&A **674**, A1 A1, arXiv: 2208.00211 [astro-ph.GA] (cit. on p. 121).

Gaia Collaboration et al. (2024), *Discovery of a dormant 33 solar-mass black hole in pre-release Gaia astrometry*, A&A **686**, L2 L2, arXiv: 2404.10486 [astro-ph.GA] (cit. on p. 121).

Ganguly, S., Walch, S., Clarke, S. D. & Seifried, D. (2024a), *SILCC-Zoom: the dynamic balance in molecular cloud substructures*, MNRAS, arXiv: 2204.02511 [astro-ph.GA] (cit. on p. 53).

Ganguly, S., Walch, S., Clarke, S. D. & Seifried, D. (2024b), *SILCC-Zoom: the dynamic balance in molecular cloud substructures*, MNRAS **528** 3630, arXiv: 2204.02511 [astro-ph.GA] (cit. on p. 70).

Ganguly, S., Walch, S., Seifried, D., Clarke, S. D. & Weis, M. (2023), *Unravelling the structure of magnetized molecular clouds with SILCC-Zoom: sheets, filaments, and fragmentation*, MNRAS **525** 721, arXiv: 2307.08746 [astro-ph.GA] (cit. on pp. 30, 43, 100).

Getman, K. V. et al. (2014), *Age Gradients in the Stellar Populations of Massive Star Forming Regions Based on a New Stellar Chronometer*, ApJ **787**, 108 108, arXiv: 1403.2741 [astro-ph.SR] (cit. on p. 94).

Geyer, M. P. & Burkert, A. (2001), *The effect of gas loss on the formation of bound stellar clusters*, MNRAS **323** 988, arXiv: astro-ph/0007413 [astro-ph] (cit. on p. 95).

Giacobbo, N., Mapelli, M. & Spera, M. (2018), *Merging black hole binaries: the effects of progenitor's metallicity, mass-loss rate and Eddington factor*, MNRAS **474** 2959, arXiv: 1711.03556 [astro-ph.SR] (cit. on p. 96).

Gómez, G. C. & Vázquez-Semadeni, E. (2014), *Filaments in Simulations of Molecular Cloud Formation*, ApJ **791**, 124 124, arXiv: 1308.6298 [astro-ph.GA] (cit. on pp. 2, 14, 18).

Gómez, G. C., Vázquez-Semadeni, E. & Palau, A. (2021), *Density profile evolution during prestellar core collapse: collapse starts at the large scale*, MNRAS **502** 4963, arXiv: 2009.14151 [astro-ph.GA] (cit. on p. 18).

Gong, H. & Ostriker, E. C. (2011), *Dense Core Formation in Supersonic Turbulent Converging Flows*, ApJ **729**, 120 120, arXiv: 1101.2650 [astro-ph.SR] (cit. on pp. 53, 63).

Goodwin, S. P. & Kroupa, P. (2005), *Limits on the primordial stellar multiplicity*, A&A **439** 565, arXiv: astro-ph/0505470 [astro-ph] (cit. on p. 8).

Goodwin, S. P. & Bastian, N. (2006), *Gas expulsion and the destruction of massive young clusters*, MNRAS **373** 752, arXiv: astro-ph/0609477 [astro-ph] (cit. on pp. 5, 119).

Gouliermis, D. A. et al. (2015), *Hierarchical star formation across the ring galaxy NGC 6503*, MNRAS **452** 3508, arXiv: 1506.03928 [astro-ph.GA] (cit. on p. 6).

Gouliermis, D. A. et al. (2017), *Hierarchical star formation across the grand-design spiral NGC 1566*, MNRAS **468** 509, arXiv: 1702.06006 [astro-ph.GA] (cit. on p. 142).

Grasha, K. et al. (2017), *The Hierarchical Distribution of the Young Stellar Clusters in Six Local Star-forming Galaxies*, ApJ **840**, 113 113, arXiv: 1704.06321 [astro-ph.GA] (cit. on p. 6).

Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S. & Finkbeiner, D. (2019), *A 3D Dust Map Based on Gaia, Pan-STARRS 1, and 2MASS*, ApJ **887**, 93 93, arXiv: 1905.02734 [astro-ph.GA] (cit. on p. 57).

Gunawardhana, M. L. P. et al. (2011), *Galaxy and Mass Assembly (GAMA): the star formation rate dependence of the stellar initial mass function*, MNRAS **415** 1647, arXiv: 1104.2379 [astro-ph.CO] (cit. on p. 136).

Hacar, A., Alves, J., Tafalla, M. & Goicoechea, J. R. (2017), *Gravitational collapse of the OMC-1 region*, A&A **602**, L2 L2, arXiv: 1703.03464 [astro-ph.GA] (cit. on p. 18).

Hacar, A., Kainulainen, J., Tafalla, M., Beuther, H. & Alves, J. (2016), *The Musca cloud: A 6 pc-long velocity-coherent, sonic filament*, A&A **587**, A97 A97, arXiv: 1511.06370 [astro-ph.GA] (cit. on p. 2).

Hacar, A. & Tafalla, M. (2011), *Dense core formation by fragmentation of velocity-coherent filaments in L1517*, A&A **533**, A34 A34, arXiv: 1107.0971 [astro-ph.GA] (cit. on p. 2).

Hartmann, L., Ballesteros-Paredes, J. & Heitsch, F. (2012), *Rapid star formation and global gravitational collapse*, MNRAS **420** 1457, arXiv: 1111.2582 [astro-ph.SR] (cit. on pp. 2, 4).

Haslbauer, M. et al. (2024), *The effect of the environment-dependent stellar initial mass function on the photometric properties of star-forming galaxies*, A&A **689**, A221 A221, arXiv: 2405.05313 [astro-ph.GA] (cit. on p. 7).

Hassani, H. et al. (2023), *PHANGS-JWST First Results: The 21 μm Compact Source Population*, ApJ **944**, L21 L21, arXiv: 2212.01526 [astro-ph.GA] (cit. on p. 41).

He, Z.-Z., Li, G.-X. & Burkert, A. (2023), *Mapping gravity in stellar nurseries - establishing the effectiveness of 2D acceleration maps*, MNRAS **526** L20, arXiv: 2307.13409 [astro-ph.GA] (cit. on pp. 53, 63).

Heggie, D. & Hut, P. (2003), *The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics* (cit. on pp. 94, 126).

Heggie, D. C., Trenti, M. & Hut, P. (2006), *Star clusters with primordial binaries - I. Dynamical evolution of isolated models*, MNRAS **368** 677, arXiv: astro-ph/0602408 [astro-ph] (cit. on p. 5).

Heitsch, F., Hartmann, L. W., Slyz, A. D., Devriendt, J. E. G. & Burkert, A. (2008), *Cooling, Gravity, and Geometry: Flow-driven Massive Core Formation*, ApJ **674** 316, arXiv: 0709.2451 [astro-ph] (cit. on pp. 72, 73).

Hennebelle, P. & Falgarone, E. (2012), *Turbulent molecular clouds*, A&A Rev. **20**, 55 55, arXiv: 1211.0637 [astro-ph.GA] (cit. on p. 4).

Henshaw, J. D., Caselli, P., Fontani, F., Jiménez-Serra, I. & Tan, J. C. (2014), *The dynamical properties of dense filaments in the infrared dark cloud G035.39-00.33*, MNRAS **440** 2860, arXiv: 1403.1444 [astro-ph.SR] (cit. on p. 2).

Henshaw, J. D., Longmore, S. N. & Kruijssen, J. M. D. (2016), *Seeding the Galactic Centre gas stream: gravitational instabilities set the initial conditions for the formation of protocluster clouds*, MNRAS **463** L122, arXiv: 1609.01721 [astro-ph.GA] (cit. on p. 2).

Henshaw, J. D. et al. (2020), *Ubiquitous velocity fluctuations throughout the molecular interstellar medium*, Nature Astronomy **4** 1064, arXiv: 2007.01877 [astro-ph.GA] (cit. on pp. 1, 6, 14, 15, 141).

Henshaw, J. D. et al. (2022), *A wind-blown bubble in the Central Molecular Zone cloud G0.253+0.016*, MNRAS **509** 4758, arXiv: 2110.11367 [astro-ph.GA] (cit. on pp. 3, 78).

Heyer, M., Krawczyk, C., Duval, J. & Jackson, J. M. (2009), *Re-Examining Larson's Scaling Relationships in Galactic Molecular Clouds*, ApJ **699** 1092, arXiv: 0809.1397 [astro-ph] (cit. on pp. 4, 72).

Heyer, M. H. & Brunt, C. M. (2004), *The Universality of Turbulence in Galactic Molecular Clouds*, ApJ **615** L45, arXiv: astro-ph/0409420 [astro-ph] (cit. on p. 4).

Hosokawa, T. & Omukai, K. (2009), *Evolution of Massive Protostars with High Accretion Rates*, ApJ **691** 823, arXiv: 0806.4122 [astro-ph] (cit. on p. 85).

Hottier, C., Babusiaux, C. & Arenou, F. (2020), *FEDReD. II. 3D extinction map with 2MASS and Gaia DR2 data*, A&A **641**, A79 A79, arXiv: 2007.03734 [astro-ph.GA] (cit. on p. 57).

Hsieh, T.-H. & Lai, S.-P. (2013), *Populations of Young Stellar Objects in Nearby Molecular Clouds*, ApJS **205**, 5 5, arXiv: 1301.5426 [astro-ph.GA] (cit. on p. 85).

Hunt, E. L. & Reffert, S. (2023), *Improving the open cluster census. II. An all-sky cluster catalogue with Gaia DR3*, A&A **673**, A114 A114, arXiv: 2303.13424 [astro-ph.GA] (cit. on p. 83).

Hunt, E. L. & Reffert, S. (2024), *Improving the open cluster census. III. Using cluster masses, radii, and dynamics to create a cleaned open cluster catalogue*, A&A **686**, A42 A42, arXiv: 2403.05143 [astro-ph.GA] (cit. on pp. 83, 98, 106, 108, 111, 117, 120).

Hunter, D. A., Elmegreen, B. G. & Baker, A. L. (1998), *The Relationship between Gas, Stars, and Star Formation in Irregular Galaxies: A Test of Simple Models*, ApJ **493** 595, arXiv: astro-ph/9712353 [astro-ph] (cit. on p. 61).

Hurley, J. R., Pols, O. R. & Tout, C. A. (2000), *Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity*, MNRAS **315** 543, arXiv: astro-ph/0001295 [astro-ph] (cit. on p. 96).

Hurley, J. R., Tout, C. A. & Pols, O. R. (2002), *Evolution of binary stars and the effect of tides on binary populations*, MNRAS **329** 897, arXiv: astro-ph/0201220 [astro-ph] (cit. on p. 96).

Ibáñez-Mejıa, J. C., Mac Low, M.-M. & Klessen, R. S. (2022), *Gravity versus Magnetic Fields in Forming Molecular Clouds*, ApJ **925**, 196 196, arXiv: 2108.04967 [astro-ph.GA] (cit. on pp. 30, 43).

Ibáñez-Mejıa, J. C., Mac Low, M.-M., Klessen, R. S. & Baczynski, C. (2016), *Gravitational Contraction versus Supernova Driving and the Origin of the Velocity Dispersion-Size Relation in Molecular Clouds*, ApJ **824**, 41 41, arXiv: 1511.05602 [astro-ph.GA] (cit. on p. 4).

Inutsuka, S.-i., Inoue, T., Iwasaki, K. & Hosokawa, T. (2015), *The formation and destruction of molecular clouds and galactic star formation. An origin for the cloud mass function and star formation efficiency*, A&A **580**, A49 A49, arXiv: 1505.04696 [astro-ph.GA] (cit. on p. 100).

Irabor, T. et al. (2023), *The coordinated radio and infrared survey for high-mass star formation - V. The CORNISH-South survey and catalogue*, MNRAS **520** 1073, arXiv: 2301.01988 [astro-ph.GA] (cit. on pp. 34, 82).

Issac, N. et al. (2019), *Initial phases of high-mass star formation: a multiwavelength study towards the extended green object G12.42+0.50*, MNRAS **485** 1775, arXiv: 1902.03724 [astro-ph.GA] (cit. on p. 2).

Izquierdo, A. F. et al. (2018), *Radiative transfer modelling of W33A MM1: 3D structure and dynamics of a complex massive star-forming region*, MNRAS **478** 2505, arXiv: 1804.09204 [astro-ph.GA] (cit. on pp. 8, 15).

Jadhav, V. V., Kroupa, P., Wu, W., Pflamm-Altenburg, J. & Thies, I. (2024), *The spin, expansion, and contraction of open star clusters*, A&A **687**, A89 A89, arXiv: 2404.05327 [astro-ph.GA] (cit. on pp. 5, 102).

Jeffreson, S. M. R., Sun, J. & Wilson, C. D. (2022), *On the scale height of the molecular gas disc in Milky Way-like galaxies*, MNRAS **515** 1663, arXiv: 2206.14818 [astro-ph.GA] (cit. on p. 63).

Jerabkova, T. et al. (2018), *Impact of metallicity and star formation rate on the time-dependent, galaxy-wide stellar initial mass function*, A&A **620**, A39 A39, arXiv: 1809.04603 [astro-ph.GA] (cit. on p. 7).

Kauffmann, J., Bertoldi, F., Bourke, T. L., Evans N. J., I. & Lee, C. W. (2008), *MAMBO mapping of Spitzer c2d small clouds and cores*, A&A **487** 993, arXiv: 0805.4205 [astro-ph] (cit. on p. 22).

Kauffmann, J., Pillai, T. & Goldsmith, P. F. (2013), *Low Virial Parameters in Molecular Clouds: Implications for High-mass Star Formation and Magnetic Fields*, ApJ **779**, 185 185, arXiv: 1308.5679 [astro-ph.GA] (cit. on p. 69).

Kendrew, S. et al. (2012), *The Milky Way Project: A Statistical Study of Massive Star Formation Associated with Infrared Bubbles*, ApJ **755**, 71 71, arXiv: 1203.5486 [astro-ph.GA] (cit. on p. 3).

Kennicutt, R. C. & Evans, N. J. (2012), *Star Formation in the Milky Way and Nearby Galaxies*, ARA&A **50** 531, arXiv: 1204.3552 [astro-ph.GA] (cit. on p. 1).

Keto, E. R. & Myers, P. C. (1986), *CO Observations of Southern High-Latitude Clouds*, ApJ **304** 466 (cit. on p. 70).

Keto, E., Zhang, Q. & Kurtz, S. (2008), *The Early Evolution of Massive Stars: Radio Recombination Line Spectra*, ApJ **672** 423, arXiv: 0708.3388 [astro-ph] (cit. on p. 22).

Kim, J. et al. (2021), *On the duration of the embedded phase of star formation*, MNRAS **504** 487, arXiv: 2012.00019 [astro-ph.GA] (cit. on p. 145).

Kim, J. et al. (2022), *Environmental dependence of the molecular cloud lifecycle in 54 main-sequence galaxies*, MNRAS **516** 3006, arXiv: 2206.09857 [astro-ph.GA] (cit. on p. 146).

Kim, J. et al. (2023), *PHANGS-JWST First Results: Duration of the Early Phase of Massive Star Formation in NGC 628*, ApJ **944**, L20 L20, arXiv: 2211.15698 [astro-ph.GA] (cit. on pp. 40, 145).

Kim, J.-G., Kim, W.-T. & Ostriker, E. C. (2016), *Disruption of Molecular Clouds by Expansion of Dusty H II Regions*, ApJ **819**, 137 137, arXiv: 1601.03035 [astro-ph.GA] (cit. on p. 7).

Kim, J.-G., Ostriker, E. C. & Filippova, N. (2021), *Star Formation Efficiency and Dispersal of Giant Molecular Clouds with UV Radiation Feedback: Dependence on Gravitational Boundedness and Magnetic Fields*, ApJ **911**, 128 128, arXiv: 2011.07772 [astro-ph.GA] (cit. on p. 30).

Kirk, H., Johnstone, D. & Di Francesco, J. (2006), *The Large- and Small-Scale Structures of Dust in the Star-forming Perseus Molecular Cloud*, ApJ **646** 1009, arXiv: astro-ph/0602089 [astro-ph] (cit. on p. 57).

Kirk, H. & Myers, P. C. (2011), *Young Stellar Groups and Their Most Massive Stars*, ApJ **727**, 64 64, arXiv: 1011.1416 [astro-ph.GA] (cit. on p. 94).

Kirk, H. et al. (2013), *Filamentary Accretion Flows in the Embedded Serpens South Protocluster*, ApJ **766**, 115 115, arXiv: 1301.6792 [astro-ph.GA] (cit. on p. 2).

Kirk, H. et al. (2017), *The Green Bank Ammonia Survey: Dense Cores under Pressure in Orion A*, ApJ **846**, 144 144, arXiv: 1708.05426 [astro-ph.GA] (cit. on pp. 43, 57, 70).

Koch, E. W. & Rosolowsky, E. W. (2015a), *Filament identification through mathematical morphology*, MNRAS **452** 3435, arXiv: 1507.02289 [astro-ph.GA] (cit. on pp. 13, 15, 19, 25).

Koch, E. W. & Rosolowsky, E. W. (2015b), *Filament identification through mathematical morphology*, MNRAS **452** 3435, arXiv: 1507.02289 [astro-ph.GA] (cit. on p. 32).

Kohno, M. et al. (2021), *FOREST unbiased Galactic plane imaging survey with the Nobeyama 45 m telescope (FUGIN). VI. Dense gas and mini-starbursts in the W 43 giant molecular cloud complex*, PASJ **73** S129, arXiv: 2001.10693 [astro-ph.GA] (cit. on p. 100).

Konig, C. et al. (2017), *ATLASGAL-selected massive clumps in the inner Galaxy. III. Dust continuum characterization of an evolutionary sample*, A&A **599**, A139 A139, arXiv: 1610.09055 [astro-ph.GA] (cit. on p. 82).

Könyves, V. et al. (2010), *The Aquila prestellar core population revealed by Herschel*, A&A **518**, L106 L106, arXiv: 1005.2981 [astro-ph.SR] (cit. on p. 8).

Könyves, V. et al. (2015), *A census of dense cores in the Aquila cloud complex: SPIRE/PACS observations from the Herschel Gould Belt survey*, A&A **584**, A91 A91, arXiv: 1507.05926 [astro-ph.GA] (cit. on pp. 7, 8).

Kroupa, P. (2005), "The Fundamental Building Blocks of Galaxies", *The Three-Dimensional Universe with Gaia*, ed. by C. Turon, K. S. O'Flaherty & M. A. C. Perryman, vol. 576, ESA Special Publication 629, arXiv: astro-ph/0412069 [astro-ph] (cit. on pp. 1, 7).

Kroupa, P. (1995a), *Inverse dynamical population synthesis and star formation*, MNRAS **277** 1491, arXiv: astro-ph/9508117 [astro-ph] (cit. on pp. 1, 7).

Kroupa, P. (1995b), *Inverse dynamical population synthesis and star formation*, MNRAS **277** 1491, arXiv: astro-ph/9508117 [astro-ph] (cit. on pp. 94, 95).

Kroupa, P. (1995c), *The dynamical properties of stellar systems in the Galactic disc*, MNRAS **277** 1507, arXiv: astro-ph/9508084 [astro-ph] (cit. on pp. 1, 7).

Kroupa, P. (1995d), *The dynamical properties of stellar systems in the Galactic disc*, MNRAS **277** 1507, arXiv: astro-ph/9508084 [astro-ph] (cit. on pp. 94, 95).

Kroupa, P. (2001), *On the variation of the initial mass function*, MNRAS **322** 231, arXiv: astro-ph/0009005 [astro-ph] (cit. on pp. 5, 94, 97).

Kroupa, P. (2008), "Initial Conditions for Star Clusters", *The Cambridge N-Body Lectures*, ed. by S. J. Aarseth, C. A. Tout & R. A. Mardling, vol. 760 181 (cit. on pp. 94, 95, 103, 126).

Kroupa, P. (2015), *Galaxies as simple dynamical systems: observational data disfavor dark matter and stochastic star formation*, Canadian Journal of Physics **93** 169, arXiv: 1406.4860 [astro-ph.GA] (cit. on pp. 6, 147).

Kroupa, P., Aarseth, S. & Hurley, J. (2001), *The formation of a bound star cluster: from the Orion nebula cluster to the Pleiades*, MNRAS **321** 699, arXiv: astro-ph/0009470 [astro-ph] (cit. on pp. 5, 94, 95, 119).

Kroupa, P., Gjergo, E., Jerabkova, T. & Yan, Z. (2024), *The initial mass function of stars*, arXiv e-prints, arXiv:2410.07311 arXiv:2410.07311, arXiv: 2410.07311 [astro-ph.GA] (cit. on pp. 132, 136).

Kroupa, P. & Weidner, C. (2003), *Galactic-Field Initial Mass Functions of Massive Stars*, ApJ **598** 1076, arXiv: astro-ph/0308356 [astro-ph] (cit. on p. 7).

Kroupa, P. et al. (2013), "The Stellar and Sub-Stellar Initial Mass Function of Simple and Composite Populations", *Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations*, ed. by T. D. Oswalt & G. Gilmore, vol. 5 115 (cit. on pp. 6, 7, 83, 85, 147).

Kruijssen, J. M. D. & Longmore, S. N. (2014), *An uncertainty principle for star formation - I. Why galactic star formation relations break down below a certain spatial scale*, MNRAS **439** 3239, arXiv: 1401.4459 [astro-ph.GA] (cit. on p. 40).

Kruijssen, J. M. D. et al. (2014), *What controls star formation in the central 500 pc of the Galaxy?*, MNRAS **440** 3370, arXiv: 1303.6286 [astro-ph.GA] (cit. on p. 61).

Kruijssen, J. M. D. et al. (2018), *An uncertainty principle for star formation - II. A new method for characterizing the cloud-scale physics of star formation and feedback across cosmic history*, MNRAS **479** 1866, arXiv: 1805.00012 [astro-ph.GA] (cit. on p. 40).

Kruijssen, J. M. D. et al. (2019), *Fast and inefficient star formation due to short-lived molecular clouds and rapid feedback*, Nature **569** 519, arXiv: 1905.08801 [astro-ph.GA] (cit. on p. 7).

Krumholz, M. R. et al. (2014), "Star Cluster Formation and Feedback", *Protostars and Planets VI*, ed. by H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning 243, arXiv: 1401.2473 [astro-ph.GA] (cit. on pp. 3, 73).

Krumholz, M. R., Klein, R. I. & McKee, C. F. (2007), *Radiation-Hydrodynamic Simulations of Collapse and Fragmentation in Massive Protostellar Cores*, ApJ **656** 959, arXiv: astro-ph/0609798 [astro-ph] (cit. on pp. 2, 4).

Krumholz, M. R., McKee, C. F. & Bland-Hawthorn, J. (2019), *Star Clusters Across Cosmic Time*, ARA&A **57** 227, arXiv: 1812.01615 [astro-ph.GA] (cit. on pp. 5, 90, 147).

Kuhn, M. A., Hillenbrand, L. A., Carpenter, J. M. & Avelar Menendez, A. R. (2020), *The Formation of a Stellar Association in the NGC 7000/IC 5070 Complex: Results from Kinematic Analysis of Stars and Gas*, ApJ **899**, 128 128, arXiv: 2006.08622 [astro-ph.GA] (cit. on pp. 5, 102).

Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D. & Getman, K. V. (2019), *Kinematics in Young Star Clusters and Associations with Gaia DR2*, ApJ **870**, 32 32, arXiv: 1807.02115 [astro-ph.GA] (cit. on pp. 5, 79, 93, 102).

Kuhn, M. A. et al. (2014), *The Spatial Structure of Young Stellar Clusters. I. Subclusters*, ApJ **787**, 107 107, arXiv: 1403.4252 [astro-ph.GA] (cit. on pp. 93, 97).

Kuhn, M. A. et al. (2015), *The Spatial Structure of Young Stellar Clusters. III. Physical Properties and Evolutionary States*, ApJ **812**, 131 131, arXiv: 1507.05653 [astro-ph.GA] (cit. on pp. 91, 92, 103, 147).

Kuhn, M. A. et al. (2017), *The Structure of the Young Star Cluster NGC 6231. I. Stellar Population*, AJ **154**, 87 87, arXiv: 1706.00017 [astro-ph.GA] (cit. on p. 83).

Kumar, M. S. N., Keto, E. & Clerkin, E. (2006), *The youngest stellar clusters. Clusters associated with massive protostellar candidates*, A&A **449** 1033, arXiv: astro-ph/0512266 [astro-ph] (cit. on pp. 82, 91, 103–105).

Kumar, M. S. N., Palmeirim, P., Arzoumanian, D. & Inutsuka, S. I. (2020), *Unifying low- and high-mass star formation through density-amplified hubs of filaments. The highest mass stars (>100 $M_\odot$) form only in hubs*, A&A **642**, A87 A87, arXiv: 2008.00295 [astro-ph.GA] (cit. on pp. 2, 6, 23, 41, 95, 132, 141).

Küpper, A. H. W., Maschberger, T., Kroupa, P. & Baumgardt, H. (2011), *Mass segregation and fractal substructure in young massive clusters - I. The McLuster code and method calibration*, MNRAS **417** 2300, arXiv: 1107.2395 [astro-ph.GA] (cit. on pp. 94, 96, 133).

Kuznetsova, A., Hartmann, L. & Ballesteros-Paredes, J. (2015), *Signatures of Star Cluster Formation by Cold Collapse*, ApJ **815**, 27 27, arXiv: 1510.08478 [astro-ph.SR] (cit. on p. 18).

Kuznetsova, A., Hartmann, L. & Ballesteros-Paredes, J. (2018), *Kinematics and structure of star-forming regions: insights from cold collapse models*, MNRAS **473** 2372, arXiv: 1709.07754 [astro-ph.GA] (cit. on p. 18).

Lada, C. J. & Lada, E. A. (2003), *Embedded Clusters in Molecular Clouds*, ARA&A **41** 57, arXiv: astro-ph/0301540 [astro-ph] (cit. on pp. 1, 5, 7, 82, 83, 87, 89–91, 95, 103–105, 147).

Lada, C. J., Muench, A. A., Rathborne, J., Alves, J. F. & Lombardi, M. (2008), *The Nature of the Dense Core Population in the Pipe Nebula: Thermal Cores Under Pressure*, ApJ **672** 410, arXiv: 0709.1164 [astro-ph] (cit. on pp. 57, 70).

Lada, E. A., Depoy, D. L., Evans Neal J., I. & Gatley, I. (1991), *A 2.2 Micron Survey in the L1630 Molecular Cloud*, ApJ **371** 171 (cit. on pp. 82, 91, 103, 104).

Lallement, R., Vergely, J. L., Babusiaux, C. & Cox, N. L. J. (2022), *Updated Gaia-2MASS 3D maps of Galactic interstellar dust*, A&A **661**, A147 A147, arXiv: 2203.01627 [astro-ph.GA] (cit. on p. 120).

Lallement, R. et al. (2019), *Gaia-2MASS 3D maps of Galactic interstellar dust within 3 kpc*, A&A **625**, A135 A135, arXiv: 1902.04116 [astro-ph.GA] (cit. on p. 57).

Lamers, H. J., Zwart, S. P. & Zwart, S. F. P. (2008), *The Evolution of Star Clusters*, The Astrophysical Journal **674** 1091 (cit. on p. 5).

Lancaster, L., Ostriker, E. C., Kim, J.-G. & Kim, C.-G. (2021a), *Efficiently Cooled Stellar Wind Bubbles in Turbulent Clouds. II. Validation of Theory with Hydrodynamic Simulations*, ApJ **914**, 90 90, arXiv: 2104.07722 [astro-ph.GA] (cit. on p. 3).

Lancaster, L., Ostriker, E. C., Kim, J.-G. & Kim, C.-G. (2021b), *Star Formation Regulation and Self-pollution by Stellar Wind Feedback*, ApJ **922**, L3 L3, arXiv: 2110.05508 [astro-ph.GA] (cit. on p. 3).

Lane, J. et al. (2016), *The JCMT Gould Belt Survey: Dense Core Clusters in Orion A*, ApJ **833**, 44 44, arXiv: 1609.06366 [astro-ph.GA] (cit. on p. 94).

Lang, P. et al. (2020), *PHANGS CO Kinematics: Disk Orientations and Rotation Curves at 150 pc Resolution*, ApJ **897**, 122 122, arXiv: 2005.11709 [astro-ph.GA] (cit. on pp. 23, 36).

Larson, K. L. et al. (2023), *Multiscale stellar associations across the star formation hierarchy in PHANGS-HST nearby galaxies: methodology and properties*, MNRAS **523** 6061, arXiv: 2212.11425 [astro-ph.GA] (cit. on pp. 130, 142, 143).

Larson, R. B. (1981), *Turbulence and star formation in molecular clouds.*, MNRAS **194** 809 (cit. on pp. 4, 76).

Larson, R. B. (2003), "The Stellar Initial Mass Function and Beyond (Invited Review)", *Galactic Star Formation Across the Stellar Mass Spectrum*, ed. by J. M. De Buizer & N. S. van der Bliek, vol. 287, Astronomical Society of the Pacific Conference Series 65, arXiv: astro-ph/0205466 [astro-ph] (cit. on pp. 6, 93).

Lee, H.-T. & Chen, W. P. (2007), *Triggered Star Formation by Massive Stars*, ApJ **657** 884, arXiv: astro-ph/0509315 [astro-ph] (cit. on p. 73).

Lee, J. C. et al. (2009), *Comparison of Hα and UV Star Formation Rates in the Local Volume: Systematic Discrepancies for Dwarf Galaxies*, ApJ **706** 599, arXiv: 0909.5205 [astro-ph.CO] (cit. on p. 136).

Lee, J. C. et al. (2022), *The PHANGS-HST Survey: Physics at High Angular Resolution in Nearby Galaxies with the Hubble Space Telescope*, ApJS **258**, 10 10, arXiv: 2101.02855 [astro-ph.GA] (cit. on pp. 130, 141).

Lee, J. C. et al. (2023), *The PHANGS-JWST Treasury Survey: Star Formation, Feedback, and Dust Physics at High Angular Resolution in Nearby GalaxieS*, ApJ **944**, L17 L17, arXiv: 2212.02667 [astro-ph.GA] (cit. on pp. 12, 131, 141).

Leike, R. H., Glatzle, M. & Enßlin, T. A. (2020), *Resolving nearby dust clouds*, A&A **639**, A138 A138, arXiv: 2004.06732 [astro-ph.GA] (cit. on p. 57).

Leroy, A. K. et al. (2015), *ALMA Reveals the Molecular Medium Fueling the Nearest Nuclear Starburst*, ApJ **801**, 25 25, arXiv: 1411.2836 [astro-ph.GA] (cit. on p. 70).

Leroy, A. K. et al. (2021a), *PHANGS-ALMA Data Processing and Pipeline*, ApJS **255**, 19 19, arXiv: 2104.07665 [astro-ph.IM] (cit. on pp. 12, 131, 141).

Leroy, A. K. et al. (2021b), *PHANGS-ALMA: Arcsecond CO(2-1) Imaging of Nearby Star-forming Galaxies*, ApJS **257**, 43 43, arXiv: 2104.07739 [astro-ph.GA] (cit. on pp. 12, 40, 63, 131, 141, 145).

Levy, R. C. et al. (2024), *JWST Observations of Starbursts: Massive Star Clusters in the Central Starburst of M82*, ApJ **973**, L55 L55, arXiv: 2408.04135 [astro-ph.GA] (cit. on p. 142).

Lewis, S. C. et al. (2023), *Early-forming Massive Stars Suppress Star Formation and Hierarchical Cluster Assembly*, ApJ **944**, 211 211, arXiv: 2212.01465 [astro-ph.GA] (cit. on pp. 3, 107).

Li, G.-X. (2024), *Tides in clouds: control of star formation by long-range gravitational force*, MNRAS **528** L52, arXiv: 2309.16125 [astro-ph.GA] (cit. on pp. 2, 43, 53, 57).

Li, H.-B. & Henning, T. (2011), *The alignment of molecular cloud magnetic fields with the spiral arms in M33*, Nature **479** 499, arXiv: 1111.2745 [astro-ph.GA] (cit. on pp. 30, 43).

Li, S. et al. (2020), *ALMA Observations of NGC 6334S. I. Forming Massive Stars and Clusters in Subsonic and Transonic Filamentary Clouds*, ApJ **896**, 110 110, arXiv: 2003.13534 [astro-ph.GA] (cit. on pp. 43, 57, 70).

Li, S. et al. (2023), *The ALMA Survey of 70 μm Dark High-mass Clumps in Early Stages (ASHES). VIII. Dynamics of Embedded Dense Cores*, ApJ **949**, 109 109, arXiv: 2304.01718 [astro-ph.GA] (cit. on p. 25).

Li, S. et al. (2024), *Observations of high-order multiplicity in a high-mass stellar protocluster*, Nature Astronomy **8** 472, arXiv: 2401.06545 [astro-ph.GA] (cit. on p. 8).

Lim, B. et al. (2020), *The Origin of a Distributed Stellar Population in the Star-forming Region W4*, ApJ **899**, 121 121, arXiv: 2006.15262 [astro-ph.SR] (cit. on pp. 5, 102).

Lim, B. et al. (2022), *A Gaia View on the Star Formation in the Monoceros OB1 and R1 Associations*, AJ **163**, 266 266, arXiv: 2204.00444 [astro-ph.SR] (cit. on pp. 5, 102).

Lindner, R. R. et al. (2015), *Autonomous Gaussian Decomposition*, AJ **149**, 138 138, arXiv: 1409.2840 [astro-ph.IM] (cit. on pp. 19, 23, 36).

Littlefair, S. P., Naylor, T., Jeffries, R. D., Devey, C. R. & Vine, S. (2003), *Mass segregation in the young open cluster NGC 2547*, MNRAS **345** 1205, arXiv: astro-ph/0308320 [astro-ph] (cit. on p. 94).

Liu, H. B. et al. (2015), *ALMA Resolves the Spiraling Accretion Flow in the Luminous OB Cluster-forming Region G33.92+0.11*, ApJ **804**, 37 37, arXiv: 1505.04255 [astro-ph.SR] (cit. on p. 15).

Liu, H.-L. et al. (2017), *Herschel observations of the Galactic H II region RCW 79*, A&A **602**, A95 A95, arXiv: 1702.01924 [astro-ph.GA] (cit. on p. 3).

Liu, H.-L. et al. (2022), *ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions - IX. A pilot study towards IRDC G034.43+00.24 on multi-scale structures and gas kinematics*, MNRAS **511** 4480, arXiv: 2202.11307 [astro-ph.GA] (cit. on p. 2).

Liu, H.-L. et al. (2023), *Evidence of high-mass star formation through multiscale mass accretion in hub-filament-system clouds*, MNRAS **522** 3719, arXiv: 2301.03144 [astro-ph.GA] (cit. on p. 2).

Liu, M., Tan, J. C., Cheng, Y. & Kong, S. (2018), *The Core Mass Function across Galactic Environments. II. Infrared Dark Cloud Clumps*, ApJ **862**, 105 105, arXiv: 1806.02213 [astro-ph.GA] (cit. on p. 7).

Liu, S.-Y. et al. (2020), *ALMA View of the Infalling Envelope around a Massive Protostar in S255IR SMA1*, ApJ **904**, 181 181, arXiv: 2010.09199 [astro-ph.GA] (cit. on p. 22).

Liu, T. et al. (2016a), *Discovery of an Extremely Wide-angle Bipolar Outflow in AFGL 5142*, ApJ **824**, 31 31, arXiv: 1604.03548 [astro-ph.GA] (cit. on p. 2).

Liu, T. et al. (2016b), *Star Formation Laws in Both Galactic Massive Clumps and External Galaxies: Extensive Study with Dust Coninuum, HCN (4-3), and CS (7-6)*, ApJ **829**, 59 59, arXiv: 1605.08979 [astro-ph.GA] (cit. on pp. 18, 34, 73).

Liu, T. et al. (2020), *ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions - I. Survey description and a first look at G9.62+0.19*, MNRAS **496** 2790, arXiv: 2006.01549 [astro-ph.GA] (cit. on pp. 11–13).

Liu, X.-L. et al. (2021), *Kinematics and star formation toward W33: a central hub as a hub-filament system*, A&A **646**, A137 A137, arXiv: 2102.10827 [astro-ph.GA] (cit. on p. 2).

Looney, L. W., Mundy, L. G. & Welch, W. J. (2000), *Unveiling the Circumstellar Envelope and Disk: A Subarcsecond Survey of Circumstellar Structures*, ApJ **529** 477, arXiv: astro-ph/9908301 [astro-ph] (cit. on p. 8).

Lopez, L. A. et al. (2014), *The Role of Stellar Feedback in the Dynamics of H II Regions*, ApJ **795**, 121 121, arXiv: 1309.5421 [astro-ph.SR] (cit. on pp. 3, 78).

Louvet, F. et al. (2024), *ALMA-IMF: XV. Core mass function in the high-mass star formation regime*, A&A **690**, A33 A33, arXiv: 2407.18719 [astro-ph.GA] (cit. on pp. 7, 130, 132, 135).

Lu, X. et al. (2018), *Filamentary Fragmentation and Accretion in High-mass Star-forming Molecular Clouds*, ApJ **855**, 9 9, arXiv: 1801.05955 [astro-ph.GA] (cit. on p. 2).

Lu, X. et al. (2020), *ALMA Observations of Massive Clouds in the Central Molecular Zone: Jeans Fragmentation and Cluster Formation*, ApJ **894**, L14 L14, arXiv: 2004.09532 [astro-ph.GA] (cit. on p. 7).

Lumsden, S. L. et al. (2013), *The Red MSX Source Survey: The Massive Young Stellar Population of Our Galaxy*, ApJS **208**, 11 11, arXiv: 1308.0134 [astro-ph.GA] (cit. on p. 87).

Mac Low, M.-M. & Klessen, R. S. (2004), *Control of star formation by supersonic turbulence*, Reviews of Modern Physics **76** 125, arXiv: astro-ph/0301093 [astro-ph] (cit. on p. 4).

Machida, M. N. & Matsumoto, T. (2012), *Impact of protostellar outflow on star formation: effects of the initial cloud mass*, MNRAS **421** 588, arXiv: 1108.3564 [astro-ph.SR] (cit. on pp. 89, 95).

Malinen, J. et al. (2012), *Profiling filaments: comparing near-infrared extinction and submillimetre data in TMC-1*, A&A **544**, A50 A50, arXiv: 1206.4837 [astro-ph.GA] (cit. on p. 94).

Marın Pina, D. et al. (2024), *Dynamical formation of Gaia BH3 in the progenitor globular cluster of the ED-2 stream*, A&A **688**, L2 L2, arXiv: 2404.13036 [astro-ph.GA] (cit. on p. 121).

Marks, M. & Kroupa, P. (2012), *Inverse dynamical population synthesis. Constraining the initial conditions of young stellar clusters by studying their binary populations*, A&A **543**, A8 A8, arXiv: 1205.1508 [astro-ph.GA] (cit. on pp. 94, 126).

Marks, M., Kroupa, P., Dabringhausen, J. & Pawlowski, M. S. (2012), *Evidence for top-heavy stellar initial mass functions with increasing density and decreasing metallicity*, MNRAS **422** 2246, arXiv: 1202.4755 [astro-ph.GA] (cit. on p. 83).

Matzner, C. D. (2002), *On the Role of Massive Stars in the Support and Destruction of Giant Molecular Clouds*, ApJ **566** 302, arXiv: astro-ph/0110278 [astro-ph] (cit. on p. 3).

Maud, L. T. et al. (2017), *The ALMA view of W33A: a spiral filament feeding the candidate disc in MM1-Main*, MNRAS **467** L120, arXiv: 1701.06958 [astro-ph.SR] (cit. on p. 15).

Mazumdar, P. et al. (2021a), *High resolution LAsMA ^{12}CO and ^{13}CO observation of the G305 giant molecular cloud complex. II. Effect of feedback on clump properties*, A&A **656**, A101 A101, arXiv: 2109.09615 [astro-ph.GA] (cit. on pp. 4, 71).

Mazumdar, P. et al. (2021b), *High-resolution LAsMA ^{12}CO and ^{13}CO observation of the G305 giant molecular cloud complex. I. Feedback on the molecular gas*, A&A **650**, A164 A164, arXiv: 2105.11703 [astro-ph.GA] (cit. on pp. 4, 11, 71).

McKee, C. F. (1989), *Photoionization-regulated Star Formation and the Structure of Molecular Clouds*, ApJ **345** 782 (cit. on pp. 57, 69).

McKee, C. F. & Ostriker, E. C. (2007), *Theory of Star Formation*, ARA&A **45** 565, arXiv: 0707.3514 [astro-ph] (cit. on pp. 1, 2, 4).

McKee, C. F. & Tan, J. C. (2003a), *The Formation of Massive Stars from Turbulent Cores*, ApJ **585** 850, arXiv: astro-ph/0206037 [astro-ph] (cit. on pp. 2, 4).

McKee, C. F. & Tan, J. C. (2003b), *The Formation of Massive Stars from Turbulent Cores*, ApJ **585** 850, arXiv: astro-ph/0206037 [astro-ph] (cit. on p. 85).

McLeod, A. F. et al. (2015), *The Pillars of Creation revisited with MUSE: gas kinematics and high-mass stellar feedback traced by optical spectroscopy*, MNRAS **450** 1057, arXiv: 1504.03323 [astro-ph.SR] (cit. on p. 79).

Megeath, S. T., Gutermuth, R. A. & Kounkel, M. A. (2022), *Low Mass Stars as Tracers of Star and Cluster Formation*, PASP **134**, 042001 042001, arXiv: 2203.03655 [astro-ph.GA] (cit. on p. 85).

Megeath, S. T. et al. (2016), *The Spitzer Space Telescope Survey of the Orion A and B Molecular Clouds. II. The Spatial Distribution and Demographics of Dusty Young Stellar Objects*, AJ **151**, 5 5, arXiv: 1511.01202 [astro-ph.GA] (cit. on pp. 1, 5, 7, 8, 83, 87, 89, 95, 105).

Meidt, S. E. et al. (2018), *A Model for the Onset of Self-gravitation and Star Formation in Molecular Gas Governed by Galactic Forces. I. Cloud-scale Gas Motions*, ApJ **854**, 100 100, arXiv: 1712.06364 [astro-ph.GA] (cit. on p. 29).

Meidt, S. E. et al. (2020), *A Model for the Onset of Self-gravitation and Star Formation in Molecular Gas Governed by Galactic Forces. II. The Bottleneck to Collapse Set by Cloud-Environment Decoupling*, ApJ **892**, 73 73, arXiv: 2001.07459 [astro-ph.GA] (cit. on p. 29).

Mezger, P. G. & Henderson, A. P. (1967), *Galactic H II Regions. I. Observations of Their Continuum Radiation at the Frequency 5 GHz*, ApJ **147** 471 (cit. on pp. 86, 87).

Misugi, Y., Inutsuka, S.-i. & Arzoumanian, D. (2019), *An Origin for the Angular Momentum of Molecular Cloud Cores: A Prediction from Filament Fragmentation*, ApJ **881**, 11 11, arXiv: 1905.08071 [astro-ph.GA] (cit. on p. 2).

Miville-Deschênes, M.-A., Murray, N. & Lee, E. J. (2017), *Physical Properties of Molecular Clouds for the Entire Milky Way Disk*, ApJ **834**, 57 57, arXiv: 1610.05918 [astro-ph.GA] (cit. on pp. 6, 101, 107, 130, 138, 139, 148).

Mok, A., Chandar, R. & Fall, S. M. (2020), *Mass Functions of Giant Molecular Clouds and Young Star Clusters in Six Nearby Galaxies*, ApJ **893**, 135 135, arXiv: 2004.02698 [astro-ph.GA] (cit. on p. 147).

Molinari, S. et al. (2008), *The evolution of the spectral energy distribution in massive young stellar objects*, A&A **481** 345 (cit. on p. 34).

Molinari, S. et al. (2010), *Clouds, filaments, and protostars: The Herschel Hi-GAL Milky Way*, A&A **518**, L100 L100, arXiv: 1005.3317 [astro-ph.GA] (cit. on p. 1).

Molinari, S. et al. (2014), "The Milky Way as a Star Formation Engine", *Protostars and Planets VI*, ed. by H. Beuther, R. S. Klessen, C. P. Dullemond & T. Henning 125, arXiv: 1402.6196 [astro-ph.GA] (cit. on p. 3).

Molinari, S. et al. (2025), *ALMAGAL I. The ALMA evolutionary study of high-mass protocluster formation in the Galaxy. Presentation of the survey and early results*, arXiv e-prints, arXiv:2503.05555 arXiv:2503.05555, arXiv: 2503.05555 [astro-ph.GA] (cit. on p. 1).

Mor, R., Robin, A. C., Figueras, F., Roca-Fàbrega, S. & Luri, X. (2019), *Gaia DR2 reveals a star formation burst in the disc 2-3 Gyr ago*, A&A **624**, L1 L1, arXiv: 1901.07564 [astro-ph.GA] (cit. on p. 136).

Motte, F. et al. (2018a), *The unexpectedly large proportion of high-mass star-forming cores in a Galactic mini-starburst*, Nature Astronomy **2** 478, arXiv: 1804.02392 [astro-ph.GA] (cit. on pp. 7, 130).

Motte, F. et al. (2022), *ALMA-IMF. I. Investigating the origin of stellar masses: Introduction to the Large Program and first results*, A&A **662**, A8 A8, arXiv: 2112.08182 [astro-ph.GA] (cit. on p. 130).

Motte, F., Bontemps, S. & Louvet, F. (2018b), *High-Mass Star and Massive Cluster Formation in the Milky Way*, ARA&A **56** 41, arXiv: 1706.00118 [astro-ph.GA] (cit. on pp. 1, 3, 5, 6, 23, 41, 95, 141).

Motte, F., Bontemps, S. & Louvet, F. (2018c), *High-Mass Star and Massive Cluster Formation in the Milky Way*, ARA&A **56** 41, arXiv: 1706.00118 [astro-ph.GA] (cit. on p. 5).

Mužić, K. et al. (2022), *Stellar population of the Rosette Nebula and NGC 2244. Application of the probabilistic random forest*, A&A **668**, A19 A19, arXiv: 2209.13302 [astro-ph.GA] (cit. on pp. 5, 102).

Myers, P. C. (2009), *Filamentary Structure of Star-forming Complexes*, ApJ **700** 1609, arXiv: 0906.2005 [astro-ph.GA] (cit. on p. 2).

Nagakura, T., Hosokawa, T. & Omukai, K. (2009), *Star formation triggered by supernova explosions in young galaxies*, MNRAS **399** 2183, arXiv: 0907.3667 [astro-ph.CO] (cit. on p. 73).

Ngoc, N. B. et al. (2023), *B-fields and Dust in Interstellar Filaments Using Dust Polarization (BALLAD-POL). I. The Massive Filament G11.11-0.12 Observed by SOFIA/HAWC+*, ApJ **953**, 66 66, arXiv: 2302.10543 [astro-ph.GA] (cit. on p. 30).

Nony, T. et al. (2021), *Mass segregation and sequential star formation in NGC 2264 revealed by Herschel*, A&A **645**, A94 A94, arXiv: 2011.05939 [astro-ph.GA] (cit. on p. 94).

Nony, T. et al. (2023), *ALMA-IMF. V. Prestellar and protostellar core populations in the W43 cloud complex*, A&A **674**, A75 A75, arXiv: 2301.07238 [astro-ph.GA] (cit. on pp. 7, 8, 130, 131).

O'Neill, T. J., Cosentino, G., Tan, J. C., Cheng, Y. & Liu, M. (2021), *The Core Mass Function across Galactic Environments. III. Massive Protoclusters*, ApJ **916**, 45 45, arXiv: 2104.08861 [astro-ph.GA] (cit. on p. 7).

O'Neill, T. J., Zucker, C., Goodman, A. A. & Edenhofer, G. (2024), *The Local Bubble Is a Local Chimney: A New Model from 3D Dust Mapping*, ApJ **973**, 136 136, arXiv: 2403.04961 [astro-ph.GA] (cit. on p. 52).

Oey, M. S. & Clarke, C. J. (2005), *Statistical Confirmation of a Stellar Upper Mass Limit*, ApJ **620** L43, arXiv: astro-ph/0501135 [astro-ph] (cit. on p. 6).

Offner, S. S. R. et al. (2023), "The Origin and Evolution of Multiple Star Systems", *Protostars and Planets VII*, ed. by S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida & M. Tamura, vol. 534, Astronomical Society of the Pacific Conference Series 275, arXiv: 2203.10066 [astro-ph.SR] (cit. on p. 8).

Oh, S. & Kroupa, P. (2016), *Dynamical ejections of massive stars from young star clusters under diverse initial conditions*, A&A **590**, A107 A107, arXiv: 1604.00006 [astro-ph.GA] (cit. on p. 94).

Oh, S. & Kroupa, P. (2018), *Very massive stars in not so massive clusters*, MNRAS **481** 153, arXiv: 1809.04596 [astro-ph.GA] (cit. on pp. 94, 117).

Oh, S., Kroupa, P. & Pflamm-Altenburg, J. (2015), *Dependency of Dynamical Ejections of O Stars on the Masses of Very Young Star Clusters*, ApJ **805**, 92 92, arXiv: 1503.08827 [astro-ph.GA] (cit. on pp. 94, 95).

Ohashi, N., Hayashi, M., Ho, P. T. P. & Momose, M. (1997), *Interferometric Imaging of IRAS 04368+2557 in the L1527 Molecular Cloud Core: A Dynamically Infalling Envelope with Rotation*, ApJ **475** 211 (cit. on p. 22).

Olguin, F. A. et al. (2021), *Digging into the Interior of Hot Cores with ALMA (DIHCA). I. Dissecting the High-mass Star-forming Core G335.579-0.292 MM1*, ApJ **909**, 199 199, arXiv: 2101.08284 [astro-ph.GA] (cit. on p. 8).

Olguin, F. A. et al. (2022), *Digging into the Interior of Hot Cores with ALMA (DIHCA). II. Exploring the Inner Binary (Multiple) System Embedded in G335 MM1 ALMA1*, ApJ **929**, 68 68, arXiv: 2203.04333 [astro-ph.GA] (cit. on p. 8).

Olivier, G. M. et al. (2021), *Evolution of Stellar Feedback in H II Regions*, ApJ **908**, 68 68, arXiv: 2009.10079 [astro-ph.GA] (cit. on pp. 3, 78).

Pabst, C. et al. (2019), *Disruption of the Orion molecular core 1 by wind from the massive star θ^1 Orionis C*, Nature **565** 618, arXiv: 1901.04221 [astro-ph.GA] (cit. on pp. 3, 78).

Padoan, P., Pan, L., Juvela, M., Haugbølle, T. & Nordlund, Å. (2020a), *The Origin of Massive Stars: The Inertial-inflow Model*, ApJ **900**, 82 82, arXiv: 1911.04465 [astro-ph.GA] (cit. on pp. 2, 4).

Padoan, P., Pan, L., Juvela, M., Haugbølle, T. & Nordlund, Å. (2020b), *The Origin of Massive Stars: The Inertial-inflow Model*, ApJ **900**, 82 82, arXiv: 1911.04465 [astro-ph.GA] (cit. on pp. 14, 41).

Pang, X. et al. (2021), *3D Morphology of Open Clusters in the Solar Neighborhood with Gaia EDR 3: Its Relation to Cluster Dynamics*, ApJ **912**, 162 162, arXiv: 2102.10508 [astro-ph.GA] (cit. on p. 96).

Pattle, K. et al. (2015), *The JCMT Gould Belt Survey: first results from the SCUBA-2 observations of the Ophiuchus molecular cloud and a virial analysis of its prestellar core population*, MNRAS **450** 1094, arXiv: 1502.05858 [astro-ph.GA] (cit. on p. 57).

Pavlik, V., Kroupa, P. & Šubr, L. (2019), *Do star clusters form in a completely mass-segregated way?*, A&A **626**, A79 A79, arXiv: 1905.09289 [astro-ph.GA] (cit. on p. 94).

Peretto, N. et al. (2013), *Global collapse of molecular clouds as a formation mechanism for the most massive stars*, A&A **555**, A112 A112, arXiv: 1307.2590 [astro-ph.GA] (cit. on p. 2).

Peretto, N. et al. (2023), *Star cluster progenitors are dynamically decoupled from their parent molecular clouds*, MNRAS **525** 2935, arXiv: 2305.02701 [astro-ph.GA] (cit. on pp. 5, 73, 107).

Peters, T., Banerjee, R., Klessen, R. S. & Mac Low, M.-M. (2011), *The Interplay of Magnetic Fields, Fragmentation, and Ionization Feedback in High-mass Star Formation*, ApJ **729**, 72 72, arXiv: 1010.5905 [astro-ph.SR] (cit. on p. 6).

Pfalzner, S. et al. (2016), *Observational constraints on star cluster formation theory. I. The mass-radius relation*, A&A **586**, A68 A68, arXiv: 1512.00334 [astro-ph.SR] (cit. on pp. 91, 103–105, 126).

Pillai, T. G. S. et al. (2020), *Magnetized filamentary gas flows feeding the young embedded cluster in Serpens South*, Nature Astronomy **4** 1195, arXiv: 2009.14100 [astro-ph.GA] (cit. on p. 2).

Plummer, H. C. (1911), *On the problem of distribution in globular star clusters*, MNRAS **71** 460 (cit. on p. 126).

Plunkett, A. L. et al. (2018), *Distribution of Serpens South protostars revealed with ALMA*, A&A **615**, A9 A9, arXiv: 1804.02405 [astro-ph.SR] (cit. on p. 94).

Pokhrel, R. et al. (2020), *Star-Gas Surface Density Correlations in 12 Nearby Molecular Clouds. I. Data Collection and Star-sampled Analysis*, ApJ **896**, 60 60, arXiv: 2005.05466 [astro-ph.GA] (cit. on p. 85).

Portegies Zwart, S. F., McMillan, S. L. W. & Gieles, M. (2010a), *Young Massive Star Clusters*, Annual Review of Astronomy and Astrophysics **48** 431 (cit. on p. 5).

Portegies Zwart, S. F., McMillan, S. L. W. & Gieles, M. (2010b), *Young Massive Star Clusters*, ARA&A **48** 431, arXiv: 1002.1961 [astro-ph.GA] (cit. on pp. 90, 94).

Pouteau, Y. et al. (2022), *ALMA-IMF. III. Investigating the origin of stellar masses: top-heavy core mass function in the W43-MM2&MM3 mini-starburst*, A&A **664**, A26 A26, arXiv: 2203.03276 [astro-ph.GA] (cit. on pp. 7, 130, 131).

Purcell, C. R. et al. (2013), *The Coordinated Radio and Infrared Survey for High-mass Star Formation. II. Source Catalog*, ApJS **205**, 1 1, arXiv: 1211.7116 [astro-ph.GA] (cit. on p. 82).

Rahner, D., Pellegrini, E. W., Glover, S. C. O. & Klessen, R. S. (2017), *Winds and radiation in unison: a new semi-analytic feedback model for cloud dissolution*, MNRAS **470** 4453, arXiv: 1704.04240 [astro-ph.GA] (cit. on pp. 3, 96).

Ramırez-Galeano, L., Ballesteros-Paredes, J., Smith, R. J., Camacho, V. & Zamora-Avilés, M. (2022), *Why most molecular clouds are gravitationally dominated*, MNRAS **515** 2822, arXiv: 2206.09187 [astro-ph.GA] (cit. on pp. 2, 29, 43).

Rastello, S. et al. (2023), *Dynamical formation of Gaia BH1 in a young star cluster*, MNRAS **526** 740, arXiv: 2306.14679 [astro-ph.SR] (cit. on p. 121).

Rawat, V. et al. (2024), *Understanding the relative importance of magnetic field, gravity, and turbulence in star formation at the hub of the giant molecular cloud G148.24+00.41*, MNRAS **528** 1460, arXiv: 2401.05310 [astro-ph.GA] (cit. on p. 30).

Rezaei Kh., S., Bailer-Jones, C. A. L., Hogg, D. W. & Schultheis, M. (2018), *Detection of the Milky Way spiral arms in dust from 3D mapping*, A&A **618**, A168 A168, arXiv: 1808.00015 [astro-ph.GA] (cit. on p. 57).

Rezaei Kh., S. & Kainulainen, J. (2022), *Three-dimensional Shape Explains Star Formation Mystery of California and Orion A*, ApJ **930**, L22 L22, arXiv: 2204.10892 [astro-ph.GA] (cit. on p. 100).
Riener, M. et al. (2019), *GAUSSPY+: A fully automated Gaussian decomposition package for emission line spectra*, A&A **628**, A78 A78, arXiv: 1906.10506 [astro-ph.IM] (cit. on pp. 19, 23, 36).
Rogers, H. & Pittard, J. M. (2013), *Feedback from winds and supernovae in massive stellar clusters - I. Hydrodynamics*, MNRAS **431** 1337, arXiv: 1302.2443 [astro-ph.SR] (cit. on p. 3).
Romine, G., Feigelson, E. D., Getman, K. V., Kuhn, M. A. & Povich, M. S. (2016), *Young Stellar Populations in MYStIX Star-forming Regions: Candidate Protostars*, ApJ **833**, 193 193, arXiv: 1609.06650 [astro-ph.SR] (cit. on p. 83).
Röser, S., Schilbach, E., Piskunov, A. E., Kharchenko, N. V. & Scholz, R. .-. (2011), *A deep all-sky census of the Hyades*, A&A **531**, A92 A92, arXiv: 1105.6093 [astro-ph.GA] (cit. on pp. 94, 126).
Röser, S. & Schilbach, E. (2019), *Praesepe (NGC 2632) and its tidal tails*, A&A **627**, A4 A4, arXiv: 1903.08610 [astro-ph.SR] (cit. on pp. 94, 126).
Rosolowsky, E. W., Pineda, J. E., Kauffmann, J. & Goodman, A. A. (2008), *Structural Analysis of Molecular Clouds: Dendrograms*, ApJ **679** 1338, arXiv: 0802.2944 [astro-ph] (cit. on pp. 25, 69).
Rosolowsky, E. et al. (2021), *Giant molecular cloud catalogues for PHANGS-ALMA: methods and initial results*, MNRAS **502** 1218, arXiv: 2101.04697 [astro-ph.GA] (cit. on p. 6).
Rubin, R. H. (1968), *A Discussion of the Sizes and Excitation of H II Regions*, ApJ **154** 391 (cit. on pp. 86, 87).
Rugel, M. R. et al. (2019), *Feedback in W49A diagnosed with radio recombination lines and models*, A&A **622**, A48 A48, arXiv: 1812.00758 [astro-ph.GA] (cit. on p. 4).
Sadaghiani, M. et al. (2020), *Physical properties of the star-forming clusters in NGC 6334. A study of the continuum dust emission with ALMA*, A&A **635**, A2 A2, arXiv: 1911.06579 [astro-ph.SR] (cit. on p. 7).
Sakai, N. et al. (2014), *Change in the chemical composition of infalling gas forming a disk around a protostar*, Nature **507** 78 (cit. on p. 22).
Sana, H. et al. (2012), *Binary Interaction Dominates the Evolution of Massive Stars*, Science **337** 444, arXiv: 1207.6397 [astro-ph.SR] (cit. on p. 95).
Sánchez-Monge, Á. et al. (2013), *Different evolutionary stages in massive star formation. Centimeter continuum and H_2O maser emission with ATCA*, A&A **550**, A21 A21, arXiv: 1211.0847 [astro-ph.GA] (cit. on pp. 86, 87).
Sanhueza, P. et al. (2012), *Chemistry in Infrared Dark Cloud Clumps: A Molecular Line Survey at 3 mm*, ApJ **756**, 60 60, arXiv: 1206.6500 [astro-ph.GA] (cit. on p. 74).
Sanhueza, P. et al. (2019), *The ALMA Survey of 70 μm Dark High-mass Clumps in Early Stages (ASHES). I. Pilot Survey: Clump Fragmentation*, ApJ **886**, 102 102, arXiv: 1909.07985 [astro-ph.GA] (cit. on pp. 7, 8).
Sanhueza, P. et al. (2021), *Gravity-driven Magnetic Field at 1000 au Scales in High-mass Star Formation*, ApJ **915**, L10 L10, arXiv: 2106.03866 [astro-ph.GA] (cit. on p. 15).
Sanna, A. et al. (2019), *Discovery of a sub-Keplerian disk with jet around a 20 $M_\odot$ young star. ALMA observations of G023.01-00.41*, A&A **623**, A77 A77, arXiv: 1805.09842 [astro-ph.SR] (cit. on p. 22).
Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M. & Molinari, S. (1996), *An evolutionary diagram for young stellar objects.*, A&A **309** 827 (cit. on p. 34).
Schinnerer, E. & Leroy, A. K. (2024), *Molecular Gas and the Star-Formation Process on Cloud Scales in Nearby Galaxies*, ARA&A **62** 369, arXiv: 2403.19843 [astro-ph.GA] (cit. on p. 1).
Schneider, N. et al. (2012), *Cluster-formation in the Rosette molecular cloud at the junctions of filaments*, A&A **540**, L11 L11, arXiv: 1203.6472 [astro-ph.GA] (cit. on p. 2).

Schuller, F. et al. (2009), *ATLASGAL - The APEX telescope large area survey of the galaxy at 870 μm*, A&A **504** 415, arXiv: 0903.1369 [astro-ph.GA] (cit. on p. 1).

Schulz, C., Pflamm-Altenburg, J. & Kroupa, P. (2015), *Mass distributions of star clusters for different star formation histories in a galaxy cluster environment*, A&A **582**, A93 A93, arXiv: 1507.00860 [astro-ph.GA] (cit. on pp. 6, 83, 147).

Seifried, D. et al. (2020), *From parallel to perpendicular - On the orientation of magnetic fields in molecular clouds*, MNRAS **497** 4196, arXiv: 2003.00017 [astro-ph.GA] (cit. on pp. 30, 43).

Shetty, R. & Ostriker, E. C. (2006), *Global Modeling of Spur Formation in Spiral Galaxies*, ApJ **647** 997, arXiv: astro-ph/0605142 [astro-ph] (cit. on p. 100).

Shimoikura, T., Dobashi, K., Matsumoto, T. & Nakamura, F. (2016), *Discovery of Infalling Motion with Rotation of the Cluster-forming Clump S235AB and Its Implication to the Clump Structures*, ApJ **832**, 205 205, arXiv: 1610.00834 [astro-ph.GA] (cit. on p. 22).

Shimoikura, T., Dobashi, K., Nakamura, F., Matsumoto, T. & Hirota, T. (2018), *A Statistical Study of Massive Cluster-forming Clumps*, ApJ **855**, 45 45, arXiv: 1802.00589 [astro-ph.GA] (cit. on p. 22).

Shimoikura, T. et al. (2022), *Cluster Formation in GGD 12-15: Infall Motion with Rotation of the Natal Clump*, ApJ **928**, 76 76, arXiv: 2202.12492 [astro-ph.GA] (cit. on p. 22).

Smith, M. C., Sijacki, D. & Shen, S. (2018), *Supernova feedback in numerical simulations of galaxy formation: separating physics from numerics*, MNRAS **478** 302, arXiv: 1709.03515 [astro-ph.GA] (cit. on p. 3).

Smith, R. J., Glover, S. C. O., Klessen, R. S. & Fuller, G. A. (2016), *On the nature of star-forming filaments - II. Subfilaments and velocities*, MNRAS **455** 3640, arXiv: 1509.03321 [astro-ph.GA] (cit. on p. 2).

Smith, R. J., Longmore, S. & Bonnell, I. (2009), *The simultaneous formation of massive stars and stellar clusters*, MNRAS **400** 1775, arXiv: 0908.3910 [astro-ph.SR] (cit. on p. 5).

Solomon, P. M., Rivolo, A. R., Barrett, J. & Yahil, A. (1987), *Mass, Luminosity, and Line Width Relations of Galactic Molecular Clouds*, ApJ **319** 730 (cit. on p. 4).

Spitzer, L. (1978), *Physical processes in the interstellar medium* (cit. on pp. 43, 57).

Spurzem, R., Wang, L. & Kroupa, P. (2018), *The Dynamical Evolution of Star Clusters: A Brief History and Recent Developments*, Journal of Physics: Conference Series **990** 012041 (cit. on p. 5).

Stark, A. A. & Blitz, L. (1978), *On the masses of giant molecular cloud complexes*, ApJ **225** L15 (cit. on pp. 59, 62).

Stark, A. A. & Lee, Y. (2005), *The Scale Height of Giant Molecular Clouds Is Less than That of Smaller Clouds*, ApJ **619** L159, arXiv: astro-ph/0403631 [astro-ph] (cit. on p. 63).

Stephens, I. W. et al. (2016), *Linking Dense Gas from the Milky Way to External Galaxies*, ApJ **824**, 29 29, arXiv: 1603.09341 [astro-ph.GA] (cit. on p. 34).

Stephens, I. W. et al. (2017), *Stellar Clusterings around "Isolated" Massive YSOs in the LMC*, ApJ **834**, 94 94, arXiv: 1609.04399 [astro-ph.GA] (cit. on p. 93).

Stephens, I. W. et al. (2022), *The Magnetic Field in the Milky Way Filamentary Bone G47*, ApJ **926**, L6 L6, arXiv: 2201.11933 [astro-ph.GA] (cit. on p. 30).

Sun, J. et al. (2018), *Cloud-scale Molecular Gas Properties in 15 Nearby Galaxies*, ApJ **860**, 172 172, arXiv: 1805.00937 [astro-ph.GA] (cit. on p. 73).

Sun, J. et al. (2020), *Dynamical Equilibrium in the Molecular ISM in 28 Nearby Star-forming Galaxies*, ApJ **892**, 148 148, arXiv: 2002.08964 [astro-ph.GA] (cit. on p. 43).

Sun, J. et al. (2022), *Molecular Cloud Populations in the Context of Their Host Galaxy Environments: A Multiwavelength Perspective*, AJ **164**, 43 43, arXiv: 2206.07055 [astro-ph.GA] (cit. on p. 63).

Swiggum, C. et al. (2021), *Evidence for Radial Expansion at the Core of the Orion Complex with Gaia EDR3*, ApJ **917**, 21 21, arXiv: 2101.10380 [astro-ph.GA] (cit. on pp. 5, 102).

Tanikawa, A., Cary, S., Shikauchi, M., Wang, L. & Fujii, M. S. (2024), *Compact binary formation in open star clusters - I. High formation efficiency of Gaia BHs and their multiplicities*, MNRAS **527** 4031, arXiv: 2303.05743 [astro-ph.GA] (cit. on p. 121).

Tanikawa, A., Yoshida, T., Kinugawa, T., Takahashi, K. & Umeda, H. (2020), *Fitting formulae for evolution tracks of massive stars under extreme metal-poor environments for population synthesis calculations and star cluster simulations*, MNRAS **495** 4170, arXiv: 1906.06641 [astro-ph.HE] (cit. on p. 96).

Tanikawa, A. et al. (2023), *Search for a Black Hole Binary in Gaia DR3 Astrometric Binary Stars with Spectroscopic Data*, ApJ **946**, 79 79, arXiv: 2209.05632 [astro-ph.SR] (cit. on p. 121).

Thilker, D. A. et al. (2022), *PHANGS-HST: new methods for star cluster identification in nearby galaxies*, MNRAS **509** 4094, arXiv: 2106.13366 [astro-ph.GA] (cit. on pp. 130, 142).

Thilliez, E., Maddison, S. T., Hughes, A. & Wong, T. (2014), *Tidal Stability of Giant Molecular Clouds in the Large Magellanic Cloud*, PASA **31**, e003 e003, arXiv: 1310.6568 [astro-ph.GA] (cit. on pp. 29, 61, 62).

Thompson, M. A., Urquhart, J. S., Moore, T. J. T. & Morgan, L. K. (2012), *The statistics of triggered star formation: an overdensity of massive young stellar objects around Spitzer bubbles*, MNRAS **421** 408, arXiv: 1111.0972 [astro-ph.GA] (cit. on p. 3).

Thompson, R. I. (1984), *Lyman and Balmer continuum ionization in zero-age main-sequence stars: applications to the line excess phenomenon.*, ApJ **283** 165 (cit. on pp. 86, 124).

Tobin, J. J. et al. (2022), *The VLA/ALMA Nascent Disk And Multiplicity (VANDAM) Survey of Orion Protostars. V. A Characterization of Protostellar Multiplicity*, ApJ **925**, 39 39, arXiv: 2111.05801 [astro-ph.GA] (cit. on p. 8).

Traficante, A. et al. (2018a), *Massive 70 μm quiet clumps - II. Non-thermal motions driven by gravity in massive star formation?*, MNRAS **473** 4975, arXiv: 1710.04904 [astro-ph.GA] (cit. on pp. 4, 25).

Traficante, A. et al. (2018b), *Testing the Larson relations in massive clumps*, MNRAS **477** 2220, arXiv: 1803.08929 [astro-ph.GA] (cit. on p. 4).

Turner, J. A. et al. (2022), *PHANGS: constraining star formation time-scales using the spatial correlations of star clusters and giant molecular clouds*, MNRAS **516** 4612, arXiv: 2209.02872 [astro-ph.GA] (cit. on p. 142).

Urquhart, J. S. et al. (2013), *ATLASGAL - properties of compact H II regions and their natal clumps*, MNRAS **435** 400, arXiv: 1307.4105 [astro-ph.GA] (cit. on p. 87).

Urquhart, J. S. et al. (2018), *ATLASGAL - properties of a complete sample of Galactic clumps*, MNRAS **473** 1059, arXiv: 1709.00392 [astro-ph.GA] (cit. on pp. 18, 34, 73, 106, 132).

Urquhart, J. S. et al. (2022), *ATLASGAL - evolutionary trends in high-mass star formation*, MNRAS **510** 3389, arXiv: 2111.12816 [astro-ph.GA] (cit. on pp. 1, 6, 34, 82–84, 92, 95, 105–107, 130, 132, 133, 138, 139, 148, 153).

Utreras, J. et al. (2020), *When Gas Dynamics Decouples from Galactic Rotation: Characterizing ISM Circulation in Disk Galaxies*, ApJ **892**, 94 94, arXiv: 2003.01131 [astro-ph.GA] (cit. on p. 29).

Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. .-., Ballesteros-Paredes, J. & Klessen, R. S. (2009), *High- and Low-Mass Star-Forming Regions from Hierarchical Gravitational Fragmentation. High Local Star Formation Rates with Low Global Efficiencies*, ApJ **707** 1023, arXiv: 0904.4515 [astro-ph.GA] (cit. on pp. 2, 4).

Vázquez-Semadeni, E., González-Samaniego, A. & Colın, P. (2017), *Hierarchical star cluster assembly in globally collapsing molecular clouds*, MNRAS **467** 1313, arXiv: 1611.00088 [astro-ph.GA] (cit. on pp. 2, 4).

Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C. & Zamora-Avilés, M. (2019), *Global hierarchical collapse in molecular clouds. Towards a comprehensive scenario*, MNRAS **490** 3061, arXiv: 1903.11247 [astro-ph.GA] (cit. on pp. 2, 4, 6, 14, 23, 41, 50, 95, 141).

von Steiger, R. & Zurbuchen, T. H. (2016), *Solar Metallicity Derived from in situ Solar Wind Composition*, ApJ **816**, 13 13 (cit. on p. 94).

Wainer, T. M. et al. (2022), *The Panchromatic Hubble Andromeda Treasury: Triangulum Extended Region (PHATTER). III. The Mass Function of Young Stellar Clusters in M33*, ApJ **928**, 15 15, arXiv: 2201.04161 [astro-ph.GA] (cit. on p. 147).

Wang, L., Iwasawa, M., Nitadori, K. & Makino, J. (2020), *PETAR: a high-performance N-body code for modelling massive collisional stellar systems*, MNRAS **497** 536, arXiv: 2006.16560 [astro-ph.IM] (cit. on p. 96).

Wang, L., Kroupa, P. & Jerabkova, T. (2019), *Complete ejection of OB stars from very young star clusters and the formation of multiple populations*, MNRAS **484** 1843, arXiv: 1810.07697 [astro-ph.SR] (cit. on pp. 94, 133).

Wang, P., Li, Z.-Y., Abel, T. & Nakamura, F. (2010), *Outflow Feedback Regulated Massive Star Formation in Parsec-Scale Cluster-Forming Clumps*, ApJ **709** 27, arXiv: 0908.4129 [astro-ph.SR] (cit. on p. 2).

Watkins, E. J., Peretto, N., Marsh, K. & Fuller, G. A. (2019a), *Feedback from OB stars on their parent cloud: gas exhaustion rather than gas ejection*, A&A **628**, A21 A21, arXiv: 1906.09275 [astro-ph.GA] (cit. on p. 4).

Watkins, E. J., Peretto, N., Marsh, K. & Fuller, G. A. (2019b), *Feedback from OB stars on their parent cloud: gas exhaustion rather than gas ejection*, A&A **628**, A21 A21, arXiv: 1906.09275 [astro-ph.GA] (cit. on p. 73).

Watkins, E. J. et al. (2023), *Quantifying the energetics of molecular superbubbles in PHANGS galaxies*, A&A **676**, A67 A67, arXiv: 2302.03699 [astro-ph.GA] (cit. on p. 7).

Weidner, C. & Kroupa, P. (2004), *Evidence for a fundamental stellar upper mass limit from clustered star formation*, MNRAS **348** 187, arXiv: astro-ph/0310860 [astro-ph] (cit. on p. 6).

Weidner, C. & Kroupa, P. (2005), *The Variation of Integrated Star Initial Mass Functions among Galaxies*, ApJ **625** 754, arXiv: astro-ph/0502525 [astro-ph] (cit. on p. 148).

Weidner, C., Kroupa, P. & Bonnell, I. A. D. (2010), *The relation between the most-massive star and its parental star cluster mass*, MNRAS **401** 275, arXiv: 0909.1555 [astro-ph.SR] (cit. on p. 6).

Weidner, C., Kroupa, P. & Larsen, S. S. (2004), *Implications for the formation of star clusters from extragalactic star formation rates*, MNRAS **350** 1503, arXiv: astro-ph/0402631 [astro-ph] (cit. on p. 148).

Weidner, C., Kroupa, P. & Pflamm-Altenburg, J. (2013a), *The m_{max}-M_{ecl} relation, the IMF and IGIMF: probabilistically sampled functions*, MNRAS **434** 84, arXiv: 1306.1229 [astro-ph.GA] (cit. on pp. 86, 88, 93, 94, 110, 117, 120).

Weidner, C., Kroupa, P., Pflamm-Altenburg, J. & Vazdekis, A. (2013b), *The galaxy-wide initial mass function of dwarf late-type to massive early-type galaxies*, MNRAS **436** 3309, arXiv: 1309.6634 [astro-ph.CO] (cit. on p. 148).

Weidner, C. & Kroupa, P. (2006), *The maximum stellar mass, star-cluster formation and composite stellar populations*, MNRAS **365** 1333, arXiv: astro-ph/0511331 [astro-ph] (cit. on pp. 6, 86, 93).

Weis, M., Walch, S., Seifried, D. & Ganguly, S. (2022), *The Virial Balance of Molecular Clumps and Cores in Colliding Magnetized Flows*, arXiv e-prints, arXiv:2208.11705 arXiv:2208.11705, arXiv: 2208.11705 [astro-ph.GA] (cit. on p. 70).

Wells, M. R. A. et al. (2022), *ATLASGAL - star forming efficiencies and the Galactic star formation rate*, MNRAS **516** 4245, arXiv: 2208.11121 [astro-ph.GA] (cit. on pp. 1, 89).

Whitmore, B. C. et al. (2021), *Star cluster classification in the PHANGS-HST survey: Comparison between human and machine learning approaches*, MNRAS **506** 5294, arXiv: 2107.13049 [astro-ph.GA] (cit. on pp. 130, 142).

Williams, G. M., Peretto, N., Avison, A., Duarte-Cabral, A. & Fuller, G. A. (2018), *Gravity drives the evolution of infrared dark hubs: JVLA observations of SDC13*, A&A **613**, A11 A11, arXiv: 1801.07253 [astro-ph.GA] (cit. on p. 2).

Williams, T. G. et al. (2024), *PHANGS-JWST: Data-processing Pipeline and First Full Public Data Release*, ApJS **273**, 13 13, arXiv: 2401.15142 [astro-ph.GA] (cit. on p. 131).

Wright, N. J. et al. (2019), *The Gaia-ESO Survey: asymmetric expansion of the Lagoon Nebula cluster NGC 6530 from GES and Gaia DR2*, MNRAS **486** 2477, arXiv: 1903.12176 [astro-ph.SR] (cit. on pp. 5, 102).

Wright, N. J. et al. (2024), *The Gaia-ESO Survey: 3D dynamics of young groups and clusters from GES and Gaia EDR3*, MNRAS **533** 705, arXiv: 2311.08358 [astro-ph.GA] (cit. on pp. 5, 102).

Wünsch, R., Dale, J. E., Palouš, J. & Whitworth, A. P. (2010), *The fragmentation of expanding shells - II. Thickness matters*, MNRAS **407** 1963, arXiv: 1005.4399 [astro-ph.GA] (cit. on p. 96).

Xu, F.-W. et al. (2023), *ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions - XV. Steady accretion from global collapse to core feeding in massive hub-filament system SDC335*, MNRAS **520** 3259, arXiv: 2301.01895 [astro-ph.GA] (cit. on p. 2).

Xu, F. et al. (2024), *The ALMA Survey of Star Formation and Evolution in Massive Protoclusters with Blue Profiles (ASSEMBLE): Core Growth, Cluster Contraction, and Primordial Mass Segregation*, ApJS **270**, 9 9, arXiv: 2309.14684 [astro-ph.GA] (cit. on p. 94).

Yan, Z., Jerabkova, T. & Kroupa, P. (2017), *The optimally sampled galaxy-wide stellar initial mass function. Observational tests and the publicly available GalIMF code*, A&A **607**, A126 A126, arXiv: 1707.04260 [astro-ph.GA] (cit. on pp. 6, 7, 83, 132, 147).

Yan, Z., Jerabkova, T. & Kroupa, P. (2023), *The most massive stars in very young star clusters with a limited mass: Evidence favours significant self-regulation in the star formation processes*, A&A **670**, A151 A151, arXiv: 2211.13244 [astro-ph.GA] (cit. on pp. 6, 83–88, 93, 110, 117, 120, 147).

Yang, D. et al. (2023), *Direct Observational Evidence of the Multi-scale, Dynamical Mass Accretion Toward a High-mass Star-forming Hub-filament System*, ApJ **953**, 40 40, arXiv: 2306.10332 [astro-ph.GA] (cit. on p. 2).

Yuan, J. et al. (2018), *High-mass Star Formation through Filamentary Collapse and Clump-fed Accretion in G22*, ApJ **852**, 12 12, arXiv: 1711.08951 [astro-ph.GA] (cit. on p. 2).

Zavagno, A. et al. (2006), *Triggered massive-star formation on the borders of Galactic H II regions. II. Evidence for the collect and collapse process around RCW 79*, A&A **446** 171, arXiv: astro-ph/0509289 [astro-ph] (cit. on p. 3).

Zhang, C. et al. (2023), *ATOMS: ALMA three-millimetre observations of massive star-forming regions - XIV. Properties of resolved ultra-compact H II regions*, MNRAS **520** 3245, arXiv: 2301.01876 [astro-ph.GA] (cit. on pp. 19, 22).

Zhang, Q., Wang, K., Lu, X. & Jiménez-Serra, I. (2015), *Fragmentation of Molecular Clumps and Formation of a Protocluster*, ApJ **804**, 141 141, arXiv: 1503.03017 [astro-ph.SR] (cit. on p. 2).

Zhang, X., Green, G. M. & Rix, H.-W. (2023), *Parameters of 220 million stars from Gaia BP/RP spectra*, MNRAS **524** 1855, arXiv: 2303.03420 [astro-ph.SR] (cit. on p. 52).

Zhang, Y. et al. (2022), *Massive Protostars in a Protocluster - A Multi-scale ALMA View of G35.20-0.74N.*, ApJ **936**, 68 68, arXiv: 2207.11320 [astro-ph.GA] (cit. on p. 94).

Zhou, J. W. (2025), *External tides: an important driver of velocity dispersion in molecular clouds*, arXiv: 2503.16937 [astro-ph.GA], URL: https://arxiv.org/abs/2503.16937 (cit. on pp. 52, 57, 59).

Zhou, J. W. & Dib, S. (2025), *An examination of large-scale galactic effects on molecular cloud properties in NGC 628: the significant impact of tidal effects from neighbouring material on the evolution of molecular clouds*, MNRAS **537** 2232, arXiv: 2501.18249 [astro-ph.GA] (cit. on pp. 52, 60, 63).

Zhou, J. W., Dib, S. & Davis, T. A. (2024a), *Molecular clouds as hubs in spiral galaxies: gas inflow and evolutionary sequence*, MNRAS **534** 683, arXiv: 2409.03696 [astro-ph.GA] (cit. on pp. 6, 11, 36, 44, 46, 51, 60, 63, 65, 141, 142, 144, 145).

Zhou, J. W., Dib, S. & Kroupa, P. (2024b), *The Star Formation Histories, Star Formation Efficiencies and Ionizing Sources of ATLASGAL Clumps with H II Regions*, PASP **136**, 094302 094302, arXiv: 2409.03234 [astro-ph.GA] (cit. on pp. 6, 82, 83, 86, 88, 93, 96, 105, 106, 120, 132, 135, 149).

Zhou, J. W., Kroupa, P. & Dib, S. (2024c), *The post-gas expulsion coalescence of embedded clusters as an origin of open clusters (merger)*, submitted to MNRAS (cit. on p. 109).

Zhou, J. W., Kroupa, P. & Dib, S. (2024d), *Physical Properties of Embedded Clusters in ATLASGAL Clumps with H II Regions*, PASP **136**, 094301 094301, arXiv: 2408.09867 [astro-ph.GA] (cit. on pp. 6, 82, 83, 85–87, 89, 95, 105, 106, 125, 132).

Zhou, J. W., Kroupa, P. & Dib, S. (2025), *Inversely synthesizing the core mass function of high-mass star-forming regions from the canonical initial mass function*, A&A **695**, L17 L17, arXiv: 2503.12251 [astro-ph.GA] (cit. on pp. 130, 131, 135).

Zhou, J. W. & Li, G.-X. (2025), *Evolution of molecular clouds on galaxy-cloud scale revealed by gravitational network analysis: high-mass clouds may deplete nearby gas via accretion or merging*, MNRAS **537** 2630, arXiv: 2501.18241 [astro-ph.GA] (cit. on pp. 6, 11, 43).

Zhou, J. W. et al. (2023), *High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation of the G333 giant molecular cloud complex. I. Evidence for gravitational acceleration in hub-filament systems*, A&A **676**, A69 A69, arXiv: 2305.12573 [astro-ph.GA] (cit. on pp. 2, 6, 11, 15, 19, 22, 23, 25, 27, 30, 32, 34, 36, 41, 49, 68–70, 73, 74, 77, 100, 107, 132, 141).

Zhou, J. W. et al. (2024e), *Feedback from protoclusters does not significantly change the kinematic properties of the embedded dense gas structures*, A&A **682**, A173 A173 (cit. on pp. 7, 25, 27, 36, 49, 53, 68, 74, 78, 90, 93).

Zhou, J. W. et al. (2024f), *Gas inflows from cloud to core scales in G332.83-0.55: Hierarchical hub-filament structures and tide-regulated gravitational collapse*, A&A **686**, A146 A146, arXiv: 2403.13442 [astro-ph.GA] (cit. on pp. 2, 5, 6, 11, 19, 35, 43, 52, 53, 55, 59, 63, 73, 95, 132, 141).

Zhou, J. W. et al. (2024g), *High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation of the G333 giant molecular cloud complex. II. Survival and gravitational collapse of dense gas structures under feedback*, A&A **682**, A128 A128, arXiv: 2309.04260 [astro-ph.GA] (cit. on pp. 5, 25, 30, 31, 36, 43, 49, 55–57, 68, 71, 72, 75, 76, 81, 90, 93).

Zhou, J.-W. & Davis, T. (2024), *Hierarchical hub-filament structures and gas inflows on galaxy-cloud scales*, PASA **41**, e076 e076, arXiv: 2404.15862 [astro-ph.GA] (cit. on pp. 6, 11, 23, 36, 39, 41–43, 50, 65, 132, 141).

Zhou, J.-w., Dib, S. & Kroupa, P. (2024a), *The currently observed clumps cannot be the "direct" precursors of the currently observed open clusters*, A&A **691**, A293 A293, arXiv: 2410.01002 [astro-ph.GA] (cit. on pp. 82, 93, 94, 105, 107–110, 112, 118, 120, 121, 127, 133, 148).

Zhou, J.-w., Dib, S. & Kroupa, P. (2025), *The post-gas expulsion coalescence of embedded clusters as an origin of open clusters*, MNRAS **537** 845, arXiv: 2501.08831 [astro-ph.GA] (cit. on pp. 82, 93, 101, 102, 105, 107–110, 112, 120, 121, 127).

Zhou, J.-w., Kroupa, P. & Dib, S. (2024b), *Self-similar cluster structures in massive star-forming regions: Isolated evolution from clumps to embedded clusters*, A&A **688**, L19 L19, arXiv: 2407.20150 [astro-ph.GA] (cit. on pp. 82, 83, 90, 95, 99, 100, 105, 107, 126, 132, 148).

Zhou, J.-W., Kroupa, P. & Wu, W. (2024), *Evolution of the mass-radius relation of expanding very young star clusters*, A&A **691**, A204 A204, arXiv: 2409.20271 [astro-ph.GA] (cit. on pp. 82, 91, 92, 94, 100, 102, 105, 106, 117).

Zhou, J.-W. et al. (2022), *ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions - XI. From inflow to infall in hub-filament systems*, MNRAS **514** 6038, arXiv: 2206.08505 [astro-ph.GA] (cit. on pp. 2, 6, 11, 13, 15, 18, 22, 23, 25, 32, 34, 39, 41, 73, 74, 76–78, 95, 107, 132, 141).

Zhou, J.-W. et al. (2023), *Formation of hub-filament structure triggered by a cloud-cloud collision in the W33 complex*, MNRAS **519** 2391, arXiv: 2212.02627 [astro-ph.GA] (cit. on p. 100).

Zinnecker, H. & Yorke, H. W. (2007), *Toward Understanding Massive Star Formation*, ARA&A **45** 481, arXiv: 0707.1279 [astro-ph] (cit. on pp. 3, 5).

Acknowledgements

I will be forever grateful to Karl Menten, Friedrich Wyrowski, the Max Planck Institute, and Germany for giving me the opportunity to pursue research in astronomy. I still vividly remember how excited I was when I received the invitation letter. My sincere thanks to Frank Bigiel for being my second supervisor, and to Petra Mutzel and Akaki Rusetsky for serving on my doctoral committee. Special thanks to Sami Dib for the timely feedback during my research. Thanks to the IMPRS school for organizing the retreats, which gave me the opportunity to enjoy the beautiful scenery in different places. I am grateful to Wenjie Wu for teaching me how to run N-body simulations, to Sudeep Neupane for guiding me in using CLASS for data processing, and to Vikrant Jadhav for helping me work with Gaia data. Thanks to Dirk Muders for resolving many of my software and computer-related issues. Thanks to the many individuals and research teams who made their codes and data publicly available, greatly facilitating my research. I would like to thank Huanchen Hu for her help during the writing of my thesis. Many thanks to Leo and Dong for their companionship, which helped me feel less lonely while living abroad. I must also thank Bathooven and his music, which gave me great spiritual strength and accompanied me through the days and nights of my research journey. During my PhD, I learned a lot from Pavel Kroupa. I am deeply grateful for his generosity and support. He is truly a teacher—always willing to open doors for someone like me who is eager to learn. My heartfelt thanks to my family for their unwavering support. Although my mother does not fully agree with my decision to study abroad, she still supports me wholeheartedly and constantly worries about me. My father and younger brother often remind me not to worry about home and to take care of myself. My sister's constant care always fills me with warmth. My beloved constantly longs to be with me, and I feel the same. Since we met, however, our lives have followed a pattern of brief reunions and long separations. This longing and sorrow are the price of my pursuit of academic goals. She loves me—though she sometimes shows her frustration, she has always supported me unconditionally. She brings comfort and peace to my heart and soul, like someone holding the string of a drifting kite. Though the future is uncertain and occasionally makes me anxious, knowing she is there, anchoring me, gives me strength and calm.

Scientific contributions

First-author papers

During the PhD

1. Zhou, J.-W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
   High-resolution APEX/LAsMA 12CO and 13CO (3–2) observation of the G333 giant molecular cloud complex I. Evidence for gravitational acceleration in hub-filament systems.
   DOI: https://doi.org/10.1051/0004-6361/202346500
2. Zhou, J.-W., Dib, S., Juvela, M., et al. 2024, A&A, 686, A146
   Gas inflows from cloud to core scales in G332.83-0.55: Hierarchical hub-filament structures and tide-regulated gravitational collapse.
   DOI: https://doi.org/10.1051/0004-6361/202449514
3. Zhou, J.-W., & Davis, T. 2024, Publ. Astron. Soc. Australia, 41, e076
   Hierarchical hub-filament structures and gas inflows on galaxy-cloud scales.
   DOI: https://doi.org/10.1017/pasa.2024.47
4. Zhou, J.-W., Urquhart, J. S., Wyrowski, F., Liu, T., and Juvela, M., submitted
   High-resolution APEX/LAsMA 12CO and 13CO (3-2) observation of the G333 giant molecular cloud complex : III. Decomposition of molecular clouds into multi-scale hub-filament structures.
   DOI: https://doi.org/10.48550/arXiv.2506.16664
5. Zhou, J.-W., Dib, S., & Davis, T. A. 2024, MNRAS, 534, 683
   Molecular clouds as hubs in spiral galaxies: gas inflow and evolutionary sequence.
   DOI: https://doi.org/10.1093/mnras/stae2101
6. Zhou, J.-W., & Li, G.-X. 2025, MNRAS
   Evolution of molecular clouds on galaxy-cloud scale revealed by gravitational network analysis: high-mass clouds may deplete nearby gas via accretion or merging.
   DOI: https://doi.org/10.1093/mnras/staf199
7. Zhou, J. W. 2025, A&A, 696, L5
   External tides: an important driver of velocity dispersion in molecular clouds.
   DOI: https://doi.org/10.1051/0004-6361/202453040
8. Zhou, J.-W., & Dib, S. 2025, MNRAS, 537, 2232
   An examination of large-scale galactic effects on molecular cloud properties in NGC 628: the significant impact of tidal effects from neighbouring material on the evolution of molecular clouds.
   DOI: https://doi.org/10.1093/mnras/staf201
9. Zhou, J.-W., Wyrowski, F., Neupane, S., et al. 2024, A&A, 682, A128
   High-resolution APEX/LAsMA 12CO and 13CO (3–2) observation of the G333 giant molecular cloud complex II. Survival and gravitational collapse of dense gas structures under feedback.
   DOI: https://doi.org/10.1051/0004-6361/202347377

10. Zhou, J.-W., Dib, S., Wyrowski, F., et al. 2024, A&A, 682, A173
Feedback from protoclusters does not significantly change the kinematic properties of the embedded dense gas structures.
DOI: https://doi.org/10.1051/0004-6361/202348108

11. Zhou, J.-W., Kroupa, P., & Dib, S. 2024, PASP, 136, 094301
Physical Properties of Embedded Clusters in ATLASGAL Clumps with H ii Regions.
DOI: https://doi.org/10.1088/1538-3873/ad6f44

12. Zhou, J.-W., Dib, S., & Kroupa, P. 2024, PASP, 136, 094302
The Star Formation Histories, Star Formation Efficiencies and Ionizing Sources of ATLASGAL Clumps with H ii Regions.
DOI: https://doi.org/10.1088/1538-3873/ad77f4

13. Zhou, J.-W., Kroupa, P., & Dib, S. 2024, A&A, 688, L19
Self-similar cluster structures in massive star-forming regions: Isolated evolution from clumps to embedded clusters.
DOI: https://doi.org/10.1051/0004-6361/202450412

14. Zhou, J.-W., Kroupa, P., & Wu, W. 2024, A&A, 691, A204
Evolution of the mass-radius relation of expanding very young star clusters.
DOI: https://doi.org/10.1051/0004-6361/202452064

15. Zhou, J.-W., Dib, S., & Kroupa, P. 2024, A&A, 691, A293
The currently observed clumps cannot be the "direct" precursors of the currently observed open clusters.
DOI: https://doi.org/10.1051/0004-6361/202451728

16. Zhou, J.-W., Dib, S., & Kroupa, P. 2025, MNRAS, 537, 845
The post-gas expulsion coalescence of embedded clusters as an origin of open clusters.
DOI: https://doi.org/10.1093/mnras/staf076

17. Zhou, J.-W., Kroupa, P., under revision
Origin of open clusters revealed by the evolution of the $m_{\rm max} - M_{\rm ecl}$ relation

18. Zhou, J.-W., Jadhav, V., Kroupa, P., submitted
Constraining the gas expulsion modes of embedded clusters by analyzing the expansion speeds of open clusters

19. Zhou, J.-W., under revision
The dependence of black hole formation in open clusters on the cluster formation process

20. Zhou, J.-W., Kroupa, P., & Dib, S. 2025, A&A, 695, L17
Inversely synthesizing the core mass function of high-mass star-forming regions from the canonical initial mass function.
DOI: https://doi.org/10.1051/0004-6361/202452924

21. Zhou, J.-W., Kroupa, P., under revision
Clump populations in molecular clouds

22. Zhou, J.-W., Kroupa, P., & Dib, S. 2025, MNRAS, 541, 1276
The most massive star clusters in molecular clouds: Insights from the integrated cloud-wide initial mass function (ICIMF) theory.
DOI: https://doi.org/10.1093/mnras/staf1070

23. Zhou, J.-W., Han, A.-A., finished
Automated bubble identification by Blendmask: from hierarchical molecular gas to hierarchical bubbles in NGC 628

24. Zhou, J.-W., finished
Connecting Star Formation in the Milky Way and Nearby Galaxies. I. Comparability of Molecular Cloud Physical Properties

Before the PhD

1. Zhou, J.-W., Li, X.-X., Gao, R., et al. 2019, Chinese Physics Letters, 36, 090302
Modulational Instability of Trapped Two-Component Bose–Einstein Condensates.
DOI: https://doi.org/10.1088/0256-307X/36/9/090302

2. Zhou, J.-W., Liu, T., Li, J.-Z., et al. 2021, MNRAS, 508, 4639
ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions–VI. On the formation of the 'L' type filament in G286.21+0.17.
DOI: https://doi.org/10.1093/mnras/stab2801

3. Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
ATOMS: ALMA Three-millimeter Observations of Massive Star-forming regions – XI. From inflow to infall in hub-filament systems.
DOI: https://doi.org/10.1093/mnras/stac1735

4. Zhou, J.-W., Li, S., Liu, H.-L., et al. 2023, MNRAS, 519, 2391
Formation of hub–filament structure triggered by a cloud–cloud collision in the W33 complex.
DOI: https://doi.org/10.1093/mnras/stac3559

Original papers of Chapter.2

The original paper of Sec.2.3,

> Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
> *High-resolution APEX/LAsMA 12CO and 13CO (3–2) observation of the G333 giant molecular cloud complex I. Evidence for gravitational acceleration in hub-filament systems*

The original paper of Sec.2.4,

> Zhou, J. W., Dib, S., Juvela, M., et al. 2024, A&A, 686, A146
> *Gas inflows from cloud to core scales in G332.83-0.55: Hierarchical hub-filament structures and tide-regulated gravitational collapse*

The original paper of Sec.2.5,

> Zhou, J.-W., & Davis, T. 2024, Publ. Astron. Soc. Australia, 41, e076,
> *Hierarchical hub-filament structures and gas inflows on galaxy-cloud scales*

The original paper of Sec.2.6,

> Zhou, J.-W., Urquhart, J. S., Wyrowski, F., Liu, T., and Juvela, M., submitted
> *High-resolution APEX/LAsMA 12CO and 13CO (3-2) observation of the G333 giant molecular cloud complex : III. Decomposition of molecular clouds into multi-scale hub-filament structures*

The original paper of Sec.2.7,

> Zhou, J. W., Dib, S., & Davis, T. A. 2024, MNRAS, 534, 683
> *Molecular clouds as hubs in spiral galaxies: gas inflow and evolutionary sequence*

The original paper of Sec.2.8,

> Zhou, J. W., & Li, G.-X. 2025, MNRAS
> *Evolution of molecular clouds on galaxy-cloud scale revealed by gravitational network analysis: high-mass clouds may deplete nearby gas via accretion or merging*

A&A 676, A69 (2023)
https://doi.org/10.1051/0004-6361/202346500
© The Authors 2023

Astronomy & Astrophysics

High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3–2) observation of the G333 giant molecular cloud complex

I. Evidence for gravitational acceleration in hub-filament systems*

J. W. Zhou[1], F. Wyrowski[1], S. Neupane[1], J. S. Urquhart[2], N. J. Evans II[3], E. Vázquez-Semadeni[4], K. M. Menten[1], Y. Gong[1], and T. Liu[5]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de
[2] Centre for Astrophysics and Planetary Science, University of Kent, Canterbury, CT2 7NH, UK
[3] Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, Texas 78712-1205, USA
[4] Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro # 8701, Ex-Hda. San José de la Huerta, Morelia, Michoacán, 58089, Mexico
[5] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, PR China

Received 24 March 2023 / Accepted 17 May 2023

ABSTRACT

Context. Hub-filament systems are suggested to be the birth cradles of high-mass stars and clusters.
Aims. We investigate the gas kinematics of hub-filament structures in the G333 giant molecular cloud complex using ^{13}CO (3−2) observed with the APEX/LAsMA heterodyne camera.
Methods. We applied the FILFINDER algorithm to the integrated intensity maps of the ^{13}CO J = 3−2 line to identify filaments in the G333 complex, and we extracted the velocity and intensity along the filament skeleton from moment maps. Clear velocity and density fluctuations are seen along the filaments, allowing us to fit velocity gradients around the intensity peaks.
Results. The velocity gradients we fit to the LAsMA and ALMA data agree with each other over the scales covered by ALMA observations in the ATOMS survey (<5 pc). Changes in velocity gradient with scale indicate a funnel structure of the velocity field in position-position-velocity (PPV) space. This is indicative of a smooth, continuously increasing velocity gradient from large to small scales, and thus is consistent with gravitational acceleration. The typical velocity gradient corresponding to a 1 pc scale is ∼1.6 km s^{-1} pc^{-1}. Assuming freefall, we estimate a kinematic mass within 1 pc of ∼1190 $M_\odot$, which is consistent with typical masses of clumps in the ATLASGAL survey of massive clumps in the inner Galaxy. We find direct evidence for gravitational acceleration from a comparison of the observed accelerations to those predicted by freefall onto dense hubs with masses from millimeter continuum observations. On large scales, we find that the inflow may be driven by the larger-scale structure, consistent with the hierarchical structure in the molecular cloud and gas inflow from large to small scales. The hub-filament structures at different scales may be organized into a hierarchical system extending up to the largest scales probed through the coupling of gravitational centers at different scales.
Conclusions. We argue that the funnel structure in PPV space can be an effective probe for the gravitational collapse motions in molecular clouds. The large-scale gas inflow is driven by gravity, implying that the molecular clouds in the G333 complex may be in a state of global gravitational collapse.

Key words. ISM: kinematics and dynamics – methods: data analysis – ISM: clouds – surveys – stars: formation – submillimeter: ISM

1. Introduction

To understand the formation of hierarchical structures in high-mass star formation regions, it is critical to measure the dynamical coupling between density enhancements in giant molecular clouds and the gas motion in their local environment (McKee & Ostriker 2007; Motte et al. 2018; Henshaw et al. 2020). These studies may distinguish between competing concepts for high-mass star formation, such as monolithic collapse of turbulent cores in virial equilibrium (McKee & Tan 2003; Krumholz et al. 2007), competitive accretion in a protocluster environment through Bondi-Hoyle accretion (Bonnell et al. 1997, 2001), turbulence-driven inertial-inflow (Padoan et al. 2020) and gravity-dominated global hierarchical collapse (Vázquez-Semadeni et al. 2009, 2017, 2019; Ballesteros-Paredes et al. 2011; Hartmann et al. 2012), as discussed in Zhou et al. (2022). High-resolution observations show that density enhancements are organized in filamentary gas networks, especially in hub-filament systems. In these systems, converging flows are funneling matter into the hub through the filaments. Many case studies have suggested that hub-filament systems are the birth cradles of high-mass stars and clusters

* The data are only available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/676/A69

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

(Peretto et al. 2013; Henshaw et al. 2014; Zhang et al. 2015; Liu et al. 2016b, 2022; Yuan et al. 2018; Lu et al. 2018; Issac et al. 2019; Dewangan et al. 2020; Zhou et al. 2022). Numerical simulations of colliding flows and collapsing clumps reveal velocity gradients along the dense filamentary streams converging toward the hubs (Wang et al. 2010; Gómez & Vázquez-Semadeni 2014; Smith et al. 2016; Padoan et al. 2020). In observations, velocity gradients along filaments are often interpreted as evidence for gas inflow along filaments (Kirk et al. 2013; Liu et al. 2016b; Yuan et al. 2018; Williams et al. 2018; Chen et al. 2019, 2020; Pillai et al. 2020; Zhou et al. 2022).

Zhou et al. (2022) studied the physical properties and evolution of hub–filament systems in a large sample of protoclusters that were observed in the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Liu et al. 2020). They found that hub-filament structures can exist not only in small-scale dense cores (∼0.1 pc), but also in large-scale clumps/clouds (∼1–10 pc), suggesting that multiscale hub-filament systems at various scales are common in regions forming massive stellar clusters in various Galactic environments. The filaments in clumps observed in the ATOMS program show clear velocity gradients. The approximately symmetric distribution of positive and negative velocity gradients strongly indicates the existence of converging gas inflows along filaments. The observations confirm that high-mass stars in protoclusters may accumulate most of their mass through longitudinal inflow along filaments. Velocity and density fluctuations are discussed in detail in Henshaw et al. (2020), who detected ubiquitous velocity fluctuations across all spatial scales and Galactic environments and discovered oscillatory gas flows with wavelengths ranging from 0.3 to 400 pc that are coupled to regularly spaced density enhancements that probably formed via gravitational instabilities (Henshaw et al. 2016; Elmegreen et al. 2018). Furthermore, the locations of some of these density enhancements spatially correlate with velocity gradient extrema, which is indicative of either convergent motion or of collapse-induced rotation (Clarke et al. 2016; Misugi et al. 2019).

In Zhou et al. (2022), we measured the velocity gradients around the intensity peaks for all observed cores. The statistical analysis found that velocity gradients are very small at scales larger than ∼1 pc, probably suggesting the dominance of pressure-driven inertial inflow, which can originate either from large-scale turbulence or from cloud-scale gravitational contraction. Below ∼1 pc, the velocity gradients dramatically increase as the filament lengths decrease, indicating that the hub or core gravity dominates gas infall on small scales. Due to the field-of-view (FOV) limitation of the ALMA observation, our previous work was restricted to scales up to about 5 pc. In this paper, following a similar method as described in Zhou et al. (2022), we can generalize the results from the clump-core scale to cloud-clump scale.

Figure 1b displays an overview three-color map of the observed field by combining ATLASGAL+*Planck* 870 μm and GLIMPSE 8.0 and 4.5 μm emission. Figure 1a shows the distribution of newly observed ^{13}CO (3–2) emission. The interesting individual subregions are marked and highlighted below. The main structures of the observed field are the G331 and G333 giant molecular clouds (GMCs) and the G332 ring structure between them.

The GMC G331 is one of the most massive molecular clouds in the southern Galaxy, in the tangent region of the Norma spiral arm (Bronfman et al. 1989). Using the C^{18}O (1–0) integrated emission, Merello et al. (2013) defined the central region of the G331 GMC at l = 331.523°, b = −0.099°, with a distance of 7.5 kpc. It may harbor one of the most extended and luminous regions of massive star formation in the Galactic disk. Caswell & Haynes (1987) determined that the line central velocity of the ionized gas is approximately −89 km s^{-1} from observations of the H109α and H110α hydrogen recombination lines, which is similar to the peak velocity of the molecular gas in the GMC. The detection of OH and methanol maser emission further provides evidence of active star formation in the GMC G331 (Goss et al. 1970; Caswell et al. 1980; Caswell 1998; Pestalozzi et al. 2005).

The GMC G333, centered at l ∼ 333.2°, b ∼ −0.4°, is a 1.2°× 0.6° region in the fourth quadrant of the Galaxy at a distance of 3.6 kpc (Lockman 1979; Bains et al. 2006). Its gas emission takes the form of a string of knots. As a part of the Galactic ring of molecular clouds at a galactocentric radius of 3–5 kpc, the G333 molecular cloud complex contains a diverse sample of molecular regions as well as a range of high-mass star-forming clouds, bright HII regions, and infrared point sources, all surrounded by diffuse atomic and molecular gas (Fujiyoshi et al. 2006; Cunningham et al. 2008; Wong et al. 2008; Lo et al. 2009; Jordan et al. 2013; Lowe et al. 2014). OH, H_2O, and CH_3OH maser lines as tracers of high-mass star formation have been detected from various sources in the region (Caswell et al. 1980, 1995, 2011; Breen et al. 2007, 2012). Toward the GMC G333, Bains et al. (2006) used the Mopra 22 m radio telescope to identify five distinct velocity features at −105, −90, −70, −50, and −10 km s^{-1} with ^{13}CO (1–0) maps. They also found at least three velocity components (−55, −50, and −42 km s^{-1}) between −65 and −35 km s^{-1}, with the brightest feature at −50 km s^{-1}. Several spiral arms intersect at l ∼ 333° along the line of sight, that is, the Sagittarius-Carina arm, the Scutum-Crux arm, and the Norma-Cygnus arm (Russeil 2003; Russeil et al. 2005; Vallée 2008). This region may therefore be suitable for studying the formation of molecular clouds under Galaxy dynamics.

Between the G333 and G331 GMCs, there is a cavity at l ∼ 332°. The upper part presents a ring structure within Galactic coordinates 332.0° < l < 332.8° and −0.3° < b < 0.4°. The principal emission of this ring structure is concentrated between the velocity range −55 < $v_{\rm LSR}$ < −44 km s^{-1} (Romano et al. 2019), and the average spectral profiles of the CO lines and CI show a peak at $v_{\rm LSR}$ = −50 km s^{-1}. The distance of the ring is determined to be ∼3.7 kpc from the Sun (Romano et al. 2019), thus the ring and GMC G333 are located in the same region. This region is referred to as the G333 complex in this work.

2. Observation

2.1. APEX/LAsMA data

We mapped a 3.4° × 1.2° area centered on (l, b) = (332.33°, −0.29°). The observations were conducted between March and August 2022 using the APEX telescope (Güsten et al. 2006)[1]. The 7 pixel Large APEX sub-Millimeter Array (LAsMA) receiver was used to observe the J = 3–2 transitions of ^{12}CO ($\nu_{\rm rest}$ ∼ 345.796 GHz) and ^{13}CO ($\nu_{\rm rest}$ ∼ 330.588 GHz) simultaneously in the upper and lower sideband, respectively. More details about the receiver are given in Mazumdar et al. (2021). The local oscillator frequency was set at 338.190 GHz in order to avoid contamination of the ^{13}CO (3–2) lines due

[1] This publication is based on data acquired with the Atacama Pathfinder EXperiment (APEX). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory and the Onsala Space Observatory.

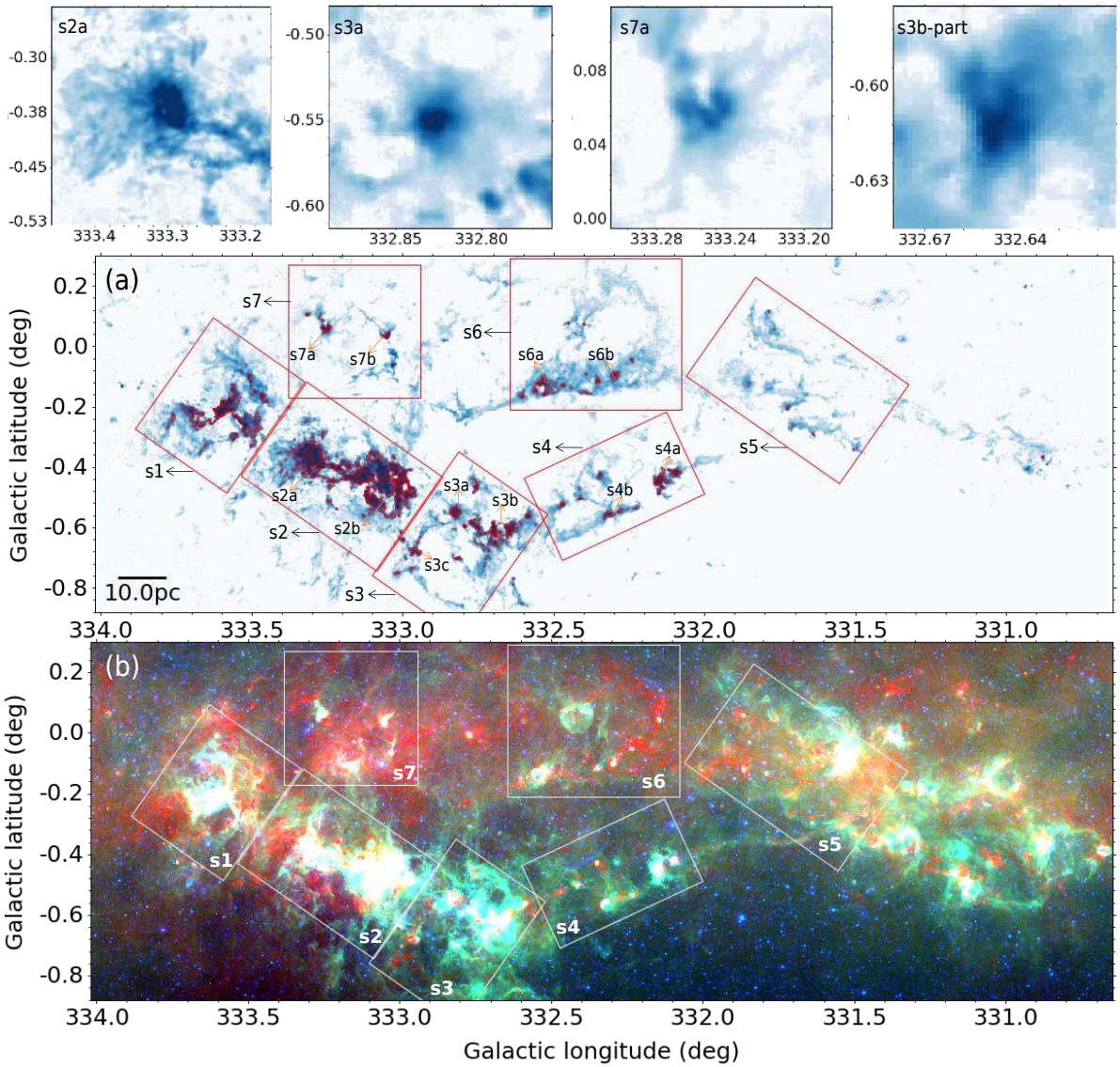


Fig. 1. Overview of the entire observed field. The enlarged maps in the first row display several typical hub-filament structures. (a): The background is the integrated intensity map of ^{13}CO (3–2) for the G333 complex, and the boxes show several divided subregions. The letters indicate the substructures studied in this work. The red contours show the peak emission of ^{13}CO (3–2). (b): Overview three-color map of the observed field by combining ATLASGAL+*Planck* 870 μm (red) and GLIMPSE 8.0 (green) and 4.5 μm (blue) emission, including the G333 complex and GMC G331.

to bright ^{12}CO (3−2) emission from the image band. The whole observed region was divided into two parts. The first part is from $l \sim 332°$ to $l \sim 334°$, and the second part is from $l \sim 330.6°$ to $l \sim 332°$. Each part was further divided into submaps of size $10' \times 10'$ each. Observations were performed in a position-switching on-the-fly (OTF) mode. Although the reference positions were carefully chosen, emission is still present in the reference position of the first part in the velocity range of −140 to 10 km s^{-1} (see Appendix B for details about how the issue was solved).

The data were calibrated using the three load chopper wheel method, which is an extension of the standard method used for millimeter observations (Ulich & Haas 1976) to calibrate the data to the antenna temperature T_A^* scale. The data were reduced using the GILDAS package[2]. A velocity range of −190 to 60 km s^{-1} was extracted for each spectrum and resampled to an adequate velocity resolution of 0.25 km s^{-1} to reduce the noise. The velocity range −140 to 10 km s^{-1} was masked before a first-order baseline was fit to each spectrum. The reduced calibrated data obtained from the different scans were then combined and gridded using a 6″ cell size. The gridding process included a convolution with a Gaussian kernel with a full width at half maximum (FWHM) size of one-third the telescope FWHM beam width. The data cubes we obtained have a final angular resolution of 19.5″ that is comparable with the resolution of the ATLASGAL survey.

After the online calibration, the intensities were obtained on the T_A^* scale (corrected antenna temperature). In addition to the atmospheric attenuation, this also corrects for rear spillover, blockage, scattering, and ohmic losses. A beam efficiency value η_{mb} = 0.71 (Mazumdar et al. 2021) was used to convert intensities from T_A^* into the main-beam brightness temperature T_{mb}. The final noise levels of the ^{12}CO (3–2) and ^{13}CO (3–2) data cubes are ∼0.32 and ∼0.46 K, respectively. In the rest of paper,

[2] http://www.iram.fr/IRAMFR/GILDAS

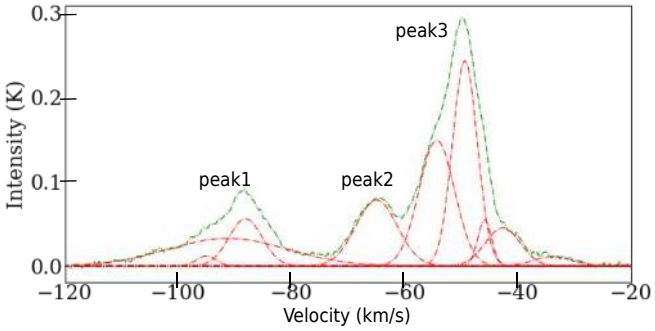


Fig. 2. Average spectra of ^{13}CO (3−2) for the entire region. The dashed red profiles represent the components of multi-Gaussian fitting.

we focus on the analysis of the ^{13}CO (3–2) emission. The information from both lines will be combined in a forthcoming paper.

2.2. Archival continuum emission data

The observed region was covered in the infrared by the GLIMPSE survey (Benjamin et al. 2003). The images of the *Spitzer* Infrared Array Camera (IRAC) at 4.5 and 8.0 μm were retrieved from the *Spitzer* Archive. The angular resolutions of the images in the IRAC bands are ∼ 2″. We also used ATLASGAL+*Planck* 870 μm data (ATLASGAL combined with *Planck* data), which are sensitive to a wide range of spatial scales at a resolution of ∼21″ (Csengeri et al. 2016).

3. Results

3.1. Velocity components

The average emission spectrum of ^{13}CO (3−2) coming from the entire observed field is shown in Fig. 2 together with Gaussian curves used to fit the average spectrum. The fitting was carried out using the GAUSSPY+ software (Riener et al. 2019), which is introduced in detail below. We divided the spectrum into three velocity components [−120, −76] km s^{-1}, [−76, −60] km s^{-1}, and [−60, −35] km s^{-1} in which emission is prominent. They are labeled peak1, peak2, and peak3. Figure 3 shows that the velocity components peak1 and peak2 represent the main two parts of the G331 GMC, called G331 GMC-blue and G331 GMC-red. Peak3 shows a very extended structure throughout the observed region with the GMC G333 and the G332 ring as the main parts. The four main ATLASGAL clump clusters are labeled A, B, C, and D in Fig. 3a (Urquhart et al. 2018, 2022). They correspond to the different velocity components shown in the moment-one maps of ^{13}CO (3−2) (Fig. 3b). Peak1 and peak2 are associated with clump clusters C and D. Clump clusters A and B are located in velocity component peak3. Peak1 is relatively independent of peak2 and peak3, but the dominant features in peak2 and peak3 appear to overlap to some extent, but they can be separated. Hence, we fit Gaussian functions in order to minimize the effects of cross-contamination. Peak2 and peak3 can be distinguished by the FWHM. We mainly focus on peak3 with the velocity range [−60, −35] km s^{-1} that represents the complete shell structure shown in Fig. 3e. The association between peak1, peak2, and peak3 will be discussed in detail in a forthcoming paper.

In order to study the structure of the molecular cloud complex in position-position-velocity (PPV) space, we applied the fully automated Gaussian decomposer GAUSSPY+ (Lindner et al. 2015; Riener et al. 2019), which can decompose the complex spectra of molecular lines into multiple Gaussian components. The most important four parameters in the algorithm are the first and second smoothing parameters α_1 and α_2, the signal-to-noise ratio (S/N) and significance criterion. To determine the smoothing parameters α_1 and α_2, sets with 500 randomly selected spectra from the ^{13}CO (3−2) cube were used to train GAUSSPY+. The default values of SNR and significance are 3 and 5, but considering that the moment-one map needs a 5σ threshold to show a clear velocity field, we adjusted the value of SNR to 5. The other parameters for the decomposition were set to the default values provided by Riener et al. (2019).

As shown in Fig. 4, the velocity components separated by the average spectrum are consistent with the decomposition of GAUSSPY+. The three main velocity components are well separated in PPV space. Different from peak1 and peak2, peak3 has a large-scale continuous structure, regardless of the projection on Galactic longitude or latitude. This indicates that the structures in peak3 are physically connected. Especially the overall structure of peak3 consists of a large-scale shell structure extended from l = 330.6° to l = 334°. It may be generated by a large-scale compression, such as the shock from supernova explosion events. This will be investigated in a forthcoming paper.

The entire G333 complex seems to have fragmented into several independent molecular clouds. They are treated as subregions in Fig. 1, such as s3. The substructures in each subregion marked by letters in Fig. 1a are divided by the bright emission of ^{13}CO (3–2) shown in red contours, such as s3a. Figures 1b and 3a show that the dense parts of the ^{13}CO (3−2) emission tightly correlate with the infrared bright regions. The overlap of different velocity components in subregions 5 and 7 is significant, as shown in Fig. 3a. However, these two subregions only include a few structures in the velocity range [−60, −35] km s^{-1}, thus they are subordinate in the G333 complex.

This analysis shows that the GMC G331 and the G333 complex overlap on the sky, but they have different velocities and lie at very different distances. Because the G333 region is closer to us, we focus the remainder of the paper on it.

3.2. Identification of filaments

The first step was to identify and characterize the filaments in the G333 complex. Following the method described in Zhou et al. (2022), we used the moment-zero maps (the integrated intensity maps) of ^{13}CO (3−2) to identify filaments in the G333 complex using the FILFINDER algorithm (Koch & Rosolowsky 2015). The velocity intervals for making moment-zero maps were determined from the averaged spectra of each subregion. They are marked by the vertical dotted lines in Fig. A.1. Moreover, a threshold of 5σ was applied to make the moment maps, which can reduce the noise contamination effectively. The skeletons of identified filaments overlaid on the moment-zero maps of the ^{13}CO (3–2) line emission are shown in Fig. 5. They are highly consistent with the gas structures traced by ^{13}CO (3–2) as seen by eye, indicating that the structures identified in FILFINDER are reliable. Then for each subregion, we broke the filamentary network into several filaments. This step is necessary to calculate the offset along the filament before fitting the velocity gradient. To show the velocity field along the filament more clearly, we tried to make each filament long enough.

Because each subregion includes many substructures, it is not surprising that their average spectra will show multiple velocity peaks. In Fig. A.1, the average spectra of G333-s3

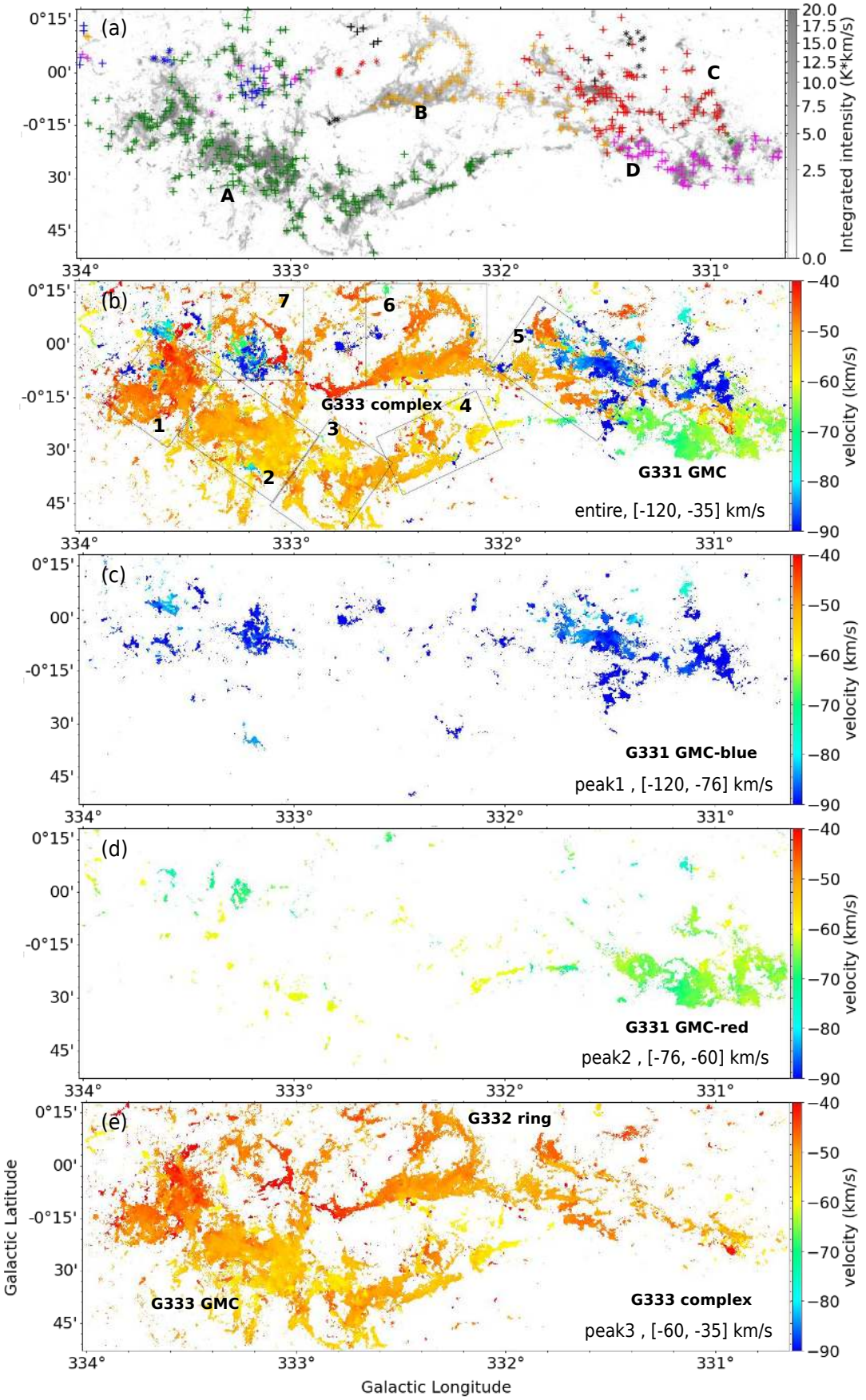


Fig. 3. ATLASGAL clump distribution and velocity components of the entire observed field. (a): The background is the integrated intensity map of ^{13}CO (3−2). The color-coded plus indicates the different ATLASGAL clump clusters A (green), B (orange), C (red), and D (magenta). (b): The moment-one map of ^{13}CO (3−2) for the entire region in the overall velocity range. Panels c–e: Moment-one map of ^{13}CO (3−2) for the three peaks marked in Fig. 2.

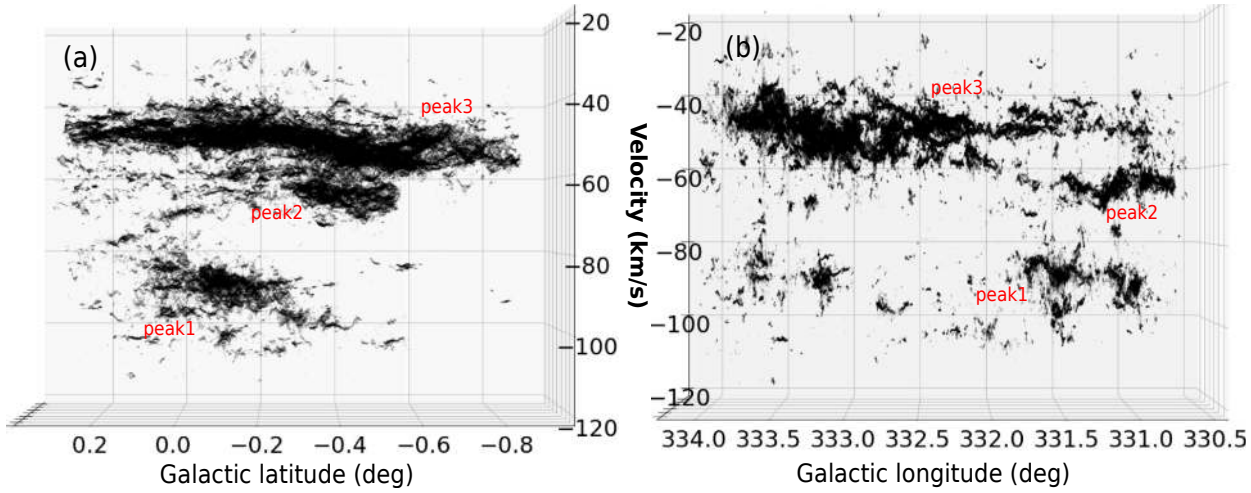


Fig. 4. Velocity distribution of the entire observed field in PPV space decomposed by GAUSSPY+ along the Galactic latitude and longitude.

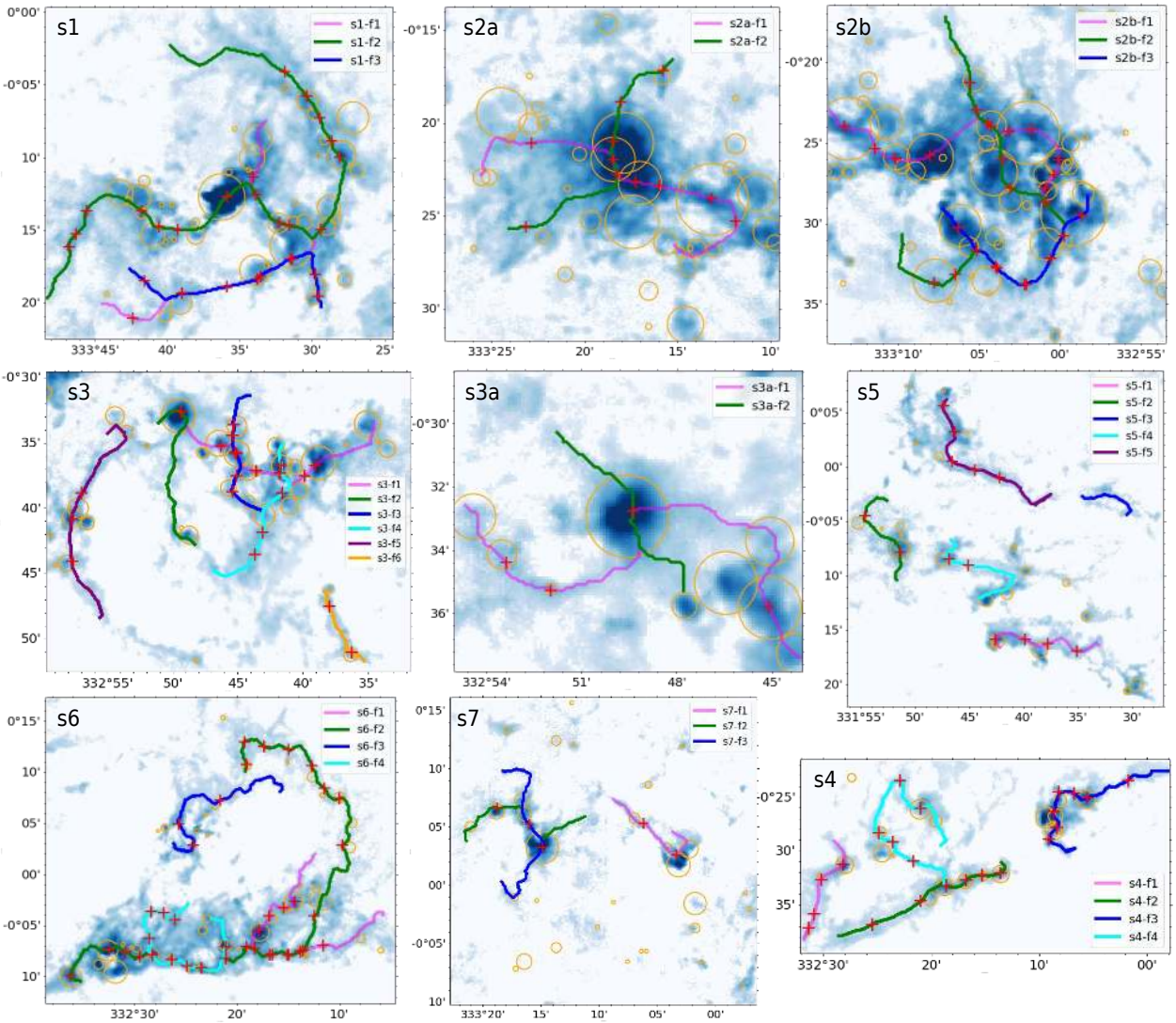


Fig. 5. Background images showing the moment-zero maps of ^{13}CO (3–2). The lines in color present the filament skeletons. The orange circles show the ATLASGAL clumps. The size of the circle represents the clump radius. The red plus marks the intensity peak of ^{13}CO (3–2) emission in Fig. 7.

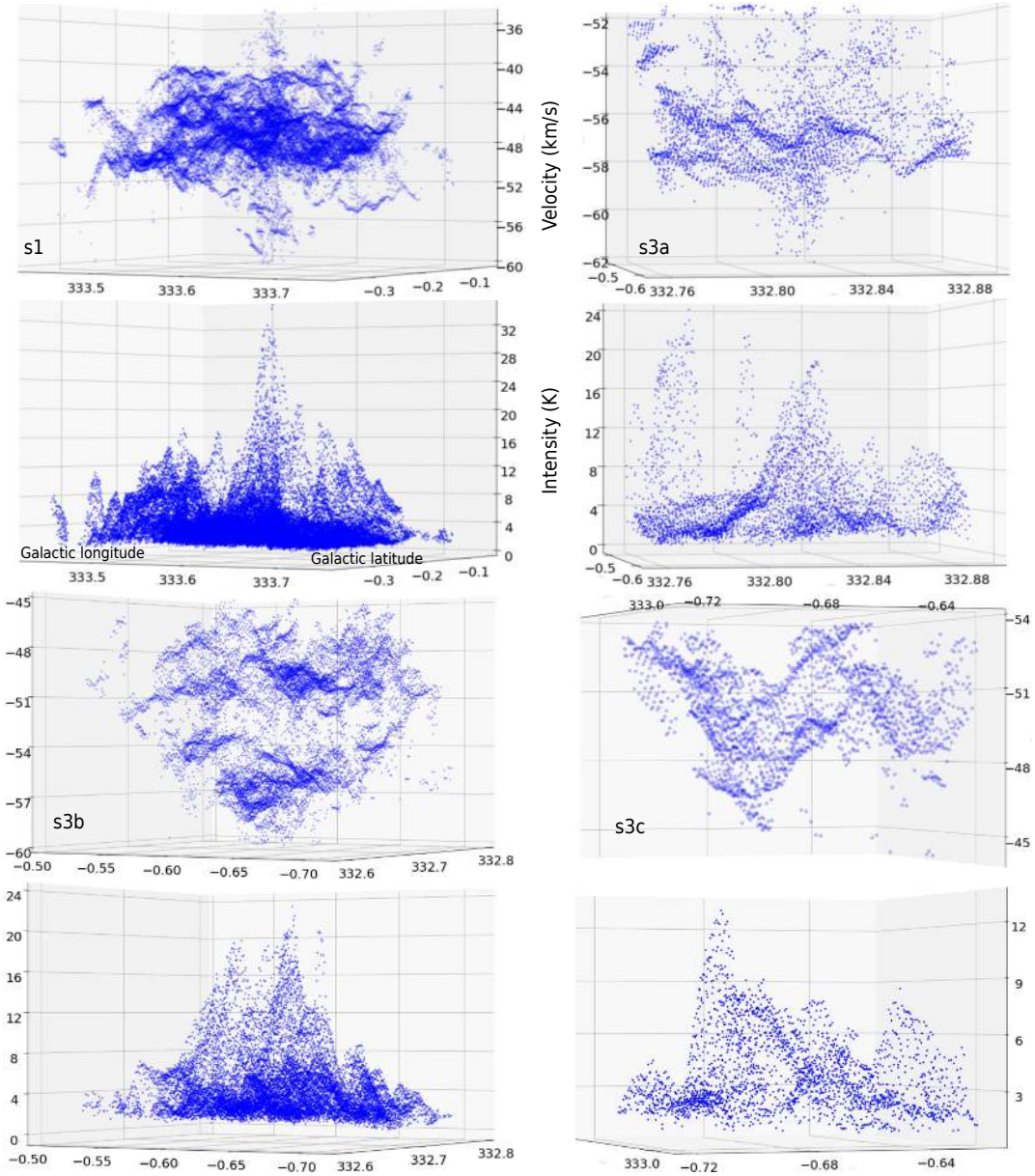


Fig. 6. Intensity and velocity distribution of the subregions in PPV space based on the decomposition of GAUSSPY+. For other subregions, see Figs. A.2 and A.3.

clearly show multiple components. We separated each peak in the line profile and found that each peak indeed corresponded to a substructure. This situation may also exist in other subregions. However, these sub-structures are generally well separated spatially and thus have no significant impact on the moment-map analysis.

Figure 6 displays the GAUSSPY+ decomposition of each subregion in PPV space. It shows that the main structures of each subregion are connected in PPV space and thus are unlikely to be contaminated by velocity components from unrelated foreground or background cloud emission. Moreover, after we extracted the velocity and intensity information along the filament skeletons, we should find anomalies in Fig. 7 if unrelated velocity components overlap. The anomalies might be a sudden break at a certain position of the skeleton. However, this is not the case in our analysis. Furthermore, we only analyze the variations of velocity and intensity along the filament skeletons with a width of one pixel, which can also effectively avoid potential overlap. In Zhou et al. (2022), we found that multiple velocity components are common in hub-filament systems, especially in hub regions. Furthermore, the observed samples in Zhou et al. (2022) with or without multiple velocity components show similar results, such as the velocity gradient estimated from the moment maps.

We can assume that most of the pixels above the filaments have a single velocity component or have a dominant velocity component. When we compare the fitting results with the freefall model and previous results from ALMA data, the hypothesis is indicated as reasonable if the results are consistent. This is the case in our work. Conversely, if the overlap of velocity components is significant, the fitted velocity gradient should be free without the regularities shown in our work because the overlap of

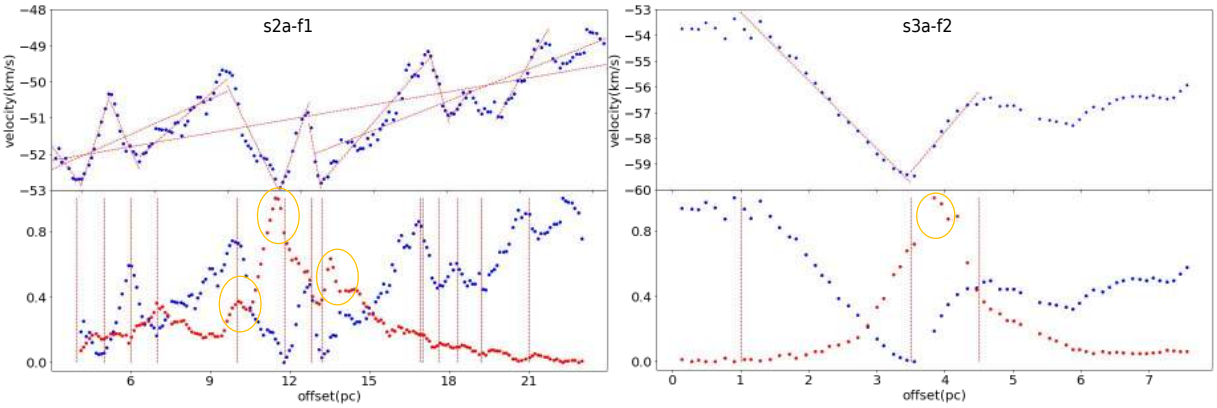


Fig. 7. Two selected filaments illustrate the fitting of the velocity gradient. Upper panels: velocity gradients are fit in the ranges defined by the vertical dashed red lines, and straight lines show the linear fitting results. Lower panels: dotted blue and red lines show the normalized velocity and intensity, respectively. The orange circles mark the intensity peaks of ^{13}CO (3−2) emission associated with ATLASGAL clumps. For other filaments, see Figs. A.4–A.8.

unrelated velocity components is random. Figure 7 shows good correlations between velocity and density fluctuations, which is consistent with previous studies (Hacar & Tafalla 2011; Henshaw et al. 2014, 2016, 2020; Liu et al. 2019). The variation in the velocity gradient with the scale shown in Fig. 8 is also consistent with the results in Zhou et al. (2022).

3.3. Velocity gradients

The kinematical features of hub-filament systems can be revealed by the velocity gradients along the filaments. However, it is difficult to directly carry out the measurement in PPV space due to the complex gas motion shown in Fig. 4. In Zhou et al. (2022), we investigated hub-filament systems in a large sample of 146 protoclusters using the $H^{13}CO^+$ J = 1–0 molecular line in the ATOMS survey. The strongest intensity peaks of the $H^{13}CO^+$ emission coincide with the brightest 3 mm cores or hub regions in these hub-filament systems. We find that filaments are ubiquitous in protoclusters. Velocity and density fluctuations along these filaments are seen. We first estimated two overall velocity gradients between the velocity peaks and valleys at the two sides of the center of the gravitational potential well and ignored local velocity fluctuation. We also derived additional velocity gradients over smaller distances around the strongest intensity peaks of the $H^{13}CO^+$ emission (see Fig. 6 in Zhou et al. 2022). In this work, the same method was used to derive the velocity gradients from the LAsMA data. Then the fitted velocity gradients derived from the LAsMA and ALMA data were combined for a statistical analysis, as shown in Fig. 8.

Figure 7 shows the intensity-weighted velocity (moment-one) and integrated intensity (moment-zero) map of the ^{13}CO (3−2) line emission along the skeletons of selected filaments with intense velocity and density fluctuations along these filaments. Generally, the peaks of the density fluctuations are associated with ATLASGAL clumps, as shown in Fig. 5. The density and velocity fluctuations along the filaments may indicate converging gas flows coupled to regularly spaced density enhancements that probably form via gravitational instabilities (Henshaw et al. 2020). Figure 7 shows the typical V-shape velocity structure around the intensity peaks, which we fit as velocity gradients on either side of the intensity peak marked by the straight lines. Figure 7a shows a more complex velocity field with more repetitive segments than the simple structure in Fig. 7b. The gas kinematics of other long filaments in Fig. 5 are similar to those in Fig. 7a. For the long filaments, except for the local velocity fluctuations or gradients, they also exhibit velocity gradients on larger scales, which were fit by ignoring the local velocity fluctuations. Generally, large-scale velocity gradients are associated with several intensity peaks.

3.4. Hub-filament structure

Each subregion marked in Fig. 1a contains many substructures that are nearly separated from each other, except for several crowded regions. The red contours from ∼50% of the integrated intensity of the ^{13}CO (3–2) emission show the high-density parts of various structures. In Fig. 1, several typical hub-filament structures are highlighted in the first row; see also Fig. 5 for more details. The filamentary structures are connected to high-density hubs. Other substructures in the entire observed region have a more or less similar morphology. Some differences are expected because the appearance of the hub-filament morphology relies on the projection angle and on the resolution of the observation. Moreover, most structures have the typical kinematic features of hub-filament systems, as discussed in the next sections. For example, the intensity peaks (the hubs) are associated with the converging velocities, indicating that the surrounding gas flows may be converging to dense structures. In particular, the velocity gradients increase at small scales.

The orange circle in Fig. 7 marks the intensity peak of ^{13}CO (3−2) emission associated with an ATLASGAL clump. The distance between the central coordinate of the ATLASGAL clump and the intensity peak of the ^{13}CO (3−2) emission is smaller than the effective radius of the ATLASGAL clump. These associated clumps are treated as hubs in this work. Their properties are listed in Tables 1 and 2. In Table 2, some of the peaks have more than one corresponding ATLASGAL clump, and we have added these together, as indicated by a space in the table. We focus on their mass, effective radius, and the corresponding velocity gradients. However, as shown in Fig. 7, many peaks of the ^{13}CO (3−2) emission lack corresponding clumps. They were therefore not included in the tables, but are similar to the peaks associated with clumps. Generally, the large-scale velocity gradients involve many intensity peaks with or without corresponding clumps. It is therefore difficult to estimate the properties of their hubs, and therefore, these are not given in the tables.

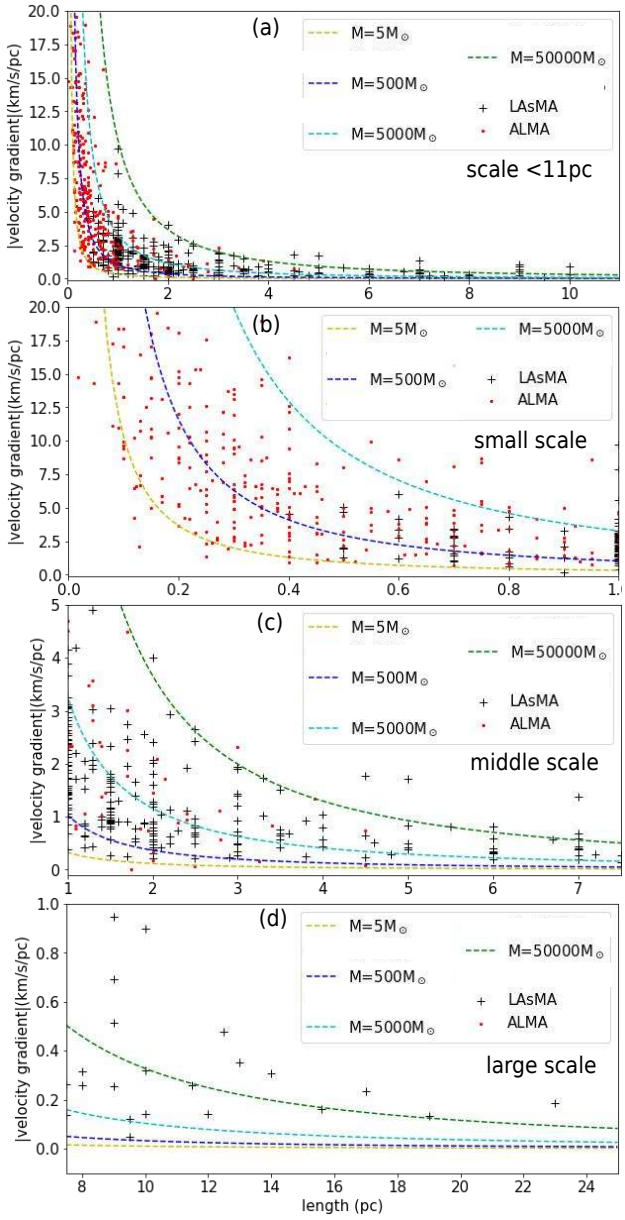


Fig. 8. Statistical analysis of all fitted velocity gradients. (a): Velocity gradient vs. the length over which the gradient has been estimated for all sources in the ATOMS survey (red dots) and current LAsMA observation (black plus). The color lines show the freefall velocity gradients for comparison. For the freefall model, yellow, blue, cyan, and green lines denote masses of 5 $M_\odot$, 500 $M_\odot$, 5000 $M_\odot$, and 50 000 $M_\odot$, respectively. Panels b–d: zoomed maps with lengths <1 pc (small scale), ~1–7.5 pc (medium scale), and >7.5 pc (large scale) in panel a.

3.5. Funnel structure

When we compare the velocity gradients of the molecular gas to iron filings, the distribution of the velocity gradients may to some extent reflect the shape of the force field that plays the dominant role in the molecular clouds, just as the iron filings can reflect the shape of a magnetic field. Below, we investigate the distribution of the molecular velocity gradients in PPV space. In the case of gravitational freefall onto a central point mass, velocities change as $v \propto r^{-1/2}$ ($r = \sqrt{x^2 + y^2}$). In the case of a collapse that is driven by the gravity of a gaseous mass with a certain density profile, the radial dependence of the infall speed depends on the slope of the density profile (Gómez et al. 2021). In either case, this results in a funnel morphology of the velocity field in PPV space, as illustrated in Fig. 9. The exterior cloud velocities show only gentle variations (a small velocity gradient that funnels material from the cloud to the clump), while the velocities in the interior part near the center change dramatically with scales (a large velocity gradient that funnels material from the clump to the core). We note that the schematic in Fig. 9 is oversimplified. A molecular cloud generally contains many clumps, and a clump also contains many cores. We chose only one clump and one core to demonstrate the gas inflow and the variation in the velocity gradient from the molecular cloud scale to the core scale. A more realistic case is shown in Fig. 6, where most of the substructures indeed show the expected funnel structures in PPV space. Particularly close to the intensity peak, the velocity gradient increases steeply, as also reflected in Fig. 7.

Moreover, the V-shape velocity structure around the intensity peaks indicates accelerated material inflowing toward the central hub (Gómez & Vázquez-Semadeni 2014; Kuznetsova et al. 2015, 2018; Hacar et al. 2017; Zhou et al. 2022), which is shown everywhere in Fig. 7. The V-shape velocity structure can be treated as the projection of the funnel structure from PPV space to the PV plane. Hence, the funnel structures in PPV space can be an effective probe for the gravitational collapse motions in molecular cloud.

4. Discussion

4.1. Characteristic scale ~1 pc

We have investigated gas motions on sufficiently large scales. The longest filament was ~50 pc. We obtained results similar to previous small-scale ALMA observations. We fit a wide range of values of the velocity gradients, as shown in Fig. 8a. The velocity gradients are small at large scales, while they become significantly larger at small scales ($\lesssim$1 pc), as is the case in the ATOMS survey (Zhou et al. 2022). For both ALMA and LAsMA data, most of the fit velocity gradients are concentrated on the scales ~1 pc, as shown in Fig. 8. This is consistent with the fact that 1 pc is considered the characteristic scale of massive clumps (Urquhart et al. 2018). In Fig. 8 the variation in the velocity gradients with the scales is comparable with the expectations from gravitational freefall, in the sense that the gradient decreases smoothly with increasing scale. Figures 10c,d show results similar to those in Fig. 8, but only consider the velocity gradients associated with ATLASGAL clumps that are listed in Tables 1 and 2. In particular, in Fig. 10c, the required central mass in the fitted freefall model is consistent with the mass distribution observed in Fig. 10b.

4.2. Evidence for gravitational acceleration

Figure 10a displays the probability distribution of velocity gradients measured around 1 pc (0.8 ~ 1.2 pc), showing that the most frequent velocity gradient is ~1.6 km s^{-1} pc^{-1}. Assuming freefall,

$$\nabla V_{\rm free} = -\frac{\rm d}{{\rm d}R}\sqrt{\frac{2GM}{R}} = \sqrt{\frac{GM}{2R^3}}, \quad (1)$$

we estimate that the kinematic mass corresponding to 1 pc is ~1190 $M_\odot$, which is comparable with the typical mass of clumps in the ATLASGAL survey (Urquhart et al. 2018). In Fig. 10b, the peak of the associated clump mass distribution is also ~1000 $M_\odot$. Thus the clump may be a gravity-dominated

Table 1. Properties of ATLASGAL clumps associated with the intensity peaks of ^{13}CO (3−2) emission (orange circles in Fig. 7), which are treated as hubs in this work.

Clump	log(M) ($M_\odot$)	Radius (pc)	$T_{\rm dust}$ (K)	log(Lum) ($L_\odot$)	$v_{\rm LSR}$ (km s^{-1})	ΔV_1 (km s^{-1} pc^{-1})	L_1 (pc)	ΔV_2 (km s^{-1} pc^{-1})	L_2 (pc)
AGAL332.094−00.421	3.155	0.865	28.4	4.789	−56.5	−2.97	1.4	1.9	1.1
AGAL332.312−00.556	2.947	0.831	21.4	3.835	−51.9	0.9	1.9	−1.76	0.6
AGAL332.351−00.436	3.034	0.935	18.3	3.483	−47.6			−0.58	6.0
AGAL332.411−00.471	2.618	0.502	19.3	2.694	−55.8	0.3	6.0	−0.49	5.0
AGAL332.467−00.522	3.162	0.848	19.8	3.864	−52.6	0.2	3.0	−2.25	0.8
AGAL332.606−00.854	2.569	0.537	19.8	3.088	−56.2	−0.53	2.0	2.0	1.0
AGAL332.647−00.609	3.549	1.35	28.3	5.195	−49.2	−2.7	2.0	3.7	1.0
AGAL332.694−00.612	3.184	0.813	26.6	4.466	−47.6	−7.84	1.0	9.7	1.0
AGAL332.694−00.612	3.184	0.813	26.6	4.466	−47.6	3.3	0.8	−4.59	1.0
AGAL332.726−00.621	2.407	0.294	24.8	3.456	−49.7	−3.2	0.7	1.5	0.8
AGAL332.751−00.597	2.964	1.038	24.7	3.937	−53.1			0.8	1.8
AGAL332.762−00.641	2.32	0.623	22.4	3.313	−50.4	−4.14	0.6	0.9	1.7
AGAL332.826−00.549	3.724	1.367	31.4	5.543	−57.3	−0.58	6.7	2.7	1.0
AGAL332.826−00.549	3.724	1.367	31.4	5.543	−57.3	−2.65	2.5	3.1	1.0
AGAL332.826−00.549	3.724	1.367	31.4	5.543	−57.3			−0.86	8.5
AGAL332.892−00.569	2.595	0.485	19.7	3.19	−57.3	−1.49	1.5	0.9	2.0
AGAL332.866−00.587	2.364	0.169	23.3	3.113	−58.1			−1.59	1.0
AGAL332.962−00.679	3.305	0.779	22.5	4.225	−48.5	2.0	3.0	−1.53	5.5
AGAL332.969−00.737	2.785	0.64	13.5	2.525	−55.6	−3.7	1.0	3.9	1.0
AGAL332.995−00.519	2.511	0.467	18.2	2.593	−52.8			3.3	1.0
AGAL333.001−00.436	2.848	0.865	35.7	4.989	−55.4	−3.4	0.7	1.0	2.5
AGAL333.013−00.466	2.821	0.606	34.5	4.968	−53.3	−1.74	1.2	3.4	0.6
AGAL333.014−00.521	2.986	1.056	22.6	4.152	−53.2	−2.58	1.0	1.3	2.0
AGAL333.053+00.029	3.017	1.056	29.2	4.398	−45.3	1.1	1.4	−0.42	2.1
AGAL333.071−00.399	3.281	0.969	19.1	3.744	−53.3			0.9	1.2
AGAL333.071−00.399	3.281	0.969	19.1	3.744	−53.3	−1.14	1.5	0.2	9.0
AGAL333.089−00.352	2.347	0.692	28.1	3.95	−52.7			0.6	2.5
AGAL333.094−00.524	2.728	0.623	26.5	3.966	−57.6	−1.64	1.5	0.8	3.5
AGAL333.103+00.087	2.071	0.169	17	2.119	−44.9	0.3	2.0	−0.26	3.0
AGAL333.103−00.502	3.251	1.367	25.5	4.601	−55.8	−0.81	6.0	1.3	1.0
AGAL333.129−00.559	3.843	1.402	16.1	3.983	−56.4	−0.8	4.0	2.4	2.0
AGAL333.134−00.431	3.959	1.506	32	5.823	−51.9	−3.3	1.3	2.0	1.7
AGAL333.169−00.431	2.959	0.831	23	4.052	−50.8	−0.63	2.2	2.0	1.3
AGAL333.248+00.054	3.27	1.454	26.3	4.706	−46.9			−0.81	5.5
AGAL333.264−00.291	2.229	0.485	27.4	3.951	−52.2			−3.4	0.8
AGAL333.284−00.387	3.678	1.211	32.1	5.435	−51.6			0.7	3.0
AGAL333.284−00.387	3.678	1.211	32.1	5.435	−51.6	−1.66	1.8	2.5	1.0
AGAL333.308−00.366	3.561	1.246	34	5.544	−50.2	−4.51	0.4	0.9	3.8
AGAL333.308−00.366	3.561	1.246	34	5.544	−50.2			2.8	1.0
AGAL333.308−00.366	3.561	1.246	34	5.544	−50.2	−2.5	1.2	2.8	0.7
AGAL333.308−00.366	3.561	1.246	34	5.544	−50.2	−0.32	6.0	0.2	6.0
AGAL333.314+00.106	2.901	0.71	23.3	3.922	−47.7	2.4	1.3	−0.88	2.7
AGAL333.466−00.164	3.427	1.038	24.6	4.591	−42.5	2.3	1.5	−1.57	5.0
AGAL333.483−00.246	3.133	0.623	18	3.255	−50.4			−1.57	5.0
AGAL333.499−00.329	2.163	0.433	23.9	3.244	−50.2			−2.18	1.0
AGAL333.506−00.101	2.653	0.744	19.7	3.096	−42.8			1.3	2.0
AGAL333.524−00.269	3.488	1.54	22.5	4.252	−48.8	−0.42	3.5	4.9	1.3
AGAL333.524−00.269	3.488	1.54	22.5	4.252	−48.8	−1.41	0.7		
AGAL333.539−00.242	2.93	0.727	22.8	3.873	−48.9	−0.79	2.5	1.5	1.0
AGAL333.566−00.296	2.982	1.09	21.7	3.881	−46.7	−2.56	1.9	1.3	3.3
AGAL333.651−00.334	2.832	0.796	18.2	3.083	−47.1	−1.51	3.5	1.6	1.0
AGAL333.683−00.256	2.83	0.883	28.3	4.18	−48.1	−2.83	1.0	0.9	1.5
AGAL333.696−00.231	2.519	0.398	29.7	3.653	−46.7			−0.29	4.8
AGAL333.774−00.256	2.993	0.692	15.5	3.061	−49.5	−2.7	1.3	1.0	0.7
AGAL331.706−00.266	2.674	0.661	16.1	2.631	−49			−0.92	1.6
AGAL331.856−00.126	3.05	0.947	18.2	3.539	−50.1			−0.39	3.2

Notes. ΔV and L represent the velocity gradient and the length of the filament. As shown in Fig. 7, there are two velocity gradients around one intensity peak, marked 1 and 2. However, here we only reserve the good fittings, thus many hubs only have a velocity gradient of one side.

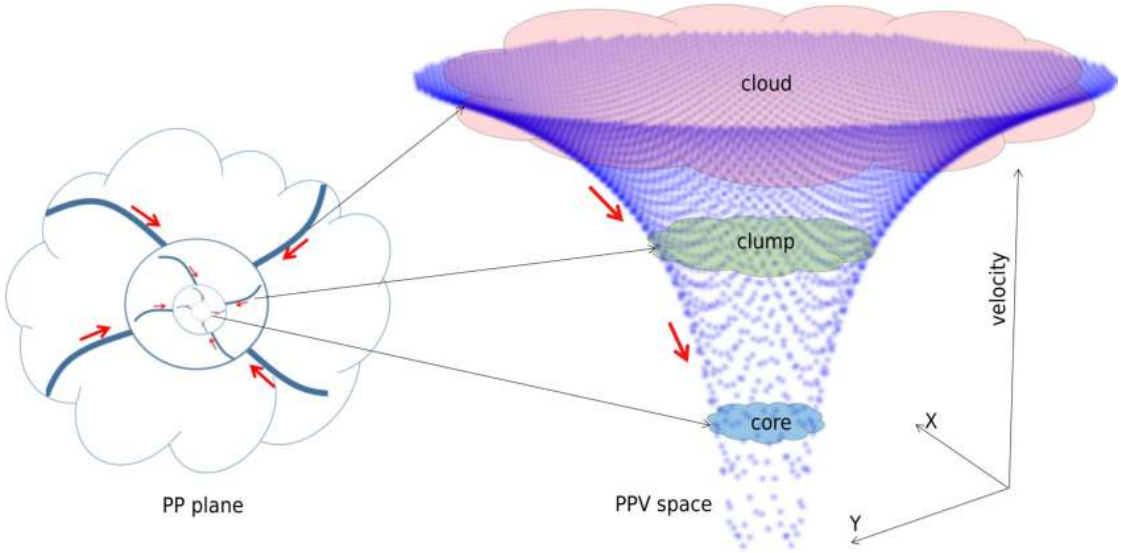


Fig. 9. Schematic diagram of the funnel structure in PPV space and on the PP plane. The PP plane shows the multiscale hub-filament structure in a molecular cloud, which can turn into a funnel structure in PPV space. The red arrow represents the gas inflow.

Table 2. Remaining part of Table 1.

Clump	log(Mass) ($M_\odot$)	Radius (pc)	$T_{\rm dust}$ (K)	log(Lum) ($L_\odot$)	$v_{\rm LSR}$ (km s^{-1})	Δv_1 (km s^{-1} pc^{-1})	L_1 (pc)	Δv_2 (km s^{-1} pc^{-1})	L_2 (pc)
AGAL332.166+00.126	2.515	0.541	14.4	2.316	−49.9	−0.65	2.0	0.7	3.5
AGAL332.226−00.124	2.892	0.556	15.4	2.579	−50.8	−2.63	1.0	1.6	2.0
AGAL332.226−00.124	2.892	0.556	15.4	2.579	−50.8	−1.57	1.5	2.9	0.7
AGAL332.241−00.044	3.226	0.917	17	3.526	−48.1	−1.45	1.0	3.0	1.0
AGAL332.254−00.056	2.862	0.421	14.3	2.675	−47.7			0.8	1.7
AGAL332.296−00.094	3.15	0.932	24	4.303	−48.2	0.6	3.0	−0.36	2.5
AGAL332.317+00.177	2.576	0.406	13.6	2.223	−48.6			2.0	0.5
AGAL332.352−00.116	2.513	0.601	27.9	3.887	−49.8	−0.39	5.0	0.5	2.0
AGAL332.442−00.139	2.723	0.421	18.3	2.669	−51.3			0.7	2.0
AGAL332.469−00.131	2.643	0.556	20.3	3.088	−50.5	1.8	1.2	−1.3	1.2
AGAL332.544−00.124	2.857	1.037	32.5	4.591	−47.5			−1.92	2.4
AGAL332.604−00.167	3.185	1.052	17.7	3.371	−46.2			−0.89	1.6
AGAL332.141−00.466	2.66	0.571	24.2	3.495	−57				
AGAL332.144−00.469	2.711	0.883	24.5	4.203	−58			3.0	0.8
AGAL332.141−00.446	2.883	0.71	27.7	4.76	−56.3				
AGAL332.147−00.439	2.371	0.169	26.9	3.974	−57.9				
AGAL332.156−00.449	3.239	1.004	30	5.078	−55.6			3.0	0.9
AGAL332.252−00.539	2.821	0.45	17.8	2.608	−52.9				
AGAL332.281−00.547	3.133	0.71	16.9	3.299	−52.2	−1.73	1.0	2.2	1.0
AGAL333.521−00.241	2.713	0.415	20	2.93	−48.8				
AGAL333.524−00.269	3.488	1.54	22.5	4.252	−48.8	2.3	1.0	−1.36	1.0
AGAL332.442−00.139	2.723	0.421	18.3	2.669	−51.3				
AGAL332.469−00.131	2.643	0.556	20.3	3.088	−50.5	−0.74	2.0	0.7	4.5
AGAL332.751−00.597	2.964	1.038	24.7	3.937	−53.1				
AGAL332.774−00.584	3.103	1.038	23	4.183	−55.2	−1.74	3.3	5.6	1.0

Notes. Usually, one clump corresponds to one hub. Here, two or three clumps separated by the gap indicate that the two or three clumps together form a hub.

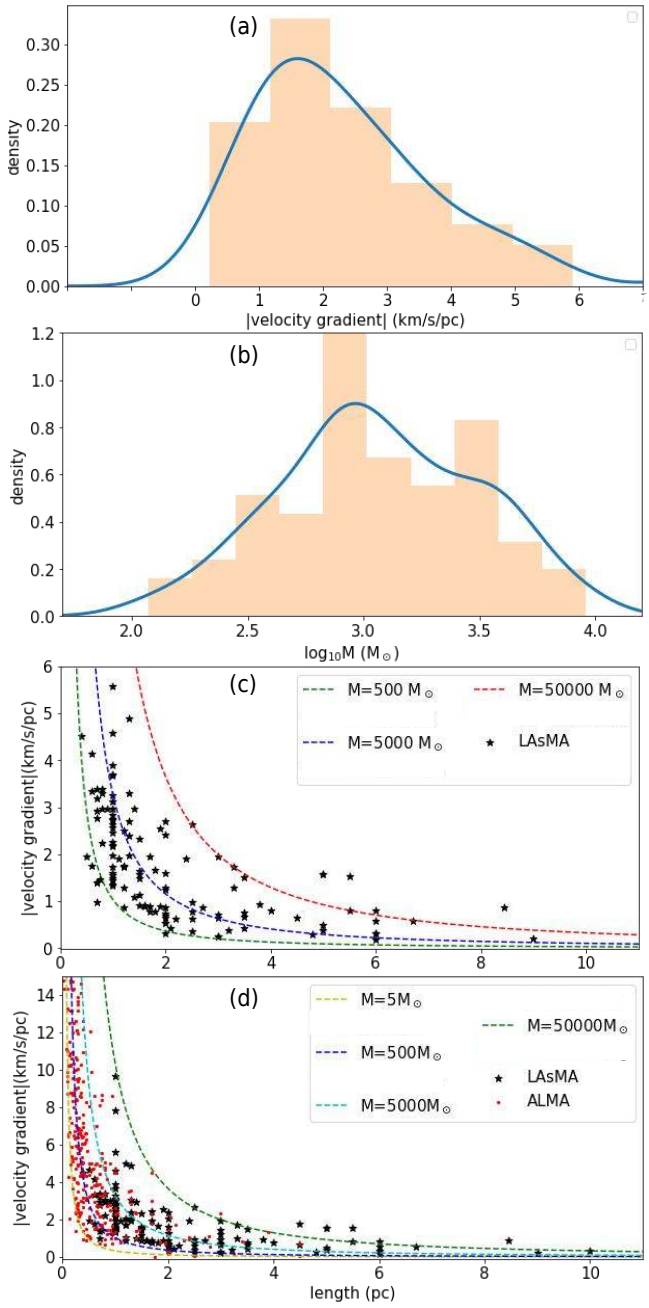


Fig. 10. Statistical analysis of the velocity gradients listed in Table 1. (a): Probability distribution of the velocity gradients measured around the 1 pc scale. (b): Mass distribution of ATLASGAL clumps listed in Tables 1 and 2. Panels c and d: same as Fig. 8, but only for the velocity gradients listed in Tables 1 and 2.

collapsing object, which is also consistent with the survey results that most Galactic parsec-scale massive clumps seem to be gravitationally bound, regardless of how evolved they are (Liu et al. 2016a; Urquhart et al. 2018; Evans et al. 2021).

Based on Eq. (1), we can use the values of M, L, and ∇V listed in Tables 1 and 2 to compare the observations to the simple theory. In Fig. 11 the observed and predicted accelerations are clearly correlated. This is strong evidence that gravity accelerates the gas inflow. The scatter in Fig. 11 is to be expected because we measured only the projection of the realistic velocity vector and hence acceleration. For the same reason, we tend to underestimate the observed acceleration. Moreover, in Tables 1

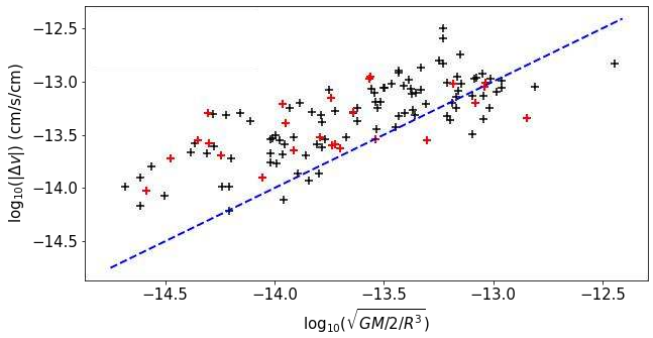


Fig. 11. Correlation between the left and right terms in Eq. (1) revealed by the values of M, L, and ∇V listed in Tables 1 and 2. The dotted blue line is the line of equality of the left and right terms. Tables 1 and 2 show two cases. A dense clump with two or more filaments leading to it, or only one filament with a gradient that can be identified (the rows with only one entry for the gradient). The number of clumps in the two cases is 49 and 25, and they marked by black and red plus, respectively.

and 2, we only selected the ATLASGAL clumps associated with the intensity peaks of the ^{13}CO (3−2) emission (a deviation within the effective radius of the clump), thus some of them are not the exact gravitational centers. Despite these caveats, the correlation between observations and predictions for the denser clumps is quite good. The less dense clumps tend to lie above the line of equality, suggesting that their masses underestimate the total mass contributing to the gravitational acceleration.

4.3. Large-scale inflow driven by gravity

As shown in Figs. 8 and 10, the velocity gradients we fit to the LAsMA and ALMA data agree with each other over the range of scales covered by ALMA observations in the ATOMS survey (<5 pc). Interestingly, the variations in velocity gradients at small scales (<1 pc), medium scales (~1–7.5 pc), and large scales (>7.5 pc) are consistent with gravitational freefall with central masses of ~500 $M_\odot$, ~5000 $M_\odot$, and ~50 000 $M_\odot$. This means that the velocity gradients on larger scales require a higher mass to be maintained, which is also revealed by the funnel structure shown in Fig. 9. Higher masses imply larger scales, that is to say, the larger-scale inflow is driven by the larger-scale structure, which may be the gravitational clustering of smaller-scale structures (Sect. 4.4), consistent with the hierarchical structure in molecular clouds and the gas inflow from large to small scales. Thus the large-scale gas inflow may also be driven by gravity, with the molecular clouds in G333 in a state of global gravitational collapse, which is consistent with the argument that these molecular clouds act as cloud-scale hub-filament structures (see the description in Sect. 3.4). Moreover, as described in Sect. 3.5, the funnel structure of the velocity field shown in PPV space also tends to support the large-scale gravitational collapse of the molecular clouds in the G333 complex.

4.4. Hierarchical gravitational structure in the molecular cloud

The scales of the hub-filament structures (cloud-scale) in this work are much larger than those (clump-scale) in Zhou et al. (2022). Each large-scale hub-filament structure will contain many sub-structures, and the kinematic structure will therefore be very complex. Nevertheless, if they are hub-filament systems, we expect that they have similar kinematic features as the clump-scale hub-filament system. A hub-filament system must have a common gravitational center, which will dominate the overall

gravitational field and thus the velocity field. Although Fig. 4 shows many substructures in each subregion, which cause many velocity and density fluctuations, the entire large-scale structure globally still displays the funnel structure. As discussed in Sects. 3.3 and 4.3, large-scale velocity gradients always involve many intensity peaks, and the larger-scale inflow is driven by the larger-scale structure, implying that the clustering of local small-scale gravitational structures can act as the gravitational center on a larger scale. To some extent, the funnel structure gives an indication of the gravitational potential well that is formed by the clustering.

Both Zhou et al. (2022) and this work find that when the scale is smaller than ∼1 pc, the gravitationally dominated gas infall can become apparent. Thus, relatively small-scale hub–filament structures will have better and more recognizable morphology than large-scale ones due to the strong local gravitational field. For the large-scale hub–filament structures, background turbulence is more likely to disturb their morphology due to weaker gravitational confinement on large scales, thus they are more strongly blurred. Zhou et al. (2023) suggested for the interpretation of the formation of the large-scale hub-filament structure in W33 complex that IRDCs and ATLASGAL clumps in W33-blue can be treated as small-scale hub–filament structures, and W33-blue itself as a huge hub–filament structure. This means that there may be a self-organization process from small-scale hub–filament structures to a large-scale one. If the large-scale gravitational center is the clustering of small-scale ones, the large-scale gravity center will be more loose compared to the latter. Its ability to control the gas motions on a large scale is therefore also weaker, which is reflected as the loose funnel structure and small velocity gradient, as shown in this work. In Fig. 6 the small-scale structures indeed have clearer funnel structures than the large-scale ones. This indicates the role of gravity in shaping the funnel structure.

4.5. G333 in the context of Galactic molecular clouds

The same region of the sky, with a slightly larger area ($332.6° < l < 333.8°$ and $-0.8° < b < 0.1°$), was also studied by Nguyen et al. (2015) using CO and ^{13}CO (1–0) emission and archival data. They called the GMC G333 the RCW 106 complex after the associated H II region. They determined the following properties for the RCW 106 complex: the total effective diameter is 183 pc, the mass is $5.9 \times 10^6\ M_\odot$, and the mass surface density is $220\ M_\odot\ \mathrm{pc}^{-2}$. The virial parameter for the whole complex is 0.35, making it strongly gravitationally bound. Because the cloud is elongated, they included a linear analysis that also implied a strongly bound cloud. These properties would mean that the GMC G333 iss one of the largest most massive complexes in the Galaxy (see Fig. 7 of Miville-Deschênes et al. 2017), and one of the few that is gravitationally bound (see Fig. 2 of Evans et al. 2021). However, Nguyen et al. (2015) took the velocity range of the RCW 106 complex as [−80, −40] km s^{-1}. As discussed in Sect. 3.1, after decomposing the velocity components, we find that the velocity range [−80, −40] km s^{-1} also includes the GMC G331. The mass of the GMC G333 is therefore overestimated in Nguyen et al. (2015). The catalog of Miville-Deschênes et al. (2017), restricted to the velocity range of peak3, yields a mass for the G333 cloud traced by CO 1–0 emission of $\sim 1.7 \times 10^6\ M_\odot$, with the surface density $\sim 120\ M_\odot\ \mathrm{pc}^{-2}$. Tracers that are biased toward higher volume densities find progressively lower masses. Karnik et al. (2001) made a far-infrared (FIR) 150 and 210 µm dust emission study of the region using the canonical gas-to-dust ratio of 100. They measured a total GMC mass of $\sim 1.8 \times 10^5\ M_\odot$. This is consistent with the total mass of $\sim 1 \times 10^5\ M_\odot$ estimated by Mookerjea et al. (2004) from a 1.2 mm cold dust continuum emission map and the same gas-to-dust ratio. These estimates show the decreasing fraction of mass in progressively denser regions.

The lower total mass from CO 1–0 based on the catalog of Miville-Deschênes et al. (2017) still makes the G333 complex one of the most massive in the Milky Way. The analysis of the RCW HII region yields a stellar mass of $48 \times 10^3\ M_\odot$. Assuming a timescale of 0.2 Myr (for an O7 star starting from its first appearance on the main-sequence diagram), Nguyen et al. (2015) derived a star formation rate so far of $0.25\ M_\odot\ \mathrm{Myr}^{-1}$ and an efficiency (M_*/M_{gas}) of 0.008. The very high star formation rate led Nguyen et al. (2015) to describe it as a mini-starburst. It is thus atypical of most molecular clouds, but an exemplar of the very small subset that forms the majority of stars in the Galaxy.

5. Summary

We investigated the gas kinematics of hub-filament structures in the G333 complex using ^{13}CO (3−2) emission. The main conclusions are listed below.

1. The G333 complex includes three main velocity components, which can be well separated by the average spectra and the velocity distribution in PPV space. We derived the velocity distribution from the decomposition of ^{13}CO (3–2) emission by `GAUSSPY+`. The velocity components with velocity ranges [−120, −76] km s^{-1} and [−76, −60] km s^{-1} represent the main two parts of GMC G331. The third velocity component with the velocity range [−60, −35] km s^{-1} shows a very extended shell structure from $l = 330.6°$ to $l = 334°$ with GMC G333 and the G332 ring as the main parts.

2. The entire G333 complex seems to have fragmented into several independent molecular clouds. Some of the substructures in these molecular clouds show typical hub-filament structures, which also have the typical kinematic features of hub-filament systems. The broken morphology of some very infrared bright structures indicates that the feedback is disrupting the star-forming regions.

3. We used the integrated intensity maps of ^{13}CO J = 3–2 to identify filaments in the G333 complex by FILFINDER algorithm and extracted the velocity and intensity along the filament skeleton from moment maps. We find a good correlation between velocity and density fluctuations, and we fit the velocity gradients around the intensity peaks. The change in velocity gradients with the scale indicates that the morphology of the velocity field in PPV space resembles a funnel structure. The funnel structure can be explained as accelerated material inflowing toward the central hub and gravitational contraction of star-forming clouds or clumps. The typical V-shape velocity structure along the filament skeleton can be treated as the projection of the funnel structure from PPV space onto the PV plane. Hence, the funnel structure in PPV space can be an effective probe for the gravitational collapse motion in molecular cloud.

4. We investigated gas motions on sufficiently large scales. The longest filament was ∼50 pc. We obtained results similar to small-scale ALMA observations, however. The typical velocity gradient corresponding to a one-parsec scale is $\sim 1.6\ \mathrm{km\ s}^{-1}\ \mathrm{pc}^{-1}$. Assuming the freefall model, we can predict the gravitational acceleration onto each hub. The observed accelerations correlate well with the predicted ones and the values arecomparable to the prediction. Because we observe only one component of the acceleration and because the masses of the hubs are uncertain,

the result provides strong evidence for gravitational acceleration of material flowing into hubs from filaments.

5. The velocity gradients fit with LAsMA data and ALMA data agree with each other over the scales covered by ALMA observations in the ATOMS survey (<5 pc). On large scales, we find that the larger-scale inflow is driven by the larger-scale structure, indicating the hierarchical structure in the molecular cloud and the gas inflow from large scale to small scale. The large-scale gas inflow is driven by gravity, implying that the molecular clouds in G333 may be in a state of global gravitational collapse. The funnel structure of the velocity field shown in PPV space also tends to support the large-scale gravitational collapse.

6. Although there are many substructures in each molecular cloud that cause ubiquitous velocity and density fluctuations, the entire large-scale structure in total still displays a loose funnel structure. Thus the hub-filament structures at different scales may be organized into a hierarchical macroscopic system through the coupling of gravitational centers at different scales.

In short, the changes in velocity gradient with scale indicate a funnel structure of the velocity field in PPV space. This is indicative of a smooth, continuously increasing velocity gradient from large to small scales, and it is thus consistent with gravitational acceleration.

Acknowledgements. This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under programme ID M-0109.F-9514A-2022. APEX has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. This work has been supported by the National Key R&D Program of China (No. 2022YFA1603100). T.L. acknowledges the supports by National Natural Science Foundation of China (NSFC) through grants No.12122307 and No.12073061, the international partnership program of Chinese Academy of Sciences through grant No.114231KYSB20200009, and Shanghai Pujiang Program 20PJ1415500.

References

Bains, I., Wong, T., Cunningham, M., et al. 2006, MNRAS, 367, 1609
Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F., & Zamora-Avilés, M. A. 2011, MNRAS, 411, 65
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953
Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 1997, MNRAS, 285, 201
Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 2001, MNRAS, 323, 785
Breen, S. L., Ellingsen, S. P., Johnston-Hollitt, M., et al. 2007, MNRAS, 377, 491
Breen, S. L., Ellingsen, S. P., Caswell, J. L., et al. 2012, MNRAS, 421, 1703
Bronfman, L., Alvarez, H., Cohen, R. S., & Thaddeus, P. 1989, ApJS, 71, 481
Caswell, J. L. 1998, MNRAS, 297, 215
Caswell, J. L., & Haynes, R. F. 1987, A&A, 171, 261
Caswell, J. L., Haynes, R. F., & Goss, W. M. 1980, Australian J. Phys., 33, 639
Caswell, J. L., Vaile, R. A., & Forster, J. R. 1995, MNRAS, 277, 210
Caswell, J. L., Fuller, G. A., Green, J. A., et al. 2011, MNRAS, 417, 1964
Chen, H.-R. V., Zhang, Q., Wright, M. C. H., et al. 2019, ApJ, 875, 24
Chen, M. C.-Y., Di Francesco, J., Rosolowsky, E., et al. 2020, ApJ, 891, 84
Clarke, S. D., Whitworth, A. P., & Hubber, D. A. 2016, MNRAS, 458, 319
Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, A&A, 585, A104
Cunningham, M., Lo, N., Kramer, C., et al. 2008, EAS Publ. Ser., 31, 9
Dewangan, L. K., Ojha, D. K., Sharma, S., et al. 2020, ApJ, 903, 13
Elmegreen, B. G., Elmegreen, D. M., & Efremov, Y. N. 2018, ApJ, 863, 59
Evans, N. J., II, Heyer, M., Miville-Deschênes, M.-A., Nguyen-Luong, Q., & Merello, M. 2021, ApJ, 920, 126
Fujiyoshi, T., Smith, C. H., Caswell, J. L., et al. 2006, MNRAS, 368, 1843
Gómez, G. C., & Vázquez-Semadeni, E. 2014, ApJ, 791, 124
Gómez, G. C., Vázquez-Semadeni, E., & Palau, A. 2021, MNRAS, 502, 4963
Goss, W. M., Manchester, R. N., & Robinson, B. J. 1970, Australian J. Phys., 23, 559
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13
Hacar, A., & Tafalla, M. 2011, A&A, 533, A34
Hacar, A., Alves, J., Tafalla, M., & Goicoechea, J. R. 2017, A&A, 602, L2
Hartmann, L., Ballesteros-Paredes, J., & Heitsch, F. 2012, MNRAS, 420, 1457
Henshaw, J. D., Caselli, P., Fontani, F., Jiménez-Serra, I., & Tan, J. C. 2014, MNRAS, 440, 2860
Henshaw, J. D., Longmore, S. N., & Kruijssen, J. M. D. 2016, MNRAS, 463, L122
Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nat. Astron., 4, 1064
Issac, N., Tej, A., Liu, T., et al. 2019, MNRAS, 485, 1775
Jordan, C. H., Walsh, A. J., Lowe, V., et al. 2013, MNRAS, 429, 469
Karnik, A. D., Ghosh, S. K., Rengarajan, T. N., & Verma, R. P. 2001, MNRAS, 326, 293
Kirk, H., Myers, P. C., Bourke, T. L., et al. 2013, ApJ, 766, 115
Koch, E. W., & Rosolowsky, E. W. 2015, MNRAS, 452, 3435
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, ApJ, 656, 959
Kuznetsova, A., Hartmann, L., & Ballesteros-Paredes, J. 2015, ApJ, 815, 27
Kuznetsova, A., Hartmann, L., & Ballesteros-Paredes, J. 2018, MNRAS, 473, 2372
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Liu, T., Kim, K.-T., Yoo, H., et al. 2016a, ApJ, 829, 59
Liu, T., Zhang, Q., Kim, K.-T., et al. 2016b, ApJ, 824, 31
Liu, H.-L., Stutz, A., & Yuan, J.-H. 2019, MNRAS, 487, 1259
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, MNRAS, 496, 2790
Liu, H.-L., Tej, A., Liu, T., et al. 2022, MNRAS, 511, 4480
Lo, N., Cunningham, M. R., Jones, P. A., et al. 2009, MNRAS, 395, 1021
Lockman, F. J. 1979, ApJ, 232, 761
Lowe, V., Cunningham, M. R., Urquhart, J. S., et al. 2014, MNRAS, 441, 256
Lu, X., Zhang, Q., Liu, H. B., et al. 2018, ApJ, 855, 9
Mazumdar, P., Wyrowski, F., Colombo, D., et al. 2021, A&A, 650, A164
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
McKee, C. F., & Tan, J. C. 2003, ApJ, 585, 850
Merello, M., Bronfman, L., Garay, G., et al. 2013, ApJ, 774, 38
Misugi, Y., Inutsuka, S.-i., & Arzoumanian, D. 2019, ApJ, 881, 11
Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57
Mookerjea, B., Kramer, C., Nielbock, M., & Nyman, L. Å. 2004, A&A, 426, 119
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Nguyen, H., Nguyen Lu'o'ng, Q., Martin, P. G., et al. 2015, ApJ, 812, 7
Padoan, P., Pan, L., Juvela, M., Haugbølle, T., & Nordlund, Å. 2020, ApJ, 900, 82
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Pestalozzi, M. R., Minier, V., & Booth, R. S. 2005, A&A, 432, 737
Pillai, T. G. S., Clemens, D. P., Reissl, S., et al. 2020, Nat. Astron., 4, 1195
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Romano, D., Burton, M. G., Ashley, M. C. B., et al. 2019, MNRAS, 484, 2089
Russeil, D. 2003, A&A, 397, 133
Russeil, D., Adami, C., Amram, P., et al. 2005, A&A, 429, 497
Smith, R. J., Glover, S. C. O., Klessen, R. S., & Fuller, G. A. 2016, MNRAS, 455, 3640
Ulich, B. L., & Haas, R. W. 1976, ApJS, 30, 247
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
Vallée, J. P. 2008, AJ, 135, 1301
Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., Ballesteros-Paredes, J., & Klessen, R. S. 2009, ApJ, 707, 1023
Vázquez-Semadeni, E., González-Samaniego, A., & Colín, P. 2017, MNRAS, 467, 1313
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Wang, P., Li, Z.-Y., Abel, T., & Nakamura, F. 2010, ApJ, 709, 27
Williams, G. M., Peretto, N., Avison, A., Duarte-Cabral, A., & Fuller, G. A. 2018, A&A, 613, A11
Wong, T., Ladd, E. F., Brisbin, D., et al. 2008, MNRAS, 386, 1069
Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, ApJ, 852, 12
Zhang, Q., Wang, K., Lu, X., & Jiménez-Serra, I. 2015, ApJ, 804, 141
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J.-W., Li, S., Liu, H.-L., et al. 2023, MNRAS, 519, 2391

A&A, 686, A146 (2024)
https://doi.org/10.1051/0004-6361/202449514
© The Authors 2024

Astronomy & Astrophysics

Gas inflows from cloud to core scales in G332.83-0.55: Hierarchical hub-filament structures and tide-regulated gravitational collapse

J. W. Zhou[1], S. Dib[2], M. Juvela[3], P. Sanhueza[4], F. Wyrowski[1], T. Liu[5], and K. M. Menten[1]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de
[2] Max-Planck-Institue für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[3] Department of Physics, PO Box 64, 00014, University of Helsinki, Finland
[4] National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[5] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, PR China

Received 6 February 2024 / Accepted 9 March 2024

ABSTRACT

The massive star-forming region G332.83-0.55 contains at least two levels of hub-filament structures. The hub-filament structures may form through the "gravitational focusing" process. High-resolution LAsMA and ALMA observations can directly trace the gas inflows from cloud to core scales. We investigated the effects of shear and tides from the protocluster on the surrounding local dense gas structures. Our results seem to deny the importance of shear and tides from the protocluster. However, for a gas structure, it bears the tidal interactions from all external material, not only the protocluster. To fully consider the tidal interactions, we derived the tide field according to the surface density distribution. Then, we used the average strength of the external tidal field of a structure to measure the total tidal interactions that are exerted on it. For comparison, we also adopted an original pixel-by-pixel computation to estimate the average tidal strength for each structure. Both methods give comparable results. After considering the total tidal interactions, for the scaling relation between the velocity dispersion σ, the effective radius R, and the column density N of all the structures, the slope of the $\sigma - N * R$ relation changes from 0.20 ± 0.04 to 0.52 ± 0.03, close to 0.5 of the pure free-fall gravitational collapse, and the correlation also becomes stronger. Thus, the deformation due to the external tides can effectively slow down the pure free-fall gravitational collapse of gas structures. The external tide tries to tear up the structure, but the external pressure on the structure prevents this process. The counterbalance between the external tide and external pressure hinders the free-fall gravitational collapse of the structure, which can also cause the pure free-fall gravitational collapse to be slowed down. These mechanisms can be called "tide-regulated gravitational collapse".

Key words. stars: formation – stars: imaging – stars: protostars – ISM: clouds – ISM: kinematics and dynamics – ISM: structure

1. Introduction

High-resolution observations of high-mass star-forming regions show that density enhancements are organized in filamentary gas networks, particularly in hub-filament systems. Hub-filament systems are known as a junction of three or more filaments. Filaments have lower column densities compared to the hubs (Myers 2009; Schneider et al. 2012; Kumar et al. 2020; Zhou et al. 2022). In such systems, converging flows funnel matter into the hub through the filaments. Previous studies have suggested that hub-filament systems are birth cradles of high-mass stars and clusters (Peretto et al. 2013; Henshaw et al. 2014; Zhang et al. 2015; Liu et al. 2016, 2021b, 2022, 2023; Contreras et al. 2016; Yuan et al. 2018; Lu et al. 2018; Issac et al. 2019; Dewangan et al. 2020; Kumar et al. 2020; Zhou et al. 2022, 2023; Xu et al. 2023; Dib 2023; Yang et al. 2023). Zhou et al. (2022) studied the physical properties and evolution of hub–filament systems by spectral lines that were observed in the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Liu et al. 2020b). We found that hub-filament structures can exist at various scales from 0.1 parsec to several parsec in very different Galactic environments. Below the 0.1 parcsec scale, slender structures similar to filaments, such as spiral arms, have also been detected in the surroundings of high-mass protostars (Liu et al. 2015; Maud et al. 2017; Izquierdo et al. 2018; Chen et al. 2020; Sanhueza et al. 2021; Olguin et al. 2023). Therefore, a scenario was proposed in Zhou et al. (2022): Self-similar hub-filament structures and filamentary accretion seem to exist at all scales (from several thousand au to several parsec) in high-mass star-forming regions. This paradigm of hierarchical and multi-scale hub-filament structures was generalized from clump-core scale to cloud-clump scale in Zhou et al. (2023). Hierarchical collapse and hub-filament structures feeding the central regions are also described in previous works, such as Motte et al. (2018), Vázquez-Semadeni et al. (2019), Kumar et al. (2020) and references therein.

As discussed in Zhou et al. (2023), the change in velocity gradients with the scale in the G333 complex indicates that the morphology of the velocity field in PPV space resembles a "funnel" structure. The funnel structure can be explained as accelerated material flowing toward the central hub and gravitational contraction of star-forming clouds or clumps. To some extent, the funnel structure gives an indication of the gravitational potential well that is formed by the clustering.

In this work, we focus on a local region G333-s3a in the G333 complex, called G332.83-0.55 hereinafter, which is marked in Fig. 1 of Zhou et al. (2023), also shown in Fig. 1.

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

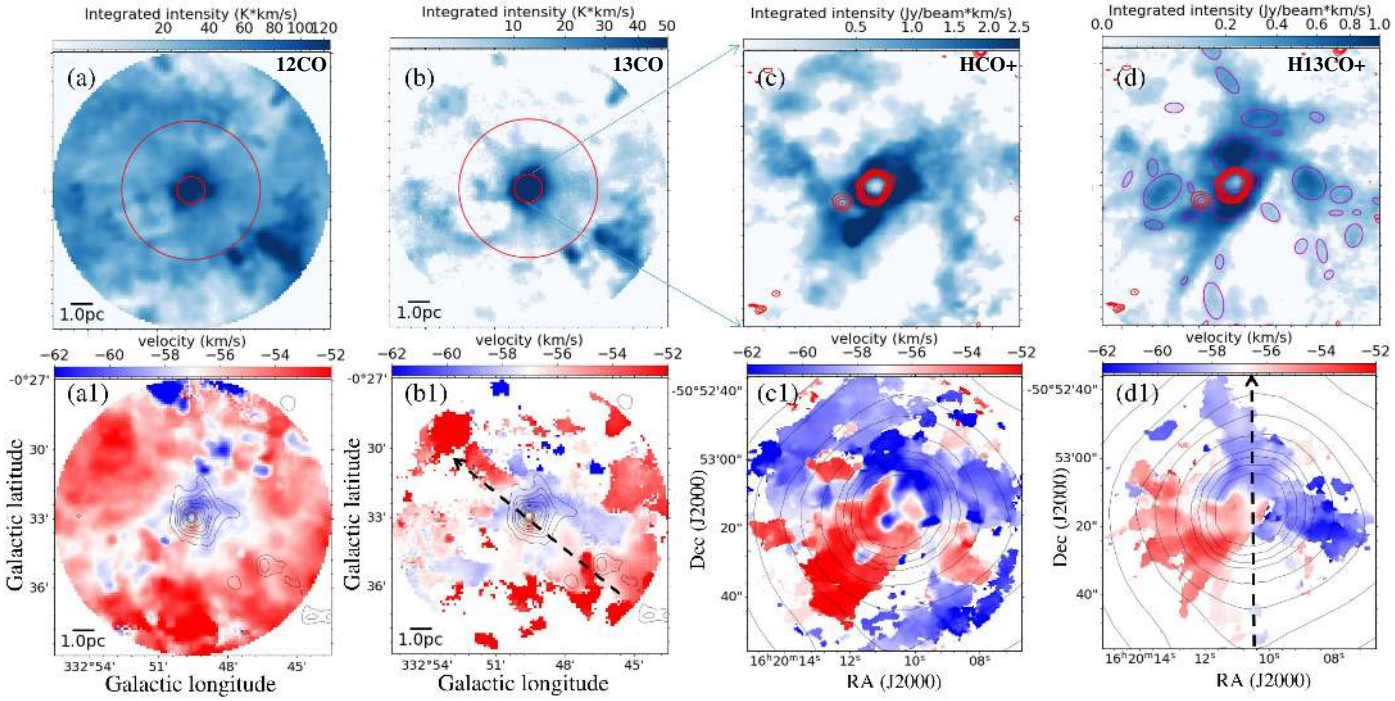


Fig. 1. First and second columns: Moment maps of LAsMA ^{12}CO (3−2) and ^{13}CO (3−2) emission. The radius of the entire region is 0.1° (∼6.28 pc). The big red circle marks the region with a 0.05° (∼3.14 pc) radius, and the small red circle shows the region with a 0.01° (∼0.63 pc) radius. Third and fourth columns: moment maps of ALMA 7 m+12 m HCO$^+$ J = 1–0 and H^{13}CO$^+$ J = 1–0 emission. The radius of the entire region is 0.01° (∼0.63 pc). Red contours in panels c and d show the 3 mm continuum emission. Black contours in the second row show ATLASGAL 870 µm continuum emission. Dashed black lines in panels b1 and d1 show the paths of the PV diagrams in Fig. 4. Ellipses in panel d represent the dendrogram structures identified in Sect. 3.3.

G332.83-0.55 presents a typical hub-filament morphology at ∼10 pc. As a cloud-scale hub-filament system, its hub corresponds to the ATLASGAL clump AGAL332.826-00.549. High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation allowed us to trace the gas kinematics and gas inflows from cloud to clump scales in G332.83-0.55. We also zoomed in on the hub region of G332.83-0.55 using high-resolution ALMA HCO$^+$ (1–0) and H^{13}CO$^+$ (1–0) observations to investigate gas motions from clump to core scales. By combing LAsMA and ALMA observations, we could directly recover the hierarchical or multi-scale hub-filament structures and gas inflows in G332.83-0.55. In addition, we were also able to characterize the details of the funnel structure at different scales.

2. Observations and data

2.1. LAsMA data

The seven pixel Large APEX sub-Millimeter Array (LAsMA) receiver of the APEX telescope was used to observe the J = 3−2 transitions of ^{12}CO and ^{13}CO in the G333 molecular cloud complex. A detailed description of data calibration, imaging, and product creation procedures is presented in Zhou et al. (2023). The final data cubes have a velocity resolution of 0.25 km s^{-1} and an angular resolution of 19.5″ and the pixel size is 6″. The final noise levels of the ^{12}CO (3–2) and ^{13}CO (3–2) data cubes are ∼0.32 K and ∼0.46 K, respectively. In this paper, we only focus on the subregion G332.83-0.55.

2.2. ALMA data

We selected one source from the ATOMS survey (Project ID: 2019.1.00685.S; PI: Tie Liu). The details of the 12 m array and 7 m array ALMA observations were summarized in Liu et al. (2020b, 2021a). Calibration and imaging were carried out using the CASA software package version 5.6 (McMullin et al. 2007); more details can be found in Zhou et al. (2021). All images have been primary-beam corrected. The continuum image reaches a typical 1σ rms noise of ∼0.2 mJy in a synthesized beam FWHM size of ∼2.2″. In this work, we also use the H^{13}CO$^+$ J = 1–0 and HCO$^+$ J = 1–0 molecular lines. The final noise levels of the H^{13}CO$^+$ J = 1–0 and HCO$^+$ J = 1–0 data cubes are ∼0.008 Jy beam^{-1} and ∼0.012 Jy beam^{-1}, respectively.

2.3. Archival data

We also used ATLASGAL+*Planck* 870 µm data (ATLASGAL combined with *Planck* data), which are sensitive to a wide range of spatial scales at a resolution of ∼21″ (Csengeri et al. 2016).

3. Results

3.1. Connection of gas inflows in LAsMA and ALMA observations

Moment maps of various molecular lines are shown in Fig. 1. The Moment 0 maps of ^{13}CO (3−2), H^{13}CO$^+$ J = 1–0, and HCO$^+$ J = 1–0 emission show a typical hub-filament morphology. The FILFINDER algorithm (Koch & Rosolowsky 2015) was used to identify and characterize filaments traced by H^{13}CO$^+$ (1–0) and ^{13}CO (3−2) emission. Following the method described in Zhou et al. (2022, 2023), we used the integrated intensity (Moment 0) maps to identify filaments. The skeletons of the identified filaments, overlaid on the Moment 0 maps in Fig. 2,

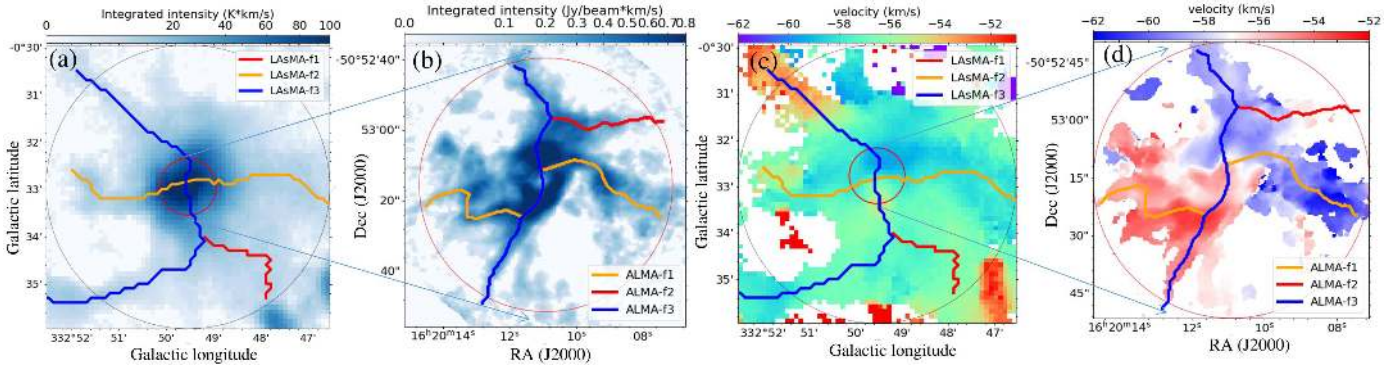


Fig. 2. Hierarchical hub-filament structures in G332.83-0.55. (a) and (b) Moment 0 maps of LAsMA ^{13}CO (3−2) and ALMA H^{13}CO$^+$ (1–0) emission. (c) and (d) Moment 1 maps of LAsMA ^{13}CO (3−2) and ALMA H^{13}CO$^+$ (1–0) emission. The hub region marked by a red circle in panels a and c are enlarged in panels b and d. More detailed structures can be seen in panels b and d, thanks to the high resolution of ALMA observation. Lines in color present the filament skeletons identified by the FILFINDER algorithm.

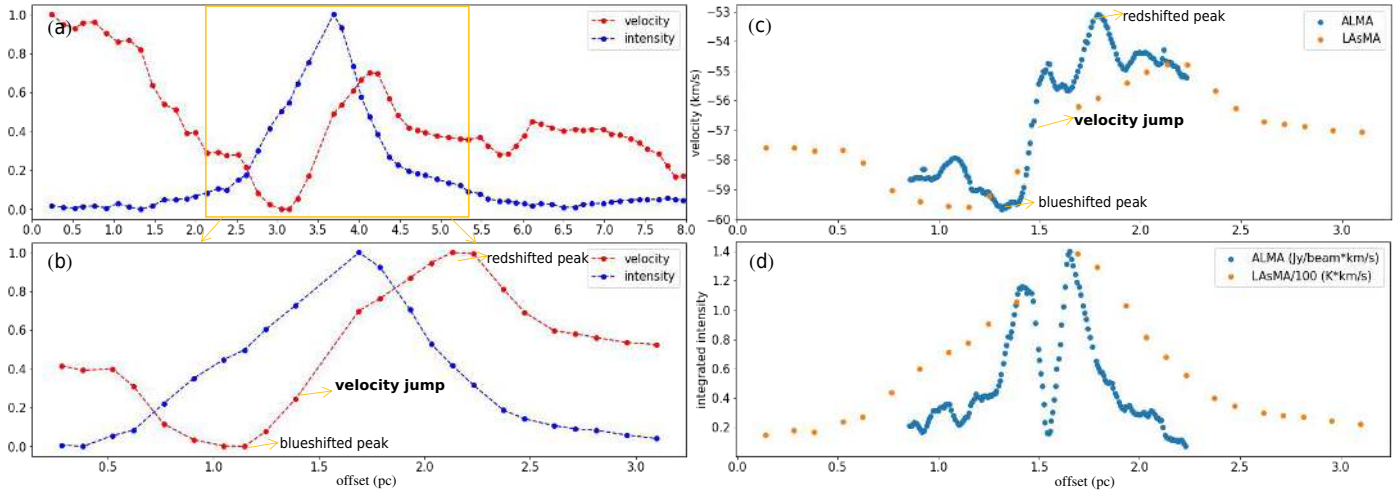


Fig. 3. (a) Normalized velocity and intensity along the filament LAsMA-f3 in Fig. 2a; here we can see a blueshifted V-shape velocity structure. (b) Segment of panel a around the central hub; here we can see a V+Λ-shape velocity structure. (c) and (d) Velocity and intensity along the segment of LAsMA-f3 in panel b compared to the velocity and intensity along the filament ALMA-f3 in Fig. 2b.

are highly consistent with the gas structures. We selected the filaments through the redshifted and blueshifted peaks (LAsMA-f3 and ALMA-f3), and extracted the velocity and intensity along the filaments, as shown in Fig. 3.

Figure 4 displays the position-velocity (PV) diagrams along the dashed lines in panels b1 and d1 of Fig. 1. For H^{13}CO$^+$ $J = 1$–0 emission, feedback from the central UC-HII region identified in Zhang et al. (2023) has destroyed the surrounding molecular gas structure (Fig. 1d). Thus, the diagram looks broken, but it still has a clear double-peak velocity structure, indicating that the surrounding gas is still infalling onto the protocluster (Sect. 3.2). In Fig. 4b, the central part of the PV diagram has a similar morphology to Fig. 4a. Large-scale gas structures with clear velocity gradients are connected to the central velocity jump. Correspondingly, large-scale filamentary structures connected to the central clump can be seen in Figs. 1a and b; small-scale filamentary structures down to dense core scales are also presented in Figs. 1c and d. These features indicate that cloud-scale gas inflows extend down to core scales. The inheritance of gas flows on large scales to small scales is clearly visible in Fig. 3c.

Figure 5a displays the PPV map of ^{13}CO (3−2) emission decomposed by the fully automated Gaussian decomposer `GAUSSPY+` (Lindner et al. 2015; Riener et al. 2019); the decomposition was carried out in Zhou et al. (2023). We can see a funnel structure that can be a probe of gravitational collapse as discussed in Zhou et al. (2023). At small scales, we can roughly identify a velocity jump comprised of the redshifted and blueshifted peaks in Fig. 2c, which is clearly shown in Fig. 2d, thanks to the higher resolution of the ALMA data.

To conclude, G332.83-0.55 shows at least two levels of hub-filament structures. Figures 1a and b present the first level from cloud to clump scales. The second level from clump to core scales is shown in Figs. 1c and d. In a statistical study of hub-filament systems in Zhou et al. (2022), we found filaments are ubiquitous in proto-clusters, and hub-filament systems are very common from dense core scales (∼0.1 pc) to clump and cloud scales (∼1–10 pc). These observed features are reproduced in G332.83-0.55 well.

The structures marked by small red circles in Figs. 1a and b present an approximately spherical morphology due to the limited resolution in single-dish observation, usually referred to as a “clump”. However, in the eyes of ALMA, the spherical clump displays anisotropic and filamentary accretion structures (hub-filament) in Figs. 1c and d. Thus, relative to the traditional “infall

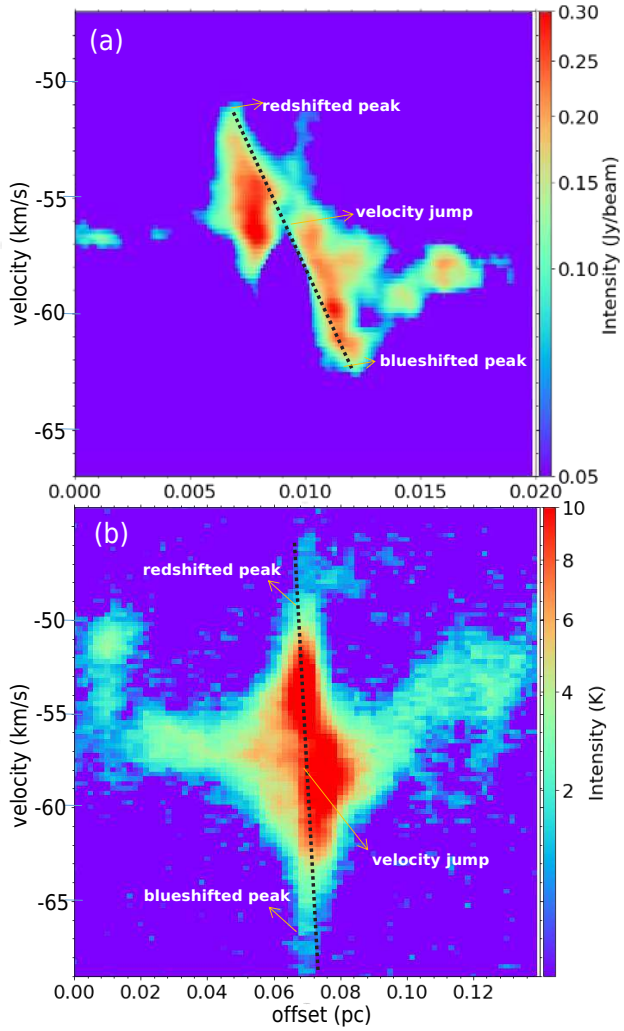


Fig. 4. PV diagrams of ALMA $H^{13}CO^+$ (1–0) and LAsMA ^{13}CO (3–2) along the paths marked in Figs. 1d1 and b1.

envelope", filamentary accretion and inflow structures may be more realistic in high-mass star-forming clumps.

3.2. Infall+rotation motions

In Figs. 1c and d, on small scales, $H^{13}CO^+$ J = 1–0 and HCO^+ J = 1–0 emission present spiral-like structures, indicating the rotation of gas around the protocluster. PV diagrams in Fig. 4 also reveal rotation motions. The PV diagrams show two peaks separated by several km s^{-1}, and they are symmetrically located with respect to the gravitational center. The double-peaked feature on the PV map is often seen in cluster-forming clumps, and this feature can only be reproduced when the clump has both infall motion and rotation at the same time, thus strongly implying that the clump is collapsing with rotation (Ohashi et al. 1997; Shimoikura et al. 2016, 2018, 2022; Sanna et al. 2019; Liu et al. 2020a). Around the gravitational center, we can see the blueshifted and redshifted extrema at the end of the velocity jump in Fig. 3, which is comparable with Fig. 2 of Sakai et al. (2014). In the schematic diagram of an infalling and rotating envelope model for IRAS 04368+2557 (Sakai et al. 2014), the shift of blueshifted and redshifted peaks can be used to estimate the radius of the centrifugal barrier r_c, where the rotation velocity $V_{rotation}$ has its maximum value. According to Eqs. (1) and (2) in Sakai et al. (2014), we can estimate the mass of the gravitational center according to r_c by

$$M_* = \frac{V^2_{\rm rotation} r_c}{2G}. \quad (1)$$

The initial physical quantities in this equation can be estimated from the velocity jump in Fig. 3.

Both the velocity fields traced by LAsMA and ALMA data show a velocity jump around the gravitational center. In Fig. 3c, the velocity fields along the large- and small-scale filaments can be roughly connected, despite the significant differences in their resolutions (2.5″ and 19.5″). According to the velocity difference and the separation between blueshifted and redshifted peaks in Fig. 3c, we roughly estimate $V_{\rm rotation,ALMA} \approx 3.29$ km s^{-1} and $r_{\rm c,ALMA} \approx 0.21$ pc. Then according to Eq. (1), the mass of the gravitational center (i.e., the protocluster) is ∼264 $M_\odot$.

As a further test, following the method described in Kauffmann et al. (2008), the mass can be computed by dust emission using

$$M = 0.12\, M_\odot \left(e^{1.439(\lambda/\rm mm)^{-1}(T/10\ K)^{-1}} - 1\right) \left(\frac{\kappa_\nu}{0.01\ \rm cm^2\ g^{-1}}\right)^{-1} \left(\frac{F_\nu}{\rm Jy}\right) \left(\frac{d}{100\ \rm pc}\right)^2 \left(\frac{\lambda}{\rm mm}\right)^3, \quad (2)$$

where the opacity is

$$\kappa_\nu = 0.1(\nu/1000\ \rm GHz)^\beta\ cm^2 g^{-1} \quad (3)$$

and β is the dust emissivity spectral index. Here, we consider β = 1.5, F_ν is the total flux of 3 mm dust emission, d is the distance to the source, and λ is the wavelength. The temperature T takes the value ∼31.4 K listed in Table A.1 of Liu et al. (2020b). A mass ∼264 $M_\odot$ corresponds to the flux of 3 mm dust emission ∼0.12 Jy. This value is ∼3% of the total flux of 3 mm continuum emission in ALMA observations. As discussed in Keto et al. (2008) and Zhang et al. (2023), 3 mm continuum emission of UC HII regions in high-resolution interferometric observations should be mostly free–free emission rather than from dust, and thus the proportion of ∼3% for the dust emission is reasonable.

3.3. Local dense gas structures

The issues of identifying the structures in a PPV cube by the dendrogram algorithm are described in Zhou et al. (2024b), and thus we adopted the same method as in Zhou et al. (2024a,b) to identify dense gas structures. We directly identified hierarchical (sub)structures based on the 2D integrated intensity (Moment 0) map of $H^{13}CO^+$ (1–0) emission, and then extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. More details can be found in Zhou et al. (2024a). In this work, we focus on leaf structures, which are local gravitational centers. According to their averaged spectra, the structures with absorption features were eliminated first. Following Zhou et al. (2024a), we only consider type 1 (single velocity component) and type 3 (blended velocity component) structures, as marked in Fig. 1d by ellipses.

We derived the column density by combining $H^{13}CO^+$ J = 1–0 and HCO^+ J = 1–0 emission using the same method described in Sect. 3.3 of Zhou et al. (2024a) by the local thermodynamic equilibrium (LTE) analysis. Then the masses and virial ratios of all leaf structures were calculated.

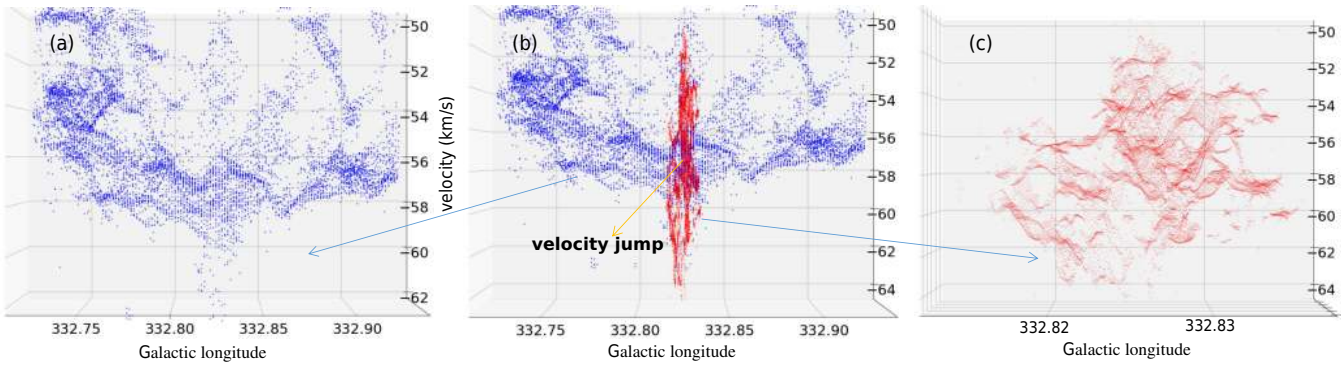


Fig. 5. Linkage of large- and small-scale velocity fields in PPV space. (a) Velocity field revealed by LAsMA data ^{13}CO (3–2) on large scales decomposed by GAUSSPY+. (b) Combined velocity fields traced by ^{13}CO (3–2) and H^{13}CO$^+$ (1–0) at different scales, i.e., panel a plus panel c. (c) Velocity field revealed by ALMA data H^{13}CO$^+$ (1–0) on small scales decomposed by GAUSSPY+.

3.4. Shear and tide from the protocluster

The very large velocity difference (velocity jump) around the protocluster in Figs. 3–5 may indicate that the surrounding gas structures bear considerable shear and tides from the protocluster, apart from infall and rotation motions. Figure 6c shows the scaling relations between the velocity dispersion σ, the effective radius R, and the column density N of all leaf structures, similar to Zhou et al. (2024a,b), the slope of the $\sigma_s - N * R$ relation is ∼0.2, which is much smaller than 0.5 and may indicate a pure free-fall gravitational collapse is slowing down.

To investigate the shear and tides from the protocluster, here we analogy to the methods that are adopted on galaxy-cloud scales. For cloud-scale gas structures in the galaxy, as described in Hunter et al. (1998) and Dib et al. (2012), the shear parameter S_g is defined as

$$S_g = \frac{\Sigma_{\rm cri}}{\Sigma}. \tag{4}$$

We note that Σ is the local gas surface density and $\Sigma_{\rm cri}$ is the critical surface density, which represents the minimal surface density needed to resist the shear. Following Eq. (7) of Dib et al. (2012),

$$\Sigma_{\rm cri} = \frac{A\sigma \ln(C)}{2\pi G}. \tag{5}$$

The C represents the density contrast between leaf structures and their surroundings. The ratio between the average surface density derived from ^{13}CO (3–2) emission in the FOV of the ALMA observation and the highest surface density of leaf structures is ∼100, and thus we consider $C = 100$. We note that σ is the local gas velocity dispersion and A is the Oort constant, which measures the local shear level. For a spherical cloud of radius r, with a rotational velocity V at a galactocentric distance R_G, A can be calculated as

$$A = 0.5\left(\frac{V}{R_G} - \frac{|\,V(R_G + r) - V(R_G - r)\,|}{2r}\right), \tag{6}$$

in the context of molecular clouds, V is determined by the rotation curve of the galaxy. In our case, for gas structures in the clump around the protocluster, V was estimated by the *Kepler* rotation curve, as discussed in Sect.3.2. The $V = \sqrt{GM_0/R'}$, with M_0 being the mass of the protocluster and R' being the distances between gas structures and the protocluster. Shear is effective at tearing apart the gas structure if $S_g > 1$ and it can be ineffective if $S_g < 1$.

In a galactic context, the overall gas and stellar distribution can contribute to confine or disrupt the gas structures. For a cloud at a distance R_G from the galactic center, the edge of the cloud from the galactic center is at a distance of R_G+r. The tidal acceleration in the cloud is $\sim T * r$, following Stark & Blitz (1978),

$$T = \frac{V^2}{R_G^2} - \frac{\partial}{\partial R_G}\left(\frac{V^2}{R_G}\right). \tag{7}$$

Inserting the Keplerian rotation curve, $T = 3GM_0/R'^3$, which is similar to Eq. (8). Then the effective velocity dispersion from the tide for a structure is $\sigma_t^2 \sim T * R^2$, where R is the effective radius of the structure.

In Fig. 6, $S_g << 1$ and $\sigma_t^2/\sigma^2 << 1$; except for two structures adjacent to the protocluster, the results seem to deny the importance of the shear and tide from the protocluster. However, the analysis here is actually not entirely appropriate. For a leaf structure, it also bears the tides from other gas structures, not only the protocluster. In fact, it should be N-body tidal interactions; therefore, a complete tidal calculation would be more appropriate.

3.5. Tidal field

To fully consider the tidal interactions, we directly derived the tidal field according to the density distribution. Only the projected surface density map is available in the observations. To align with other projected physical quantities derived from the observations, we also calculated the 2D projected tide. We note that 3D calculations can refer to Ganguly et al. (2024) and Li (2024). Calculating the gravitational potential field $\phi(x, y)$ based on the surface density map $\Sigma(x, y)$ has been discussed in Gong & Ostriker (2011); He et al. (2023). The tidal tensor described in Appendix A has only two eigenvalues in the case of 2D, that is to say λ_i (i = 1,2). As described in Appendix A, the tidal tensor of a structure includes both self-gravity and external tides, that is, $T = T_{\rm ext} + T_{\rm int}$. The decomposition of the tidal tensor is discussed in detail in Appendix A. For three limiting cases, the external tides can be measured by λ_1, λ_2, and $(\lambda_2 - \lambda_1)/2$ and they are called $T_{\rm d,1}$, $T_{\rm d,2}$, and $T_{\rm d,3}$, respectively. Figure 7 displays different components of the tidal tensor.

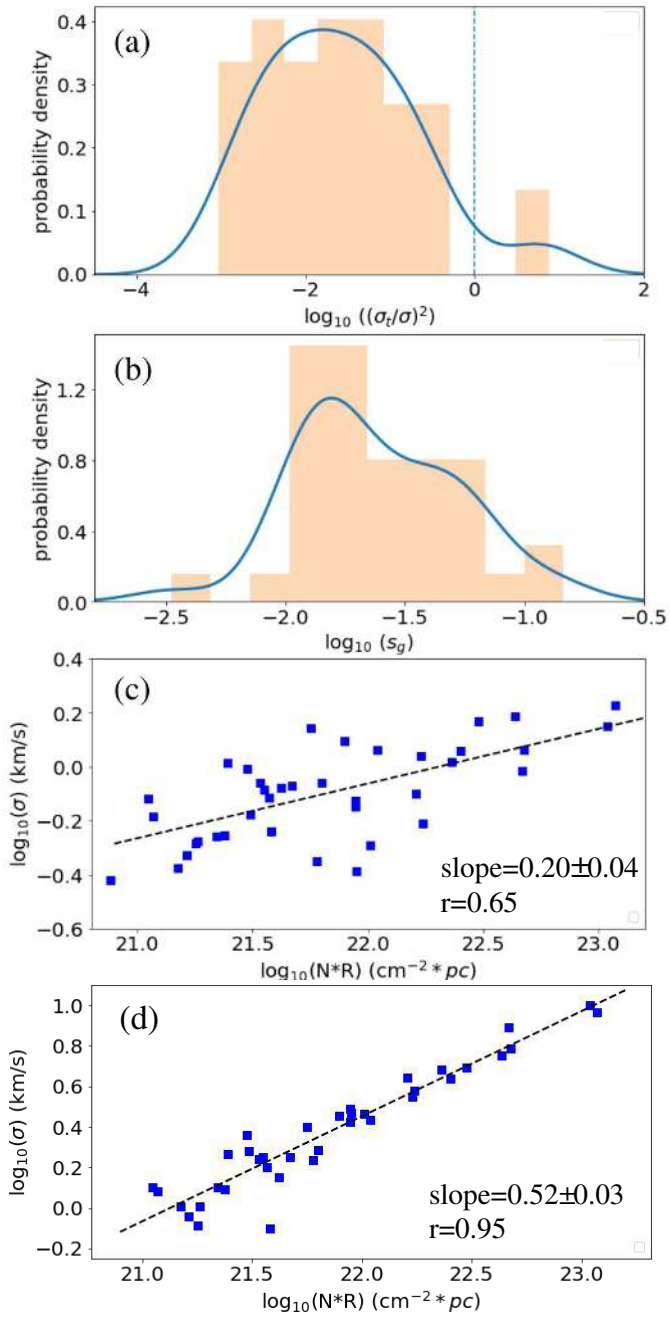


Fig. 6. (a) Distribution of the shear parameter S_g. (b) Distribution of the velocity dispersion ratio between σ_t and σ. (c) $\sigma - N * R$ relation, where σ is the local gas velocity dispersion. (d) $\sigma_s - N * R$ relation, where $\sigma_s = \sqrt{\sigma_t^2 + \sigma^2}$. We note that "r" represents the Pearson correlation coefficient.

3.6. Structure of G332.83-0.55

As shown in Fig. 1 of Zhou et al. (2023), G332.83-0.55 is part of a large-scale shell structure. Thus, it may have a sheet structure. In combining Figs. 1b and 5a, the single direction gas inflows on large scales also indicate that the structure on large scales should be approximately 2D, and should be face-on. Thus, we can observe the unidirectional converging flows. Close to the center, the structure becomes thicker and thicker. The redshifted and blueshifted components may originate from the forward and backward gas motions in the hub region. However, strong rotation motions on small scales may also explain their origin. The 3D and 2D structures of G332.83-0.55 may not be very different, given its limited thickness. Therefore, 2D-projected tidal field strength should be comparable to those for 3D ones, and Eq. (A.8) derived from the 3D tidal tensor in Appendix A should still be effective for a 2D situation, which is confirmed in Fig. 8a.

3.7. Pixel-by-pixel computation

Below we explain how we tested the above tidal tensor analysis using an original method.

3.7.1. Maximum computation

For a structure A, located at a distance R from the center of another structure B, of mass M, the tidal strength sustained by structure A due to the external gravity from structure B is

$$T \approx \frac{2GM}{R^3}. \tag{8}$$

The total tidal strength at a pixel measures the total deformation due to the external tides at that point. A gas structure is neither a rigid body nor point mass, but it is a parcel of gas with malleability. Most gas structures are also not spherically symmetric, and their morphologies can be highly complex. Although the external gravity from all directions may cancel each other out in a structure, the tides inside the structure due to the external gravity do not. Therefore, we used the scalar superposition to measure the overall tidal strength in the structure from the external gravity. Tidal tearing may be an important cause for the deformation and complex morphology of gas structures, as discussed in Sect. 4.1.

Using the mask of the dendrogram structure, we divided the density map into two parts, that is, the structure itself and all other materials apart from this structure. For the external tides from all outside material sustained by a point in structure A, a simple calculation involves using Eq. (8) to directly calculate the tidal strength at that point pixel by pixel, as illustrated in Fig. 9b. The average tidal strength of all the pixels in structure A can then represent the average tidal strength sustained by the structure.

3.7.2. Minimal computation

As illustrated in Fig. 9b, the above computation is limited to a 2D plane. In Eq. (8), the tidal strength is very sensitive to the distance. If the external structure and the calculated structure are not in the same plane, the 2D calculation underestimates the distance between them and thus overestimates the tidal strength (maximum computation, $T_{\rm max}$). However, we only have the 2D projection in the observation. The 3D structure of the molecular gas is unknown, but we can consider an extreme case. Using the average effective radius (~0.03 pc) of all leaf structures as the characteristic scale height of G332.83-0.55, all external material apart from the calculated structure was moved vertically to 0.03 pc away, as illustrated in Fig. 10c. This is unrealistic, given the continuity of the matter distribution in a cloud. As a limiting case, the estimated tidal strength should be smaller than the actual value (minimal computation, $T_{\rm min,s}$).

The other minimal computation ($T_{\rm min,v}$) is the vector superposition of tides, which allows the tidal interactions to directly cancel each other out. Using the calculated point in the target structure as the coordinate origin to build a 2D coordinate system, all external tides were decomposed to the coordinate axes.

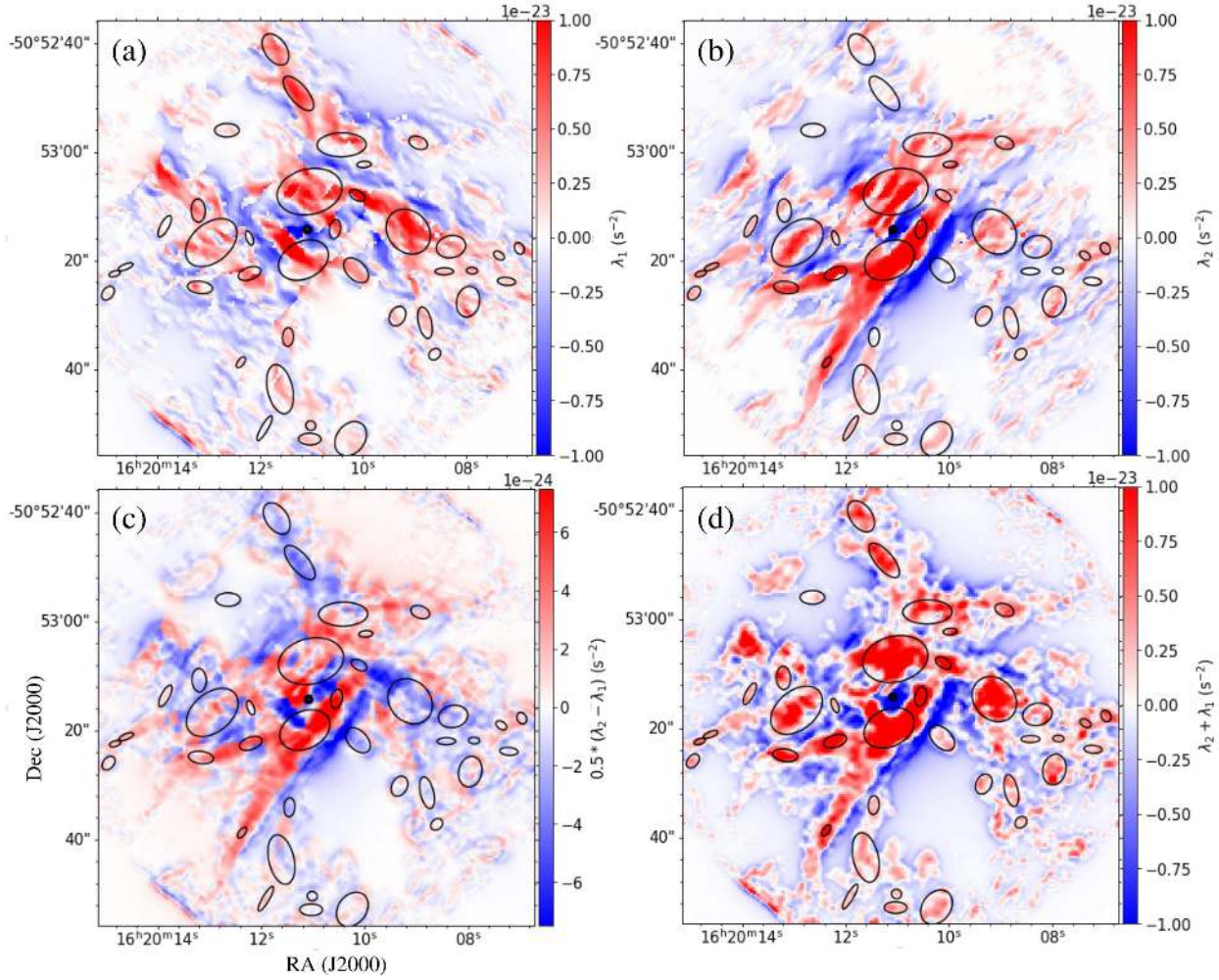


Fig. 7. Different components of the projected tidal field in the FOV of the ALMA observation.

Then we calculated the overall tidal strength at the calculated point after performing vector superposition along each axis.

3.7.3. Tide from the envelope

As shown in Fig. 2, the dendrogram structures are embedded in a larger-scale gas environment. The material outside of the FOV of ALMA observation may also produce considerable external tides ($T_{\rm out}$) on small-scale dendrogram structures. Using the H_2 column density derived from ATLASGAL+*Planck* 870 μm in Zhou et al. (2024b), the matter in the FOV of ALMA observation was first masked, as shown in Fig. 9a. Then we calculated the external tides at each pixel in the FOV. The averaged tidal strength of all pixels is $\sim 1.4 \times 10^{-25}$ s^{-2}. Compared with the tidal strength in Fig. 8b, $T_{\rm out}$ is very small and therefore can be ignored.

3.7.4. Comparison

Figure 8b shows that the components $T_{\rm d,1}$, $T_{\rm d,2}$, and $T_{\rm d,3}$ (whose derivation is detailed in Appendix A) are comparable. Especially, they are also comparable with $T_{\rm min,s}$ and $T_{\rm min,v}$. As expected, $T_{\rm max}$ is significantly larger than others. For $T_{\rm d,1}$, $T_{\rm d,2}$, Td, 3, $T_{\rm min,s}$, and $T_{\rm min,v}$, $T_{\rm min,v}$ is slightly larger than others and the value of $T_{\rm min,s}$ is in the middle. As a compromise, we use the value of $T_{\rm min,s}$ to represent the external tides in the subsequent analysis.

4. Discussion

4.1. Tide-regulated gravitational collapse

The average strength of the external tides $T_{\rm min,s}$ of a structure was used to measure the deformation of the structure due to the external gravity. The effective velocity dispersion from the external tides of the structure is $\sigma_t^2 \approx T_{\rm min,s} * R^2$, and the corresponding energy

$$E_{\rm t} \approx \frac{1}{2} M * T_{\rm min,s} * R^2. \tag{9}$$

The gravitational potential energy $E_{\rm g}$ and internal kinetic energy $E_{\rm k}$ were calculated in the same way as presented in Zhou et al. (2024b).

The slowing down of a pure free-fall gravitational collapse as revealed by the shallow $\sigma_{\rm s} - N * R$ relation means that there are some physical processes working against gravitational collapse, such as tidal tearing from external gravity. The effect of the external tides on the gravitational collapse of gas structures can also be reflected in the scaling relation. After adding the contribution of the tidal interactions to the observed velocity dispersion, the effective velocity dispersion $\sigma_{\rm s} = \sqrt{\sigma_{\rm t}^2 + \sigma^2}$, then the $\sigma_{\rm s} - N * R$ relation was fitted again. As shown in Fig. 6d, now the slope is close to 0.5, and the correlation also gets stronger. Thus, the

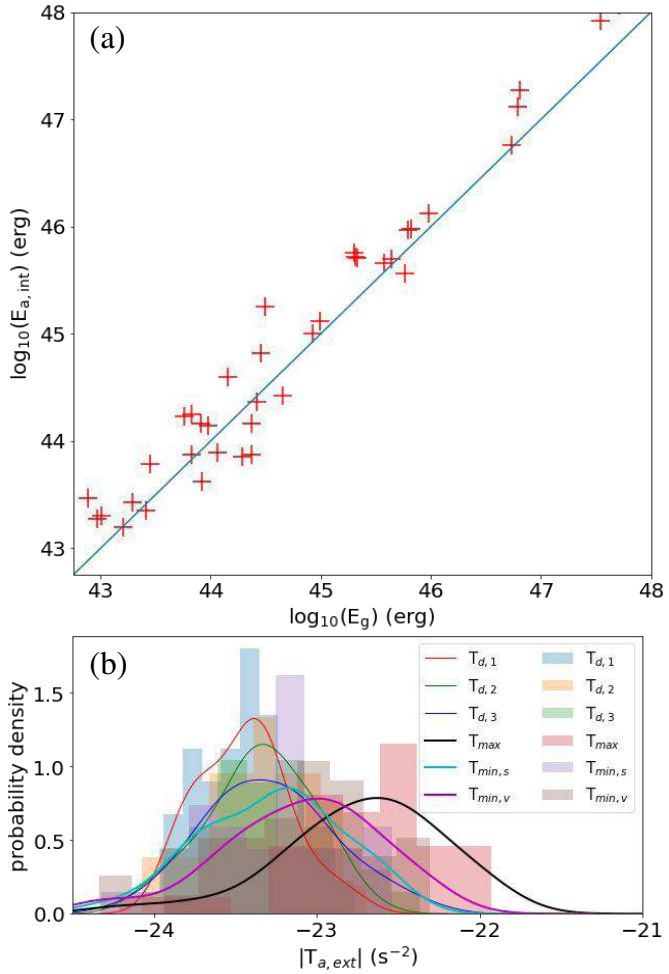


Fig. 8. (a) Comparison of $E_{\rm a,int}$ and $E_{\rm g}$ of dendrogram structures. (b) Distribution of the average tidal strength in each computation.

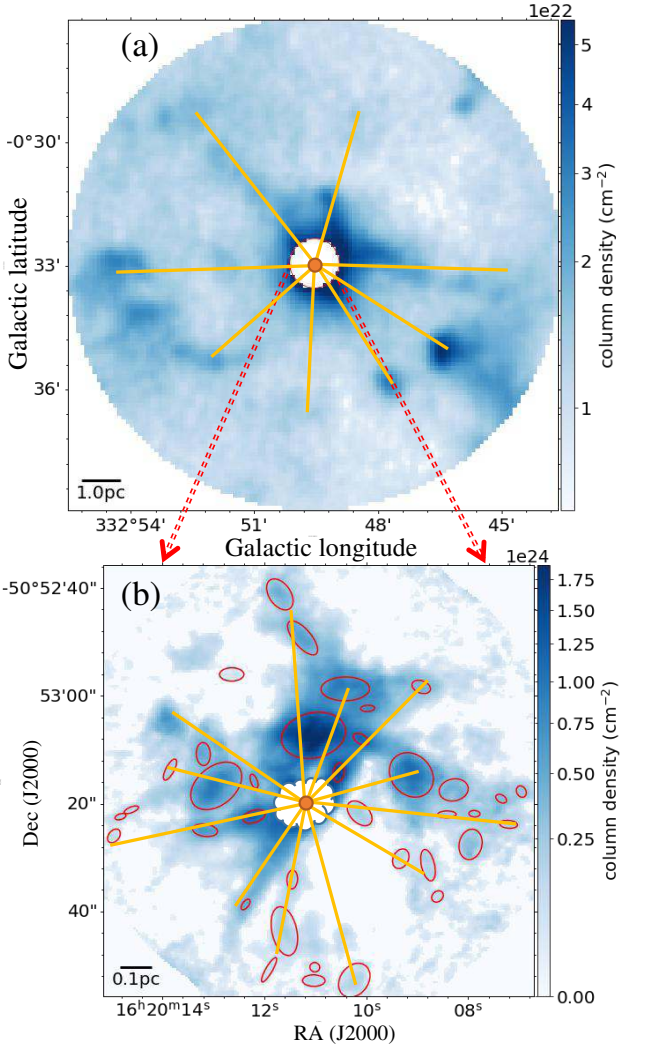


Fig. 9. Computed structure first masked from the density map. The external tides from all outside material sustained by a point in the computed structure was calculated pixel by pixel and is illustrated by orange lines. (a) Background representing the column density map derived from ATLASGAL+*Planck* 870 μm data. The material in the FOV of the ALMA observation (white region) was masked. (b) Background representing the column density map derived from $H^{13}CO^+$ J = 1–0 and HCO^+ J = 1–0 lines from the ALMA observation by the LTE analysis.

deformation due to the external tides may effectively slow down the pure free-fall gravitational collapse of gas structures.

However, as shown in Fig. 11a, the sum of $E_{\rm k}$ and $E_{\rm t}$ is notably larger than $E_{\rm g}$. The external tidal strength may be overestimated. The velocity dispersion is a manifestation of gas perturbations due to many physical processes, including gravitation collapse, tidal tearing, and turbulence. Thus, $E_{\rm k}$ should be overestimated. Moreover, when $E_{\rm k} + E_{\rm t} > E_{\rm g}$, it does not mean that the structure cannot be gravitationally bound. A structure is deformed by external tides from the surrounding material, and at the same time, the surrounding material also provides the structure with a confining pressure.

Previous studies have suggested that the external pressure provided by the larger-scale molecular gas might help to confine dense structures in molecular clouds (Spitzer 1978; McKee 1989; Elmegreen 1989; Ballesteros-Paredes 2006; Kirk et al. 2006, 2017; Dib et al. 2007; Lada et al. 2008; Pattle et al. 2015; Li et al. 2020; Zhou et al. 2024b). External pressure can have various origins; here we mainly consider the external pressure from the ambient cloud for each decomposed structure using

$$P_{\rm cl} = \pi G \bar{\Sigma} \Sigma_r, \tag{10}$$

where $P_{\rm cl}$ is the gas pressure, $\bar{\Sigma}$ is the mean surface density of the cloud, and $\Sigma_{\rm r}$ is the surface density at the location of each structure (McKee 1989; Kirk et al. 2017). The external pressure energy can be calculated as

$$E_{\rm p} = -4\pi P_{\rm cl} R^3. \tag{11}$$

As shown in Fig. 11b, after considering the external pressure, $E_{\rm k} + E_{\rm t} < E_{\rm g} + E_{\rm p}$. The external tide tries to tear up the structure, but the external pressure on the structure prevents this process from occurring. The counterbalance between the external tide and external pressure hinders the gravitational collapse of the structure, which can also cause a pure free-fall gravitational collapse to slow down.

Gravity can play multiple roles in cloud evolution. Inside dense regions, it drives fragmentation and collapse, and outside these regions, it can suppress the low-density gas from collapsing through extensive tidal forces and drive accretion (Li 2024). This work confirms that tidal interactions can significantly influence the dynamics and kinematics of gas structures and regulate their gravitational collapse. In Figs. 14c and 15c of

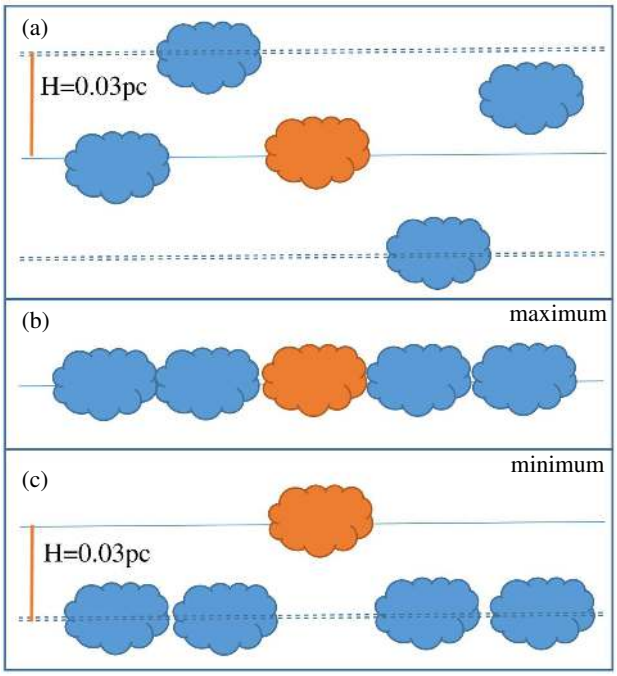


Fig. 10. Edge-on view of G332.83-0.55. The calculated gas structure is in red, and all external material apart from the calculated structure is in blue. (a) The gas distribution in G332.83-0.55 can be everywhere. (b) In 2D computation, all molecular gases are assumed to be concentrated on the same plane. Thus, we obtained a maximal estimate for the tidal strength. (c) Taking a characteristic scale height of the molecular gas as 0.03 pc, all external material apart from the calculated structure were moved vertically to 0.03 pc away. Thus we obtained a minimal estimate of the tidal strength.

Zhou et al. (2024b), in the G333 complex, for 3608 gas structures on cloud-clump scales, the slopes of the $\sigma - N * R$ relations are ~0.19 and ~0.26, respectively. In Fig. 5a of Zhou et al. (2024a), for 2420 dense gas structures on clump-core scales in 64 ATOMS sources (clumps) at different evolutionary stages, the slope of the $\sigma - N * R$ relation is ~0.27. According to the above analysis, these shallow slopes (significantly smaller than 0.5) may indicate extensive tide-regulated gravitational collapse from cloud to core scales, which will be discussed in detail in future work.

4.2. Gravitational focusing

In Figs. 4b and 5b, on large scales, the funnel structure is consistent with the schematic diagram in Fig. 9 of Zhou et al. (2023). However, it looks as though the bottom of the funnel splits into two parts in Figs. 4b and 5b. A velocity jump appears in the center comprised of the redshifted and blueshifted components, which may be attributed to the infalling and rotating motions around the protocluster. If we turn over the redshifted component and let it merge with the blueshifted component, then we obtain the funnel structure in Fig. 9 of Zhou et al. (2023).

Similar to lights converging to the focus, now the gas inflows are converging to the gravitational center, that is, the hub region of G332.83-0.55. Thus, there is a vivid "gravitational focusing" process, which also provides a clue as to the formation of the hub-filament structure. As shown in Fig. 1 of Zhou et al. (2023), the entire G333 complex has a large-scale shell structure. G332.83-0.55 is a fragment on the shell, which should have a sheet configuration initially; for more details, readers can also refer to Sect. 3.6. Then the subsequent gravitational

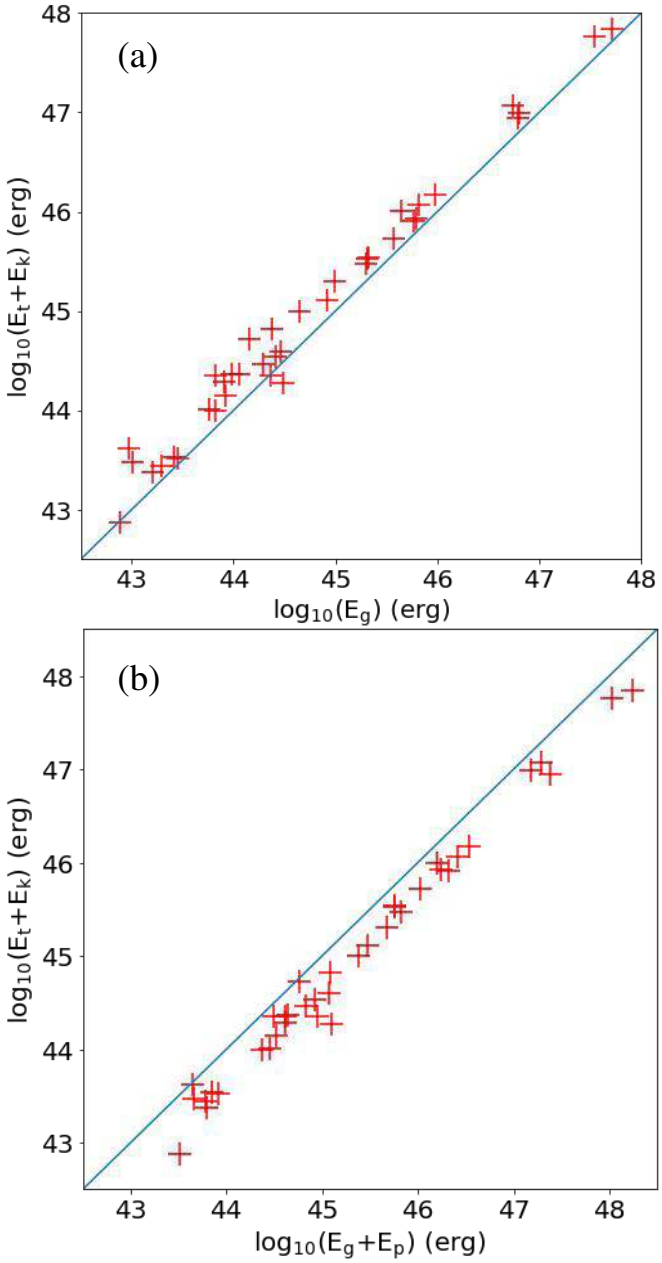


Fig. 11. Comparison between the structural energies, i.e., the gravitational potential energy E_g, the kinetic energy E_k, the tidal energy E_t, and the external pressure energy E_p.

focusing process yields the hub-filament structure. In Fig. 1b, the hub-filament structure is decoupling from the surrounding gas environment due to gravitational collapse.

5. Summary

1. In high-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observations, G332.83-0.55 has a clear hub-filament morphology at ~10 pc, and its hub corresponds to the ATLASGAL clump AGAL332.826-00.549. We zoomed in on the hub region of G332.83-0.55 using higher-resolution ALMA HCO^+ (1–0) and $H^{13}CO^+$ (1–0) observations, and found that it still presents a hub-filament structure. Thus, G332.83-0.55 contains at least two levels of hub-filament structures. High-resolution single-dish and ALMA observations provide the opportunity to trace gas inflows from cloud to core scales.

2. G332.83-0.5 is collapsing with rotation. Around the gravitational center (i.e., the protocluster), both the velocity fields traced by LAsMA and ALMA observations show velocity jumps comprised of redshifted and blueshifted peaks. According to the velocity difference and the separation between blueshifted and redshifted peaks, we roughly estimated the rotation velocity $V_{\rm rotation,ALMA} \approx 3.29$ km s^{-1} and the centrifugal barrier $r_{\rm c,ALMA} \approx 0.21$ pc. Then the mass of the gravitational center (i.e., the protocluster) was estimated to be $\approx$264 $M_{\odot}$, which agrees with the mass estimated by 3 mm dust emission.

3. We directly identified local dense gas structures based on the 2D integrated intensity (Moment 0) map of $H^{13}CO^+$ $J = 1–0$ emission using the dendrogram algorithm, and then extracted the average spectra to fit the velocity dispersion. The slope of the $\sigma - N * R$ relation for leaf structures is ~0.2, which is much smaller than 0.5 of the pure free-fall gravitational collapse. This may indicate that a pure free-fall gravitational collapse is slowing down.

4. We investigated the effect of shear and tides from the protocluster on the surrounding local dense gas structures. The results seem to deny the importance of shear and tides from the protocluster. However, for a gas structure, it also bears the tides from other gas structures and the general gas distribution, not only the protocluster. To fully consider the tidal interactions, we derived the tide field according to the density distribution, and then used the average strength of the external tides of a structure to measure the total tidal interactions on the structure. For comparison, we also adopted an original pixel-by-pixel computation to estimate the average tidal strength of each structure. Both methods give comparable calculation results.

5. After considering the tidal interactions for all structures, the slope of the $\sigma - N * R$ relation is close to 0.5, and the correlation also gets stronger. Thus, the deformation due to the external tides can effectively slow down the pure free-fall gravitational collapse of gas structures. The external tide tries to tear up the structure, but the external pressure on the structure prevents this process from occurring. The counterbalance between an external tide and external pressure hinders the free-fall gravitational collapse of the structure, which can also cause a pure free-fall gravitational collapse to slow down. These mechanisms can be called "tide-regulated gravitational collapse".

Acknowledgements. This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under programme ID M-0109.F-9514A-2022. APEX has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. This research made use of astrodendro, a Python package to compute dendrograms of Astronomical data (`http://www.dendrograms.org/`). This paper makes use of the following ALMA data: ADS/JAO.ALMA♯2019.1.00685.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO. T.L. acknowledges the supports by the National Key R&D Program of China (No. 2022YFA1603101), National Natural Science Foundation of China (NSFC) through grants No.12073061 and No.12122307, the international partnership program of Chinese Academy of Sciences through grant No.114231KYSB20200009, and the Tianchi Talent Program of Xinjiang Uygur Autonomous Region.

References

Ballesteros-Paredes, J. 2006, MNRAS, 372, 443
Chen, X., Sobolev, A. M., Ren, Z.-Y., et al. 2020, Nat. Astron., 4, 1170
Contreras, Y., Garay, G., Rathborne, J. M., & Sanhueza, P. 2016, MNRAS, 456, 2041
Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, A&A, 585, A104
Dewangan, L. K., Ojha, D. K., Sharma, S., et al. 2020, ApJ, 903, 13
Dib, S. 2023, MNRAS, 524, 1625
Dib, S., Kim, J., Vázquez-Semadeni, E., Burkert, A., & Shadmehri, M. 2007, ApJ, 661, 262
Dib, S., Helou, G., Moore, T. J. T., Urquhart, J. S., & Dariush, A. 2012, ApJ, 758, 125
Elmegreen, B. G. 1989, ApJ, 338, 178
Ganguly, S., Walch, S., Clarke, S. D., & Seifried, D. 2024, MNRAS, 528, 3630
Gong, H., & Ostriker, E. C. 2011, ApJ, 729, 120
He, Z.-Z., Li, G.-X., & Burkert, A. 2023, MNRAS, 526, L20
Henshaw, J. D., Caselli, P., Fontani, F., Jiménez-Serra, I., & Tan, J. C. 2014, MNRAS, 440, 2860
Hunter, D. A., Elmegreen, B. G., & Baker, A. L. 1998, ApJ, 493, 595
Issac, N., Tej, A., Liu, T., et al. 2019, MNRAS, 485, 1775
Izquierdo, A. F., Galván-Madrid, R., Maud, L. T., et al. 2018, MNRAS, 478, 2505
Kauffmann, J., Bertoldi, F., Bourke, T. L., Evans, N. J., I., & Lee, C. W. 2008, A&A, 487, 993
Keto, E., Zhang, Q., & Kurtz, S. 2008, ApJ, 672, 423
Kirk, H., Johnstone, D., & Di Francesco, J. 2006, ApJ, 646, 1009
Kirk, H., Friesen, R. K., Pineda, J. E., et al. 2017, ApJ, 846, 144
Koch, E. W., & Rosolowsky, E. W. 2015, MNRAS, 452, 3435
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87
Lada, C. J., Muench, A. A., Rathborne, J., Alves, J. F., & Lombardi, M. 2008, ApJ, 672, 410
Li, G.-X. 2024, MNRAS, 528, L52
Li, S., Zhang, Q., Liu, H. B., et al. 2020, ApJ, 896, 110
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Liu, H. B., Galván-Madrid, R., Jiménez-Serra, I., et al. 2015, ApJ, 804, 37
Liu, T., Zhang, Q., Kim, K.-T., et al. 2016, ApJ, 824, 31
Liu, S.-Y., Su, Y.-N., Zinchenko, I., et al. 2020a, ApJ, 904, 181
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020b, MNRAS, 496, 2790
Liu, H.-L., Liu, T., Evans, Neal J., I., et al. 2021a, MNRAS, 505, 2801
Liu, X.-L., Xu, J.-L., Wang, J.-J., et al. 2021b, A&A, 646, A137
Liu, H.-L., Tej, A., Liu, T., et al. 2022, MNRAS, 511, 4480
Liu, H.-L., Tej, A., Liu, T., et al. 2023, MNRAS, 522, 3719
Lu, X., Zhang, Q., Liu, H. B., et al. 2018, ApJ, 855, 9
Maud, L. T., Hoare, M. G., Galván-Madrid, R., et al. 2017, MNRAS, 467, L120
McKee, C. F. 1989, ApJ, 345, 782
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, ASP Conf. Ser., 376, 127
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Myers, P. C. 2009, ApJ, 700, 1609
Ohashi, N., Hayashi, M., Ho, P. T. P., & Momose, M. 1997, ApJ, 475, 211
Olguin, F. A., Sanhueza, P., Chen, H.-R. V., et al. 2023, ApJ, 959, L31
Pattle, K., Ward-Thompson, D., Kirk, J. M., et al. 2015, MNRAS, 450, 1094
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Sakai, N., Sakai, T., Hirota, T., et al. 2014, Nature, 507, 78
Sanhueza, P., Girart, J. M., Padovani, M., et al. 2021, ApJ, 915, L10
Sanna, A., Kölligan, A., Moscadelli, L., et al. 2019, A&A, 623, A77
Schneider, N., Csengeri, T., Hennemann, M., et al. 2012, A&A, 540, L11
Shimoikura, T., Dobashi, K., Matsumoto, T., & Nakamura, F. 2016, ApJ, 832, 205
Shimoikura, T., Dobashi, K., Nakamura, F., Matsumoto, T., & Hirota, T. 2018, ApJ, 855, 45
Shimoikura, T., Dobashi, K., Hirano, N., et al. 2022, ApJ, 928, 76
Spitzer, L. 1978, Physical Processes in the Interstellar Medium (Hoboken: Wiley)
Stark, A. A., & Blitz, L. 1978, ApJ, 225, L15
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Xu, F.-W., Wang, K., Liu, T., et al. 2023, MNRAS, 520, 3259
Yang, D., Liu, H.-L., Tej, A., et al. 2023, ApJ, 953, 40
Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, ApJ, 852, 12
Zhang, Q., Wang, K., Lu, X., & Jiménez-Serra, I. 2015, ApJ, 804, 141
Zhang, C., Zhu, F.-Y., Liu, T., et al. 2023, MNRAS, 520, 3245
Zhou, J.-W., Liu, T., Li, J.-Z., et al. 2021, MNRAS, 508, 4639
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
Zhou, J. W., Dib, S., Wyrowski, F., et al. 2024a, A&A, 682, A173
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2024b, A&A, 682, A128

Astronomy & Astrophysics manuscript no. output
February 8, 2025
©ESO 2025

High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation of the G333 giant molecular cloud complex : III. Decomposition of molecular clouds into multi-scale hub-filament structures

J. W. Zhou[1], J. S. Urquhart[2], F. Wyrowski[1], T. Liu[3], and M. Juvela[4]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de
[2] Centre for Astrophysics and Planetary Science, University of Kent, Canterbury, CT2 7NH, UK

[3] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, Peoples Republic of China

[4] Department of Physics, P.O.Box 64, FI-00014, University of Helsinki, Finland

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We decomposed the G333 complex and the G331 GMC into multi-scale hub-filament systems (HFs) using the high-resolution ^{13}CO (3−2) data from LAsMA observations. We employed the filfinder algorithm to identify and characterize filaments within HFs. Compared with non-HFs, HFs have significantly higher density contrast, larger masses and lower virial ratios. Velocity gradient measurements around intensity peaks provide evidence of gas inflow within these structures. There may be an evolutionary sequence from non-HFs to HFs. Currently, non-HFs lack a distinct gravitational focusing process that would result in significant density contrast. The density contrast can effectively measure the extent of gravitational collapse and the strength of the gravitational center of the structure that definitively shape the hub-filament morphology. Combined with the kinematic evidence in our previous studies, we suggest that molecular clouds are network structures formed by the gravitational coupling of multi-scale hub-filament structures. The knots in the networks are the hubs, they are the local gravitational centers and the main star-forming sites. Actually, clumps in molecular clouds are equivalent to the hubs. The network structure of molecular clouds can naturally explain that feedback from protoclusters does not significantly change the kinematic properties of the surrounding embedded dense gas structures, as concluded in our previous studies.

Key words. Submillimeter: ISM – ISM:structure – ISM: evolution – Stars: formation – methods: analytical – techniques: image processing

1. Introduction

Assessing the dynamic interaction between density fluctuations in molecular clouds and the surrounding gas motions offers insight into the development of hierarchical structures within high-mass star-forming regions (McKee & Ostriker 2007; Motte et al. 2018; Henshaw et al. 2020; Hacar et al. 2023). High-resolution observations of high-mass star-forming regions show that density enhancements are organized in filamentary gas networks, particularly in hub-filament systems. Hub-filament systems are known as a junction of three or more filaments. Filaments have lower column densities compared to the hubs (Myers 2009; Schneider et al. 2012; Kumar et al. 2020; Zhou et al. 2022). In these systems, converging flows channel matter into the hub via filaments. Earlier research has proposed that hub-filament structures serve as the nurseries for the formation of high-mass stars and clusters (Peretto et al. 2013; Henshaw et al. 2014; Zhang et al. 2015; Liu et al. 2016a; Yuan et al. 2018; Lu et al. 2018; Issac et al. 2019; Dewangan et al. 2020; Liu et al. 2021, 2022; Kumar et al. 2020; Zhou et al. 2022; Liu et al. 2023; Xu et al. 2023; Zhou et al. 2023; Yang et al. 2023; Zhou et al. 2024b). Statistically, Kumar et al. (2020) identified the candidate hub-filament systems in the Milky Way inner disc by searching for filamentary structures around 35 000 clumps in the Hi-GAL catalogue (Elia et al. 2017). ∼11% of the clumps are the candidates of hub-filament systems. The column densities of hubs are found to be enhanced from a factor of approximately two (pre-stellar sources) up to about ten (proto-stellar sources). High-mass stars preferentially form in the density-enhanced hubs of hub-filament systems. This amplification can drive the observed longitudinal flows along filaments providing further mass accretion.

Specifically, Zhou et al. (2022) investigated the physical characteristics and evolution of hub-filament systems in approximately 140 protoclusters using spectral line data from the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Liu et al. 2020). They suggested that hub-filament structures, characterized by self-similarity and filamentary accretion, remain consistent across various scales in high-mass star-forming regions, ranging from several thousand astronomical units to several parsecs. This model of hierarchical, multi-scale hub-filament structures was extended from the clump-core scale to the cloud-clump scale by Zhou et al. (2023), and further developed to the galaxy-cloud scale in Zhou & Davis (2024); Zhou et al. (2024a). Earlier studies have also illustrated hierarchical collapse and hub-filament structures feeding central

regions, as described by Motte et al. (2018), Vázquez-Semadeni et al. (2019) and Kumar et al. (2020), and references therein.

Kinematically, the results in Zhou et al. (2023) have revealed the multi-scale hub-filament structures in the G333 complex. Numerous structures within the G333 complex exhibit the characteristic kinematic properties of hub-filament systems. For instance, the intensity peaks, which represent the hubs, are linked to converging velocities, indicating that the surrounding gas is flowing towards and accumulating in these dense regions. In particular, the velocity gradients increase at small scales. The filaments in both LAsMA and ALMA observations show clear velocity gradients. The velocity gradients we fit to the LAsMA and ALMA data agree with each other over the range of scales covered by ALMA observations in the ATOMS survey (< 5 pc). In Zhou et al. (2023), larger scale gas motions were investigated, the longest filament being ~50 pc, but we obtained similar results to small-scale ALMA observations. The variations in velocity gradients at small scales (< 1 pc), medium scales (~ 1-7.5 pc), and large scales (> 7.5 pc) are consistent with gravitational free-fall with central masses of ~ 500 $M_\odot$, ~ 5000 $M_\odot$ and ~ 50000 $M_\odot$. This means that the velocity gradients on larger scales require a higher mass to be maintained. Higher masses correspond to larger scales, indicating that large-scale inflow is driven by larger-scale structures. This may result from the gravitational clustering of smaller-scale structures, aligning with the hierarchical nature of molecular clouds and the gas inflow from larger to smaller scales. The large-scale gas inflow is likely driven by gravity, suggesting that the molecular clouds in G333 are undergoing global gravitational collapse. Additionally, the funnel-like velocity field observed in position-position-velocity (PPV) space further supports the concept of large-scale gravitational collapse within the molecular clouds of the G333 complex (Zhou et al. 2023). The change in velocity gradients with the scale in the G333 complex indicates that the morphology of the velocity field in PPV space resembles a "funnel" structure (Zhou et al. 2023). The funnel structure can be explained as accelerated material flowing towards the central hub and gravitational contraction of star-forming clouds/clumps. Large-scale velocity gradients always involve many intensity peaks, and the larger-scale inflow is driven by the larger-scale structure, implying that the clustering of local small-scale gravitational structures can act as the gravitational center on a larger scale. To some extent, the funnel structure gives an indication of the gravitational potential well that is formed by the clustering (Zhou et al. 2023). The “funnel” behaviour in terms of gravitational collapse has also been described in previous works, such as Kuznetsova et al. (2015) and Hacar et al. (2016, 2017).

Based on the evidence from the kinematics, the main goal of this work is to directly recover the multi-scale hub-filament structures in the G333 complex and the G331 giant molecular cloud (GMC) using the high-resolution ^{13}CO (3−2) data in LAsMA observation.

2. Data

The observations and data reduction are described in detail in Zhou et al. (2023). We mapped a 3.4° × 1.2° area centered at (l, b) = (332.33°, −0.29°) using the APEX telescope (Güsten et al. 2006), which includes the G333 complex and the G331 GMC. [1] The 7 pixel Large APEX sub-Millimeter Array (LAsMA) receiver was used to observe the J = 3 − 2 transitions of ^{12}CO ($\nu_{\rm rest}$ ~ 345.796 GHz) and ^{13}CO ($\nu_{\rm rest}$ ~ 330.588 GHz) simultaneously. Observations were performed in a position switching on-the-fly (OTF) mode. The data were reduced using the GILDAS package[2]. The final data cubes have a velocity resolution of 0.25 km s^{-1} and an angular resolution of 19.5″. A main-beam efficiency value $\eta_{\rm mb}$ = 0.71 (Mazumdar et al. 2021) was used to convert intensities from the T_A^* scale to main beam brightness temperatures, $T_{\rm mb}$.

The observed region was covered in the IR range by the Galactic Legacy Infrared Midplane Survey (GLIMPSE, Benjamin et al. 2003). GLIMPSE images, obtained with the Spitzer Infrared Array Camera (IRAC) at 8.0 μm, were retrieved from the Spitzer archive. The angular resolution of the images in the IRAC bands is ~ 2″.

3. Results

3.1. Dendrogram analysis

The most likely hub candidates are the intensity peaks. Following the procedure described in Zhou et al. (2024d), we directly identify the intensity peaks based on the integrated intensity (Moment 0) map of ^{13}CO (3−2) emission. Using the *astrodendro* package [3], there are three major input parameters for the dendrogram algorithm: *min_value* for the minimum value to be considered in the dataset, *min_delta* for a leaf that can be considered as an independent entity, and *min_npix* for the minimum area of a structure. For the Moment 0 map, a 5σ threshold has been set. In Zhou et al. (2024d), we only require the minimum structural area to exceed the size of a single beam and avoid setting additional parameters, thus minimizing the algorithm’s reliance on specific parameter settings for structure identification. In this work, we focus on the hubs and are mainly interested in the strongest local intensity peaks. In addition to setting *min_npix* as one beam, a threshold 15 K km s^{-1} was also adopted for the parameter *min_value* to filter the relatively low-density structures. We tried different values of *min_value*, and found the value of 15 K km s^{-1} can best recover all high intensity peaks, as shown in Fig.1. The dendrogram algorithm decomposes the intensity data (Moment 0 map) into hierarchical structures, called leaves and branches. Finally, a total of 612 structures are identified, including 438 leaves and 174 branches. In this work, we only focus on leaf structures (leaves). Then, we applied the fully automated Gaussian decomposer algorithm, `GAUSSPY+` (Lindner et al. 2015; Riener et al. 2019), to fit the averaged spectra of leaf structures individually. The parameter settings for the decomposition were identical to those used by Zhou et al. (2023).

astrodendro approximates the morphology of each structure as an ellipse. In the dendrogram, the long and short axes of an ellipse, a and b, represent the root mean square (rms) sizes (second moments) of the intensity distribution along the two spatial dimensions. As outlined in Zhou et al. (2024d), these rms sizes result in a smaller ellipse compared to the actual size of the identified structure. Therefore, it is necessary to multiply these dimensions by a factor of 2 to properly scale the ellipse to match the structure’s size. The effective physical radius of an ellipse is then $R_{\rm eff} = \sqrt{2a \times 2b} \times d$, with d = 3.6 kpc for the distance to the G333 complex (Lockman 1979; Bains et al. 2006).

In Fig.2 of Zhou et al. (2023), the average spectrum of ^{13}CO (3−2) for the entire region (the G333 complex and the G331

[1] This publication is based on data acquired with the Atacama Pathfinder EXperiment (APEX) while it has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory and the Onsala Space Observatory.

[2] http://www.iram.fr/IRAMFR/GILDAS

[3] https://dendrograms.readthedocs.io/en/stable/index.html

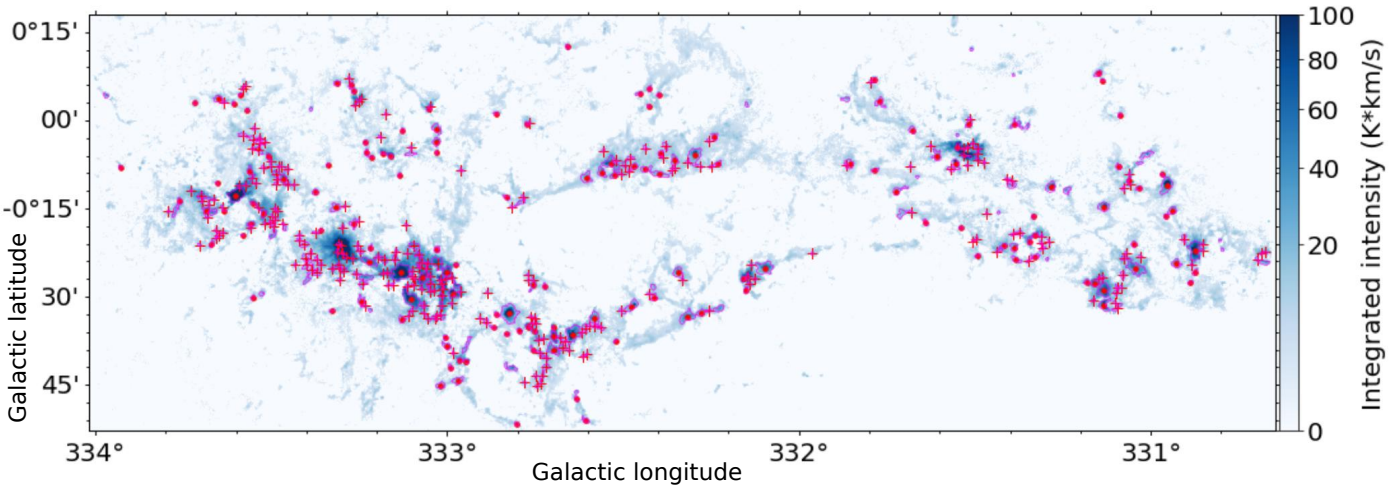


Fig. 1. The background is the integrated intensity map of ^{13}CO (3−2) in the full velocity range [−120, −20] km s^{-1}, that is covered by the masks of leaf structures (magenta contours) identified by the dendrogram algorithm. Red pluses and dots mark the positions of non-HFs and HFs classified in Sec.3.3.

GMC) shows that the emission is concentrated in the velocity range [−120, −20] km s^{-1}, and there are three main velocity components peak1, peak2 and peak3 defined in Zhou et al. (2023). In Zhou et al. (2023), we only considered the peak3, which is regarded as the G333 molecular cloud complex. However, in Fig.4(b) of Zhou et al. (2023), a large-scale velocity gradient that links peak3 with peak2, and thus peak3 and peak2 make up a loop structure. In Nguyen et al. (2015), both peak2 and peak3 were assigned to the same molecular cloud complex in the Scutum-Centaurus arm with the distance ∼ 3.6 kpc, while peak1 is thought to be located in the Norma arm at a distance ∼ 5 kpc. However, the associations between peak1, peak2 and peak3 are still unclear. Four ATLASGAL clump clusters marked in Fig.3 of Zhou et al. (2023) were distributed to similar distances ∼ 3.6 kpc in Urquhart et al. (2022), although one of them belongs to peak1. In Fig.2 and Fig.3 of Zhou et al. (2023), peak1 and peak2 were called G331 GMC-blue and -red, because they seem to be two symmetrical halves of the G331 GMC. However, as shown in Fig.2 of Zhou et al. (2023), their velocity difference is ∼ 25 km s^{-1}, thus they are put in different spiral arms in Nguyen et al. (2015). In this work, we consider the entire velocity range [−120, −20] km s^{-1}, i.e. both the G333 complex and the G331 GMC, and take the distance d = 3.6 kpc of the G333 complex (Lockman 1979; Bains et al. 2006) for all identified structures. Considering the emission of peak2 and peak3 is dominant in the entire observed region, therefore, the potential distance bias of peak1 will not significantly affect the statistical results. Moreover, there is no big difference whether the distance is ∼ 5 kpc or ∼ 3.6 kpc for peak1.

3.2. Velocity components

For the hub-filament systems, as presented in the literature listed in Sec.1, hubs are essentially structures with higher density compared to the surrounding diffuse gas. As local gravitational centers, these hubs can accrete the surrounding diffuse gas, forming hub-filament structures. The filamentary structures are actually gas inflows converging towards the hubs. A hub-filament structure essentially consists of a gravitational center and gas inflows converging towards it. Therefore, all identified intensity peaks are thought to be potential hubs. Around the intensity peaks, we extended the spatial ranges to investigate the filamentary structures connected with them. As shown in the first row of Fig.2, extending to 2.5 times the hub size (the effective radius of the corresponding leaf structure) can well recover a relatively complete hub-filament morphology. However, the hub-filament structures are identified directly according to the integrated intensity map of the full velocity range. Therefore, the issue of the overlapping of different velocity components is inevitable, as shown in Fig.3 of Zhou et al. (2023). To address this issue, we extracted the average spectra for all structures in the 2.5 times extended regions to decompose their velocity components.

The velocity components within the same structure that exhibit large velocity differences are generally independent of each other, which can be attributed to projection effects. Thus, for each structure, we mainly focus on the main emission marked by the red dashed box in Fig.3. As shown in Fig.3(a) and (c), the main emission can be one velocity component or multiple blended velocity components. Using the names in Zhou et al. (2024d), they are called type1 and type3 structures. For type2 structures with separated velocity components (such as Fig.3(b)), we only consider the dominant component, thus they are reduced to type1 structures.

The velocity range of a structure is essential for accurately estimating its fundamental physical properties, including velocity dispersion and mass. As discussed in Zhou et al. (2024d), we take the velocity range of each velocity component as [v_c-FWHM,v_c+FWHM]. For a type3 structure, the velocity range for the leftmost and rightmost velocity components are [v_{c1}-FWHM$_1$,v_{c1}+FWHM$_1$] and [v_{c2}-FWHM$_2$,v_{c2}+FWHM$_2$], respectively. Then the total velocity range of the type3 structure is [v_{c1}-FWHM$_1$,v_{c2}+FWHM$_2$].

In the first row of Fig.2, the hub-filament structures are connected to the gas environment by the filamentary structures. They are not independent, as the filamentary structures lack a clear boundary with the surrounding gas. Therefore, accurately estimating the physical properties of the entire hub-filament structures, such as their scale and density, is challenging. But the hubs are relatively isolated, and as local structures, it is feasible to calculate their physical parameters. However, the velocity range of the hub may be different from that of the entire hub-filament structure. Thus, we repeat the procedures described

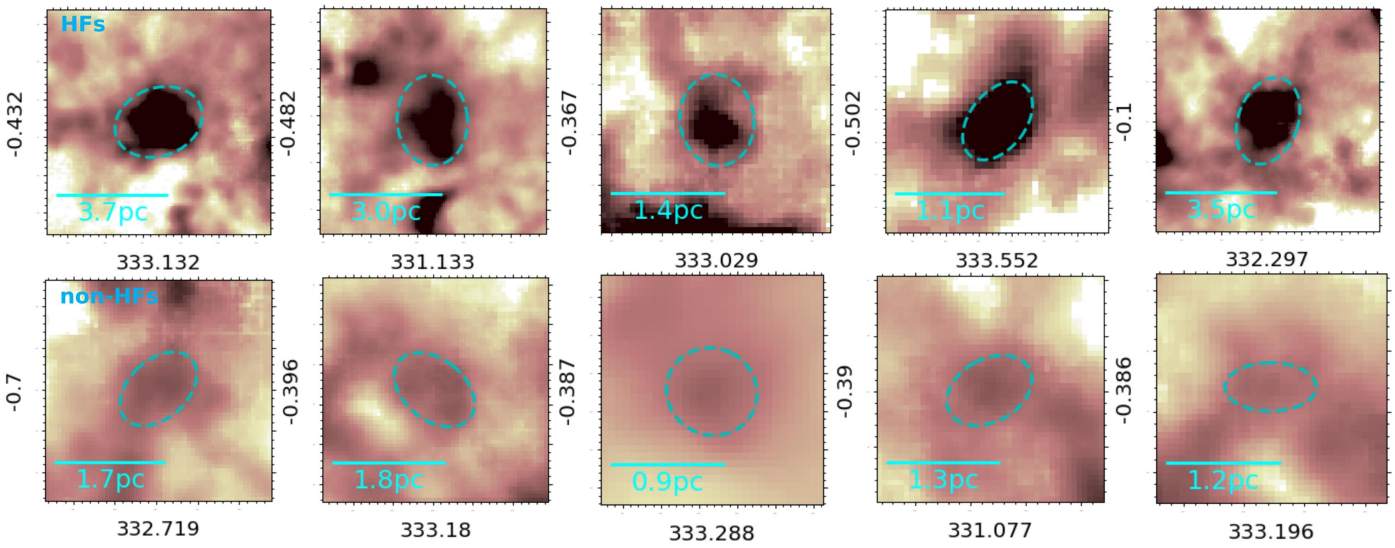


Fig. 2. Some examples of HFs and non-HFs classified in Sec.3.3. The dashed ellipses represent the equivalent ellipses of the identified leaf structures.

above to estimate the velocity ranges and velocity dispersion of the hubs. Following the method outlined in Zhou et al. (2024d), we extract the physical quantities for each hub by utilizing its velocity range and effective ellipse based on the column density (N) and temperature cubes derived from ^{13}CO (3-2) and ^{12}CO (3-2) lines by the local thermodynamic equilibrium (LTE) analysis in the full velocity range [−120, −20] km s^{-1} (Zhou et al. 2023). Then we calculate the mass and virial ratio of each structure, as done in Zhou et al. (2024d).

In this work, we consider type3 structures to be independent entities, similar to type1 structures, but characterized by more complex gas motions. If the different velocity components within a type3 structure are parts of the same structure, the velocity dispersion of the type3 structure can be determined by calculating an intensity-weighted average of the velocity dispersion of all its components.

Evidence supporting the hypothesis that a type3 structure is a single entity, rather than a result of overlapping uncorrelated gas components, includes the following points: (1) In Fig.3, the significant overlap in the line profiles of type3 structures may indicate a correlation between different velocity components within the structure. The greater the overlap between these components, the more likely they are part of the same structure. In extreme cases, this overlap results in a single-peak line profile (type1 structure); (2) This study focuses on hub-filament structures, where gas inflows along the filaments exhibit varying velocities, leading to multiple velocity components within the same structure; (3) In Sec.3.4, for type3 structures, the velocity variation along the filament is continuous, without abrupt changes.

3.3. Classification

Restricted to the fixed velocity ranges of all structures (2.5 times extended) in Sec.3.2, their Moment-0 maps were reproduced. In hub-filament systems (HFs), hubs exhibit substantially higher densities compared to filaments. Based on this, we categorized the structures into two groups: those with a distinct high-density central region and those without. The two categories comprise 195 and 243 structures (non-HFs), respectively. For the 195 structures featuring a distinct high-density center, we further identified the associated filaments as described in Sec.3.4. Structures in which filaments converge at the high-density center were classified as hub-filament systems, totaling 148 structures (HFs). A small number of remaining structures, which were challenging to classify, were excluded to maintain the clarity of comparison in the subsequent analysis. Fig.2 illustrates examples of non-HFs and HFs. In Fig.4, we present a comparison of the physical properties of HFs and non-HFs, revealing significant differences between the two groups.

3.4. Filamentary structures

3.4.1. Identification

As done in Zhou et al. (2023), we also use the filfinder algorithm (Koch & Rosolowsky 2015) to characterize the filamentary structures around the hub. In Zhou et al. (2023), our focus was on large-scale filaments within molecular clouds. Along these filaments, we observed clear fluctuations in both velocity and intensity, with each intensity peak indicating a local hub. In this study, we began by identifying local intensity peaks (i.e. leaf structures), and subsequently searched for hub-filament structures centered around these intensity peaks. Our current objective is to identify and analyze small-scale filamentary structures within the local hub-filament systems. However, due to limitations in observational resolution, the number of pixels is insufficient to directly identify extended filamentary structures using the filfinder algorithm. To address this, we regridded the images, increasing the pixel count from $N_x * N_y$ to $2N_x * 2N_y$ [4]. To minimize the introduction of artificial structures, we only doubled the number of pixels. As shown in Fig. 2 and Fig. 5, the morphology of the structures remains unchanged. The identified filamentary structures also align well with the background.

3.4.2. Velocity gradient

A kinematic feature of hub-filament structures is that the gas flow along the filament converges towards the central hub, result-

[4] This was achieved using the **zoom** function from the **scipy.ndimage** module, with an interpolation order of 3 for regridding.

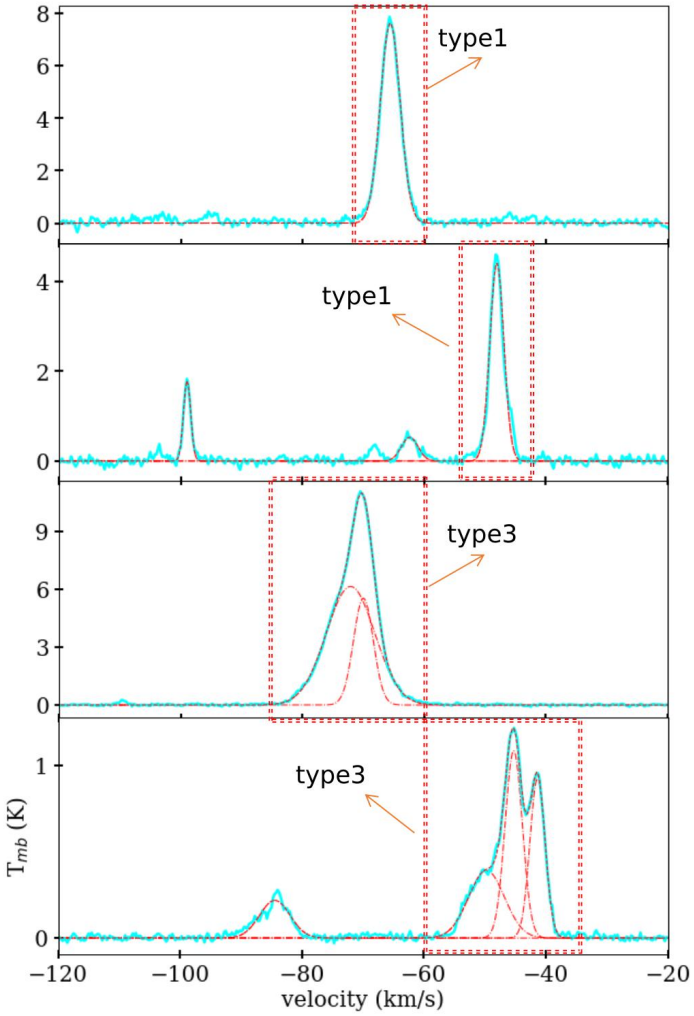


Fig. 3. Typical ^{13}CO (3−2) line profiles of the identified structures. Red dashed box marks the limited velocity range of each structure.

ing in a measurable velocity gradient along the filament. Since we are now studying each local hub-filament structure, we follow the same analysis presentend in Zhou et al. (2022) to fit the velocity gradients around the intensity peaks. In Fig. 6(a), the velocity gradients fitted in those local hub-filament structures are quite comparable to those presented in Zhou et al. (2022) and Zhou et al. (2023). For the velocity gradients measured in this work, almost all of them can be fitted with masses in the range $\sim$ 100−10000 $M_\odot$, which is consistent with the mass distribution of leaf structures shown in Fig.4(b), indicating that local dense structures as gravitational centers can accrete the surrounding diffuse gas and produce the observed velocity gradients. Measurements of velocity gradients provide evidence for gas inflow within these structures, which can serve as kinematic evidence that these structures can be regarded as hub-filament systems.

3.5. Density contrast

As illustrated in Fig. 2, the presence of a distinct central hub indicates a significant density contrast between the hub and the surrounding diffuse gas. We define the density contrast C as the ratio of the average column density in the hub region, $N_{\rm hub}$, to the average column density within an elliptical ring surrounding the hub, which has a width equal to the size of the hub (the effective radius of the leaf structure), $N_{\rm a}$,

$$C = N_{\rm hub}/N_{\rm a}. \quad (1)$$

As expected, in Fig. 4(a), HFs display the highest density contrast. Moreover, HFs also possess the largest masses and the lowest virial ratios. Nearly half of the leaf structures are non-HFs. We divided non-HFs into two groups, using the median density contrast of non-HFs (C = 1.2) as the threshold. In Fig.7, the structures in the group with the higher density contrast have larger scale, larger mass and lower virial ratio. We note that the virial parameter α_{vir} is only a ratio of the gravitational energy (E_g) and the kinetic energy (E_k) in this work, which only provides relative measurements of gravitational strength. The value of α_{vir} should not be a criterion for determining whether a structure undergoes gravitational collapse.

In Fig. 4(a) and (d), non-HFs have significantly smaller density contrasts and scales. Therefore, one would argue that only when the structural scale is large enough, the density contrast can clearly manifest. In other words, the hub size of non-HFs could be too small to be resolved. Moreover, beam smearing effect may create an illusion of uniform density distribution. In order to check these possibilities, in Fig. 8, we fitted the correlation between density contrast and scale of non-HFs, and found that there is no clear correlation between them. Moreover, since HFs and non-HFs have significant scale overlap (see Fig. 4(d)), if HFs can discern the hubs, non-HFs should be able to as well. Therefore, relative to HFs, the absence of the hubs in non-HFs is real and not merely a resolution issue.

3.6. ATLASGAL clumps and radio sources

The physical properties of the ATLASGAL clumps (Urquhart et al. 2022) associated with leaf structures can reveal their evolutionary stages. If the distance between the central coordinates of a clump and a leaf structure is less than the effective radius of the leaf structure, we associate the clump with the leaf structure. In total, 438 leaf structures are associated with 293 ATLASGAL clumps. The fraction of hubs associated with ATLASGAL clumps is approximately 0.83 for HFs and about 0.32 for non-HFs. We also match leaf structures with the radio sources identified in the CORNISH-South survey (Irabor et al. 2023). Approximately 29% of hubs in HFs are associated with radio sources, compared to only around 4% in non-HFs. The ratio of HFs associated with clumps and radio sources is much higher than that of non-HFs, indicating that HFs are more evolved than non-HFs. In Fig.4(f), the *Spitzer* 8 μm peak intensity of HFs is significantly higher than that of non-HFs.

The luminosity-to-mass ratios, L/M, of the clumps can be used to trace the evolutionary process of star formation (Saraceno et al. 1996; Molinari et al. 2008; Liu et al. 2016b; Stephens et al. 2016; Urquhart et al. 2018). In Fig.9, the clumps associated with HFs have significantly higher L/M ratios compared to those associated with non-HFs, which also suggests that HFs are more evolved than non-HFs.

4. Discussion

4.1. Evolutionary sequence

In Fig. 10, we compared the physical properties of two types of non-HFs, i.e. without and with associated ATLASGAL clumps. The structures associated with ATLASGAL clumps (more evolved) have higher density contrast, larger mass and

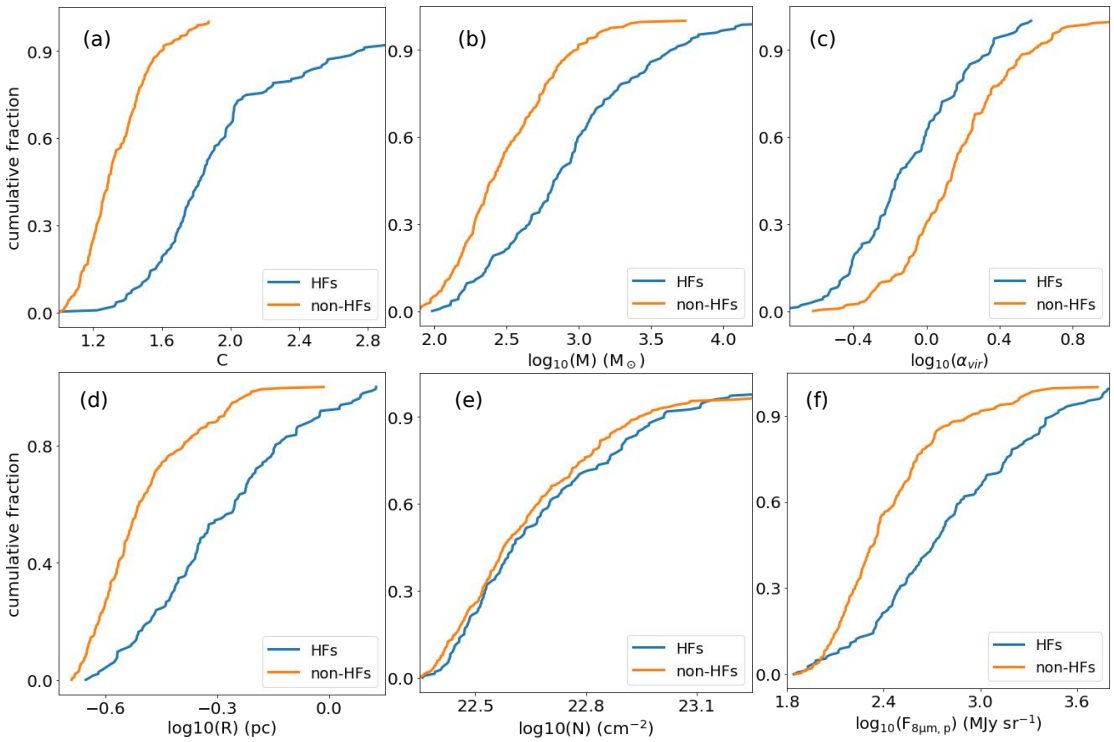


Fig. 4. Physical properties of HFs and non-HFs. (a) Column density contrast defined in Sec.3.5; (b) Mass; (c) Virial ratio; (d) Effective radius; (e) Average column density; (f) Peak intensity of *Spitzer* 8 μm emission.

lower virial ratio. We also compared the physical properties of two types of HFs, i.e. without and with associated radio sources. The structures associated with radio sources (more evolved) have higher density contrast, larger mass and lower virial ratio. As confirmed in Sec.3.6, HFs is more evolved than non-HFs. In Fig.4, compared with non-HFs, the more evolved HFs have higher density contrast, larger mass and lower virial ratio. Therefore, there may be an evolutionary sequence from non-HFs to HFs. In Fig.1, in terms of spatial distribution, HFs and non-HFs are mixed together, sharing the same physical environment. In other words, the differences between the two types of structures are not caused by their physical environment. If there is an evolutionary relationship between them, these differences can be naturally explained.

4.2. Density contrast

As shown in Fig.3.5(e), although HFs have larger density contrast, the average density of non-HFs is almost comparable with HFs. Therefore, non-HFs are not necessarily low-density structures. It is just that their density distribution is relatively uniform. Leaf structures may undergo a gravitational focusing process starting from an initially relatively uniform density distribution, ultimately forming hub-filament structures. This process has been investigated in detail for the fragment G332.83-0.55 (one of HFs in this work) in Zhou et al. (2024c), where a strong gravitational center (hub) plays a crucial role in shaping the hub-filament morphology. According to the possible evolutionary sequence from non-HFs to HFs, currently, non-HFs lack a distinct gravitational focusing process that would result in significant density contrast. Thus, their density distribution is relatively uniform. Therefore, density contrast C effectively measures the extent of gravitational collapse and the strength of the gravitational center of the structure, which definitively shape the hub-filament morphology. In terms of reflecting the development and evolutionary stage of a structure, density contrast is more crucial than the density itself.

4.3. Structure of molecular clouds

Considering the evolutionary sequence from non-HFs to HFs, non-HFs can be regarded as potential HFs, since they will evolve into HFs in the future. In this work, the molecular clouds are decomposed into hundreds of HFs and potential HFs with different scales. Combined with the kinematic evidence in Zhou et al. (2023) and Zhou et al. (2024d), we suggest that molecular clouds are network structures formed by the gravitational coupling of multi-scale hub-filament structures. Hub-filament structures serve as the basic constituent units of molecular clouds. The knots in the networks are the hubs, and they are the local gravitational centers and the main star-forming sites. Actually, clumps in molecular clouds are equal to the hubs.

This network picture can naturally explain the results in Zhou et al. (2024d) and Zhou et al. (2024e): In star-forming regions, spanning from cloud to core scales, initial-stage feedback has minimal impact on the physical properties of embedded dense gas structures. While feedback at these scales can disrupt the original cloud complex, the substructures within it can reorganize into new configurations governed by gravity around new gravitational centers. This dynamic process involves both structural destruction and formation, shifts in gravitational centers, but gravitational collapse is always ongoing.

The network structure of molecular clouds implies that the knots or local dense structures as local hubs or local gravitational centers are relatively independent of each other. Although some knots in the cloud evolve more rapidly, forming HII regions, due to their relative independence with the neighboring knots, early feedback from them does not significantly impact the physical

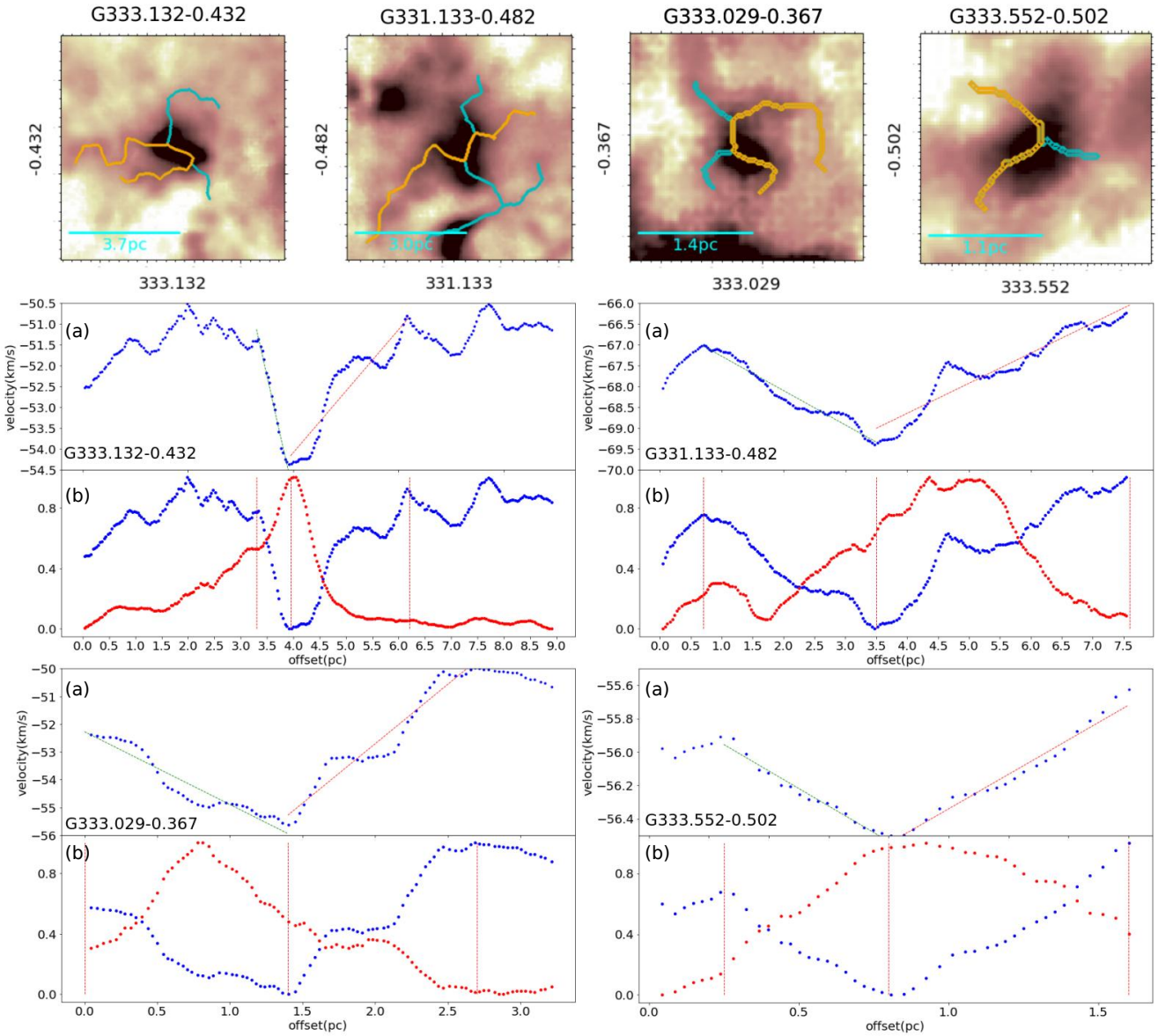


Fig. 5. Four HFs structures used to demonstrate the velocity gradient fitting. In the first row, the background is the integrated intensity map of ^{13}CO (3−2), orange and cyan lines are the filaments identified by the filfinder algorithm. The two rows below show the velocity and intensity profiles extracted along the orange filaments. In panel (a), velocity gradients are fitted in the ranges defined by the red vertical dashed lines in panel (b). Straight lines show the linear fitting results. In panel (b), blue and red dotted lines show the normalized velocity and intensity, respectively.

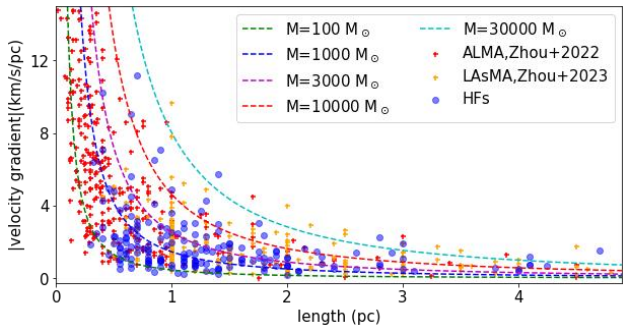


Fig. 6. Velocity gradient versus the length over which the gradient has been fitted. Red and orange "+" represent the velocity gradients fitted in Zhou et al. (2022) and Zhou et al. (2023), respectively. The dashed lines show free-fall velocity gradients for comparison. For the free-fall model, from left to right, the lines denote masses of 100 $M_\odot$, 1000 $M_\odot$, 3000 $M_\odot$, 10000 $M_\odot$ and 30000 $M_\odot$, respectively.

properties of the neighboring knots. Therefore, early feedback mainly reshapes the network structure of molecular clouds, representing a topological deformation rather than destruction of the molecular cloud, as illustrated in Fig.8 of Zhou et al. (2024e).

5. Summary

We decomposed the G333 complex and the G331 GMC into multi-scale hub-filament structures using the high-resolution ^{13}CO (3−2) data in LAsMA observation. The main conclusions are as follows:

1. The intensity peaks as the potential hubs were identified based on the integrated intensity (Moment 0) map of ^{13}CO (3−2) emission by the dendrogram algorithm. Around all intensity peaks, we extracted the average spectra in the 2.5 times extended regions to decompose their velocity components and fix the velocity range for each structure. The final identification of hub-filament systems (HFs) is based on the cleaned Moment-0

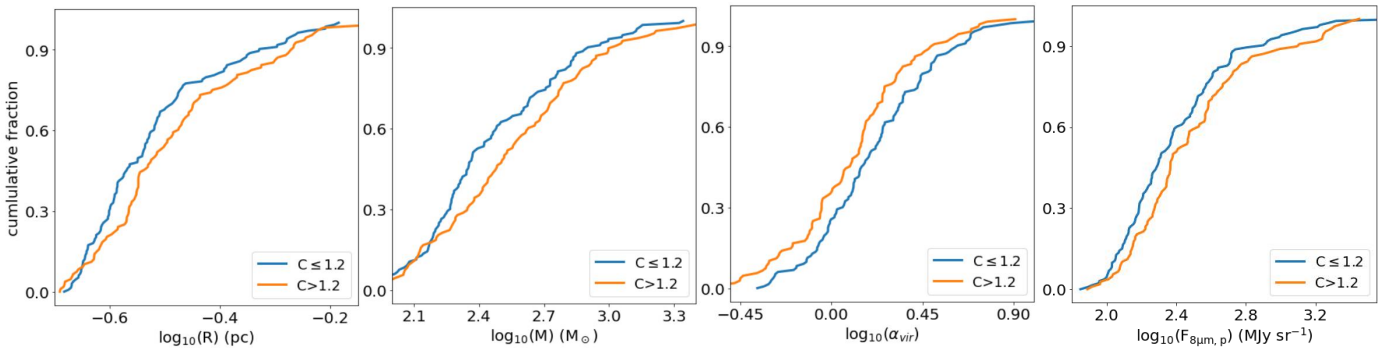


Fig. 7. Comparison of non-HFs with the density contrast C<1.2 and C>1.2. (a) Effective radius; (b) Mass; (c) Virial ratio; (d) Peak intensity of *Spitzer* 8 μm emission.

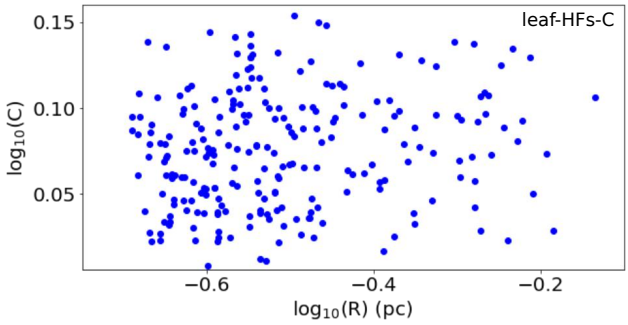


Fig. 8. The correlation between density contrast and scale of non-HFs.

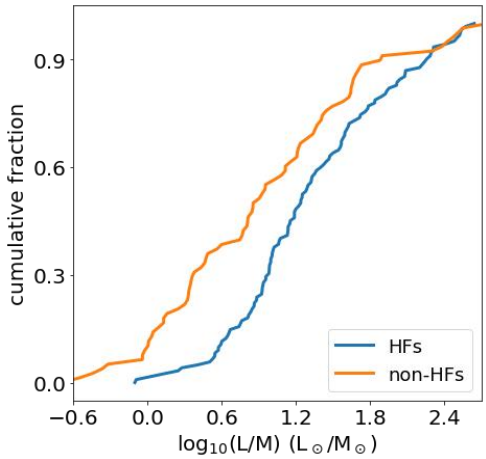


Fig. 9. The luminosity-to-mass ratio of ATLASGAL clumps associated with HFs and non-HFs.

maps made by restricting to the fixed velocity ranges to exclude the possibility that the overlaps of uncorrelated velocity components produce fake hub-filament structures.

2. Compared with non-HFs, HFs have significantly higher density contrast, larger masses and lower virial ratios.

3. We employed the filfinder algorithm to identify and characterize filaments within HFs based on integrated intensity maps. We extracted the velocity and intensity profiles along the filaments, and fitted the velocity gradients around intensity peaks. Changes in velocity gradient with scale roughly follow the free-fall model in the mass range $\sim$ 100–10000 $M_\odot$, which is consistent with the mass distribution of leaf structures, indicating that local dense structures as gravitational centers accrete the surrounding diffuse gas and produce the observed velocity gradients. Velocity gradient measurements provide evidence of gas inflow within these structures, offering kinematic support that they are hub-filament systems.

4. The ratio of HFs associated with clumps and radio sources is much higher than that of non-HFs. The *Spitzer* 8 μm peak intensity of HFs is significantly higher than that of non-HFs. The luminosity-to-mass ratios, L/M, of the associated ATLASGAL clumps from HFs to non-HFs is decreasing. These results suggest that HFs are more evolved than non-HFs.

5. In non-HFs, structures associated with clumps exhibit higher density contrast, larger mass, and lower virial ratios compared to those without clump associations. Similarly, in HFs, structures linked to radio sources show higher density contrast, larger mass, and lower virial ratios than those without such associations. Additionally, compared to non-HFs, the more evolved HFs demonstrate higher density contrast, larger mass, and lower virial ratios. These observations suggest a potential evolutionary sequence from non-HFs to HFs.

6. Although HFs have larger density contrast, the average density of non-HFs is almost comparable with HFs. Thus, non-HFs are not necessarily low-density structures. It is just that their density distribution is relatively uniform. According to the possible evolutionary sequence from non-HFs to HFs, currently, non-HFs lack a distinct gravitational focusing process that would result in significant density contrast. The density contrast C effectively measures the extent of gravitational collapse and the strength of the gravitational center of the structure that definitively shape the hub-filament morphology. In terms of reflecting the development and evolutionary stage of a structure, density contrast is more crucial than density itself.

7. Molecular clouds are decomposed into hundreds of hub-filament structures with different scales. Combined with the kinematic evidence in Zhou et al. (2023) and Zhou et al. (2024d), we suggest that molecular clouds are network structures formed by the gravitational coupling of multi-scale hub-filament structures. Hub-filament structures serve as the basic constituent units of molecular clouds. The knots in the networks are the hubs, they are the local gravitational centers and the main star-forming sites. Actually, clumps in molecular clouds are equal to the hubs. The network structure of molecular clouds implies that the knots or local hubs are relatively independent of each other. Although some knots in the cloud evolve more rapidly, forming HII regions, due to their relative independence with the neighboring knots, early feedback from them does not significantly impact the physical properties of the neighboring knots, as concluded in Zhou et al. (2024d) and Zhou et al. (2024e).

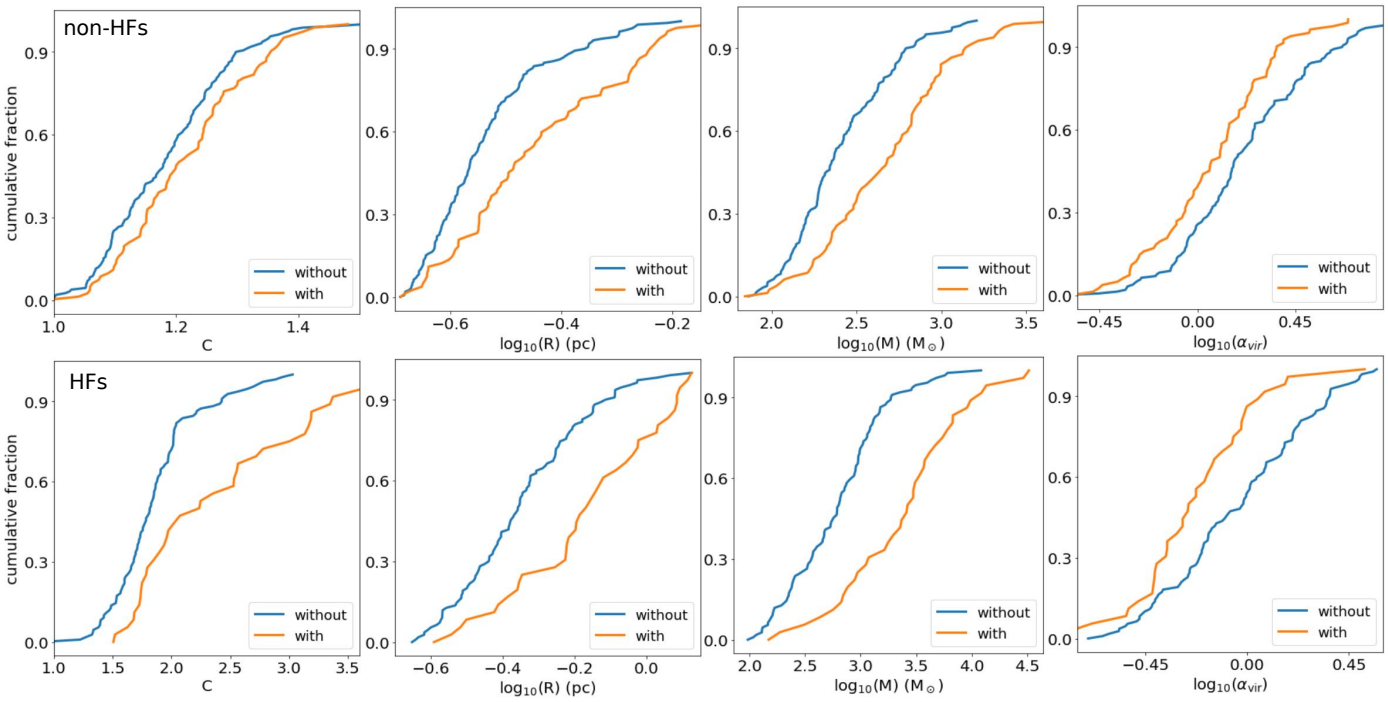


Fig. 10. The first row shows the physical properties of non-HFs without and with associated ATLASGAL clumps. For left to right, the physical parameters are the column density contrast defined in Sec.3.5, effective radius, mass and virial ratio, respectively. The second row shows the physical properties of HFs without and with associated radio sources.

Acknowledgements. Thanks to E. Vázquez-Semadeni and S. H. Li for providing helpful comments. This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under programme ID M-0109.F-9514A-2022. APEX has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. This research made use of astrodendro, a Python package to compute dendrograms of Astronomical data (http://www.dendrograms.org/).

References

Bains, I., Wong, T., Cunningham, M., et al. 2006, MNRAS, 367, 1609, doi: 10.1111/j.1365-2966.2006.10055.x
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953, doi: 10.1086/376696
Dewangan, L. K., Ojha, D. K., Sharma, S., et al. 2020, ApJ, 903, 13, doi: 10.3847/1538-4357/abb827
Elia, D., Molinari, S., Schisano, E., et al. 2017, MNRAS, 471, 100, doi: 10.1093/mnras/stx1357
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13, doi: 10.1051/0004-6361:20065420
Hacar, A., Alves, J., Forbrich, J., et al. 2016, A&A, 589, A80, doi: 10.1051/0004-6361/201527805
Hacar, A., Alves, J., Tafalla, M., & Goicoechea, J. R. 2017, A&A, 602, L2, doi: 10.1051/0004-6361/201730732
Hacar, A., Clark, S. E., Heitsch, F., et al. 2023, in Astronomical Society of the Pacific Conference Series, Vol. 534, Protostars and Planets VII, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 153, doi: 10.48550/arXiv.2203.09562
Henshaw, J. D., Caselli, P., Fontani, F., Jiménez-Serra, I., & Tan, J. C. 2014, MNRAS, 440, 2860, doi: 10.1093/mnras/stu446
Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nature Astronomy, 4, 1064, doi: 10.1038/s41550-020-1126-z
Irabor, T., Hoare, M. G., Burton, M., et al. 2023, MNRAS, 520, 1073, doi: 10.1093/mnras/stad005
Issac, N., Tej, A., Liu, T., et al. 2019, MNRAS, 485, 1775, doi: 10.1093/mnras/stz466
Koch, E. W., & Rosolowsky, E. W. 2015, MNRAS, 452, 3435, doi: 10.1093/mnras/stv1521
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87, doi: 10.1051/0004-6361/202038232
Kuznetsova, A., Hartmann, L., & Ballesteros-Paredes, J. 2015, ApJ, 815, 27, doi: 10.1088/0004-637X/815/1/27
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138, doi: 10.1088/0004-6256/149/4/138
Liu, H.-L., Tej, A., Liu, T., et al. 2022, MNRAS, 511, 4480, doi: 10.1093/mnras/stac378
—. 2023, MNRAS, 522, 3719, doi: 10.1093/mnras/stad047
Liu, T., Zhang, Q., Kim, K.-T., et al. 2016a, ApJ, 824, 31, doi: 10.3847/0004-637X/824/1/31
Liu, T., Kim, K.-T., Yoo, H., et al. 2016b, ApJ, 829, 59, doi: 10.3847/0004-637X/829/2/59
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, MNRAS, 496, 2790, doi: 10.1093/mnras/staa1577
Liu, X.-L., Xu, J.-L., Wang, J.-J., et al. 2021, A&A, 646, A137, doi: 10.1051/0004-6361/201935035
Lockman, F. J. 1979, ApJ, 232, 761, doi: 10.1086/157337
Lu, X., Zhang, Q., Liu, H. B., et al. 2018, ApJ, 855, 9, doi: 10.3847/1538-4357/aaad11
Mazumdar, P., Wyrowski, F., Colombo, D., et al. 2021, A&A, 650, A164, doi: 10.1051/0004-6361/202040205
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565, doi: 10.1146/annurev.astro.45.051806.110602
Molinari, S., Pezzuto, S., Cesaroni, R., et al. 2008, A&A, 481, 345, doi: 10.1051/0004-6361:20078661
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41, doi: 10.1146/annurev-astro-091916-055235
Myers, P. C. 2009, ApJ, 700, 1609, doi: 10.1088/0004-637X/700/2/1609
Nguyen, H., Nguyen Lu'o'ng, Q., Martin, P. G., et al. 2015, ApJ, 812, 7, doi: 10.1088/0004-637X/812/1/7
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112, doi: 10.1051/0004-6361/201321318
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78, doi: 10.1051/0004-6361/201935519
Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., & Molinari, S. 1996, A&A, 309, 827
Schneider, N., Csengeri, T., Hennemann, M., et al. 2012, A&A, 540, L11, doi: 10.1051/0004-6361/201118566
Stephens, I. W., Jackson, J. M., Whitaker, J. S., et al. 2016, ApJ, 824, 29, doi: 10.3847/0004-637X/824/1/29
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059, doi: 10.1093/mnras/stx2258
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389, doi: 10.1093/mnras/stab3511

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **534,** 683–694 (2024) https://doi.org/10.1093/mnras/stae2101
Advance Access publication 2024 September 9

Molecular clouds as hubs in spiral galaxies: gas inflow and evolutionary sequence

J. W. Zhou [1]★ Sami Dib [2]★ and Timothy A. Davis [3]

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*Max Planck Institute für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[3]*Cardiff Hub for Astrophysics Research & Technology, School of Physics & Astronomy, Cardiff University, Queens Buildings, Cardiff CF24 3AA, UK*

Accepted 2024 September 5. Received 2024 September 5; in original form 2024 June 14

ABSTRACT

We decomposed the molecular gas in the spiral galaxy NGC 628 (M74) into multiscale hub-filament structures using the CO (2−1) line by the dendrogram algorithm. All leaf structures as potential hubs were classified into three categories, i.e. leaf-HFs-A, leaf-HFs-B and leaf-HFs-C. Leaf-HFs-A exhibit the best hub-filament morphology, which also have the highest density contrast, the largest mass and the lowest virial ratio. We employed the filfinder algorithm to identify and characterize filaments within 185 leaf-HFs-A structures, and fitted the velocity gradients around the intensity peaks. Measurements of velocity gradients provide evidence for gas inflow within these structures, which can serve as a kinematic evidence that these structures are hub-filament structures. The numbers of the associated 21 μm and H α structures and the peak intensities of 7.7 μm, 21 μm, and H α emissions decrease from leaf-HFs-A to leaf-HFs-C. The spatial separations between the intensity peaks of CO and 21 μm structures of leaf-HFs-A are larger than those of leaf-HFs-C. These evidence indicate that leaf-HFs-A are more evolved than leaf-HFs-C. There may be an evolutionary sequence from leaf-HFs-C to leaf-HFs-A. Currently, leaf-HFs-C lack a distinct gravitational collapse process that would result in a significant density contrast. The density contrast can effectively measure the extent of the gravitational collapse and the depth of the gravitational potential of the structure which, in turn, shapes the hub-filament morphology. Combined with the kinematic analysis presented in previous studies, a picture emerges that molecular gas in spiral galaxies is organized into network structures through the gravitational coupling of multiscale hub-filament structures. Molecular clouds, acting as knots within these networks, serve as hubs, which are local gravitational centres and the main sites of star formation.

Key words: techniques: image processing – ISM: clouds – ISM: kinematics and dynamics – galaxies: ISM – galaxies: star formation – galaxies: structure.

1 INTRODUCTION

Analysing the dynamical interaction between density enhancements in giant molecular clouds and gas motion in their surrounding environment provides insight into the formation of hierarchical structures in high-mass star and star cluster-forming regions (McKee & Ostriker 2007; Dib et al. 2012; Motte, Bontemps & Louvet 2018; Henshaw et al. 2020). Detailed observations of high-mass star-forming regions with high resolution unveil the organized distribution of density enhancements within filamentary networks of gas, particularly evident in hub-filament systems. In these systems, converging flows channel material towards the central hub along the interconnected filaments (Peretto et al. 2013; Henshaw et al. 2014; Zhang et al. 2015; Liu et al. 2016, 2022; Lu et al. 2018; Yuan et al. 2018; Issac et al. 2019; Dewangan et al. 2020; Zhou et al. 2022, 2023; Dib 2023). In particular, Zhou et al. (2022) studied the physical properties and evolution of hub-filament systems across ∼140 protoclusters using spectral line data obtained from the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Liu et al. 2020). They proposed that hub-filament structures exhibiting self-similarity and filamentary accretion appear to persist across a range of scales within high-mass star-forming regions, spanning from several thousand astronomical units to several parsecs. This paradigm of hierarchical, multiscale hub-filament structures was generalized from clump–core scale to cloud–clump scale in Zhou et al. (2023) and from cloud–clump scale to galaxy–cloud scale in Zhou & Davis (2024). Hierarchical collapse and hub-filament structures feeding the central regions are also depicted in previous works, see Motte et al. (2018), Vázquez-Semadeni et al. (2019), Kumar et al. (2020) and references therein.

Kinematic analyses presented in Zhou et al. (2023) and Zhou & Davis (2024) demonstrate the presence of multiscale hub-filament structures within molecular clouds and spiral galaxies. The results notably show that intensity peaks, acting as hubs, are correlated with converging velocities, suggesting that surrounding gas flows are directed towards these dense regions. Filaments across various scales exhibit distinct velocity gradients, with a marked increase in these gradients at smaller scales. Interestingly, the variations in velocity gradients measured at larger scales align with expectations from gravitational free-fall with higher central masses. This correlation implies that inflows on large scales are driven by large-scale structures, potentially due to the gravitational coupling of smaller scale structures. Fig. 5 of Zhou & Davis (2024) shows a

★ E-mails: jwzhou@mpifr-bonn.mpg.de (JWZ); sami.dib@gmail.com (SD)

© 2024 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

vivid example of gravitational coupling, where multiple peaks are coupled together to form a gravitational potential well on larger scale, and each peak itself is also a local gravitational centre. This aligns with the global hierarchical collapse (GHC) scenario proposed by Vázquez-Semadeni et al. (2019), which suggests that clouds are composed of multiple nested collapses occurring across a wide range of scales. These observations agree with the hierarchical nature found in molecular clouds and spiral galaxies, and gas inflow from large to small scales. Large-scale velocity gradients are consistently associated with numerous intensity peaks, reinforcing the idea that the clustering of smaller scale structures can act as gravitational centres on larger scales.

Based on the kinematic evidence, the main goal of this work is to directly recover the multiscale hub-filament structures in a spiral galaxy NGC 628 (M74 or the Phantom Galaxy).

2 DATA

We selected a face-on spiral galaxy NGC 628 (M74) from the PHANGS–ALMA survey. We used the combined 12m+7m+TP PHANGS–ALMA CO (2−1) data cubes to investigate gas kinematics and dynamics and identify hub-filament structures in the galaxy, which have a spectral resolution of 2.5 km s^{-1} and an angular resolution $\sim$ 1.1 arcsec, corresponding to a linear resolution $\sim$50 pc at the distance 9.8 Mpc (Anand et al. 2021; Leroy et al. 2021b). We also use the *JWST* 7.7 and 21 μm maps with an angular resolution $\sim$ 0.67 arcsec or a linear resolution $\sim$30 pc from the PHANGS–*JWST* survey and the Hα emission map with an angular resolution $\sim$ 0.92 arcsec or a linear resolution $\sim$45 pc from the PHANGS–MUSE survey. Overviews of the PHANGS–ALMA, PHANGS–MUSE, and PHANGS–*JWST* surveys' science goals, sample selection, observation strategy, and data products are described in Leroy et al. (2021a,b), Emsellem et al. (2022), and Lee et al. (2023). All the data are available on the PHANGS team website.[1] The field-of-views (FOVs) of these observations are shown in Fig. 1. NGC 628 was selected for the following reasons:

(1) It is a face-on galaxy. The inclination angle of NGC 628 is only 8.9° et al. 2020). Therefore, this facilitates the identification of the hub-filament structures embedded within the galaxy.

(2) High-resolution CO data and the galaxy's relatively near distance can clearly reveal the hub-filament structures.

(3) It is covered by multiwavelength observations. As discussed below, high-resolution 21 μm and Hα emission are crucial to determine the evolutionary states of CO structures.

3 RESULTS

3.1 Dendrogram

We conducted a direct identification of hierarchical (sub-)structures based on the two-dimensional intensity maps. As described in Rosolowsky et al. (2008), the dendrogram algorithm decomposes density or intensity data into hierarchical structures called leaves, branches, and trunks. Using the ASTRODENDRO package,[2] there are three major input parameters for the dendrogram algorithm: *min_value* for the minimum value to be considered in the data set, *min_delta* for a leaf that can be considered as an independent entity, and *min_npix* for the minimum area of a structure. For the CO (2−1) data cube, there are two types of Moment 0 maps (strictly masked and broadly masked) in the data product of the PHANGS–ALMA survey.[3] The strictly masked maps only include emissions that are identified as signals with high confidence in the data cube, which might filter out the relatively faint structures. The broadly masked maps offer superior completeness and cover larger areas compared to the strictly masked maps. However, due to the inclusion of more regions with faint emissions or areas close to bright emissions, they tend to be noisier and may contain false positives. In order to ensure the reliability of the identified structures and because we are only interested in local dense structures, we select the strictly masked Moment 0 map to identify structures. Since all the retained structures on the strictly masked Moment 0 map are reliable, we only require the smallest area of the identified structure be larger than 1 beam area. We do not set additional parameters in the algorithm to minimize the dependence of the identification on parameter settings. Finally, we obtained 773 leaf structures.

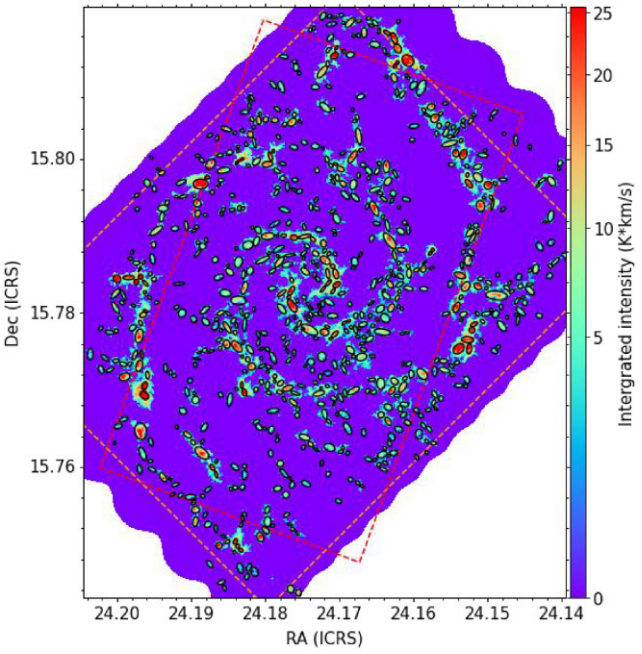


Figure 1. Background is the integrated intensity map of CO (2−1) emission. Black ellipses represent leaf structures identified by the dendrogram algorithm. Red and orange boxes mark the field-of-views in 21 μm and Hα observations, respectively. The long and short axes of the ellipse are 2*a* and 2*b*, as discussed in Section 3.1.

For the 21 μm and Hα emission, apart from *min_npix* = 1 beam area, we also take the values of *min_value* = 3∗$\sigma_{\rm rms}$, *min_delta* = 3∗$\sigma_{\rm rms}$, where $\sigma_{\rm rms}$ is the background intensity. The total numbers of 21 μm and Hα leaf structures are 1491 and 1965, respectively. We first retained as much structures as possible using these lower standards, then eliminated the diffuse structures, as described in Section 3.5.2. In Figs 1, 2, and 3, the CO, 21 μm and Hα structures identified by the dendrogram algorithm exhibit a strong correspondence with the background intensity maps.

The algorithm characterizes the morphology of each structure by approximating it as an ellipse. Within the dendrogram, the root mean square (rms) sizes (second moments) of the intensity distribution

[1] https://sites.google.com/view/phangs/home
[2] https://dendrograms.readthedocs.io/en/stable/index.html
[3] Details of the masking strategy and completeness statistics are presented in the PHANGS pipeline paper (Leroy et al. 2021a).

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

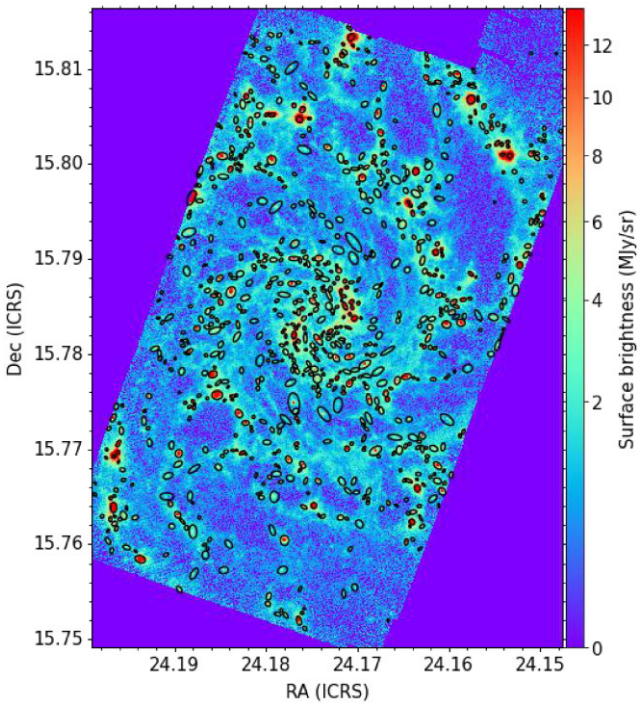


Figure 2. Background is the intensity map of *JWST* 21 μm emission. Black ellipses represent the remained (bright) 21 μm structures. The selection was done in Section 3.5.2. The long and short axes of the ellipse are 2a and 2b.

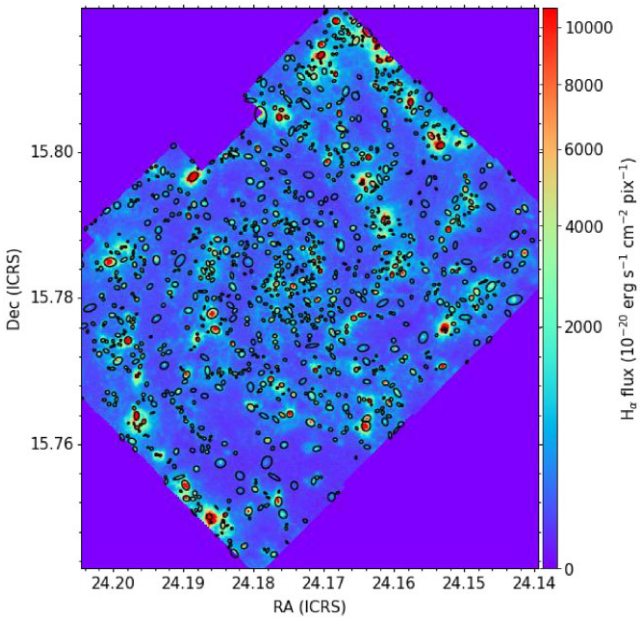


Figure 3. Background is the intensity map of Hα emission. Black ellipses represent the remained (bright) Hα structures. The selection was done in Section 3.5.2. The long and short axes of the ellipse are 2a and 2b.

along the two spatial dimensions define the long and short axes of the ellipse, denoted as a and b. As described in Zhou et al. (2024a), the initial ellipse with a and b is smaller, so a multiplication factor of two is applied to appropriately enlarge the ellipse. Then the effective physical radius of an ellipse is $R_{\rm eff} = \sqrt{2a \times 2b} * D$, where D is the distance of the galaxy. For a structure with an area A and a total integrated intensity $I_{\rm CO}$, the mass of the structure can be calculated by

$$M = \alpha_{\rm CO}^{2-1} \times I_{\rm CO} \times A, \quad (1)$$

where $\alpha_{\rm CO}^{2-1} \approx 6.7\ {\rm M}_\odot {\rm pc}^{-2} ({\rm K\ km\ s}^{-1})^{-1}$ (Leroy et al. 2022).

3.2 Velocity components

For the identified molecular structures, we extracted the average spectrum of each structure to investigate its velocity components and gas kinematics. Large-scale velocity gradients from the galaxy's rotation contribute to the non-thermal velocity dispersion. To address this, and before extracting the average spectra, we removed the bulk motion due to the rotation of galaxy by creating gas dynamical models using the Kinematic Molecular Simulation (KINMS) package (Davis et al. 2013), as shown in Fig. 4. Then, following the procedure described in Zhou et al. (2024a), we fitted the averaged spectra of 773 leaf structures individually using the fully automated Gaussian decomposer GAUSSPY+ (Lindner et al. 2015; Riener et al. 2019) algorithm. Almost all structures exhibit a single-peak profile. Only a few structures have a clear double-peak profile. Fig. 5 displays the average spectra of the central regions of the structures presented in Figs 6, 7, and 8, marked by cyan ellipses.

3.3 Classification

From the line profile, we can fix the velocity range of each structure. Then, the Moment 0 map of each structure was reproduced in the corresponding velocity range to eliminate the overlap of potential incoherent velocity components. All identified intensity peaks (leaf structures) of CO (2−1) emission are thought to be potential hubs. Around these intensity peaks, we extended the spatial ranges to investigate the filamentary structures connected with them. After trying different enlargement factors, extending to 2.5 times the hub size (the effective radius of the leaf structure), we can recover the entire hub-filament structure and at the same time, avoid including many neighbouring structures, as shown in Figs 6, 7, and 8.

All leaf structures were classified by eye into three categories only based on their morphology, i.e. leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C. The numbers of structures in the three categories are 234, 181, and 358. Some examples in the three categories are shown in Figs 6, 7, and 8. From these maps, we can observe that leaf-HFs-C do not exhibit clear central hubs, meaning that the density contrast between the hub and the surrounding diffuse gas is not pronounced. In contrast, leaf-HFs-A and leaf-HFs-B feature distinct central hubs or high-density central regions, characteristic of hub-filament structures. Compared to leaf-HFs-B, leaf-HFs-A demonstrate more defined hub-filament morphology, with filamentary diffuse gas surrounding their central hubs. The hubs of leaf-HFs-A are more outstanding. While the boundary between leaf-HFs-A and leaf-HFs-B might be less distinct, the difference between leaf-HFs-A and leaf-HFs-C is significant. Therefore, in the subsequent discussion, we focus on the comparison between leaf-HFs-A and leaf-HFs-C.

The initial classification was only based on the morphology, because other differences between the structures were unknown. In the subsequent analysis, we can see that the physical properties of the structures in three categories are also significantly different. Therefore, the morphological differences are the result of certain physical processes shaping them and are not coincidental.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

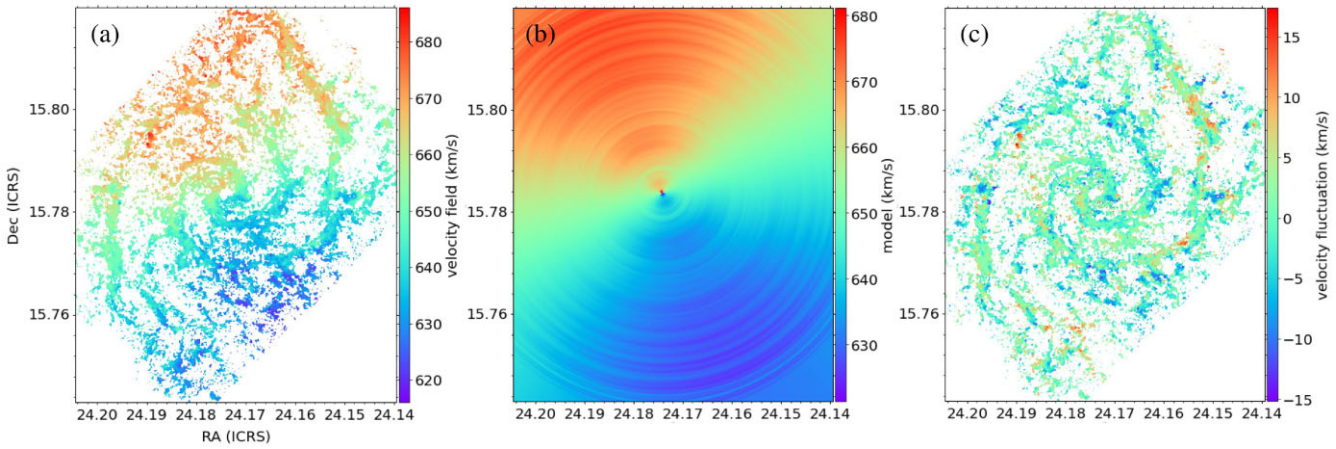


Figure 4. (a) The velocity field (Moment 1) of NGC 628 in the observation; (b) gas dynamical model created by the Kinematic Molecular Simulation (KINMS) package; (c) velocity fluctuations after subtracting the model.

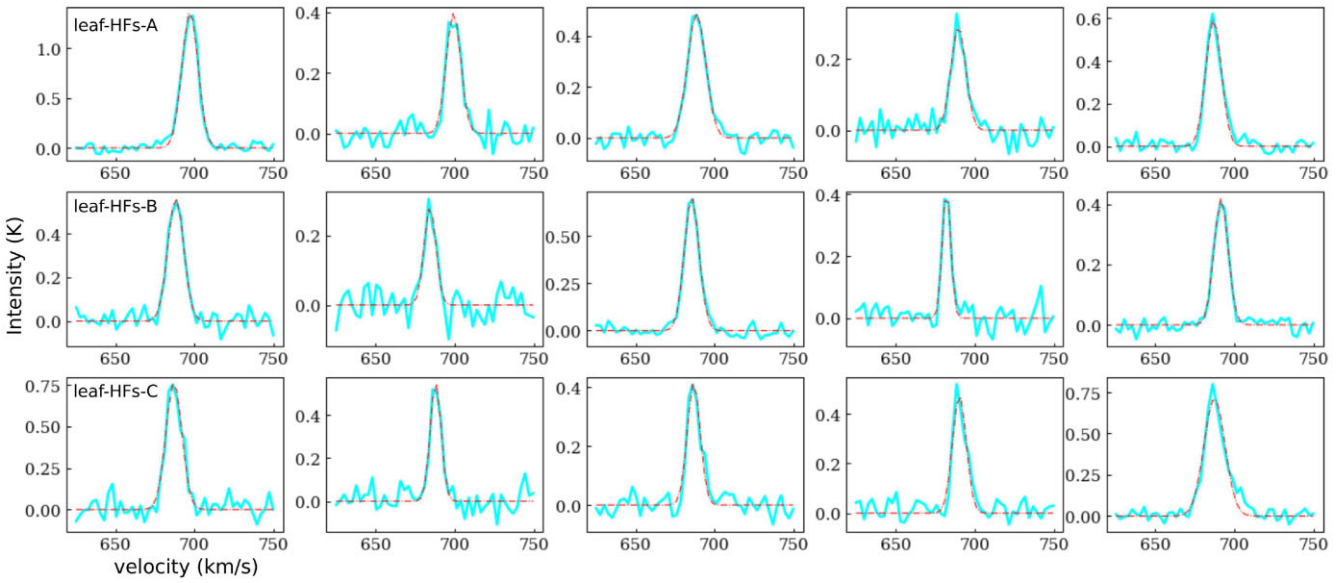


Figure 5. Average spectra and their Gaussian fitting of the CO structures marked by cyan ellipses in Fig. 6 (the first row), Fig. 7 (the second row), and Fig. 8 (the third row).

3.4 Filamentary structures

3.4.1 Identification

As done in Zhou & Davis (2024), we also use the filfinder algorithm to characterize the filamentary structures around the hub. In Zhou & Davis (2024), we only focussed on large-scale filaments along the spiral arms. Clear fluctuations in velocity and density were observed along these filaments. Each individual intensity peak reveals a local hub. In this work, we first identified the local intensity peaks (leaf structures), and then searched for hub-filament structures centred around these intensity peaks as potential hubs. Now, we need to continue identifying small-scale filamentary structures within these local hub-filament structures. Due to the limited observational resolution, even for the largest leaf-HFs-A structures, they still lack enough pixels to recognize the filaments. Therefore, we have to regrid the images and increase the number of pixels in the images from $N_x * N_y$ to $2N_x * 2N_y$.[4] To avoid introducing false structures, we only doubled the number of pixels. As shown in Fig. 6, the morphology of the structures remains unchanged. The identified filamentary structures also align well with the background. However, due to the resolution limitations, the filamentary structures do not appear very extended.

3.4.2 Velocity gradient

A kinematic feature of hub-filament structures is that the gas flow along the filament converges towards the central hub, resulting in a measurable velocity gradient along the filament. Since we are

[4]We do this by using the zoom function from the scipy.ndimage module to regrid the image by setting an interpolation order of 3.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

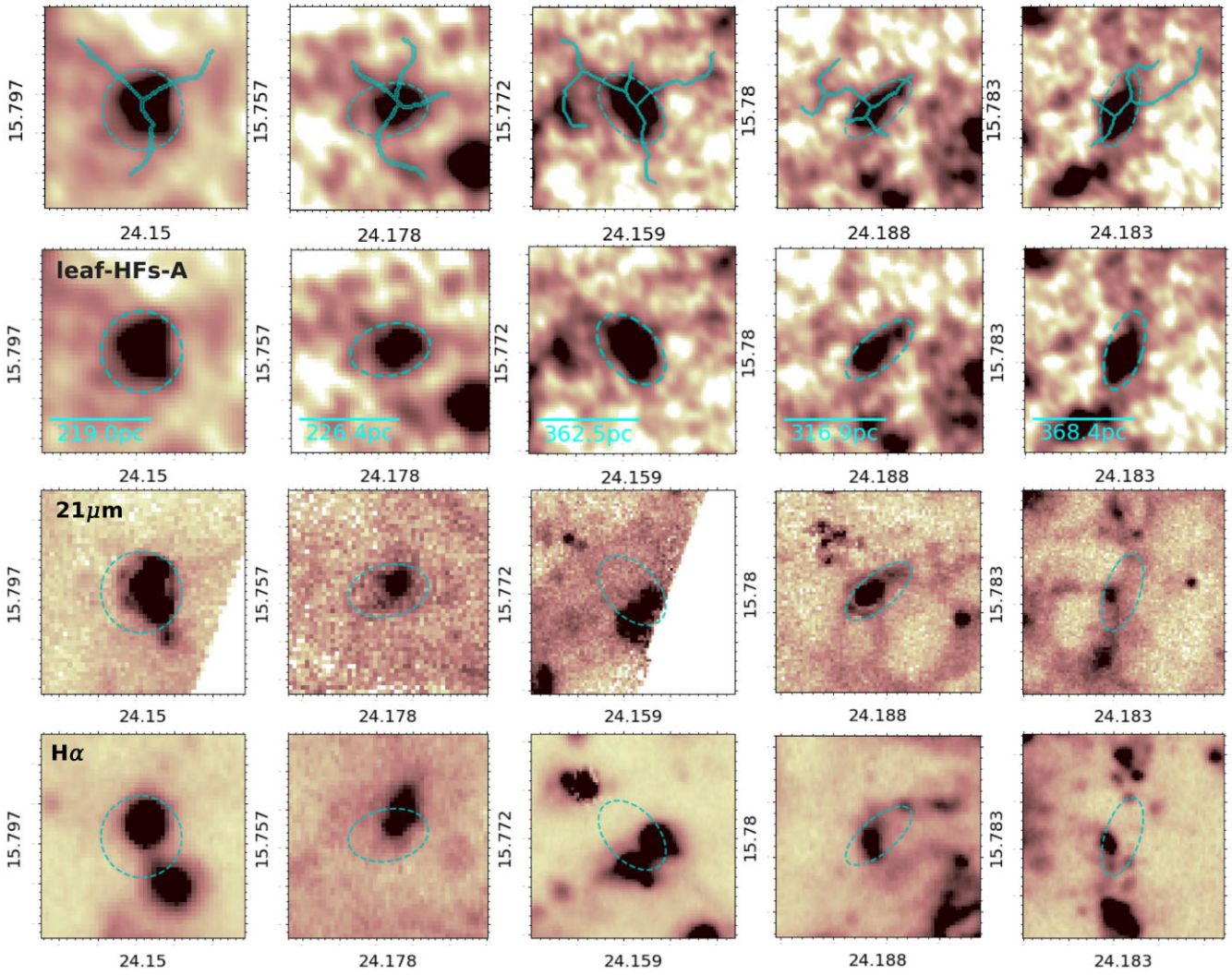


Figure 6. Some examples of leaf-HFs-A. CO emission (the first and second rows), 21 μm emission (the second row) and H$_\alpha$ emission (the third row). The long and short axes of the ellipse are 2*a* and 2*b*, as discussed in Section 3.1. In the first row, the filaments identified by the FILFINDER algorithm are overlaid on the Moment 0 maps.

now studying each local hub-filament structure, we follow the same analysis presentend in Zhou et al. (2022) to fit the velocity gradients around the intensity peaks. In Fig. 9(a), the velocity gradients fitted in NGC 628 are quite comparable to those in NGC 4321 and NGC 5236 presented in Zhou & Davis (2024). The measured velocity gradients pre-dominantly fit within the mass range of approximately 10^5 to 10^7 M$_\odot$. This range aligns with the mass distribution of leaf structures shown in Fig. 9(b). This consistency suggests that local dense structures act as gravitational centres, accreting surrounding diffuse gas and thereby generating the observed velocity gradients. Measurements of velocity gradients provide evidence for gas inflow within these structures, which can also serve as kinematic evidence that these structures can be regarded as hub-filament systems.

We note that only leaf-HFs-A structures were considered in the identification of filaments and the analysis of the velocity gradients. Furthermore, we successfully identified the filaments for only 185 leaf-HFs-A structures, and the filaments are long enough to show clear velocity gradients.

3.5 Physical properties

3.5.1 Density contrast

As shown in Fig. 6, a clear central hub implies a noticeable density contrast between the hub and the surrounding diffuse gas. We define the density contrast C as the ratio of the average column density in the hub region, $N_{\rm hub}$, to the average column density within an elliptical ring around the hub with the width equal to the hub size (the effective radius of the leaf structure), $N_{\rm a}$,

$$C = N_{\rm hub}/N_{\rm a}. \quad (2)$$

As expected, leaf-HFs-A with the best hub-filament morphology have the highest density contrast in Fig. 10(a). Moreover, leaf-HFs-A also possess the largest masses and the lowest virial ratios.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

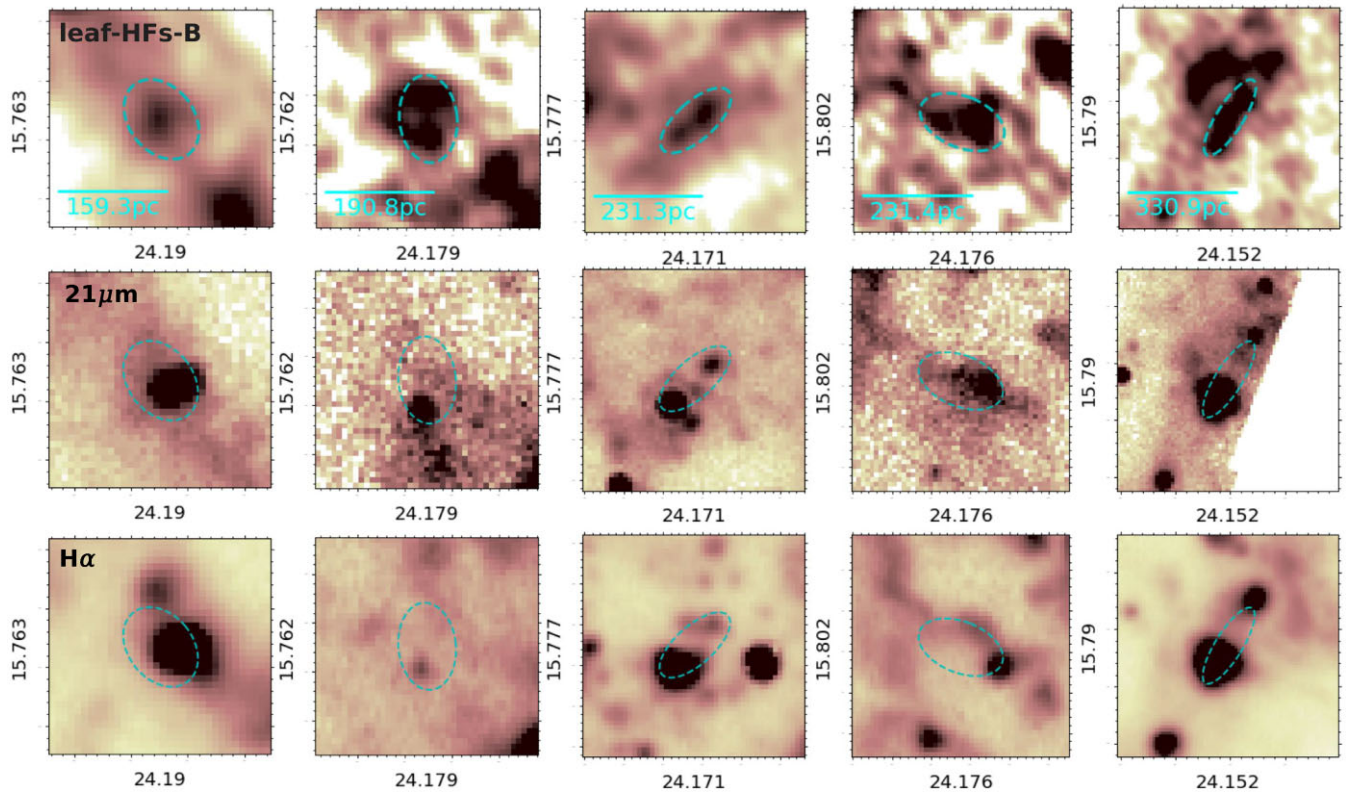


Figure 7. Same as Fig. 6, but for leaf-HFs-B.

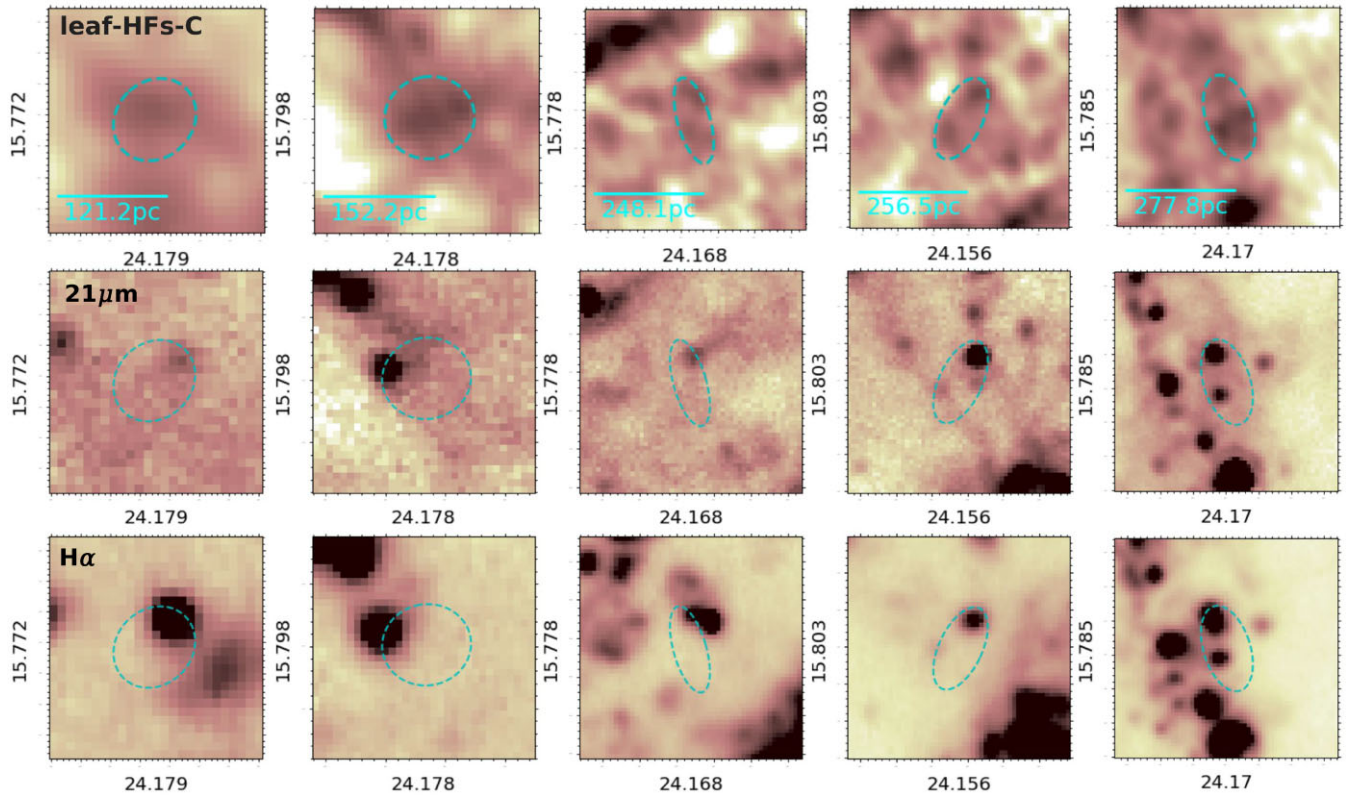


Figure 8. Same as Fig. 6, but for leaf-HFs-C.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

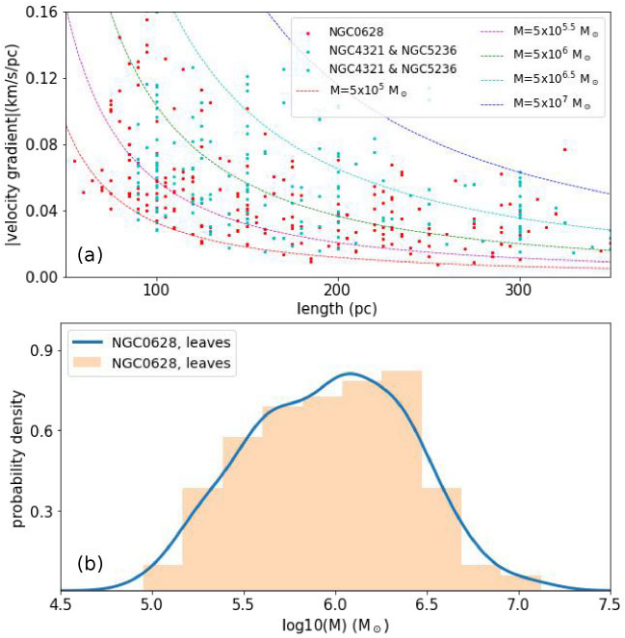


Figure 9. Statistical analysis of all fitted velocity gradients. (a) Velocity gradient versus the length. The colour lines show the freefall velocity gradients for comparison. For the freefall model, red, magenta, green, cyan, and blue lines denote masses of 10^5, $10^{5.5}$, 10^6, $10^{6.5}$, and 10^7 $M_\odot$, respectively; (b) the mass distribution of leaf structures.

3.5.2 *Association with 21 μm and H α emission*

Diffuse 21 μm and H α emission should be unrelated to recent star formation activities. Therefore, for 21 μm and H α structures identified in Section 3.1, we need to filter the faint structures. 21 μm and H α emission that originate from the clusters born in the clouds should present intensity peaks. Similar to the CO structures, we calculated the intensity contrast for all identified 21 μm and H α structures. Before the calculation, we first shift the centre of each structure to the pixel with the strongest emission. Structures with peak intensities greater than the median peak intensity of all structures are retained. Given that a large number of faint structures were identified using loose criteria in Section 3.1, this standard is not stringent. For structures that do not meet this criterion, if their intensity contrast is greater than 1.5, they will also be retained. The final numbers of 21 μm and H α leaf structures are 841 and 1142, respectively. As can be seen from Figs 2 and 3, the retained (bright) structures encompass all the significant emissions.

As shown in Fig. 1, *JWST* 21 μm emission has the smallest FOV in the observations. Thus, we discarded the CO (2−1) and H α structures which are beyond the FOV of the 21 μm observation. Generally, the structure seen in CO is irregular and extended, so the densest part may not be at the effective centre of the structure output by the dendrogram algorithm. However, the position with the highest column density truly represents the site of star formation. Therefore, we calculate the spatial separations between the position of maximum CO column density and the positions of maximum 21 μm or H α intensity. If the separations are less than the effective radius of a CO structure, we consider the corresponding 21 μm or H α structures are associated with the CO structure. We note that CO could be optically thick, which may lead to an underestimation of the column density and, consequently, affect the estimation of the density centre.

Finally, there are 313 21 μm structures (37 per cent, 313/841) and 324 H α structures (28 per cent, 324/1142) associated with CO structures, where 185 CO structures have both strong 21 μm and H α emissions. Limited to the FOV of 21 μm observations, the total numbers of leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C are 163, 119, and 251, respectively. The proportions of leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C associated with both 21 μm and H α structures are 58 per cent (95/163), 37 per cent (44/119), and 18 per cent (46/251), indicating that leaf-HFs-A are more evolved than leaf-HFs-C. We also calculated the peak intensities of 7.7 μm, 21 μm, and H α emissions for each CO structure. As shown in Fig. 10 (f)–(h), all emissions gradually increase from leaf-HFs-C to leaf-HFs-A.

3.5.3 *Scale and density*

In Fig. 10(a) and (e), leaf-HFs-C have significantly smaller density contrasts and scales. However, leaf-HFs-C and leaf-HFs-B show comparable column density distributions in Fig. 10(d), although leaf-HFs-B have larger density contrasts. Therefore, leaf-HFs-C are not necessarily low-density structures. It is just that the density distribution in leaf-HFs-C structures is relatively uniform. From these results, one would argue that only when the structural scale is large enough, the density contrast can clearly manifest. In other words, the hub size of leaf-HFs-C could be too small to be resolved. Moreover, beam smearing effect may create an illusion of uniform density distribution. In order to check these possibilities, in Fig. 11, we fitted the correlation between density contrast and scale. Individually, for leaf-HFs-A and leaf-HFs-C, there is almost no correlation between density contrast and scale. When the two categories were combined, a correlation between density contrast and scale does appear. However, since leaf-HFs-A and leaf-HFs-C have significant scale overlap, if leaf-HFs-A can discern the hubs, leaf-HFs-C should be able to as well. Therefore, relative to leaf-HFs-A, the absence of the hubs in leaf-HFs-C is real and not merely a resolution issue.

Two vertical dashed lines in Fig. 11 mark the approximate range where leaf-HFs-A and leaf-HFs-C have significant scale overlap. We confined the comparison of leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C to this scale range and obtained results consistent with previous findings, as shown in Fig. A1. Therefore, the difference in scale is not a significant factor affecting the physical properties of different types of structures.

3.6 Star formation

In this section, we examined the physical properties of embedded star clusters within molecular clouds. Following the prescriptions described in Leroy et al. (2021b), the WISE 22 μm data can be used to calculate the local star formation rate (SFR) surface density via

$$\frac{\Sigma_{\rm SFR}}{\rm M_\odot\ yr^{-1}\ kpc^{-2}} = 3.8\times10^{-3}\left(\frac{I_{22\mu m}}{\rm MJy\ sr^{-1}}\right)\cos i, \quad (3)$$

where i is the inclination angle of the galaxy ($i = 0^\circ$ is face-on). In this work, we used *JWST* 21 μm data to estimate the SFR surface density.

Kruijssen & Longmore (2014) and Kruijssen et al. (2018) have developed a statistically robust method to convert the observed spatial separations between cold gas and SFR tracers into their underlying time-scales. This approach was employed by Kim et al. (2023) to study NGC 628 using emission maps of CO, *JWST* 21 μm, and H α emission maps which respectively trace molecular clouds, embedded star formation, and exposed star formation. This analysis yielded systematic constraints on the duration of the embedded phase of star

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

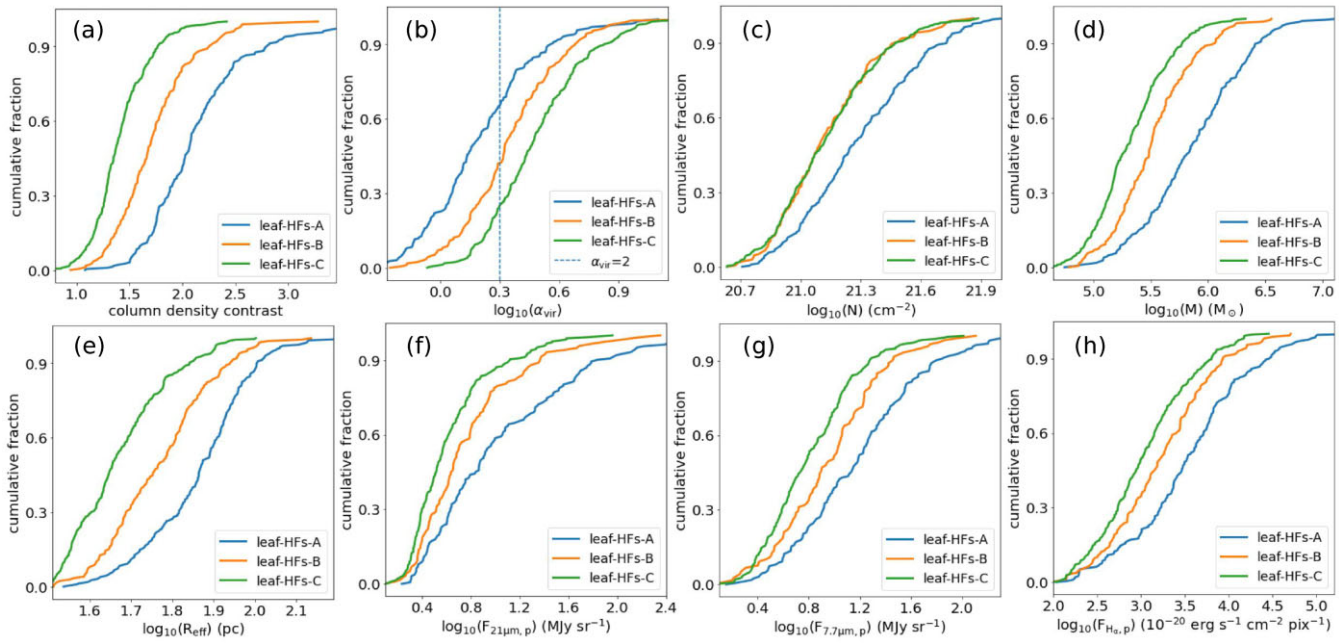


Figure 10. Physical parameters of three types of structures. (a) Column density contrast; (b) virial ratio. The vertical dashed line marks the position of $\alpha_{\rm vir} = 2$; (c) mass; (d) column density; (e) scale (the effective radius of the CO structures); (f) peak intensity of 7.7 μm emission; (g) peak intensity of 21 μm emission; and (h) peak intensity of H α emission.

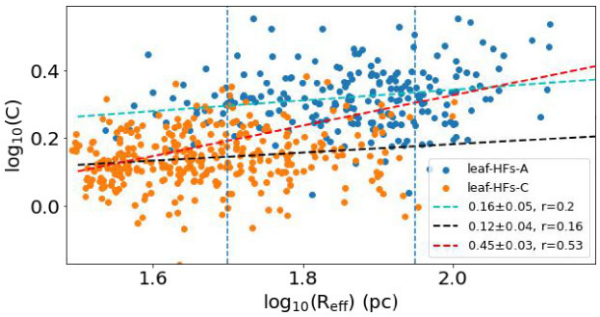


Figure 11. Correlation between density contrast and scale (the effective radius) for leaf-HFs-A and leaf-HFs-C. The black, cyan, and red dashed lines fit leaf-HFs-C, leaf-HFs-A, and both leaf-HFs-C and leaf-HFs-A, respectively. 'r' is the Pearson correlation coefficient. Two vertical dashed lines mark the range where leaf-HFs-A and leaf-HFs-C have significant scale overlap.

formation in NGC 628. They defined the duration of the embedded phase of star formation as the time during which CO and 21 μm emissions are found to be overlapping, $t_{\rm fb,21\mu m}$. In contrast, the 'heavily obscured phase' refers to the period when both CO and 21 μm emissions are present without associated H α emission, $t_{\rm obsc}$. Finally, they obtained $t_{\rm fb,21\mu m} \approx 5.1^{+2.7}_{-1.4}$ Myr and $t_{\rm obsc} \approx 2.3^{+2.7}_{-1.4}$ Myr. In this work, we directly take the typical values $t_{\rm fb,21\mu m} = 5.1$ Myr and $t_{\rm obsc} = 2.3$ Myr. Therefore, the estimate is quite rough, given the significant uncertainty regarding the duration time. For 185 CO structures associated with both 21 μm and H α emissions, the duration time should be $\sim 2.3-5.1$ Myr. The spatial separations d between the intensity peaks of CO and 21 μm structures calculated in Section 3.5.2 can be used to distribute the age of the embedded star clusters in each molecular cloud. The typical value of d is $\sim$ 0.64 arcsec, which is significantly larger than the typical positional uncertainty of each data set. Assuming $t_0 = 2.3$ Myr, the age of the embedded star cluster is

$$t = t_0 + \frac{t_{\rm fb,21\mu m} - t_{\rm obsc}}{d_{\rm max} - d_{\rm min}} \times d, \quad (4)$$

where $d_{\rm max}$ and $d_{\rm min}$ are the maximum and minimum spatial separations. By combining the local SFR of each molecular cloud, we estimated the total mass of the corresponding embedded star clusters. Fig. 12(a) shows the mass distribution of embedded star clusters. The cluster mass of leaf-HFs-A is much larger than those of leaf-HFs-B and leaf-HFs-C. The results here are quite consistent with the findings in Hassani et al. (2023). For the *JWST* 21 μm compact source population identified in Hassani et al. (2023), using spectral energy distribution (SED) fitting, they found that the 21 μm sources have SEDs consistent with stellar population masses of $10^2 < M_*/{\rm M}_\odot < 10^{4.5}$. They also found a range of mass–weighted ages spanning 2–25 Myr, though most objects are < 8 Myr.

In Fig. 12(b), the spatial separations between the intensity peaks of CO and 21 μm structures of leaf-HFs-A are also larger than leaf-HFs-C. Since the spatial separation can measure the evolutionary timescales of molecular clouds, this is further evidence supporting that leaf-HFs-A is more evolved than leaf-HFs-C. However, the sample size here is very limited. The numbers of leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C are 95, 44, and 46, respectively. More samples are necessary to quantitatively assess the correspondence between the separations and the evolutionary stages of molecular clouds.

4 DISCUSSION

4.1 Formation of hub-filament structures

The hub-filament morphology could be shaped by either gravitational collapse (Vázquez-Semadeni et al. 2019) or strong shocks due to turbulence (Padoan et al. 2020). In the second case, there is no anticipated consistency between the velocity gradients and the masses of the hubs, while in the first case, a relationship between the

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

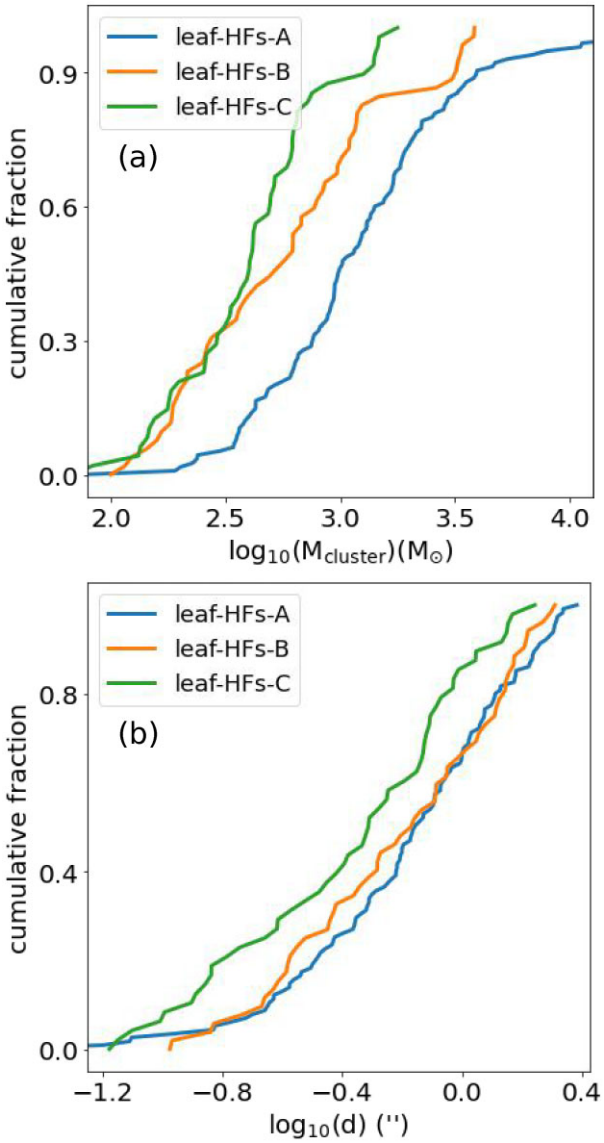


Figure 12. (a) The mass distribution of embedded star clusters; (b) the spatial separations between the intensity peaks of CO and associated 21 μm structures.

velocity gradients and the hub masses is naturally expected. Thus, the results presented in Section 3.4 clearly support the scenario in which the hub-filament structures form through gravitational contraction of gas structures on galaxy–cloud scales, which is also revealed in Zhou & Davis (2024). Similar findings on cloud–clump scales and clump–core scales are presented in Zhou et al. (2023), (2022), also see Motte et al. (2018), Vázquez-Semadeni et al. (2019), Kumar et al. (2020) and references therein. The kinematic characteristics of gas structures on galaxy–cloud scale are very similar to those on cloud–clump and clump–core scales. Therefore, a picture emerges where dense cores act like hubs within clumps, clumps act like hubs within molecular clouds, and clouds act like hubs in spiral galaxies. The interstellar medium from galaxy to dense core scales presents multiscale/hierarchical hub-filament structures. Gas structures at different scales in the galaxy may be organized into hierarchical systems through gravitational coupling. The hierarchical hub-filament structures are also present in the galaxy cluster (hub)–cosmic web (filament) picture, reflecting the self-similarity of the structural organization in the Universe. The fundamental reason may be that structures at different scales in the Universe primarily evolve under the influence of gravity.

Hubs are essentially structures with higher density compared to the surrounding more diffuse gas. As local gravitational centres, these hubs seem to accrete the surrounding diffuse gas, forming hub-filament structures. The filamentary structures seem to be associated with gas flows converging towards the hubs. A hub-filament structure essentially consists of a gravitational centre and the gas flow converging towards it. This could be considered as a fundamental structural type in the interstellar medium (ISM). In short, as long as there are local dense structures within the relatively diffuse ISM, hub-filament structures will inevitably form.

Finally, we note that although the measured velocity gradients support the gravitational collapse of gas structures on galaxy–cloud scales, the collapse is much slower than a pure free-fall gravitational collapse. In the free-fall, the velocity gradient ∇v and the scale R satisfy $\nabla v \propto R^{-1.5}$. However, Fig. 9(a) only gives $\nabla v \propto R^{-0.9}$. For NGC 4321 and NGC 5236, Zhou & Davis (2024) obtained a similar slope, i.e. $\nabla v \propto R^{-0.8}$. As discussed in Zhou & Davis (2024), the deviation from the free-fall may come from the measurement biases, the coupling with the galactic potential, the tidal forces from the galaxy or the neighbouring structures, or the turbulence and magnetic field supports (Chevance et al. 2023).

4.2 Evolutionary sequence of molecular clouds

The results of this work seem to suggest an evolutionary sequence among molecular structures. First, we focus solely on leaf-HFs-A. We compared two types of leaf-HFs-A, i.e. without and with 21 μm and Hα emissions. As shown in Fig. 13, the structures with 21 μm and Hα emissions present larger density contrasts, lower virial ratios, higher column densities, and larger scales (larger masses). In Figs 10 and A1, from leaf-HFs-C to leaf-HFs-A, we can see the same trend for each physical quantity. Leaf-HFs-A without and with 21 μm and Hα emissions represent structures from two different evolutionary stages. Leaf-HFs-C and leaf-HFs-A may have the same relation.

A typical characteristic of hub-filament structures is that the hub region has significantly higher density compared to the surrounding filamentary structures, i.e. the density contrast. The hub-filament morphology is shaped by gravitational collapse. Hub-filament structures only become apparent when gravitational collapse has progressed to a certain extent. In this work, the identified dense CO structures are potential hubs, which act as local gravitational centres. Leaf-HFs-A have significantly greater mass than C, so they serve as stronger gravitational centres. Leaf-HFs-A also present much better hub-filament morphology than leaf-HFs-C, which reflects the extent of gravitational collapse. Leaf-HFs-C have not yet exhibited a pronounced gravitational collapse process that would lead to a substantial density contrast. From leaf-HFs-C to leaf-HFs-A, gravitational collapse will increase the density and density contrast of the structure. Thus, the density contrast C effectively gauges the degree of gravitational collapse and the depth of the gravitational potential well within a structure, which decisively shape the hub-filament morphology.

Fig. 14 shows a clear correlation between the density contrast and the virial ratio. As expected, the density contrast decreases as the virial ratio increases. However, the classical virial analysis only considered the gravitational potential energy and internal kinetic energy may not completely reflect the true physical state of the identified structures. A more comprehensive virial analysis should include additional physical mechanisms, such as external pressure

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

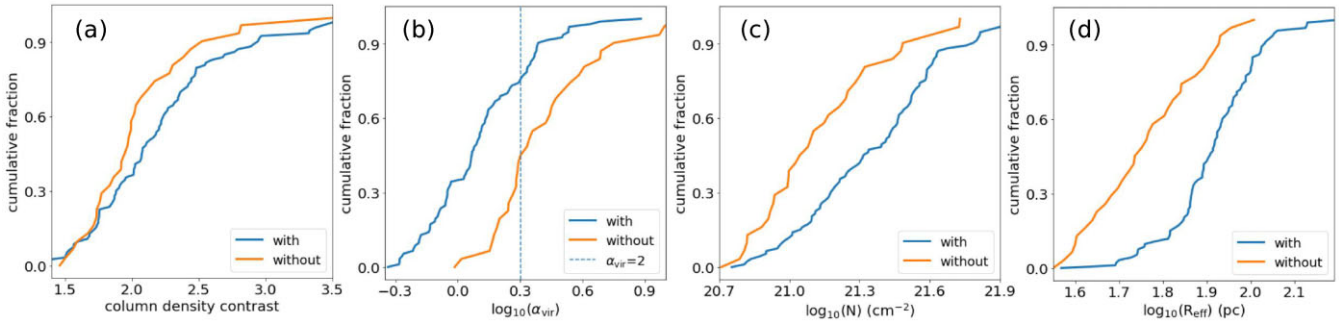


Figure 13. Comparison of leaf-HFs-A with and without 21 μm and Hα emission. (a) Column density contrast; (b) virial ratio. The vertical dashed line marks the position of $\alpha_{\rm vir} = 2$; (c) column density; (d) scale (the effective radius).

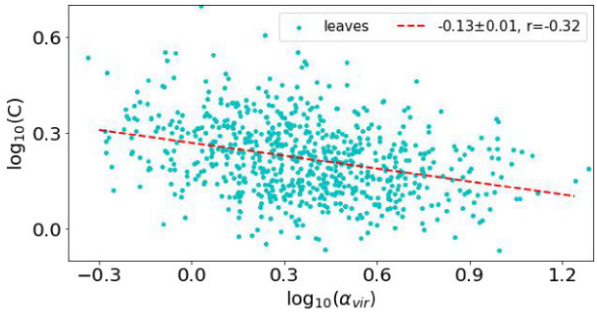


Figure 14. Correlation between the density contrast and the virial ratio for all leaf structures. 'r' is the Pearson correlation coefficient.

(Spitzer 1978; Elmegreen 1989; Kirk et al. 2017; Li et al. 2020; Sun et al. 2020; Zhou et al. 2024a), tidal forces (Ballesteros-Paredes et al. 2009a,b; Ramírez-Galeano et al. 2022; Li 2024; Zhou et al. 2024b) and magnetic fields (Dib et al. 2007; Crutcher et al. 2010; Li & Henning 2011; Crutcher 2012; Seifried et al. 2020; Ibáñez-Mejía, Mac Low & Klessen 2022; Ganguly et al. 2023). Local dense structures are embedded within larger gas environments. Their interactions with the surrounding environment may significantly impact their physical state. After calculating a more accurate virial ratio, the correlation in Fig. 14 should be stronger.

As the masses and scales of the structures also increase from leaf-HFs-C to leaf-HFs-A, the results of this work suggest that the structures inevitably accrete matter from their surrounding environment during the evolution. It is consistent with the analysis in Section 3.4 and the results presented in Zhou & Davis (2024) for NGC 5236 (M83) and NGC 4321 (M100). In these spiral galaxies, there are clear velocity gradients around intensity peaks, and the variations in velocity gradient across different scales suggest a gradual and consistent increase in velocity gradient from large to small scales, indicative of gravitational collapse and gas inflow at different scales. Local dense structures act as local gravitational centres, naturally accreting matter from the surrounding diffuse gas environment and thereby accumulating mass.

5 SUMMARY

We decomposed the spiral galaxy NGC 628 into multiscale hub-filament structures using the CO (2−1) line map. The main conclusions are as follows:

(1) The intensity peaks as potential hubs were identified based on the integrated intensity (Moment 0) map of CO (2−1) emission by the dendrogram algorithm. Around all intensity peaks, we extracted the average spectra for all structures to decompose their velocity components and fix their velocity ranges. The final identification of hub-filament structures is based on the cleaned Moment 0 map of each structure made by restricting to the fixed velocity range to exclude the potential overlap of uncorrelated velocity components. In practice, some steps might not be necessary, as almost all structures show only one velocity component.

(2) All leaf structures as potential hubs were classified into three categories, i.e. leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C. For leaf-HFs-C, the density contrast between the hub and the surrounding diffuse gas is not pronounced. Both leaf-HFs-A and leaf-HFs-B have clear central hubs. But leaf-HFs-A exhibit the best hub-filament morphology, which also have the highest density contrast, the largest mass and the lowest virial ratio.

(3) We employed the filfinder algorithm to identify and characterize filaments within 185 leaf-HFs-A structures using integrated intensity maps. We also fitted the velocity gradients around intensity peaks, a process performed after removing the global large-scale velocity gradients attributed to the galaxy's rotation. Measurements of velocity gradients provide evidence for gas inflow within these structures, which can also serve as kinematic evidence that these structures can be regarded as hub-filament structures.

(4) Leaf-HFs-C are not necessarily low-density structures. It is just that their density distribution is relatively uniform. There may be an evolutionary sequence from leaf-HFs-C to leaf-HFs-A. Currently, leaf-HFs-C lack a distinct gravitational collapse process that would result in significant density contrast. The numbers of the associated 21 μm and Hα structures and the peak intensities of 7.7 μm, 21 μm, and Hα emissions decrease from leaf-HFs-A to leaf-HFs-C. The spatial separations between the intensity peaks of CO and 21 μm structures of leaf-HFs-A are larger than leaf-HFs-C. These evidence indicate that leaf-HFs-A are more evolved than leaf-HFs-C.

(5) There is a clear correlation between the density contrast and the virial ratio, and the density contrast decreases as the virial ratio increases. The density contrast C effectively measures the extent of gravitational collapse and the depth of the gravitational potential well of the structure that shape the hub-filament morphology. In terms of reflecting the development and evolutionary stage of the structure, density contrast is more crucial than density itself.

(6) Combining the local SFR derived from the *JWST* 21 μm emission with the time-scale revealed by the spatial separation between CO and 21 μm emissions yields mass estimates of embedded star clusters in molecular clouds comparable to those obtained from SED fitting. As expected, the cluster mass of leaf-HFs-A is much larger than those of leaf-HFs-B and leaf-HFs-C.

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

(7) Combined with the kinematic evidence presented in Zhou & Davis (2024), a picture emerges where molecular gas in spiral galaxies is organized into a network of structures through gravitational coupling of multiscale hub-filament structures, consistent with the global hierarchical collapse scenario (Vázquez-Semadeni et al. 2019). Molecular clouds with sizes of hundreds of pc in NGC 628 are knots in networks of hub-filament systems, and are the local gravitational centres and the main star-forming sites.

ACKNOWLEDGEMENTS

We would like to thank the referee for the detailed comments and suggestions that significantly improve and clarify this work. It is a pleasure to thank the PHANGS team, the data cubes and other data products shared by the team make this work can be carried out easily. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00650.S and ADS/JAO.ALMA#2017.1.00886.L. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSTC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO, and NAOJ. Based on observations taken as part of the PHANGS–MUSE large programme (Emsellem et al. 2022). Based on data products created from observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme(s) 1100.B-0651, 095.C-0473, and 094.C-0623 (PHANGS-MUSE; PI Schinnerer), as well as 094.B-0321 (MAGNUM; PI Marconi), 099.B-0242, 0100.B-0116, 098.B-0551 (MAD; PI Carollo), and 097.B-0640 (TIMER; PI Gadotti). This research has made use of the services of the ESO Science Archive Facility. This work is based on observations made with the NASA/ESA/CSA *JWST*. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for *JWST*. These observations are associated with programme 2107.

DATA AVAILABILITY

All the data used in this work are available on the PHANGS team website.[5]

REFERENCES

Anand G. S. et al., 2021, MNRAS, 501, 3621
Ballesteros-Paredes J., Gómez G. C., Loinard L., Torres R. M., Pichardo B., 2009b, MNRAS, 395, L81
Ballesteros-Paredes J., Gómez G. C., Pichardo B., Vázquez-Semadeni E., 2009a, MNRAS, 393, 1563
Chevance M., Krumholz M. R., McLeod A. F., Ostriker E. C., Rosolowsky E. W., Sternberg A., 2023, in Inutsuka S., Aikawa Y., Muto T., Tomida K., Tamura M., eds, ASP Conf. Ser. Vol. 534, Protostars and Planets VII. Astron. Soc. Pac., San Francisco, p. 1
Crutcher R. M., 2012, ARA&A, 50, 29
Crutcher R. M., Wandelt B., Heiles C., Falgarone E., Troland T. H., 2010, ApJ, 725, 466
Davis T. A. et al., 2013, MNRAS, 429, 534
Dewangan L. K., Ojha D. K., Sharma S., Palacio S. d., Bhadari N. K., Das A., 2020, ApJ, 903, 13
Dib S., 2023, MNRAS, 524, 1625
Dib S., Helou G., Moore T. J. T., Urquhart J. S., Dariush A., 2012, ApJ, 758, 125
Dib S., Kim J., Vázquez-Semadeni E., Burkert A., Shadmehri M., 2007, ApJ, 661, 262
Elmegreen B. G., 1989, ApJ, 338, 178
Emsellem E. et al., 2022, A&A, 659, A191
Ganguly S., Walch S., Seifried D., Clarke S. D., Weis M., 2023, MNRAS, 525, 721
Hassani H. et al., 2023, ApJ, 944, L21
Henshaw J. D. et al., 2020, Nat. Astron., 4, 1064
Henshaw J. D., Caselli P., Fontani F., Jiménez-Serra I., Tan J. C., 2014, MNRAS, 440, 2860
Ibáñez-Mejía J. C., Mac Low M.-M., Klessen R. S., 2022, ApJ, 925, 196
Issac N., Tej A., Liu T., Varricatt W., Vig S., Ishwara Chandra C. H., Schultheis M., 2019, MNRAS, 485, 1775
Kim J. et al., 2023, ApJ, 944, L20
Kirk H. et al., 2017, ApJ, 846, 144
Kruijssen J. M. D., Longmore S. N., 2014, MNRAS, 439, 3239
Kruijssen J. M. D., Schruba A., Hygate A. P. S., Hu C.-Y., Haydon D. T., Longmore S. N., 2018, MNRAS, 479, 1866
Kumar M. S. N., Palmeirim P., Arzoumanian D., Inutsuka S. I., 2020, A&A, 642, A87
Lang P. et al., 2020, ApJ, 897, 122
Lee J. C. et al., 2023, ApJ, 944, L17
Leroy A. K. et al., 2021a, ApJS, 255, 19
Leroy A. K. et al., 2021b, ApJS, 257, 43
Leroy A. K. et al., 2022, ApJ, 927, 149
Li G.-X., 2024, MNRAS, 528, L52
Li H.-B., Henning T., 2011, Nature, 479, 499
Li S. et al., 2020, ApJ, 896, 110
Lindner R. R. et al., 2015, AJ, 149, 138
Liu H.-L. et al., 2022, MNRAS, 511, 4480
Liu T. et al., 2016, ApJ, 824, 31
Liu T. et al., 2020, MNRAS, 496, 2790
Lu X. et al., 2018, ApJ, 855, 9
McKee C. F., Ostriker E. C., 2007, ARA&A, 45, 565
Motte F., Bontemps S., Louvet F., 2018, ARA&A, 56, 41
Padoan P., Pan L., Juvela M., Haugbølle T., Nordlund Å., 2020, ApJ, 900, 82
Peretto N. et al., 2013, A&A, 555, A112
Ramírez-Galeano L., Ballesteros-Paredes J., Smith R. J., Camacho V., Zamora-Avilés M., 2022, MNRAS, 515, 2822
Riener M., Kainulainen J., Henshaw J. D., Orkisz J. H., Murray C. E., Beuther H., 2019, A&A, 628, A78
Rosolowsky E. W., Pineda J. E., Kauffmann J., Goodman A. A., 2008, ApJ, 679, 1338
Seifried D., Walch S., Weis M., Reissl S., Soler J. D., Klessen R. S., Joshi P. R., 2020, MNRAS, 497, 4196
Spitzer L., 1978, Physical Processes in the Interstellar Medium. Wiley, New York
Sun J. et al., 2020, ApJ, 892, 148
Vázquez-Semadeni E., Palau A., Ballesteros-Paredes J., Gómez G. C., Zamora-Avilés M., 2019, MNRAS, 490, 3061
Yuan J. et al., 2018, ApJ, 852, 12
Zhang Q., Wang K., Lu X., Jiménez-Serra I., 2015, ApJ, 804, 141
Zhou J. W. et al., 2023, A&A, 676, A69
Zhou J. W., Davis T. A., 2024, preprint (arXiv:2404.15862)
Zhou J. W., Dib S., Juvela M., Sanhueza P., Wyrowski F., Liu T., Menten K. M., 2024b, A&A, 686, A146
Zhou J. W., Wyrowski F., Neupane S., Barlach Christensen I., Menten K. M., Li S. H., Liu T., 2024a, A&A, 682, A128
Zhou J.-W. et al., 2022, MNRAS, 514, 6038

[5] https://sites.google.com/view/phangs/home

Downloaded from https://academic.oup.com/mnras/article/534/1/683/7754179 by guest on 15 April 2025

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **537,** 2630–2634 (2025) https://doi.org/10.1093/mnras/staf199
Advance Access publication 2025 February 3

Evolution of molecular clouds on galaxy-cloud scale revealed by gravitational network analysis: high-mass clouds may deplete nearby gas via accretion or merging

J. W. Zhou [1]★ and Guang-Xing Li[2]★

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*South-Western Institute For Astronomy Research, Yunnan University, Kunming 650600, Yunnan, P.R. China*

Accepted 2025 January 29. Received 2025 January 29; in original form 2024 December 3

ABSTRACT

Observations show that molecular gas in spiral galaxies is organized into a network of interconnected systems through the gravitational coupling of multiscale hub-filament structures. Building on this picture, we model molecular gas in the galaxy NGC 628 as a gravitational network, where molecular clouds are represented as nodes. Through analysing this network, we can characterize both the gravitational interactions and the physical properties of the clouds using geometry-based network metrics. A strong correlation is observed between the geometric and physical properties of the nodes (clouds). High-mass clouds tend to exhibit less clustering and greater average separations, suggesting that they generally have fewer neighbours. During their formation and evolution, high-mass clouds may deplete nearby gas via accretion or merging, leading to more isolated characteristics within the network. This aligns with observations showing a decrease in the virial ratio of molecular clouds as their mass increases. For clouds at different evolutionary stages, less evolved clouds with lower mass are typically found in tighter gravitational subnetworks, with closer proximity to neighbouring clouds. As a result, they are more prone to accretion or merging during evolution.

Key words: techniques: image processing – ISM: clouds – ISM: kinematics and dynamics – galaxies: ISM – galaxies: star formation – galaxies: structure.

1 INTRODUCTION

Giant molecular clouds (GMCs) are widely acknowledged as the primary reservoirs of gas that drive star formation, serving as the cradles for nearly all stars. Understanding the properties of GMCs is therefore essential for unravelling the intricate link between gas and star formation in galaxies (Kennicutt & Evans 2012; Schinnerer & Leroy 2024). Observations show that the physical characteristics of GMCs systematically change depending on their location within a galaxy, underscoring the significant interaction between these clouds and their surrounding environment, which influences their formation, structure, and evolution (Dib et al. 2012; Hughes et al. 2013; Colombo et al. 2014; Miville-Deschênes, Murray & Lee 2017; Kim et al. 2022; Pessa et al. 2022; Sun et al. 2022; Zhou, Dib & Davis 2024).

Kinematic studies reveal the existence of multiscale, hierarchical hub-filament structures (HFs) within molecular clouds and spiral galaxies (Peretto et al. 2013; Lu et al. 2018; Henshaw et al. 2020; Zhou et al. 2022, 2023; Zhou & Davis 2024). Their findings particularly highlight that intensity peaks, functioning as hubs, are linked to converging velocities, implying that the surrounding gas is flowing towards these dense regions. Additionally, the observed variations in the velocity gradient across different scales point to a gradual and consistent increase from larger to smaller scales, which is indicative of gravitational collapse. Especially, the velocity gradients on larger scales require a higher central mass to be maintained. Higher masses correspond to larger scales, suggesting that large-scale inflow is driven by these larger structures. This could be a consequence of the gravitational clustering of smaller scale structures, implying hierarchical gas structures and gas inflow from larger to smaller scales. Based on the kinematic evidence, Zhou et al. (2024) directly decomposed the molecular gas in the spiral galaxy NGC 628 into multiscale hub-filament structures using the CO (2–1) line map. The molecular clouds identified as potential hubs were categorized into three groups: leaf-HFs-A, leaf-HFs-B, and leaf-HFs-C. For leaf-HFs-C, the contrast in density between the hubs and the surrounding diffuse gas is not pronounced. Both leaf-HFs-A and leaf-HFs-B show distinct central hubs, but leaf-HFs-A exhibit the most well-defined hub-filament structure. Leaf-HFs-A have the highest density contrast, the largest mass, and the lowest virial ratio. It is possible that there is an evolutionary sequence from leaf-HFs-C to leaf-HFs-A. Currently, leaf-HFs-C do not show a clear gravitational collapse process that would lead to a high density contrast. These observations suggest that molecular gas in spiral galaxies is structured into a network of interconnected systems through the gravitational coupling of multiscale hub-filament structures. Molecular clouds act as nodes

★ E-mail: jwzhou@mpifr-bonn.mpg.de (JWZ); ligx.ngc7293@gmail.com (G-XL)

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/537/3/2630/7997691 by guest on 15 April 2025

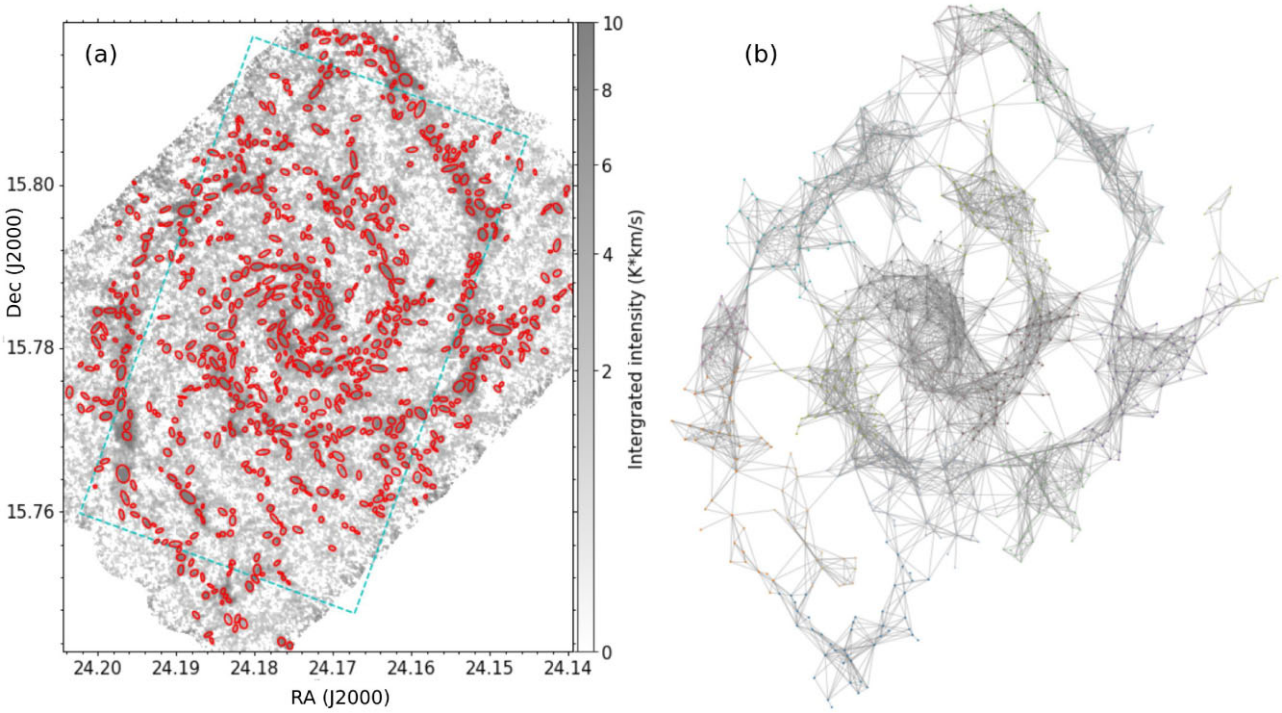


Figure 1. (a) Background is the broad-masked integrated intensity map of CO (2–1) emission. Red ellipses represent leaf structures identified by the dendrogram algorithm in Zhou et al. (2024). Cyan box marks the field of views of the *JWST* 21 μm observation. (b) The gravitational network constructed in Section 3.1, with molecular clouds as nodes.

within these hub-filament networks, serving as local gravitational centres and primary sites for star formation. In this work, we employ network-based analytical methods to gain further insights into the formation and evolution of molecular clouds in the spiral galaxy NGC 628.

2 DATA

In Zhou et al. (2024), we used the combined 12 m + 7 m + TP PHANGS-ALMA CO (2–1) data cube of the spiral galaxy NGC 628 to investigate gas kinematics and dynamics and identify hub-filament structures.The CO data have a spectral resolution of 2.5 km s^{-1} and an angular resolution ∼1.1 arcsec, corresponding to a linear resolution ∼50 pc at the distance of 9.8 Mpc (Anand et al. 2021; Leroy et al. 2021). All the data are available on the PHANGS team website.[1] The analysis in this work is based on the cloud catalogue constructed in Zhou et al. (2024).

3 RESULTS AND DISCUSSION

3.1 Gravitational network

Considering the picture of multiscale/hierarchical hub-filament structures in spiral galaxies revealed by observations, we can construct a gravitational network with molecular clouds as nodes all over the galaxy NGC 628. The elements of a network include nodes, edge weights, and connection rules. The edge weights are defined by the gravitational force between nodes, with greater weights indicating stronger gravitational interactions. The connection rules stipulate that nodes separated by a distance greater than 1 kpc are not connected, as this work focuses only on locally coupled structures. The distance is calculated in a 2D plane using Euclidean distance according to the central coordinates of nodes. We tested different values of the separation threshold and found that a threshold below 1 kpc (e.g. 500 pc) mainly results in discrete local networks, preventing the formation of the overall network shown in Fig. 1. A larger threshold (e.g. 1500 pc) leads to more connections, which eliminate locally coupled structures (local networks) but do not affect the correlations between node geometric parameters and node mass presented in the next section. Furthermore, a gravitational threshold is defined as the median gravitational force calculated for all node pairs with separations less than 1 kpc. Nodes are connected only if their gravitational force exceeds this threshold to outstand the main gravitational interactions between nodes. Fig. 1 presents the constructed network, revealing some localized cloud clusters.

3.2 Accretion or merger during the formation and evolution of molecular clouds

The physical properties of molecular clouds in NGC 628 have been calculated and discussed in Zhou et al. (2024). Now, with each cloud being considered a node, we can calculate the geometric parameters of each node and explore the correlation between geometric and physical properties of molecular clouds. In other words, since we have visualized the gravitational interactions between molecular clouds using the gravitational network, we can now characterize the gravitational interactions between the clouds or the physical properties of the clouds themselves using geometry-based network metrics.

In a network, the degree of a node (k_i) represents the number of other nodes directly connected to this node. It is a fundamental

[1] https://sites.google.com/view/phangs/home

Downloaded from https://academic.oup.com/mnras/article/537/3/2630/7997691 by guest on 15 April 2025

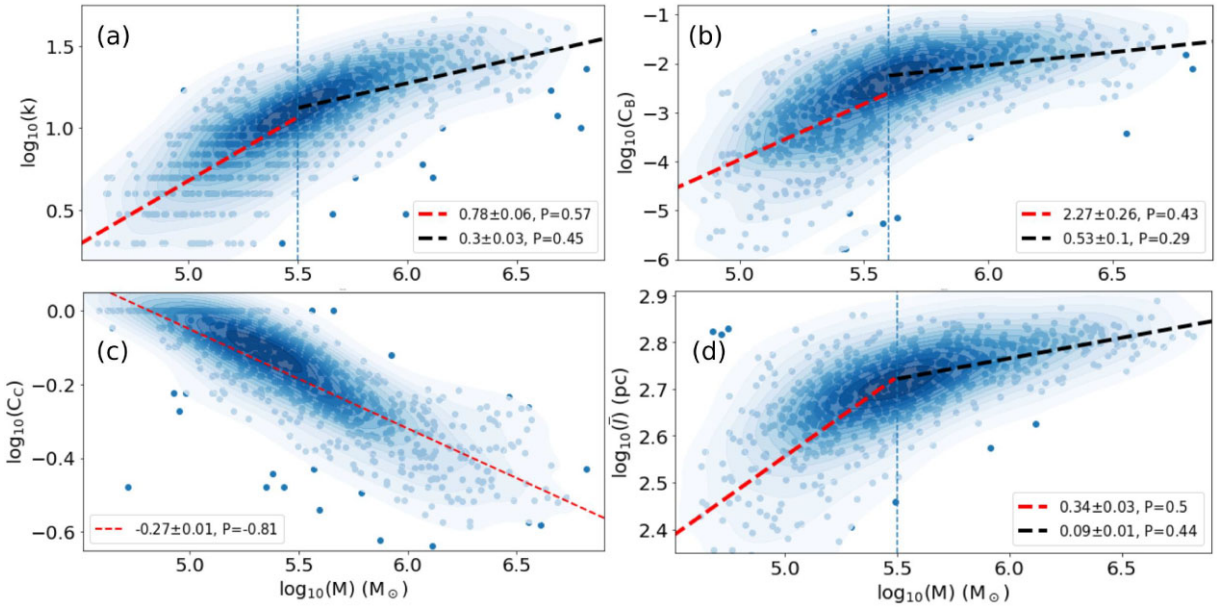


Figure 2. Correlation between the geometric parameters of nodes and node mass. (a) The degree (k). (b) The betweenness centrality ($C_{\rm B}$). (c) The clustering coefficient ($C_{\rm C}$). (d) The average physical path length ($\bar{l}$) between each node and its neighbours. The background contour in each panel shows the probability density (KDE) of the scatter points. Vertical dashed lines mark the inflection points of the segmented fit.

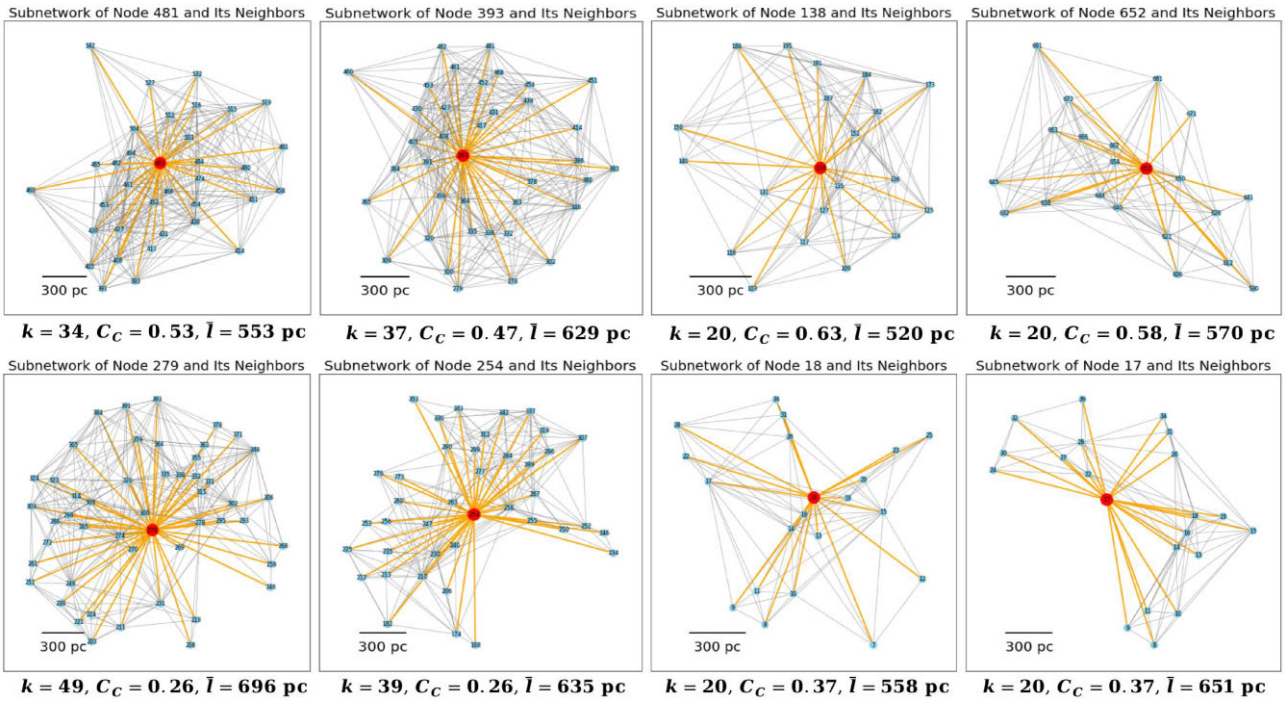


Figure 3. Some subnetworks and their geometric properties. Red dot and blue dots represent the node and its neighbours, respectively. k, $C_{\rm C}$, and $\bar{l}$ are the degree, the clustering coefficient, and the average physical path length between each node and its neighbours, respectively.

attribute for measuring the importance of a node. Particularly in a gravitational network, nodes with high degrees may represent massive objects that exert significant gravitational influence on their surroundings. Fig. 2(a) shows a strong positive correlation between the mass of the nodes and their degree.

The betweenness centrality of a node ($C_{{\rm B},i}$, see the definition of equation A1) measures the extent to which this node acts as an intermediary within the network. Nodes with high betweenness centrality have greater influence in the network. Massive clouds, serving as local hubs and gravitational centres, naturally exhibit greater influence, leading to higher values of $C_{\rm B}$, as shown in Fig. 2(b).

The clustering coefficient of a node ($C_{{\rm C},i}$, see the definition of equation A2) is an important metric for measuring the degree of connectivity among its neighbours. It reflects the tightness of the local network structure around a node. In a gravitational network,

Downloaded from https://academic.oup.com/mnras/article/537/3/2630/7997691 by guest on 15 April 2025

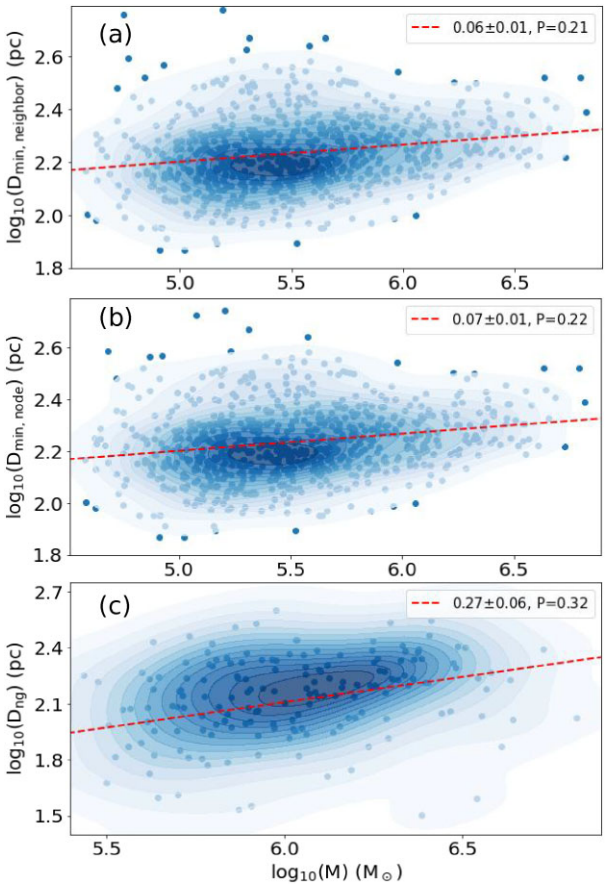


Figure 4. (a) Correlation between node mass and the distance to its nearest neighbour ($D_{\rm min,neighbour}$). (b) Correlation between node mass and the distance to its nearest node ($D_{\rm min,node}$). (c) Correlation between node mass and the displacement of the node from the geometric centre of the subnetwork comprising the node and its neighbours ($D_{\rm ng}$). To ensure more precise calculations of the subnetwork's geometric centre, we only consider nodes with a degree $\geq$20. The background contour in each panel shows the probability density (KDE) of the scatter points.

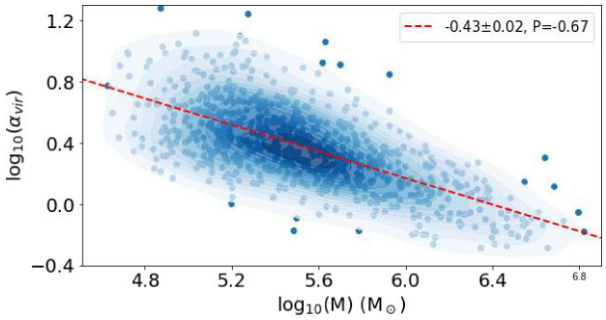


Figure 5. Correlation between the virial ratio and mass of molecular clouds. The background contour shows the probability density (KDE) of the scatter points.

a high clustering coefficient indicates that the neighbours of a node have stronger gravitational relationships with one another, forming a highly interconnected subnetwork. Surprisingly, as shown in Fig. 2(c), high-mass molecular clouds exhibit a lower clustering coefficient. We also calculate the average physical path length ($\bar{l}_i$) between each node and its neighbours (the average length of all

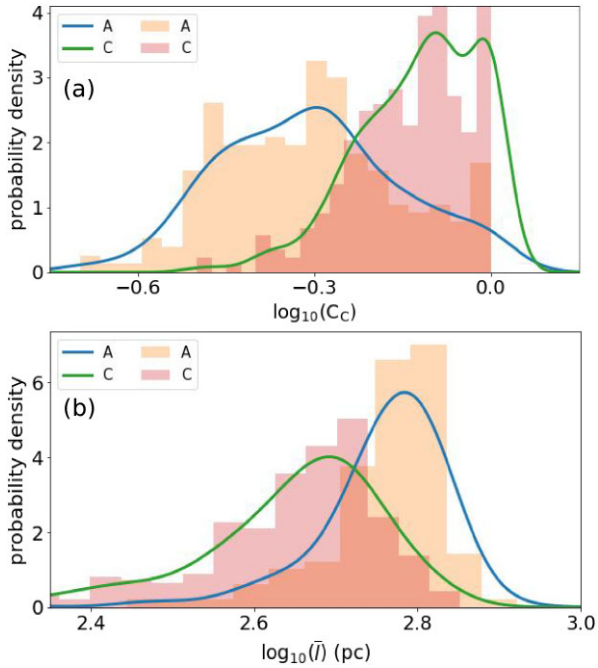


Figure 6. Comparison of the geometric parameters of A-type and C-type clouds. (a) The clustering coefficient ($C_{\rm C}$). (b) The average physical path length between each node and its neighbours ($\bar{l}$).

orange lines for each subnetwork in Fig. 3). In Fig. 2(d), $\bar{l}$ increases with the node's mass, which means an anticorrelation between $\bar{l}$ and $C_{\rm C}$. This could be due to high-mass nodes connecting to more distant nodes, which increases the average distance to their neighbours. Alternatively, it might be that there are simply fewer nodes near high-mass nodes, which also accounts for the lower $C_{\rm C}$ observed in these high-mass nodes.

Fig. 4(a) shows the relationship between node mass and the distance to its nearest neighbour ($D_{\rm min}$), demonstrating a positive correlation between node mass and its distance from the nearest neighbour. If we relax the requirement and only consider the shortest distance between two nodes, without requiring them to be neighbours, we observe a similar correlation between node mass and distance (Fig. 4b). In Fig. 4(c), for subnetworks formed by a node and its neighbours (degree $\geq$20), nodes with greater mass and longer average path lengths tend to show a more significant displacement from the geometric centre of the subnetwork ($D_{\rm ng}$). Fig. 3 displays some subnetworks and their geometric properties.

The above results suggest that high-mass clouds, during their formation and evolution, deplete nearby gas through accretion or merger, resulting in more isolated network characteristics. Molecular clouds are inherently irregular gas structures, so describing them as being merged or accreted essentially conveys a similar idea. Initially, molecular clouds may exhibit a relatively uniform distribution. Over time, however, molecular gas accumulates towards the gravitational centres, and the high-mass clouds that form eventually become decoupled from their surrounding environment, as these clouds experience stronger self-gravity. As shown in Fig. 5, the virial ratio of molecular clouds decreases as the molecular cloud mass increases.

To further verify the above scenario, now we focus on two different types of molecular clouds classified in Zhou et al. (2024), i.e. leaf-HFs-A and leaf-HFs-C (A-type and C-type). According to the evolutionary sequence proposed in Zhou et al. (2024), A-type clouds are more evolved than C-type clouds. As shown in Fig. 6, C-type

Downloaded from https://academic.oup.com/mnras/article/537/3/2630/7997691 by guest on 15 April 2025

clouds as nodes have higher C_C and smaller $\bar{l}$. This indicates that, compared to A-type clouds, C-type clouds are generally located in tighter gravitational subnetworks, with shorter distances to other molecular clouds, which makes them more prone to accretion or merging as they evolve. As shown in Zhou et al. (2024), the more evolved A-type clouds have significantly larger mass and lower virial ratio than C-type clouds. In Fig. 2(d), for the relation $M \propto (\bar{l})^{\alpha}$, low-mass clouds exhibit a significantly steeper slope ($\alpha \approx 0.34$) compared to high-mass clouds ($\alpha \approx 0.1$), suggesting that low-mass clouds are more susceptible to accretion or merging during evolution, consistent with the above results.

4 CONCLUSION

Observations reveal that molecular gas in spiral galaxies forms a network of interconnected systems through the gravitational coupling of multiscale hub-filament structures. Within these networks, molecular clouds serve as nodes, acting as local gravitational centres and primary sites for star formation. Based on this physical picture, we construct a gravitational network for the galaxy NGC 628, where molecular clouds are treated as nodes. By visualizing the gravitational interactions between molecular clouds in this network, we can characterize both the gravitational interactions and the physical properties of the clouds using geometry-based network metrics. In the gravitational network, there is a strong correlation between the geometric and physical properties of the nodes (clouds). Metrics such as node degree (k), betweenness centrality (C_B), clustering coefficient (C_C), and average physical path length ($\bar{l}$) show strong correlations with node mass. High-mass molecular clouds exhibit lower C_C and longer $\bar{l}$, indicating that high-mass nodes tend to have fewer neighbours. Additionally, there is a positive correlation between node mass and the distance to the nearest neighbour (D_{min}). For subnetworks consisting of a node and its neighbours (node degree $\geq$20), nodes with greater mass and longer $\bar{l}$ are more significantly displaced from the geometric centre of the subnetwork (D_{ng}).

These findings suggest that high-mass clouds, during their formation and evolution, deplete nearby gas through accretion or merger, resulting in more isolated network characteristics. This is consistent with the observed decrease in the virial ratio of molecular clouds as their mass increases. For molecular clouds at different evolutionary stages classified in Zhou et al. (2024) (i.e. A-type and C-type), a comparison reveals that C-type clouds, which are less evolved, have higher C_C and shorter $\bar{l}$ compared to A-type clouds. This indicates that C-type clouds are generally located in tighter gravitational subnetworks, with closer proximity to other molecular clouds. Consequently, they are more prone to accretion or merging during evolution. For the relation $M \propto (\bar{l})^{\alpha}$, the slope for low-mass clouds ($\alpha \approx 0.34$) is significantly steeper than that for high-mass clouds ($\alpha \approx 0.1$). This also indicates that low-mass clouds are more susceptible to accretion or merging as they evolve, as the more evolved A-type clouds possess significantly greater mass and a lower virial ratio compared to C-type clouds.

DATA AVAILABILITY

All the data used in this work are available on the PHANGS team website.[2]

[2]https://sites.google.com/view/phangs/home

REFERENCES

Anand G. S. et al., 2021, MNRAS, 501, 3621
Colombo D. et al., 2014, ApJ, 784, 3
Dib S., Helou G., Moore T. J. T., Urquhart J. S., Dariush A., 2012, ApJ, 758, 125
Henshaw J. D. et al., 2020, Nat. Astron., 4, 1064
Hughes A. et al., 2013, ApJ, 779, 46
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Kim J. et al., 2022, MNRAS, 516, 3006
Leroy A. K. et al., 2021, ApJS, 257, 43
Lu X. et al., 2018, ApJ, 855, 9
Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, ApJ, 834, 57
Peretto N. et al., 2013, A&A, 555, A112
Pessa I. et al., 2022, A&A, 663, A61
Schinnerer E., Leroy A. K., 2024, ARA&A, 62, 369
Sun J. et al., 2022, AJ, 164, 43
Zhou J.-W., Davis T., 2024, Publ. Astron. Soc. Aust., 41, e076
Zhou J.-W. et al., 2022, MNRAS, 514, 6038
Zhou J. W. et al., 2023, A&A, 676, A69
Zhou J. W., Dib S., Davis T. A., 2024, MNRAS, 534, 683

APPENDIX A: DEFINITION OF GEOMETRIC PARAMETERS

Betweenness centrality measures the extent to which a node acts as an intermediary or 'broker' within a network. It quantifies how often a node lies on the shortest path between two other nodes, reflecting its role in connecting different parts of the network. Formally, the betweenness centrality $C_{B,i}$ of a node i is defined as

$$C_{B,i} = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}, \tag{A1}$$

where σ_{st} is the total number of shortest paths between nodes s and t. $\sigma_{st}(i)$ is the number of shortest paths between nodes s and t that pass through node i. The sum is taken over all pairs of nodes s and t that are not equal to i. A node with high betweenness centrality plays a crucial role in connecting different parts of the network, acting as a bridge between otherwise disconnected nodes or communities. Nodes with high betweenness centrality often have significant influence because they control the flow of information or resources across the network.

For a node i, the clustering coefficient is defined as the ratio of the actual number of edges between its neighbours to the total possible number of edges between them. It quantifies the tightness of the local network around the node. Specifically, the clustering coefficient $C_{C,i}$ of node i can be calculated using the following formula:

$$C_{C,i} = \frac{2e_i}{k_i(k_i - 1)}, \tag{A2}$$

where e_i is the actual number of edges between the neighbours of node i (i.e. the number of edges that exist between the neighbours). k_i is the degree of node i, i.e. the number of neighbours of node i. If the clustering coefficient $C_{C,i}$ is close to 1, it indicates that the neighbours of node i are nearly all connected to each other, meaning that the local network is very tight. If $C_{C,i}$ is close to 0, it indicates that there are almost no connections between the neighbours of node i, meaning that the local network is sparse.

This paper has been typeset from a TEX/LATEX file prepared by the author.

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/537/3/2630/7997691 by guest on 15 April 2025

Original papers of Chapter.3

The original paper of Sec.3.2,

> Zhou, J. W., Dib, S., Juvela, M., et al. 2024, A&A, 686, A146
> *Gas inflows from cloud to core scales in G332.83-0.55: Hierarchical hub-filament structures and tide-regulated gravitational collapse*

The original paper of Sec.3.3,

> Zhou, J. W. 2025, A&A, 696, L5
> *External tides: an important driver of velocity dispersion in molecular clouds*

The original paper of Sec.3.4,

> Zhou, J. W., & Dib, S. 2025, MNRAS, 537, 2232
> *An examination of large-scale galactic effects on molecular cloud properties in NGC 628: the significant impact of tidal effects from neighbouring material on the evolution of molecular clouds*

A&A, 696, L5 (2025)
https://doi.org/10.1051/0004-6361/202453040
© The Authors 2025

Astronomy & Astrophysics

LETTER TO THE EDITOR

External tides: An important driver of velocity dispersion in molecular clouds

J. W. Zhou*

Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

Received 17 November 2024 / Accepted 20 March 2025

ABSTRACT

Using the 3D density distribution derived from the 3D dust map of the solar neighborhood, the gravitational potential is obtained by solving the Poisson equation and the tidal tensor can then be computed. In optimal decomposition, the external tidal tensor follows the same formalism as that of a point mass. The average tidal strength of the clouds, derived from both tidal tensor analysis and pixel-by-pixel computation, shows consistent results. The equivalent velocity dispersion of the clouds, estimated from the average tidal strength, is comparable in magnitude to the velocity dispersion measured from CO (1–0) line emission. This suggests that tidal effects from surrounding material may play a significant role in driving velocity dispersion within the clouds. Future studies should carefully consider these tidal effects in star-forming regions.

Key words. ISM: clouds – dust, extinction – ISM: kinematics and dynamics – ISM: structure

1. Introduction

One prevailing paradigm about star formation is that the collapse of gas is driven by gravity and balanced by processes such as turbulence and magnetic fields (McKee & Ostriker 2007). Based on this picture, the virial parameter (Bertoldi & McKee 1992) is always used to determine whether a structure is in a state of gravitational collapse. However, as pointed out in Ramírez-Galeano et al. (2022), Li (2024), studies usually focus on the role of gravity in individual, isolated parts and thus neglect long-range gravitational interactions. Diverse manifestations of gravitational effects on gas within molecular clouds were unveiled in Li (2024). Gravity influences cloud evolution in various ways. Within dense regions, it facilitates fragmentation and collapse. Outside these regions, it can suppress the collapse of low-density gas through extensive tidal forces and drive accretion.

Star-forming regions typically exhibit hierarchical gas structures. High-mass stars preferentially form in the density enhanced hubs of hub-filament systems (Myers 2009; Schneider et al. 2012; Motte et al. 2018; Kumar et al. 2020; Zhou et al. 2022). The hub can drive longitudinal gas inflows along filaments, providing further mass accretion (Peretto et al. 2013; Henshaw et al. 2014; Zhang et al. 2015; Yuan et al. 2018; Lu et al. 2018; Issac et al. 2019; Dewangan et al. 2020; Liu et al. 2016, 2021b, 2022, 2023; Kumar et al. 2020; Zhou et al. 2022; Xu et al. 2023; Zhou et al. 2023; Yang et al. 2023; Zhou et al. 2024b). Zhou et al. (2022) examined the physical properties and evolution of hub-filament systems in roughly 140 clumps using spectral line data from the ALMA Three-millimeter Observations of Massive Star-forming regions (ATOMS) survey (Liu et al. 2020). We proposed that hub-filament structures, characterized by self-similarity and filamentary accretion, are consistent across different scales in high-mass star-forming regions, spanning from a few thousand astronomical units to several parsecs. This picture of hierarchical, multi-scale hub-filament structures was expanded from the clump-core scale to the cloud-clump scale by Zhou et al. (2023) and later extended to the galaxy-cloud scale in Zhou & Davis (2024), Zhou et al. (2024a), Zhou & Li (2025). Previous works have also demonstrated hierarchical collapse and the feeding of central regions by hub-filament structures, including Motte et al. (2018), Vázquez-Semadeni et al. (2019), Kumar et al. (2020), and references therein.

Gravity is a long-range force. A local gas structure evolves under its self-gravity, but as a gravitational center, its influence can also affect neighboring structures. At the same time, it also experiences external gravity from neighboring material. Tidal and gravitational fields are mutually interdependent. Tidal forces have been proposed in previous studies as a factor that can either regulate or initiate star formation. Ballesteros-Paredes et al. (2009b,a) examined the effects of tidal forces induced by the Galactic potential on molecular clouds and demonstrated that these forces can either compress or disrupt the clouds, thereby impacting star-formation efficiency. For M 51 and NGC 4429, the models of Meidt et al. (2018), Liu et al. (2021a) investigated the influence of the host galaxy potential on molecular clouds, showing that cloud-scale gas dynamics result from the interplay between the galactic potential and gas self-gravity, which plays a key role in shaping molecular cloud properties. However, for the Large Magellanic Cloud (LMC), Thilliez et al. (2014) found that tidal instability does not hinder star formation. Furthermore, studies by Dib et al. (2012), Thilliez et al. (2014), Zhou & Dib (2025) have indicated that for molecular clouds in the Milky Way, the LMC, and NGC 628, the shear derived from the galactic rotation curve is negligible. Ramírez-Galeano et al. (2022) also suggest that tidal stresses from nearby molecular cloud com-

* Corresponding author; jwzhou@mpifr-bonn.mpg.de

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

plexes contribute more to interstellar turbulence than the overall galactic potential. The hierarchical, multi-scale hub-filament structures indicate that there are extensive tidal interactions in the interstellar medium (ISM). Regardless of whether the galactic potential influences molecular clouds and their complexes, these clouds are inevitably impacted by the cumulative tidal interactions with surrounding material. Such interactions may hinder gravitational collapse, suppress instability growth, and inhibit star formation within the cloud. Zhou & Dib (2025) carried out an examination of large-scale galactic effects on molecular cloud properties in NGC 628 and found a significant impact of tidal effects from neighboring material on the evolution of molecular clouds. The tidal effects from neighboring material may also be a significant contributing factor to the slowing down of a pure free-fall gravitational collapse for gas structures on galaxy-cloud scales revealed by velocity gradient measurements (Zhou & Davis 2024; Zhou et al. 2024a). In Zhou et al. (2024b), it was shown that the deformation due to the external tides can also effectively slow down the pure free-fall gravitational collapse of gas structures on clump-core scales. These mechanisms could be called "tide-regulated gravitational collapse."

Due to the diffuse and complex morphology of the ISM, a complete tidal calculation would be complex. One should derive the gravitational potential distribution from the 3D density distribution and then calculate the tidal field according to the gravitational potential, as presented in Li (2024). In observations, one can usually only obtain the 2D projected surface density map. However, building on the accurate distances enabled in the *Gaia* era, we have been able to infer the 3D density distribution of the ISM by the 3D dust mapping technique (Rezaei et al. 2018; Chen et al. 2019; Green et al. 2019; Lallement et al. 2019; Hottier et al. 2020; Leike et al. 2020; Edenhofer et al. 2024). In this work, we first calculate and decompose the tidal field using the 3D dust map and then explore methods to estimate the tidal strength within the clouds induced by external gravity. Finally, we assess the impact of external tides on the physical properties of molecular clouds.

2. Data

2.1. 3D dust map

Edenhofer et al. (2024) offers the most accurate 3D dust map to date for the solar neighborhood within 1.25 kpc, employing a novel Gaussian process prior approach to reduce the impact of fingers-of-god artifacts. This 3D map was generated using stellar distance and extinction estimates from Zhang et al. (2023) derived from *Gaia* spectra. There are 12 samples from their inferred 3D dust extinction distribution. In this work, our calculation is based on the posterior mean of their reconstruction, which is the average of these 12 samples. Using the method described in O'Neill et al. (2024), we converted the differential extinction in the 3D map to the number density of hydrogen nuclei, n, independent of phase ($n = n_{\rm H_I} + 2n_{\rm H_2}$). The volume density is then $\rho = 1.37 m_{\rm p} n$, where $m_{\rm p}$ represents the mass of a proton and the factor 1.37 accounts for the contribution of helium to the total mass based on cosmic abundance ratios relative to hydrogen.

2.2. CO (1–0) line emission

We also utilized the full position–position–velocity (l–b–v) cube of CO[1] from Dame et al. (2001). The spatial range of this cube is all Galactic longitudes and $\pm 30^0$ Galactic latitude. The velocity range is $\pm 320\,{\rm km\,s^{-1}}$.

3. Results

3.1. Structural identification

Fig. 1a shows the surface density map of hydrogen nuclei on the l–b plane. We applied the dendrogram algorithm to identify the local dense structures based on this surface density map. Then the CO (1–0) cube was utilized to derive the velocity dispersion of the identified structures. Since the CO (1–0) cube is in the Galactic coordinates (l–b–v), we also used the 3D dust map provided in the Galactic coordinate system (l–b–d)[2]. This allowed both cubes to be analyzed in the same way, as they have the same data structure.

As described in Rosolowsky et al. (2008), the dendrogram algorithm decomposes density data into hierarchical structures. Using the *astrodendro* package[3], there are three major input parameters for the dendrogram algorithm: *min_value* for the minimum value to be considered in the dataset, *min_delta* for a leaf that can be considered as an independent entity, and *min_npix* for the minimum area of a structure. In this work, we took the values of $min_value = 5 * \Sigma_{\rm rms}$ and $min_delta = 5 * \Sigma_{\rm rms}$, where $\Sigma_{\rm rms}$ is the average density of the background on the map shown in Fig. 1a. After testing different values, we set $min_npix = 100$ pixels to ensure that the identified structures are not too small. As shown in Fig. 1a, the structures identified by the dendrogram correspond well to the surface density distribution. In Fig. 1b, CO and dust also show a good correspondence in their distribution, but CO emission is primarily concentrated in the Galactic plane.

Using the mask of each identified structure in Fig. 1a, we extracted the average "spectra" according to the l–b–v and l–b–d cubes. Specifically, for each identified structure, we averaged the intensity or density values across all pixels within the masked region and analyzed the intensity distribution along the velocity axis and the density distribution along the distance axis, as shown in Fig. 2. This approach allowed us to investigate the characteristic velocity and distance distributions of the structures, thus providing insights into their kinematics and spatial properties. If an identified structure is a single structure without overlap, we would expect both types of "spectra" to exhibit a single peak, as shown in Fig. 2. In total, 90 structures satisfy this condition. From the "spectra," the systematic velocity, the velocity dispersion ($\sigma_{\rm g}$), the distance, and the thickness of each structure can be obtained. The thickness is the full width at half maximum of the line profile, displayed in Fig. 2a. The dendrogram algorithm approximates the morphology of each structure as an ellipse. Using the long and short axes (l_1 and l_2) of the effective ellipse and the half-thickness (l_3), we estimated a 3D effective radius for each structure, $R \approx (l_1 * l_2 * l_3)^{1/3}$.

3.2. Tidal strength

In Appendix A, for a subregion of the 3D map, based on the 3D density distribution, the gravitational potential was computed by solving the Poisson equation. The tidal tensor was then derived from the gravitational potential. For a cloud, its tidal tensor can be divided into two parts (internal and external) due to the mat-

[1] https://lweb.cfa.harvard.edu/rtdc/CO/

[2] https://zenodo.org/records/10658339

[3] https://dendrograms.readthedocs.io/en/stable/index.html

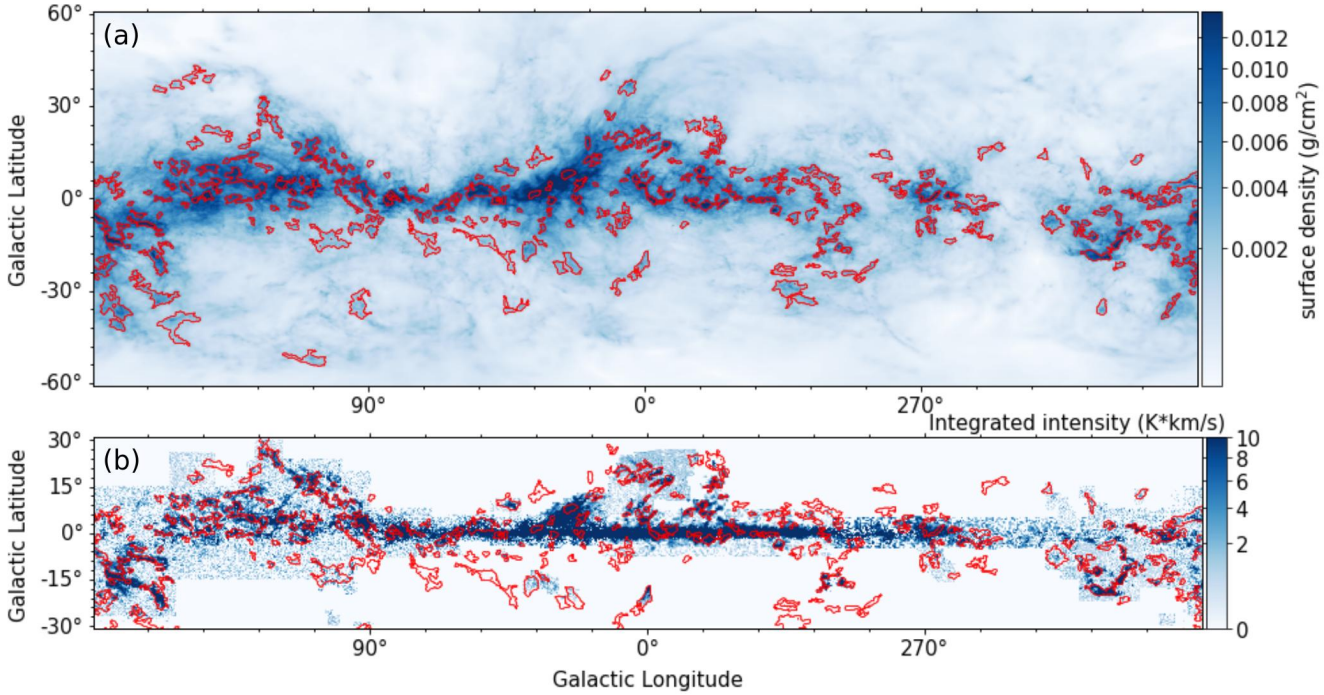


Fig. 1. Structural identification. (a) Surface density of hydrogen nuclei on the l–b plane. Red contours are the structures identified by the dendrogram algorithm. (b) Integrated intensity map of CO (1−0) line emission. Red contours are the same as in panel (a).

ter inside and outside the cloud. Different decomposition methods resulted in four possible maximum eigenvalues ($|\lambda_{\rm ext,max}|$) of the external tidal tensor ($\boldsymbol{T}_{\rm ext}$). An independent and original pixel-by-pixel computation method was used to select the optimal $|\lambda_{\rm ext,max}|$. For the eigenvalues, we found that the optimal $\boldsymbol{T}_{\rm ext}$ shows the same formalism with the tidal tensor of a point mass. Therefore, for the external tides exerted by external material at a point within a cloud, the external material can be approximated as a collection of point masses. Each point mass generates its own tidal field. The combined effect of the external tides is equivalent to the superposition of the tidal fields produced by all these point masses. Essentially, this approach is analogous to the pixel-by-pixel computation. As shown in Fig. A.4, the average tidal strength in the clouds calculated by the two methods shows comparable results.

Our current computational resources are insufficient to directly calculate the tidal tensor of the entire 3D map. Therefore, we adopted the pixel-by-pixel computation approach to calculate the average tidal strength ($\lambda_{\rm pp}$) of all structures in the subsequent analysis. This method only considers the structure being calculated and its neighboring matter each time. Typically, accounting for matter within eight to ten times the structure's radius is sufficient, as tidal strength decreases rapidly with distance (Equation (A.10)).

3.3. Velocity dispersion

We estimated an equivalent velocity dispersion for the clouds according to the average tidal strength, as done in Zhou et al. (2024b). The tidal strength defined in this work is the tidal acceleration per unit distance (Stark & Blitz 1978). Thus, for a cloud with radius R and mass M, $\lambda_{\rm pp} * R$ is the tidal acceleration. The velocity dispersion of a gas structure measures the intensity of internal gas motion within the structure. The tides induced by external gravity pervade the interior of the gas structure. We propose that these ubiquitous gravitational differences serve as the driving force for the motion of gas parcels within the structure. The work done by tidal forces is converted into kinetic energy, which corresponds to the equivalent velocity dispersion, $\sigma_{\rm t}$:

$$\frac{1}{2}M\sigma_t^2 \approx \int_0^R M\lambda_{\rm pp}R\,{\rm d}R. \quad (1)$$

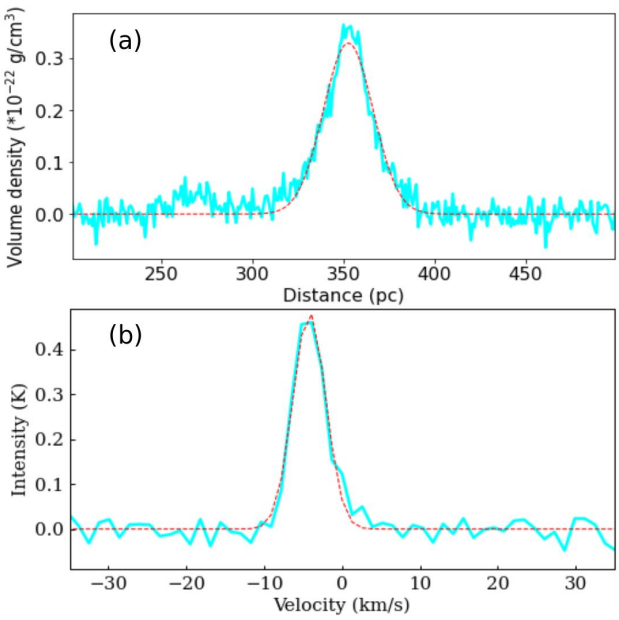


Fig. 2. Average "spectra" (cyan lines) of a single structure extracted from the l–b–v and l–b–d cubes, respectively. (a) Averaged intensity distribution along the velocity axis. (b) Averaged density distribution along the distance axis. Red dashed lines represent the Gaussian fitting.

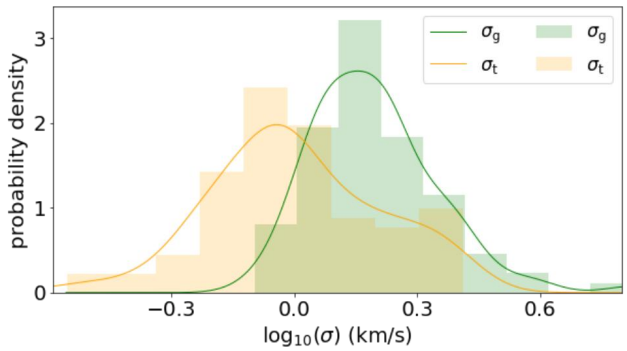


Fig. 3. Kernel density estimation and the histogram of the velocity dispersion of the clouds. The term σ_t is the equivalent velocity dispersion derived from the average tidal strength, and σ_g is the velocity dispersion measured from CO emission.

We then have

$$\sigma_t \approx \sqrt{\lambda_{pp} * R^2}. \quad (2)$$

Since σ_t is derived from the average tidal strength over the entire structure, it is isotropic and can be directly compared to the observed velocity dispersion of the structure obtained in Sec. 3.1. As shown in Fig. 3, σ_t is roughly comparable with σ_g in magnitude, which means that tidal effects from external gravity can significantly contribute to the velocity dispersion of the clouds. A relatively smaller value of σ_t is also reasonable, as other physical processes (such as gravitational collapse and feedback activities) also contribute to the overall velocity dispersion of the cloud.

4. Discussion and conclusions

The 3D density distribution derived from the 3D dust map is crucial for calculating the tidal field. However, velocity information is only available from the projected CO (1−0) line emission. To address this, we first projected the 3D density cube onto the l–b plane and then applied the dendrogram algorithm to identify the clouds. As shown in Fig. 1, the distributions of projected CO and dust emission exhibit a good correspondence. For 90 non-overlapping clouds, we fit the systematic velocity, velocity dispersion, distance, and thickness based on the average "spectra" extracted from the l–b–v and l–b–d cubes. These clouds were then used to study tidal effects. We employed two methods to estimate tidal strength, tidal tensor analysis and pixel-by-pixel computation, which produced comparable results. The equivalent velocity dispersion, derived from the average tidal strength, is roughly consistent in magnitude with the observed velocity dispersion, suggesting that tidal effects from external gravity play a significant role in shaping the velocity dispersion of the clouds. The significant impact of external gravity from the neighboring material on the physical properties of molecular clouds has also been demonstrated in Ramírez-Galeano et al. (2022), Li (2024), Zhou et al. (2024b), Zhou & Dib (2025). The extensive tides can produce turbulence, suppress fragmentation, and slow down gravitational collapse. Extended gas structures in the ISM are sensitive to the tidal effects exerted by surrounding material. Future studies should incorporate these tidal effects into star-formation theories. Star formation in molecular clouds typically occurs within locally dense clumps. It is essential to study these clumps in the context of their surrounding environment rather than in isolation, as their interactions with the surrounding material cannot be ignored.

As shown in Appendix A, the calculation of tidal strength is based on density. In this work, we adopted the method described in O'Neill et al. (2024) to convert the differential extinction in the 3D dust map to the density of hydrogen nuclei. However, this method relies on certain assumptions, such as a constant ratio of hydrogen column density to extinction. To achieve a more accurate calculation of tidal strength, better estimates of both the density distribution and magnitude will be essential.

References

Ballesteros-Paredes, J., Gómez, G. C., Loinard, L., Torres, R. M., & Pichardo, B. 2009a, MNRAS, 395, L81
Ballesteros-Paredes, J., Gómez, G. C., Pichardo, B., & Vázquez-Semadeni, E. 2009b, MNRAS, 393, 1563
Bertoldi, F., & McKee, C. F. 1992, ApJ, 395, 140
Chen, B. Q., Huang, Y., Yuan, H. B., et al. 2019, MNRAS, 483, 4277
Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792
Dewangan, L. K., Ojha, D. K., Sharma, S., et al. 2020, ApJ, 903, 13
Dib, S., Helou, G., Moore, T. J. T., Urquhart, J. S., & Dariush, A. 2012, ApJ, 758, 125
Edenhofer, G., Zucker, C., Frank, P., et al. 2024, A&A, 685, A82
Ganguly, S., Walch, S., Clarke, S. D., & Seifried, D. 2024, MNRAS, 528, 3630
Green, G. M., Schlafly, E., Zucker, C., Speagle, J. S., & Finkbeiner, D. 2019, ApJ, 887, 93
Henshaw, J. D., Caselli, P., Fontani, F., Jiménez-Serra, I., & Tan, J. C. 2014, MNRAS, 440, 2860
Hottier, C., Babusiaux, C., & Arenou, F. 2020, A&A, 641, A79
Issac, N., Tej, A., Liu, T., et al. 2019, MNRAS, 485, 1775
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87
Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, A&A, 625, A135
Leike, R. H., Glatzle, M., & Enßlin, T. A. 2020, A&A, 639, A138
Li, G.-X. 2024, MNRAS, 528, L52
Liu, T., Zhang, Q., Kim, K.-T., et al. 2016, ApJ, 824, 31
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, MNRAS, 496, 2790
Liu, L., Bureau, M., Blitz, L., et al. 2021a, MNRAS, 505, 4048
Liu, X.-L., Xu, J.-L., Wang, J.-J., et al. 2021b, A&A, 646, A137
Liu, H.-L., Tej, A., Liu, T., et al. 2022, MNRAS, 511, 4480
Liu, H.-L., Tej, A., Liu, T., et al. 2023, MNRAS, 522, 3719
Lu, X., Zhang, Q., Liu, H. B., et al. 2018, ApJ, 855, 9
McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565
Meidt, S. E., Leroy, A. K., Rosolowsky, E., et al. 2018, ApJ, 854, 100
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Myers, P. C. 2009, ApJ, 700, 1609
O'Neill, T. J., Zucker, C., Goodman, A. A., & Edenhofer, G. 2024, ApJ, 973, 136
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Ramírez-Galeano, L., Ballesteros-Paredes, J., Smith, R. J., Camacho, V., & Zamora-Avilés, M. 2022, MNRAS, 515, 2822
Rezaei, K. S., Bailer-Jones, C. A. L., Hogg, D. W., & Schultheis, M. 2018, A&A, 618, A168
Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, ApJ, 679, 1338
Schneider, N., Csengeri, T., Hennemann, M., et al. 2012, A&A, 540, L11
Stark, A. A., & Blitz, L. 1978, ApJ, 225, L15
Thilliez, E., Maddison, S. T., Hughes, A., & Wong, T. 2014, PASA, 31, e003
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Xu, F.-W., Wang, K., Liu, T., et al. 2023, MNRAS, 520, 3259
Yang, D., Liu, H.-L., Tej, A., et al. 2023, ApJ, 953, 40
Yuan, J., Li, J.-Z., Wu, Y., et al. 2018, ApJ, 852, 12
Zhang, Q., Wang, K., Lu, X., & Jiménez-Serra, I. 2015, ApJ, 804, 141
Zhang, X., Green, G. M., & Rix, H.-W. 2023, MNRAS, 524, 1855
Zhou, J.-W., & Davis, T. 2024, PASA, 41
Zhou, J. W., & Dib, S. 2025, MNRAS, 537, 2232
Zhou, J. W., & Li, G.-X. 2025, MNRAS, 537, 2630
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
Zhou, J. W., Dib, S., & Davis, T. A. 2024a, MNRAS, 534, 683
Zhou, J. W., Dib, S., Juvela, M., et al. 2024b, A&A, 686, A146

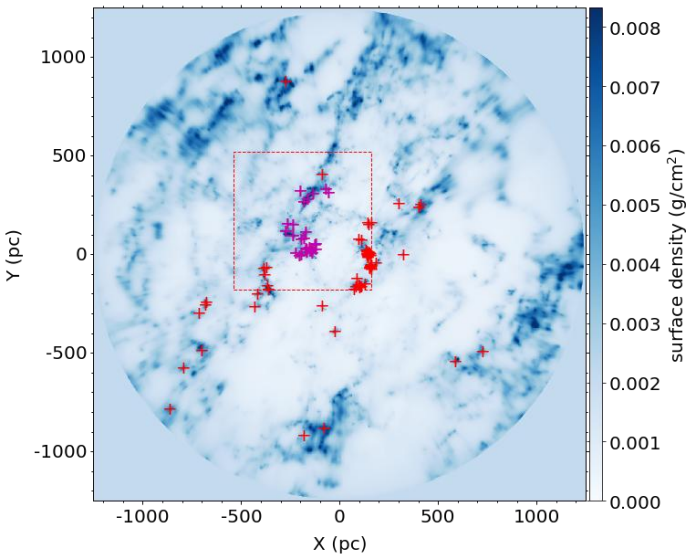


Fig. A.1. Surface density of hydrogen nuclei on the XY plane. The pluses represent the 90 non-overlapping structures in Sec. 3.1.

Appendix A: Tidal calculation

A.1. Tidal tensor

In the tidal tensor computation, we utilized the 3D dust map available in the heliocentric Galactic Cartesian coordinate system (X-Y-Z).[4] Based on the 3D density distribution derived from the 3D dust map, the gravitational potential is computed by solving the Poisson equation $\nabla^2\phi = 4\pi G\rho$ in the Fourier space,

$$\Phi_{k,3D} = -\frac{4\pi G\rho_k}{k_{3D}^2}\,, \tag{A.1}$$

where we first calculate $\rho_k = \hat{f}(\rho(x, y, z))$. The gravitational potential in the real space is then computed through $\Phi_(x, y, z) = \hat{f}^-1(\Phi_k)$. The terms $\hat{f}$ and $\hat{f}^{-1}$ are the Fourier transform and the inverse Fourier transform. For a gravitational potential field Φ, the tidal tensor $\mathbf{T}$ is defined as

$$\mathrm{T}_{ij} = \partial_i\partial_j\Phi. \tag{A.2}$$

The tidal tensor is symmetric and real-valued. Thus, it can be written in orthogonal form. The ordered and diagonalized tidal tensor is given as

$$\mathbf{T} = \begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix}. \tag{A.3}$$

The eigenvalues of the tidal tensor provide insight into gravitationally induced deformations, such as compression or disruption, at a given point within the gravitational field. The sign of an eigenvalue determines whether the corresponding mode is compressive ($\lambda_i > 0$) or disruptive ($\lambda_i < 0$), while its magnitude indicates the strength of the respective compressive or disruptive effect. The "ordered" means $|\lambda_3| > |\lambda_2| > |\lambda_1|$. The trace of the tidal tensor, $\mathrm{Tr}(\mathbf{T}) = \sum_{i=1}^{3}\lambda_i$, contains the local density information in the Poisson equation:

$$\mathrm{Tr}(\mathbf{T}) = \nabla^2\Phi = 4\pi G\rho. \tag{A.4}$$

[4] https://zenodo.org/records/10658339

This implies that the trace is 0 if the point at which the tidal tensor is evaluated outside the mass distribution.

Due to computational resource limitations, we focused our calculations on the subregion ($X \times Y \times Z = 700$ pc $\times$ 700 pc $\times$ 700 pc) outlined by the red box in Fig. A.1. Fig. A.2 displays a snapshot of the density and the maximum eigenvalue of the tidal tensor (λ_3) on XY plane, when $Z = 80$ pc.

As discussed in Ganguly et al. (2024), the tidal tensor can be decomposed into three components: the tidal tensor generated solely by the matter within the structure, $\mathbf{T}_{\rm int}$; the tidal tensor arising exclusively from the external matter distribution, $\mathbf{T}_{\rm ext}$; and their sum, which represents the contribution from the entire matter distribution:

$$\mathbf{T}_{\rm tot} = \mathbf{T}_{\rm int} + \mathbf{T}_{\rm ext}. \tag{A.5}$$

Here, $\mathbf{T}_{\rm tot}$ represents the net deformation due to both the structure itself and its external material. The traces of the three parts are

$$\mathrm{Tr}(\mathbf{T}_{\rm tot}) = 4\pi G\rho \tag{A.6}$$
$$\mathrm{Tr}(\mathbf{T}_{\rm int}) = 4\pi G\rho \tag{A.7}$$
$$\mathrm{Tr}(\mathbf{T}_{\rm ext}) = 0. \tag{A.8}$$

The trace in equation.A.8 suggests that $\mathbf{T}_{\rm ext}$ must include both compressive and disruptive modes, corresponding to positive and negative λ_i, respectively. In contrast, $\mathbf{T}_{\rm int}$ and $\mathbf{T}_{\rm tot}$ must contain at least one compressive mode but can also be entirely compressive. According to equation.A.5, the 3D tidal tensor can be decomposed as

$$\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix} = \begin{bmatrix}A_0 & 0 & 0\\ 0 & A_1 & 0\\ 0 & 0 & A_2\end{bmatrix} + \begin{bmatrix}A_3 & 0 & 0\\ 0 & A_4 & 0\\ 0 & 0 & A_5\end{bmatrix}. \tag{A.9}$$

For the six unknown parameters, they satisfy the following four equations:

$$\begin{cases} A_0 + A_3 = \lambda_1\\ A_1 + A_4 = \lambda_2\\ A_2 + A_5 = \lambda_3\\ A_3 + A_4 + A_5 = \lambda_1 + \lambda_2 + \lambda_3 \end{cases}.$$

Since the six unknown parameters are real numbers, we set $A_4 = k * A_5$ and $A_1 = h * A_2$, where k and h are arbitrary constants, then we have the solutions of the unknown parameters:

$$\begin{cases} A_0 = \frac{hk\lambda_3 - h\lambda_2 + k\lambda_3 - \lambda_2}{h-k}\\ A_1 = \frac{-hk\lambda_3 + h\lambda_2}{h-k}\\ A_2 = \frac{-k\lambda_3 + \lambda_2}{h-k}\\ A_3 = \frac{-hk\lambda_3 + h\lambda_1 + h\lambda_2 - k\lambda_1 - k\lambda_3 + \lambda_2}{h-k}\\ A_4 = \frac{hk\lambda_3 - k\lambda_2}{h-k}\\ A_5 = \frac{h\lambda_3 - \lambda_2}{h-k} \end{cases}$$

The solutions require $h \neq k$. Below, we consider several limiting cases:

1. When $k \gg 1$ and $h \approx 1$, the solutions of the six unknown parameters are $A_0 = -2\lambda_3$, $A_1 = \lambda_3$, $A_2 = \lambda_3$, $A_3 = \lambda_1 + 2\lambda_3$, $A_4 = \lambda_2 - \lambda_3$, $A_5 = 0$.

2. When $k \gg 1$ and $h \ll 1$, the solutions of the six unknown parameters are $A_0 = -\lambda_3$, $A_1 = 0$, $A_2 = \lambda_3$, $A_3 = \lambda_1 + \lambda_3$, $A_4 = \lambda_2$, $A_5 = 0$.

3. When $k \approx 1$ and $h \gg 1$, the solutions of the six unknown parameters are $A_0 = -\lambda_2 + \lambda_3$, $A_1 = \lambda_2 - \lambda_3$, $A_2 = 0$, $A_3 = \lambda_1 + \lambda_2 - \lambda_3$, $A_4 = \lambda_3$, $A_5 = \lambda_3$.

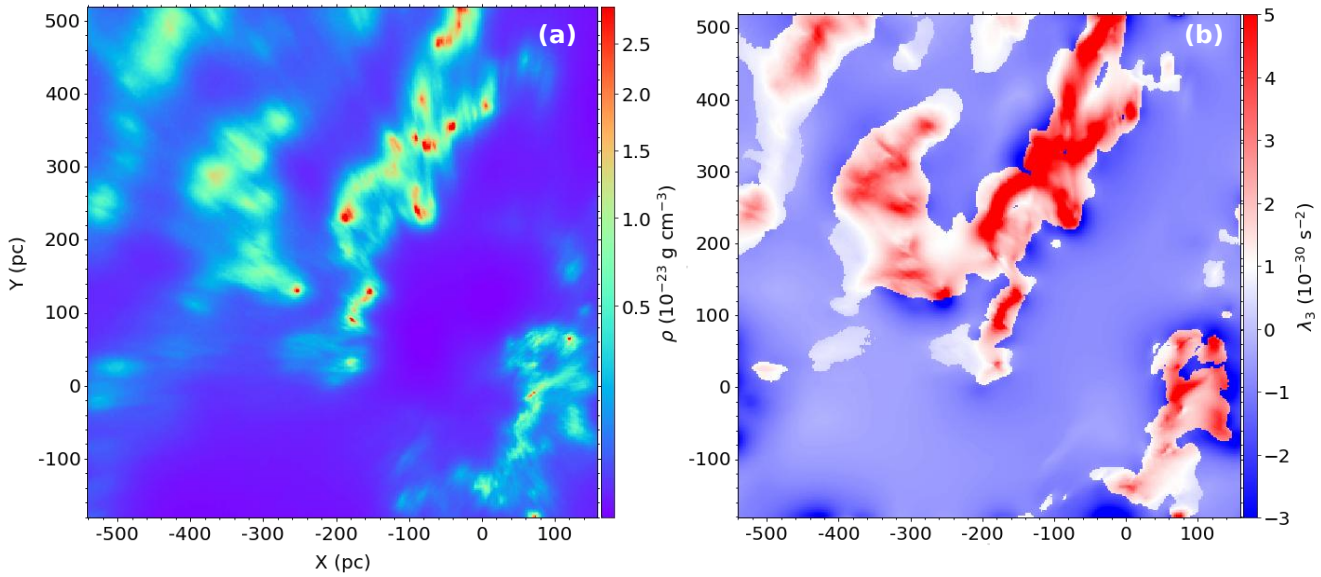


Fig. A.2. Tidal field in the heliocentric Galactic Cartesian coordinate system calculated based on the 3D dust distribution near the Sun. The Sun is located at the origin (0,0). Here, we present snapshots of the volume density and the maximum eigenvalue of the tidal tensor on the XY plane at $Z = 80$ pc.

4. When $k \approx 1$ and $h \ll 1$, the solutions of the six unknown parameters are $A_0 = \lambda_2 - \lambda_3$, $A_1 = 0$, $A_2 = -\lambda_2 + \lambda_3$, $A_3 = \lambda_1 - \lambda_2 + \lambda_3$, $A_4 = \lambda_2$, $A_5 = \lambda_2$.

5. When $k \ll 1$ and $h \gg 1$, the solutions of the six unknown parameters are $A_0 = -\lambda_2$, $A_1 = \lambda_2$, $A_2 = 0$, $A_3 = \lambda_1 + \lambda_2$, $A_4 = 0$, $A_5 = \lambda_3$.

6. When $k \ll 1$ and $h \approx 1$, the solutions of the six unknown parameters are $A_0 = -2\lambda_2$, $A_1 = \lambda_2$, $A_2 = \lambda_2$, $A_3 = \lambda_1 + 2\lambda_2$, $A_4 = 0$, $A_5 = -\lambda_2 + \lambda_3$.

Assuming that the largest eigenvalue of $\mathbf{T}_{\rm ext}$ dominates the deformation of a given structure due to the external tides, in the six limiting cases, the maximum eigenvalue of $\mathbf{T}_{\rm ext}$ (i.e., $\lambda_{\rm ext,max}$) can be $|-2\lambda_2|$, $|-2\lambda_3|$, $|\lambda_2-\lambda_3|$, $|-\lambda_2|$, and $|-\lambda_3|$. Since the eigenvalues can become negative, we consider their absolute values in order to compare the relative magnitude. Similar to the case of $A_4 = k * A_5$ and $A_1 = h * A_2$, there are eight other situations:

$A_0 = h * A_1$ and $A_3 = k * A_4$,
$A_0 = h * A_1$ and $A_3 = k * A_5$,
$A_0 = h * A_1$ and $A_4 = k * A_5$,
$A_0 = h * A_2$ and $A_3 = k * A_4$,
$A_0 = h * A_2$ and $A_3 = k * A_5$,
$A_0 = h * A_2$ and $A_4 = k * A_5$,
$A_1 = h * A_2$ and $A_3 = k * A_4$,
$A_1 = h * A_2$ and $A_3 = k * A_5$.

For all limiting cases, there are 12 possible $|\lambda_{\rm ext,max}|$: $|-2\lambda_1|$, $|-\lambda_1|$, $|-2\lambda_2|$, $|-\lambda_2|$, $|-2\lambda_3|$, $|-\lambda_3|$, $|\lambda_1-\lambda_2|$, $|\lambda_1-\lambda_3|$, $|\lambda_2-\lambda_3|$, $|\lambda_1-\lambda_2|/2$, $|\lambda_1-\lambda_3|/2$, and $|\lambda_2-\lambda_3|/2$. Since $|\lambda_3| > |\lambda_2| > |\lambda_1|$, we only need to consider four of them, namely: $|-2\lambda_3|$, $|\lambda_1-\lambda_2|$, $|\lambda_1-\lambda_3|$, and $|\lambda_2-\lambda_3|$.

A.2. Pixel-by-pixel computation

Below, we introduce an independent and original method to select the optimal $|\lambda_{\rm ext,max}|$ from four possible ones. At a distance R' from a pixel B with mass M', the tidal strength sustained by a pixel A due to the external gravity of pixel B is

$$T \approx \frac{2GM'}{R'^3}. \tag{A.10}$$

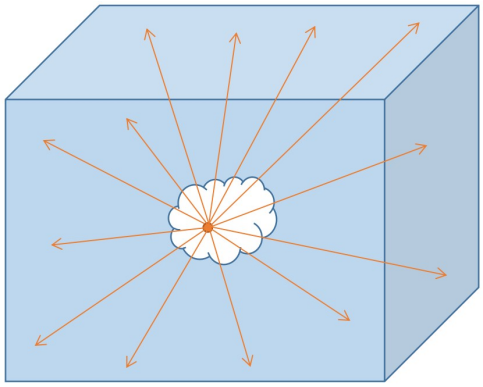


Fig. A.3. Pixel-by-pixel computation. Here, we took out a cloud from the 3D density cube and calculated the external tides from all external material at a point in the cloud.

The cumulative tidal strength at a pixel signifies the aggregate deformation resulting from external gravity at that point. Unlike a rigid body or a point-mass, a gas structure is flexible and can be deformed. Gas structures are often irregular in shape, and their morphologies can be quite intricate. While the gravitational forces acting on a gas structure from different directions might balance out, the tidal effects within the structure, driven by external gravity, do not. To quantify the collective tidal strength within a gas structure induced by external gravity, we adopt the scalar superposition. According to the central coordinates (X, Y and Z) and radius of each cloud identified in Sec. 3.1, we can divide the 3D density cube into two parts, i.e. the cloud itself and all external material apart from this cloud. For the external tides from external material sustained by a point in the cloud, a simple but heavy calculation is using equation.A.10 to directly calculate the tidal strength at that point pixel-by-pixel through the entire cube, as illustrated in Fig. A.3. Fig. A.4 shows the distribution of the average tidal strengths (four possible $|\lambda_{\rm ext,max}|$ and the pixel-by-pixel computed one, i.e. $\lambda_{\rm pp}$) for the clouds marked by magenta "+" within the box of Fig. A.1. Since the clouds at

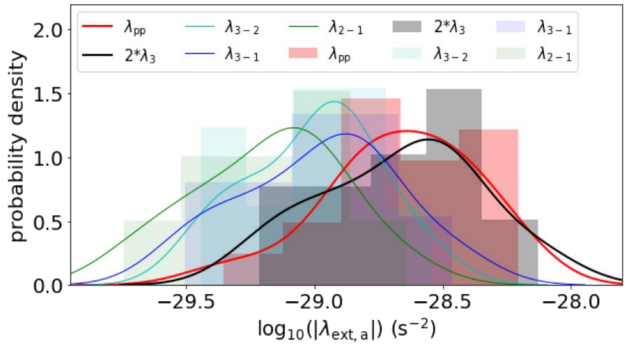


Fig. A.4. Kernel density estimation and the histogram of the average external tidal strength of the clouds. Here, $|-2\lambda_3|$, $|\lambda_1-\lambda_2|$, $|\lambda_1-\lambda_3|$, and $|\lambda_2-\lambda_3|$ are four possible maximum eigenvalues derived in Sec. A.1, and λ_{pp} is the tidal strength from the pixel-by-pixel computation. The comparison is only conducted on the clouds marked by a magenta "+" within the box of Fig. A.1.

the boundary cannot account for all neighboring material, they were excluded. The distribution of $|-2\lambda_3|$ is closest to λ_{pp}. In equation.A.9, when $A_0 = -2\lambda_3$, we obtain a decomposition of the tidal tensor:

$$\mathbf{T}_{tot} = \mathbf{T}_{ext} + \mathbf{T}_{int} = \begin{bmatrix} -2\lambda_3 & 0 & 0 \\ 0 & \lambda_3 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} + \begin{bmatrix} \lambda_1+2\lambda_3 & 0 & 0 \\ 0 & \lambda_2-\lambda_3 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{A.11}$$

For a point-mass with mass M_0, the tidal tensor at distance d is given as

$$\mathbf{T} = \begin{bmatrix} -\frac{2GM_0}{d^3} & 0 & 0 \\ 0 & \frac{GM_0}{d^3} & 0 \\ 0 & 0 & \frac{GM_0}{d^3} \end{bmatrix}. \tag{A.12}$$

As shown in Fig. A.4, $<|2\lambda_3|> \approx < 2GM'/R'^3 >$ (equation.A.10), and the eigenvalues of $\mathbf{T}_{ext}$ in equation.A.11 has the same formalism with the tidal tensor of a point-mass. Therefore, for the external tides from external material sustained by a point in a cloud, the external material can be divided into approximate point-masses. Each point-mass can produce its own tidal field. The external tides from all external material are equivalent to the superposition of the tidal fields of all divided point-masses. In essence, it is equivalent to the above pixel-by-pixel computation.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **537,** 2232–2242 (2025) https://doi.org/10.1093/mnras/staf201
Advance Access publication 2025 February 4

An examination of large-scale galactic effects on molecular cloud properties in NGC 628: the significant impact of tidal effects from neighbouring material on the evolution of molecular clouds

J. W. Zhou [1]★ and Sami Dib [2]

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*Max Planck Institute für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

Accepted 2025 January 29. Received 2025 January 29; in original form 2024 October 30

ABSTRACT

The physical factors that influence the development of molecular cloud's density contrast are connected to those that affect star formation in the galaxy. For NGC 628 (M74), the proportion of high- and low-density contrast clouds initially increases with the distance to the galactic centre (R_G) and then keeps relatively stable. Spiral arms, bubbles, and magnetic fields are not responsible for the variations in density contrast observed among molecular clouds. The effects of shear and tides calculated from the galactic rotation curve consistently decrease as R_G increases, and the shear effect can be neglected. We further studied the tidal effects of the neighbouring material on each cloud using the tidal tensor analysis and the pixel-by-pixel computation, after combining molecular gas, atomic gas, and stellar mass surface density maps. When $R_{\rm G} < 4$ kpc, the tidal strengths derived from the pixel-by-pixel computation decrease as $R_{\rm G}$ increases, and then remains relatively constant when $R_{\rm G} > 4$ kpc. This aligns well with the dependence of the proportion of high- and low-density contrast clouds on $R_{\rm G}$. Therefore, the tidal effects of neighbouring material have a significant impact on the development of molecular cloud's density contrast. A key factor contributing to the low star formation rate in the galactic centre is the excessive tidal influences from neighbouring material on molecular clouds, which hinder the gravitational collapse within these clouds, resulting in low density contrasts. The tidal effects from neighbouring material may also be a significant contributing factor to the slowing down of a pure free-fall gravitational collapse for gas structures on galaxy-cloud scales revealed in our previous works by velocity gradient measurements.

Key words: techniques: image processing – ISM: clouds – ISM: kinematics and dynamics – galaxies: ISM – galaxies: star formation – galaxies: structure.

1 INTRODUCTION

Giant molecular clouds (GMCs) are widely recognized as the primary gas reservoirs that fuel star formation and serve as the birthplaces of nearly all stars. Gaining insight into the characteristics of GMCs is therefore crucial for deciphering the relationship between gas and star formation in galaxies (Kennicutt & Evans 2012; Schinnerer & Leroy 2024). Observations show that the physical characteristics of GMCs systematically change depending on their location within a galaxy. This implies a connection between GMCs and their galactic environment, which influences their formation, structure, and evolution (Hughes et al. 2013; Colombo et al. 2014; Miville-Deschênes, Murray & Lee 2017; Kim et al. 2022; Pessa et al. 2022; Sun et al. 2022; Zhou, Dib & Davis 2024a). The range of potential physical mechanisms that influence the properties of GMCs is extensive and diverse (Jeffreson et al. 2020; Chevance et al. 2023) including galactic shear and tides (Ballesteros-Paredes et al. 2009; Dib et al. 2012; Thilliez et al. 2014; Aouad, James & Chilingarian 2020; Ramírez-Galeano et al. 2022), interactions with spiral arms (Meidt et al. 2013; Querejeta et al. 2024), epicyclic motions driven by the galactic rotation (Meidt et al. 2020; Utreras et al. 2020; Liu et al. 2021), accretion flows from galactic scales down to GMC scales (Klessen & Hennebelle 2010; Zhou & Davis 2024), collisions between clouds (Tasker & Tan 2009; Fukui et al. 2021; Sano et al. 2021; Zhou et al. 2023), stellar feedback processes (Dib et al. 2009; Krumholz et al. 2014; Barnes et al. 2022; Chevance et al. 2022), and magnetic fields (Li & Henning 2011; Crutcher 2012; Ibáñez-Mejía, Mac Low & Klessen 2022; Ganguly et al. 2023; Rawat et al. 2024). These intricate interactions result in significant, measurable correlations between the characteristics of molecular clouds and both the local and global properties of their host galaxy. Investigating these correlations between clouds and their environments presents an opportunity to deepen our understanding of the physics that drive molecular cloud evolution, and, by extension, star formation and galaxy evolution (Sun et al. 2022; Schinnerer & Leroy 2024).

In Zhou et al. (2024a), we decomposed the molecular gas in the spiral galaxy NGC 628 (M47) into multiscale hub-filament structures using the CO (2-1) line map. All molecular clouds as potential hubs were classified into three categories: leaf-HFs-A, leaf-HFs-B, and

★ E-mail: jwzhou@mpifr-bonn.mpg.de

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

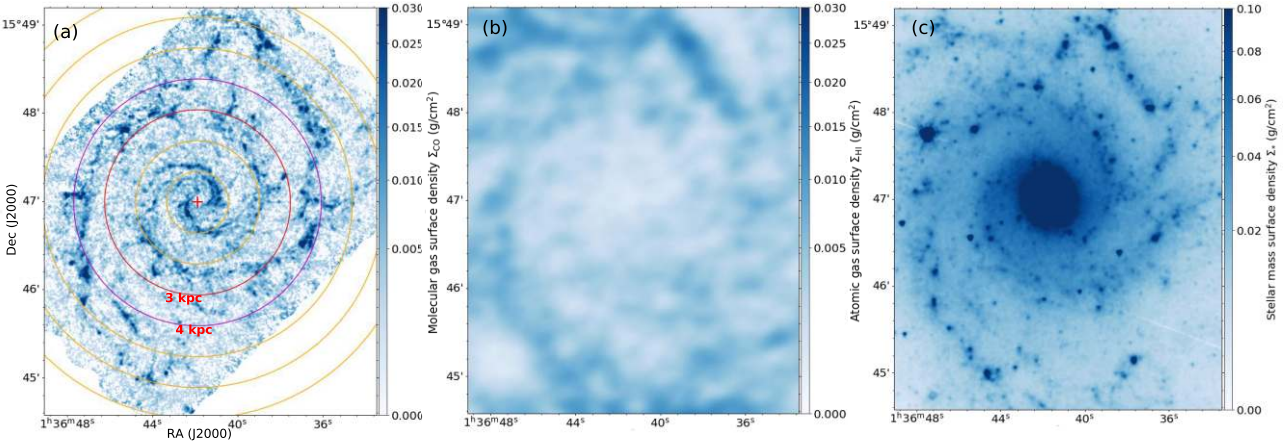


Figure 1. Surface density maps of CO (2–1), H I, and stellar mass.

leaf-HFs-C. For leaf-HFs-C, the density contrast between the hub and the surrounding diffuse gas is not significant. Both leaf-HFs-A and leaf-HFs-B feature distinct central hubs, but leaf-HFs-A exhibit the best-defined hub-filament morphology. Leaf-HFs-A also have the highest density contrast, the largest mass, and the lowest virial ratio. Leaf-HFs-C are not necessarily low-density structures, but their density is more uniformly distributed. Currently, leaf-HFs-C lack a clear gravitational collapse process that would produce a pronounced density contrast. We found a clear correlation between density contrast and the virial ratio, with the density contrast decreasing as the virial ratio increases. The density contrast (C) effectively gauges the degree of gravitational collapse and the depth of the gravitational potential well that shape the hub-filament morphology. In terms of understanding the development and evolutionary stage of molecular clouds, the density contrast is more indicative than density itself.

The study of the formation and evolution of three types of structures mentioned above is important to understand what are the main mechanisms that regulate star formation in galaxies. The interstellar medium (ISM) in galaxies ranges from diffuse gas to regions of localized clumping, followed by a further increase in density due to gravitational collapse and accretion, which eventually results in star formation. The physical factors that influence the development of molecular cloud's density contrast are the same factors that affect star formation in galaxies. As shown in Zhou et al. (2024a), as long as C is different, other physical quantities will also differ significantly. C can be used to represent the star formation capability of molecular clouds. The correlation between C and various physical quantities can reflect the impact of those physical processes on the star formation in molecular clouds. In this work, we are not concerned with the formation of molecular clouds, but rather with the physical processes that create variations in density contrast among molecular clouds in NGC 628. Particularly, we focus on the role of different types of tides.

2 DATA

In this work, we use the integrated intensity map derived from the combined 12m+7m + TP CO (2−1) data cube of NGC 628, which is the data product in the PHANGS-ALMA survey (Leroy et al. 2021a, b). We also use the integrated intensity map of H I from the H I Nearby Galaxy Survey (THINGS) (Walter et al. 2008), and the *Spitzer* IRAC 3.6 μm image in the *Spitzer* Infrared Nearby Galaxies Survey (SINGS) (Dale et al. 2009).

3 RESULTS

3.1 Density contrast distribution

In Fig. 2, we can see a significant proportion of low-density contrast clouds concentrated in the galaxy centre. For all clouds, the median value of the density contrast C is ~ 1.63. Based on the distance to the galaxy centre (R_G), we divided the galaxy into different rings. As shown in Fig. 1(a), only rings with a radius of ≤ 4 kpc are complete. In Table 1, for $R_G \leq 4$ kpc, the number of high-density contrast clouds rises notably as R_G increases. In order to extend the comparison to a wider spatial range within the galaxy, we calculate the proportion of clouds with $C > 1.63$ and $C < 1.63$ within each ring. As shown in Sections 3.2 and 3.3, the distributions of high- and low-density contrast clouds are relatively uniform within each ring, meaning that they are not biased towards any specific part of the ring. Despite the incompleteness of the outer rings, calculating the ratio of high- to low-density contrast clouds remains feasible.

In Table 1, $R_G < 3$ kpc has significantly lower proportion of clouds with $C > 1.63$, consistent with the fact that star formation is suppressed in the centre of the galaxy (Kruijssen et al. 2014). When $R_G > 3$ kpc, the proportion is relatively stable. When $R_G < 4$ kpc, there is a clear trend where the proportion of clouds with $C > 1.63$ increases significantly with increasing R_G. Next, we will explore in detail the reasons behind the dependence of the cloud's density contrast on R_G.

3.2 Spiral arm, bubble, and magnetic field

We first investigated the impact of spiral arms and bubbles on the density contrast of molecular clouds. Querejeta et al. (2021)

Table 1. The ratio of high-density contrast clouds to low-density contrast clouds at different distances from the galactic centre.

Distance (kpc)	$C > 1.63$	$C < 1.63$	Proportion
$R_G < 1$	14	39	0.358
$1 < R_G < 2$	45	63	0.714
$2 < R_G < 3$	68	77	0.883
$3 < R_G < 4$	94	68	1.382
$4 < R_G < 5$	88	75	1.173
$5 < R_G < 6$	57	42	1.357
$6 < R_G < 7$	20	18	1.111

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

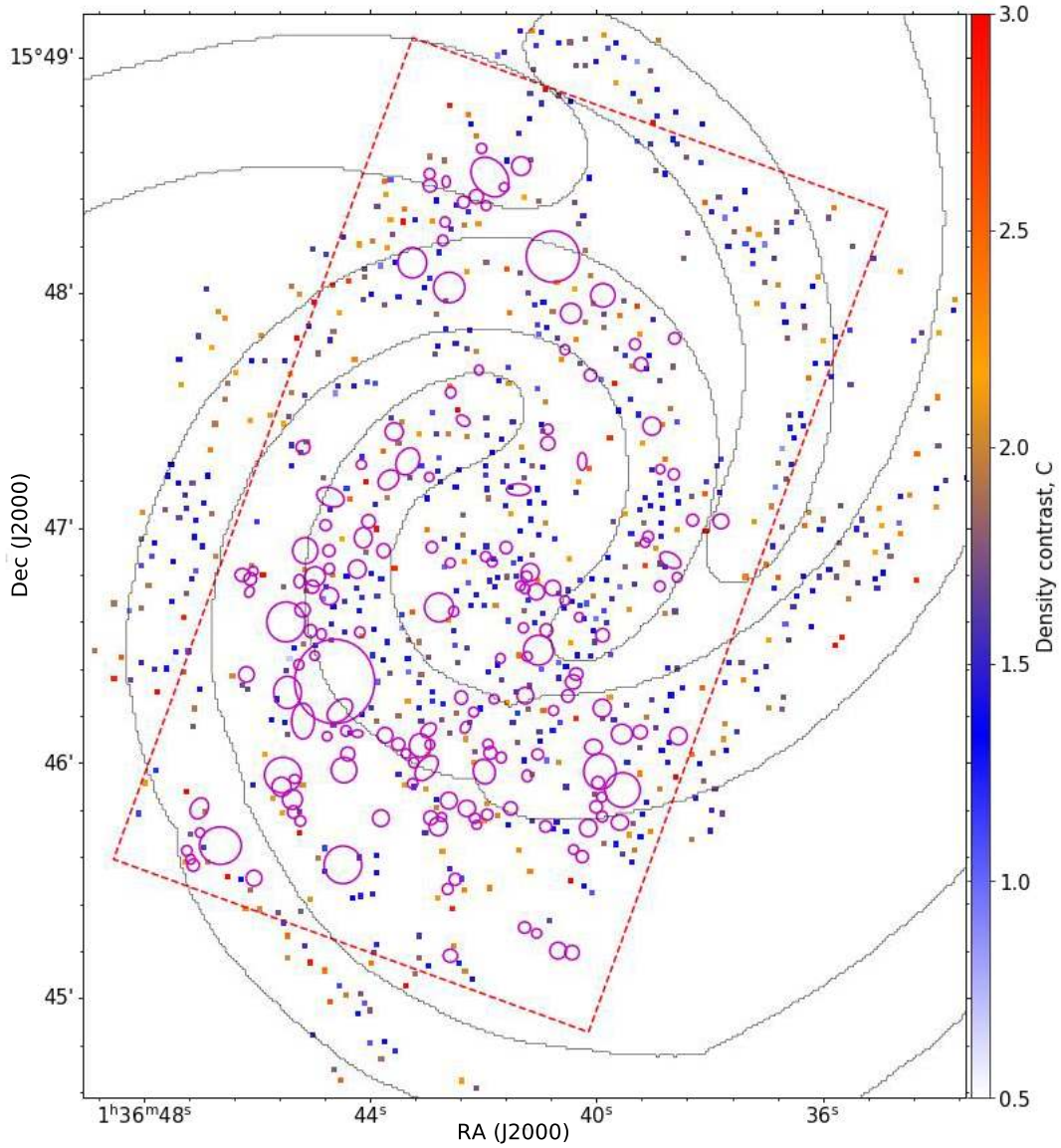


Figure 2. The column density contrast map shows the density contrast distribution of all clouds in NGC 628. Each point represents a cloud. Grey contours are the spiral arms identified in Querejeta et al. (2021). Magenta ellipses represent the bubbles identified in Watkins et al. (2023). Red dashed box shows the region observed by *JWST*.

identified arms and interarms of NGC 628 by using the log-spiral pattern, based on the *Spitzer* 3.6 μm image. The bubbles in NGC 628 were identified in Watkins et al. (2023) using *JWST* mid-infrared observation. If high-density contrast clouds are triggered by bubbles, bubbles should be larger than the clouds. The median value of the effective radii of all clouds identified in Zhou et al. (2024a) is ∼ 55 pc. Therefore, we only consider the bubbles with the effective radii larger than 55 pc. We construct a density contrast map based on the calculation in Zhou et al. (2024a), which shows the density contrast distribution of all clouds in the galaxy. In Fig. 2, each point represents a cloud. We overlay the bubbles and spiral arms on the density contrast map. Overall, high- and low-density contrast clouds are mixed together, primarily concentrated within the spiral arms. Additionally, high-density contrast clouds do not tend to be distributed around bubbles. In Mulcahy, Beck & Heald (2017), the total magnetic field strength of NGC 628 was estimated from the non-thermal emission, assuming equipartition between the energy densities of cosmic rays and the magnetic field, applying the updated formula from Beck & Krause (2005). Based on the total magnetic field strength map [fig. 12 in Mulcahy et al. (2017)], we roughly estimate the average magnetic field strength for each cloud. As shown in Fig. 3, there is no correlation between the magnetic field strength and the density contrast. Therefore, the spiral arms, bubbles, and magnetic fields are not the reasons for the variation in the density contrast of molecular clouds with galactocentric radius.

3.3 Shear and tidal fields of the galaxy

As discussed in Hunter, Elmegreen & Baker (1998), Dib et al. (2012), and Thilliez et al. (2014), the shear parameter characterizes the effect of differential rotation on the stability of molecular clouds against gravitational collapse, assessing the time available for perturbations to collapse in the presence of local rotational shear caused by the

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

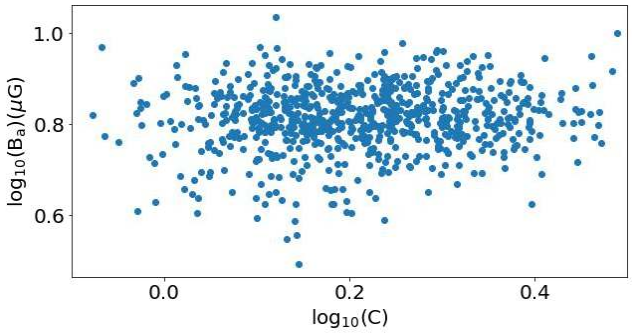


Figure 3. Correlation between the average magnetic field strength and the density contrast of the clouds.

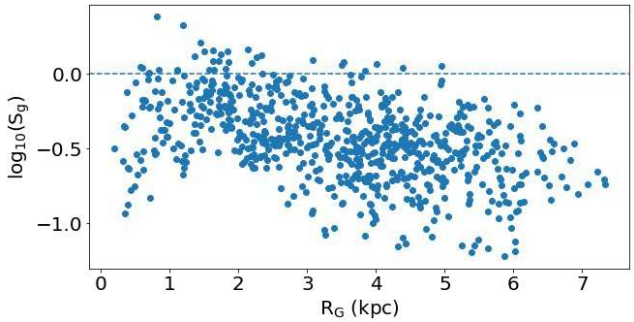


Figure 4. Correlation between the shear parameter (S_g) and the distance to the galaxy centre (R_G).

galaxy's global rotation. The shear parameter is defined as

$$S_g = \frac{\Sigma_{\rm cri}}{\Sigma}, \tag{1}$$

where Σ is the local gas surface density. $\Sigma_{\rm cri}$ is the critical surface density, which represents the minimal surface density needed to resist the shear. It is defined as

$$\Sigma_{\rm cri} = \frac{A\sigma \ln(C^{'})}{2\pi G}, \tag{2}$$

where $C^{'}$ is taken to be 100 in Hunter et al. (1998). σ is the local gas velocity dispersion. A is the Oort constant, which measures the local shear level. For a spherical cloud of radius R, with a rotational velocity V at a radius R_G, following Dib et al. (2012), A can be calculated as

$$\begin{aligned} A &= 0.5\left(\frac{V}{R_G} - \frac{{\rm d}V}{{\rm d}R_G}\right) \\ &\approx 0.5\left(\frac{V}{R_G} - \frac{\mid V(R_G+R) - V(R_G-R)\mid}{2R}\right), \end{aligned} \tag{3}$$

where V is determined by the rotation curve of the galaxy. The analytical rotation curve of NGC 628 derived in Lang et al. (2020) using CO (2–1) emission is

$$V = \frac{2V_{\rm max}}{\pi}\arctan\left(\frac{R_G}{R_{\rm t}}\right) \approx \frac{2\times 144.8}{\pi}\arctan\left(\frac{R_G}{0.56\ {\rm kpc}}\right)\ {\rm km\,s^{-1}}. \tag{4}$$

Shear is effective at tearing apart the gas structure if $S_g > 1$ and ineffective if $S_g < 1$. In Fig. 4, only a few structures have $S_g > 1$, so the effect of shear can be neglected. This is consistent with the results of Dib et al. (2012) and Thilliez et al. (2014) for molecular clouds in the Milky Way and the Large Magellanic Cloud. Moreover, there is a clear trend where S_g decreases as R_G increases, which is inconsistent with the trend in Table 1, where the proportion of high- and low-density contrast clouds initially increases with R_G and then keeps relatively stable.

In a galactic context, the overall gas and stellar distribution can contribute to the confinement or disruption of gas structures. For a cloud at a distance R_G from the galactic centre, the edge of the cloud from the galactic centre is at a distance of $R_G + R$. The tidal acceleration per unit distance (called 'tidal strength' in the later) defined in Stark & Blitz (1978) is,

$$T_G = \frac{V^2}{R_G^2} - \frac{\partial}{\partial R_G}\left(\frac{V^2}{R_G}\right). \tag{5}$$

Fig. 5 shows that the tidal strength T_G is a strong function of R_G, which is also inconsistent with the trend in Table 1.

3.4 Tidal effects of neighbouring material

3.4.1 Tidal tensor

Gravity is a long-range force. A local dense structure evolves under its self-gravity, but as a gravitational centre, its gravity can also affect neighbouring structures. At the same time, it also experiences the external gravity from neighbouring material. The tidal and gravitational fields are mutually interdependent. As described in Zhou & Davis (2024) and Zhou et al. (2024a), the hierarchical/multiscale gravitational coupling of gas structures implies extensive tidal interactions between structures in the galaxy. Whether the galactic potential has an effect on molecular clouds or not, molecular clouds should be affected by the cumulative tidal interactions of the neighbouring material, which may prevent gravitational collapse of the clouds. Ramírez-Galeano et al. (2022) shows that tidal stresses from neighbouring molecular cloud complex (rather than from the galactic potential) may increase interstellar turbulence.

However, the complex morphology of material distribution in the galaxy means that a complete tidal calculation would be also complex. One should derive the gravitational potential distribution from the observed density distribution and then calculate the tidal field according to the gravitational potential distribution. Using the same method as in Zhou et al. (2024b), we directly derived the tidal field according to the surface density distribution obtained from CO (2−1) emission. Calculating the gravitational potential field $\phi(x, y)$ based on the surface density map $\Sigma(x, y)$ has been discussed in Gong & Ostriker (2011) and He, Li & Burkert (2023). For a layer of half-thickness H, the gravitational potential component $\Phi_{k,\,2\rm D}$ of surface density component Σ_k in Fourier phase space is

$$\Phi_{k,\,2\rm D} = -\frac{2\pi G\Sigma_k}{|k|\,(1+|kH|)}, \tag{6}$$

where $|k| = \sqrt{k_x^2 + k_y^2}$. Note that for $|kH| \gg 1$, $\Phi_{k,\,2\rm D} \sim -4\pi G\rho_k/k^2$, which is the solution of the Poisson equation in three dimensions, for $\rho_k = \Sigma_k/2H$. When $|kH| \ll 1$, equation (6) is the solution of the Poisson equation for an infinitesimally thin layer. The gravitational potential $\Phi_{2\rm D}(x, y)$ is the inverse Fourier transform of $\Phi_{k,\,2\rm D}$. The tidal tensor $\mathbf{T}$, defined as $T_{ij} = \frac{\partial\Phi}{\partial i\,\partial j}$, only has two eigenvalues in the case of 2D, i.e. λ_i ($i = 1,2$). The tidal tensor of a structure includes both self-gravity and external tides, that is, $\mathbf{T} = \mathbf{T}_{\rm ext} + \mathbf{T}_{\rm int}$. The decomposition of the tidal tensor is discussed in detail in Zhou et al. (2024b). For three extreme cases, the external tides can be estimated by λ_1, λ_2, and $(\lambda_2 - \lambda_1)/2$, and they are called $T_{\rm d,1}$, $T_{\rm d,2}$, and $T_{\rm d,3}$, respectively.

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

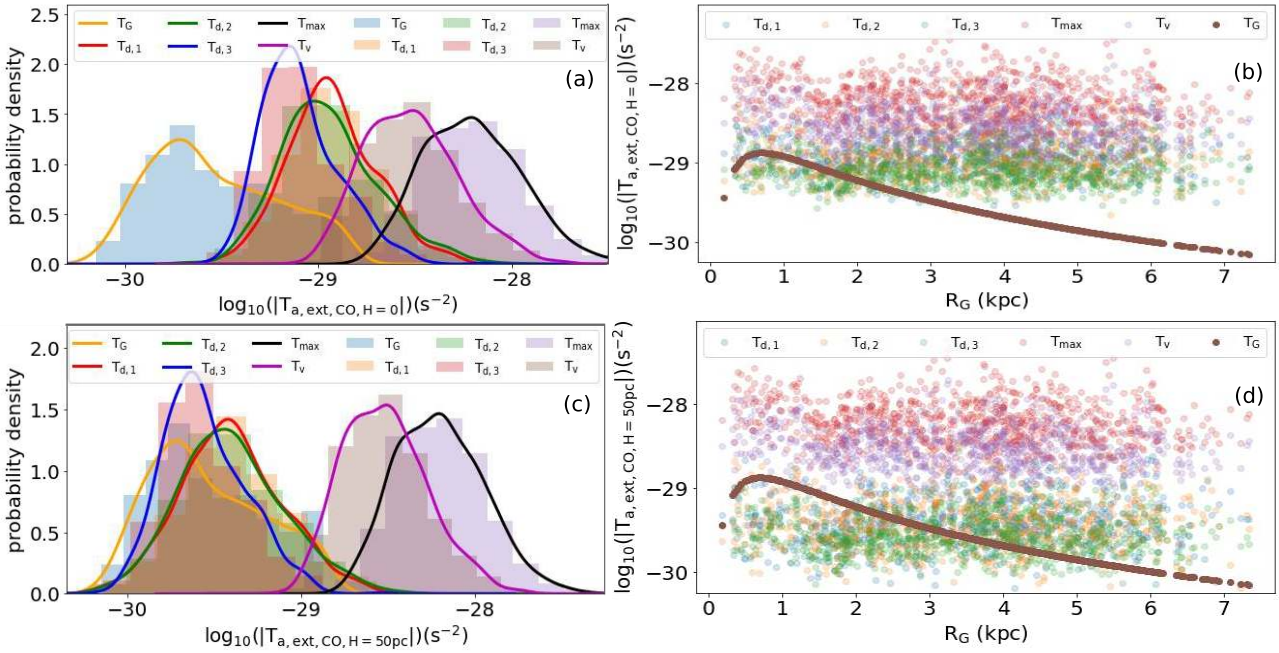


Figure 5. Comparison of the tidal strengths obtained from the tidal tensor analysis and the pixel-by-pixel computation based on the CO surface density map. (a) Distribution of the average tidal strength in molecular clouds, assuming the half-thickness of the galaxy $H = 0$. Each type of tidal strength is explained in Table 2; (b) Variation of the average tidal strength of molecular clouds with the distance of the molecular clouds to the galactic centre. Panels (c) and (d) are same as panels (a) and (b), but for $H = 50$ pc.

For the surface density map, the half-thickness H is 0. Taking a characteristic scale height of the molecular gas in the galaxy as 50 pc (Bronfman et al. 1988; Cox 2005; Stark & Lee 2005; Ellsworth-Bowers et al. 2013; Jeffreson, Sun & Wilson 2022), which is also close to the median value of the effective radii of all identified clouds (∼55 pc), then we have $H = 50$ pc. We calculated the tidal tensors for both $H = 0$ and $H = 50$ pc. Fig. A1 displays different components of the tidal tensor in each case.

In the case of 3D, for a cloud, the energy derived from $\mathbf{T}_{\rm int}$ should be comparable with its self-gravity energy ($E_{\rm g}$), i.e. equation (A8) in Zhou et al. (2024b),

$$E_{\rm a,int,H} \approx \frac{1}{2} M * {\rm Tr}(\mathbf{T}_{\rm a,int,H}) * R^2 \approx \frac{GM^2}{R} \approx E_{\rm g}. \quad (7)$$

In Fig. 6, $E_{\rm a,int,H=0}$ is comparable with $E_{\rm g}$. However, $E_{\rm a,int,H=50pc}$ is systematically smaller than $E_{\rm g}$. Thus, the computed 2D tidal field strength ($H = 0$) can roughly reflect the real 3D tidal field strength of the galaxy.

3.4.2 Pixel-by-pixel computation

In this section, we further verify the above tidal tensor analysis through an original methodology. Tidal forces can significantly contribute to the deformation and intricate shapes of the gas structures. At a distance $R^{'}$ from the centre of cloud B with mass $M^{'}$, the tidal strength sustained by cloud A due to the external gravity of cloud B is

$$T \approx \frac{2GM^{'}}{R^{'3}}. \quad (8)$$

The cumulative tidal strength at a pixel signifies the aggregate deformation resulting from external gravity at that point. Unlike a rigid body or a point mass, a gas structure is flexible and can deform. Gas structures are often irregular in shape, and their morphologies can be quite intricate. While the gravitational forces acting on a gas structure from different directions might balance out, the tidal effects within the structure, driven by external gravity, do not. To quantify the collective tidal strength within a structure induced by external gravity of the neighbouring material, we adopt the scalar superposition.

We divided the surface density map of the galaxy into two distinct regions using the cloud's mask obtained in Zhou et al. (2024a), i.e. the cloud itself and all other material outside of it. To compute the external tidal effects exerted by the surrounding material on a point within the cloud, we use equation (8) to directly calculate the tidal strength at that point pixel by pixel, as depicted in Fig. 7. The average tidal strength across all pixels within the cloud subsequently provides an estimate of the average tidal strength experienced by the cloud. Our computation is limited to a 2D plane. Actually, there is a characteristic scale height of the molecular gas in the galaxy. In equation (8), the tidal strength is very sensitive to the distance. If the external material and the calculated point in a cloud do not lie on the same plane, the 2D computation results in an underestimation of the distance between them, consequently leading to an overestimation of the tidal strength (the maximum computation, $T_{\rm max}$).

A mask of a cloud contains N pixels. Since the tidal strength T is very sensitive to the distance $R^{'}$, material located sufficiently far from a cloud can be expected to have a negligible tidal effects on the cloud. Therefore, it is not necessary to account for all external material within the entire galaxy. We only need to consider the material near the cloud. As illustrated in Fig. 7, we first calculated the average tidal strength exerted by the material within a radius of four times the effective radius of a cloud, with the centre of the cloud as the origin. Then, we extended the calculation to a radius of five times the effective radius of the cloud and performed the same calculation. As shown in Fig. 8, for all clouds in the galaxy, both calculations yield

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

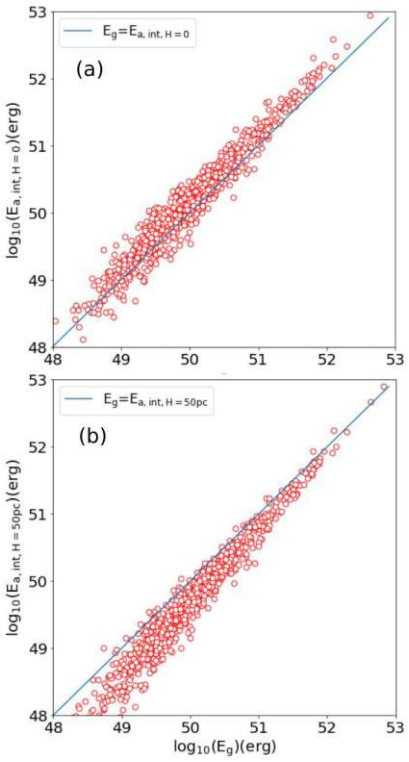


Figure 6. Comparison of $E_{\rm a,int,H}$ and $E_{\rm g}$ defined in Section 3.4.1. There are two cases, i.e. $H = 0$ and $H = 50$ pc. These physical quantities are explained in Table 2.

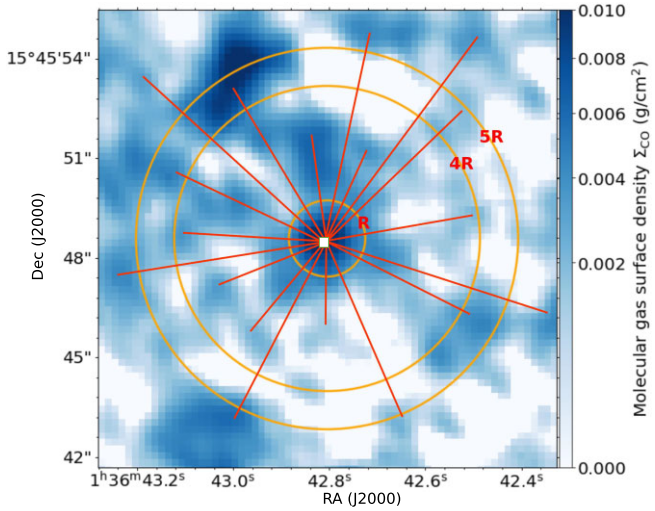


Figure 7. Illustration of the tidal effects exerted by the neighbouring material on a pixel/point within a cloud. R is the effective radius of the cloud.

similar results, indicating that the tidal effects of material beyond four times the effective radius of a cloud can be neglected.

To compare with the scalar superposition, we also calculate a vector superposition of tides, enabling direct cancellation of tidal interactions. By establishing a 2D coordinate system with the calculated pixel/point in a cloud as the origin, we decomposed all external tides along the coordinate axes. Subsequently, we calculated the aggregate tidal strength at the calculated pixel after performing vector superposition along each axis ($T_{\rm v}$).

3.4.3 Baryonic matter

Although we focus solely on molecular clouds as sites of star formation in the galaxy, it is inadequate to only consider molecular gas when accounting for external tidal effects on the clouds. Strictly speaking, we should include all baryonic matter, and even dark matter. As a rough approximation, apart from molecular gas surface density ($\Sigma_{\rm CO}$), we also consider atomic gas surface density ($\Sigma_{\rm HI}$) and stellar mass surface density (Σ_*).

Using the recipes in Leroy et al. (2021b) and Sun et al. (2022), assuming optically thin 21 cm emission, the 21 cm line integrated intensity $I_{\rm HI}$ can be converted to atomic gas surface density $\Sigma_{\rm HI}$ via

$$\frac{\Sigma_{\rm HI}}{\rm M_\odot\ pc^{-2}} = 2.0\times10^{-2}\left(\frac{I_{\rm HI}}{\rm K\ kms^{-1}}\right). \tag{9}$$

Near-IR imaging data can be used to trace the old stellar mass distribution. Using the *Spitzer* IRAC 3.6 μm image, the stellar continuum intensity at 3.6 μm can be converted to the stellar mass surface density by

$$\frac{\Sigma_*}{\rm M_\odot\ pc^{-2}} = 350\left(\frac{\gamma_{3.4\,\mu\rm m}}{0.5}\right)\left(\frac{I_{3.6\,\mu\rm m}}{\rm MJy\ sr^{-1}}\right). \tag{10}$$

Here, $\gamma_{3.4\,\mu\rm m}$ is the stellar mass-to-light ratio at 3.4 μm, we take a value of 0.5 $\rm M_\odot/L_\odot$ (Leroy et al. 2021b). As shown in Fig. 1, the stellar mass surface density is significantly higher than molecular and atomic gas surface density. We combined $\Sigma_{\rm CO}$, $\Sigma_{\rm HI}$, and Σ_* to create a total surface density map to approximate the baryonic matter distribution in the galaxy, which is dominated by Σ_*.

3.5 Comparison of different tidal calculations

Some physical quantities calculated in Section 3.4 are summarized in Table 2. Here, we will compare the different types of tidal calculations mentioned above.

3.5.1 Only the molecular gas

For the tidal strengths calculated according to the CO surface density map, Fig. 5(a) shows that the components $T_{\rm d,1}$, $T_{\rm d,2}$, and $T_{\rm d,3}$ derived from the tidal tensor analysis in Section 3.4.1 are comparable. As expected, the values of $T_{\rm max}$ are significantly larger than others. The tidal strength T_G derived in Section 3.3 according to the rotation curve is significantly smaller than others at large R_G, when $H = 0$. In Fig. 5(b), we plot the variation of the tidal strengths with R_G. Except for T_G, the other quantities do not vary with the distance to the galaxy centre. Because the calculations of the other quantities are based on the CO surface density map. In Fig. 1(a), the galactic centre does not have significantly higher CO surface density.

3.5.2 All matter

For the tidal strengths calculated based on the total surface density map, although the components $T_{\rm d,1,tot}$, $T_{\rm d,2,tot}$, and $T_{\rm d,3,tot}$ derived from the tidal tensor analysis in Section 3.4.1 are comparable, they are significantly smaller than the tidal strengths obtained from pixel-by-pixel computation ($T_{\rm v,tot}$ and $T_{\rm max,tot}$) in Section 3.4.2. In Fig. 9(b), all calculated tidal strengths change with R_G, and the scatterings are also smaller than those in Fig. 5(b). Especially, in Fig. 9(b), when $R_{\rm G} < 4$ kpc, $T_{\rm v,tot}$ and $T_{\rm max,tot}$ decrease as $R_{\rm G}$ increases, and then remains relatively constant when $R_{\rm G} > 4$ kpc. This aligns well with the dependence of the proportion of high- and low-density contrast clouds on $R_{\rm G}$, as presented in Section 3.1.

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

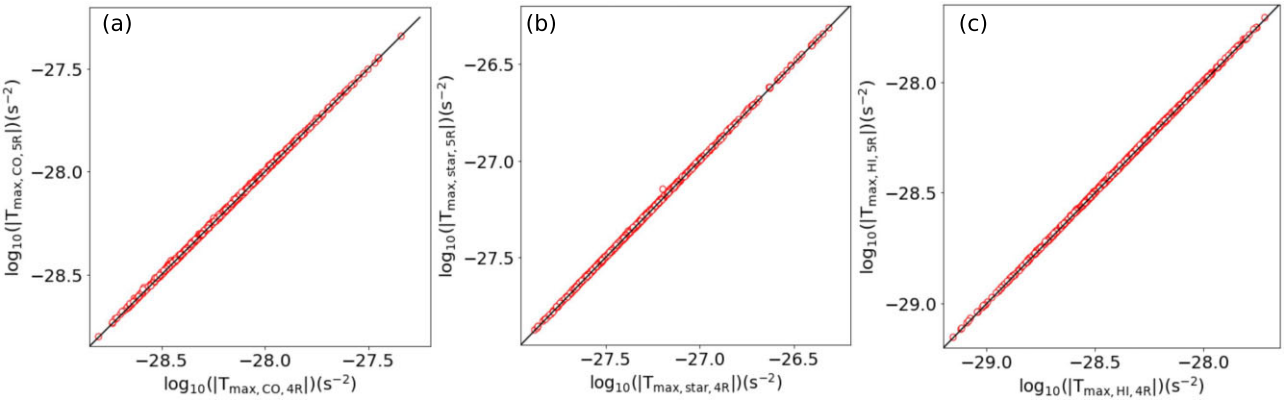


Figure 8. Comparison of the tidal strength exerted on molecular clouds by neighbouring material within four times and five times the effective radius of the clouds.

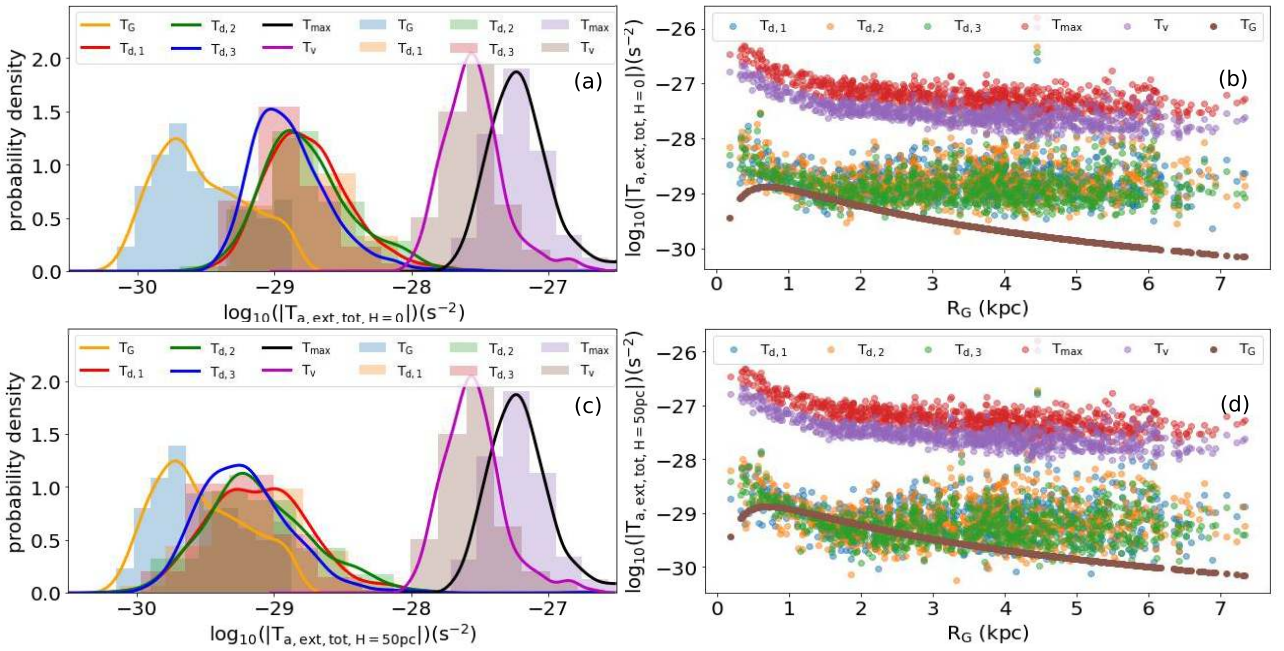


Figure 9. Same as Fig. 5, but all tidal strengths derived from the total surface density map, rather than only the CO surface density map in Fig. 5.

Table 2. Explanation of the physical quantities. The tidal strength is the tidal acceleration per unit distance. In the tidal tensor analysis, the tidal strength is the numerical value of a specific component on the tidal tensor.

R_G	The distance to the galactic centre.
$\mathrm{Tr}(\mathbf{T}_{\mathrm{int}})$	The trace of 2D tidal tensor derived from the CO surface density in Section 3.4.1.
$E_{\mathrm{a,int,H}}$	The energy corresponding to $\mathrm{Tr}(\mathbf{T}_{\mathrm{a,int,H}})$ averaged over a cloud. H is the half-thickness of the galaxy. There are two cases, i.e. $H = 0$ and $H = 50$ pc.
E_{g}	Self-gravity energy of a cloud.
T_G	The tidal strength calculated using the rotation curve by equation (5).
$T_{\mathrm{d},1}$, $T_{\mathrm{d},2}$, and $T_{\mathrm{d},3}$	Three extreme cases of the tidal strengths obtained from the decomposition of the tidal tensor.
T_{max} and T_{v}	The tidal strengths computed pixel-by-pixel on 2D plane in the cases of scalar and vector superpositions, respectively, see Section 3.4.2 for more details.

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

3.5.3 Scale height and scalar versus vector superpositions

In the two cases mentioned above, $H = 50$ pc generates lower tidal strengths than $H = 0$, but is otherwise identical to $H = 0$ in all other respects. As described in Section 3.4.2, $T_{\rm v,tot}$ and $T_{\rm max,tot}$ are calculated in ways that are quite different, but their trends of variation with respect to $R_{\rm G}$ are consistent. The only difference is in the magnitude of the tidal strengths.

3.5.4 Tidal tensor analysis versus pixel-by-pixel computation

The pixel-by-pixel computation in Section 3.4.2 can reflect the variation in tidal strength with $R_{\rm G}$ more precisely than the tidal tensor analysis in Section 3.4.1, as it has a smaller scatter in tidal strength, as shown in Fig. 9. However, to accurately estimate the tidal strength using the pixel-by-pixel computation, we need to know the specific 3D spatial distribution of the material in the galaxy, which is challenging.

3.5.5 Consequences of the tidal analysis

Above, we conducted an extensive discussion aimed at summarizing an effective tidal analysis method for future research. To explain the variation in the density contrast of molecular clouds with $R_{\rm G}$ in NGC 628, it is necessary to integrate a total (baryonic matter) surface density map that includes molecular gas, atomic gas, and stellar components. This is because the stellar mass surface density is significantly higher than the molecular gas surface density, indicating stronger tidal effects. The pixel-by-pixel computation can more precisely reflect the variation in tidal strength with $R_{\rm G}$ than the tidal tensor analysis but cannot provide a correct magnitude estimate of the tidal strength due to the lack of knowledge about the specific 3D spatial distribution of the material in the galaxy. This issue can be resolved using the tidal tensor analysis. Therefore, by combining these two methods, we can achieve both appropriate variation and magnitude estimates of the tidal strength across the galaxy.

4 DISCUSSION AND CONCLUSION

The physical factors influencing the density contrast of molecular clouds are the same ones that impact star formation in the galaxy. The proportion of high- and low-density contrast clouds initially increases with the distance to the galactic centre (R_G) and then keeps relatively stable. The spiral arms, bubbles, and magnetic fields are not responsible for the variations in density contrast observed among molecular clouds. The galactic shear and tidal strength calculated by the rotation curve consistently decrease as R_G increases, inconsistent with the variation of the proportion between high- and low-density contrast clouds with R_G.

We studied the tidal effects of neighbouring material on each cloud using two different methods, i.e. the tidal tensor analysis and the pixel-by-pixel computation. All calculations were carried out on 2D plane, but the computed 2D tidal field strength ($H = 0$) in the tidal tensor analysis can roughly reflect the real 3D tidal field strength of the galaxy. In the pixel-by-pixel computation, we consider both scalar and vector superpositions of tides. The tidal effects of material beyond four times the effective radius of a cloud can be neglected. Although we focus solely on molecular clouds as sites of star formation in the galaxy, it is inadequate to only consider molecular gas when accounting for external tidal effects on the clouds. We constructed a total surface density map by integrating molecular gas, atomic gas, and stellar mass surface density maps to estimate the distribution of baryonic matter within the galaxy.

In all cases, assuming a galaxy half-thickness of $H = 50$ pc results in lower tidal strengths compared to $H = 0$, while remaining otherwise identical to $H = 0$ in every other aspect. In both tidal tensor analysis and pixel-by-pixel computation, the tidal strengths calculated according to the CO (2–1) surface density map do not vary with R_G, because the galactic centre does not have significantly higher CO surface density. However, for the tidal strengths calculated based on the total surface density map, all of them change with R_G. The tidal strengths derived from the pixel-by-pixel computation have smaller scatter than those obtained from tidal tensor analysis and can reflect the variation in tidal strength with $R_{\rm G}$ more precisely. When $R_{\rm G} < 4$ kpc, the tidal strengths calculated pixel-by-pixel decrease as $R_{\rm G}$ increases, and then stabilize for $R_{\rm G} > 4$ kpc. This trend aligns closely with the variation in the proportion of high- and low-density contrast clouds with $R_{\rm G}$. Therefore, the tidal effects of neighbouring material have a significant impact on the development of molecular cloud's density contrast, which in turn affects star formation within molecular cloud.

In Fig. 9, T_G calculated by the rotation curve consistently decreases with increasing $R_{\rm G}$. $T_{\rm v,tot}$ and $T_{\rm max,tot}$ represent local tidal strengths from neighbouring material and are only sensitive to local material distribution. As shown in Fig. 1, the surface density is significantly higher in galactic centre. Therefore, $T_{\rm v,tot}$ and $T_{\rm max,tot}$ only decrease with increasing $R_{\rm G}$ in the central region of the galaxy. And they remain relatively constant in the outer regions of the galaxy, as the surface density distribution in the outer regions is relatively uniform. Therefore, one important reason for the low star formation rate in the galactic centre is that the tidal effects from neighbouring material acting on molecular clouds are strong, restricting the gravitational collapse of the clouds. As a result, the clouds exhibit low density contrasts.

Molecular clouds do not evolve in isolation, their gravitational collapse is modulated by various external physical processes. Although the measured velocity gradients in Zhou & Davis (2024) and Zhou et al. (2024a) support the gravitational collapse of gas structures on galaxy-cloud scales, the collapse is significantly slower than a pure free-fall gravitational collapse. In the case of free-fall, the velocity gradient (∇v) and the scale (R) satisfy $\nabla v \propto R^{-1.5}$. However, for NGC 628, the fitting in Zhou et al. (2024a) was $\nabla v \propto R^{-0.9}$. For NGC 4321 and NGC 5236, Zhou & Davis (2024) obtained a similar slope, i.e. $\nabla v \propto R^{-0.8}$. As discussed in Zhou & Davis (2024), the slowing down of a pure free-fall gravitational collapse can be caused by multiple physical processes. The results of this work suggest that the tidal effects of neighbouring material can resist the gravitational collapse of molecular clouds.

This work is merely a case study, and future work needs to apply similar methods to large galaxy samples containing different types of galaxies to verify the generality of the conclusions presented here.

ACKNOWLEDGEMENTS

Thanks to the referee for providing detailed review comments, which have helped to improve and clarify this work. We would like to thank R. Beck for sharing the non-thermal emission map of NGC 628 with us. It is a pleasure to thank the PHANGS team, the data cubes, and other data products shared by the team so that this work can be carried out easily. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00650.S and ADS/JAO.ALMA#2017.1.00886.L. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan),

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

together with NRC (Canada), NSTC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile.

DATA AVAILABILITY

All the data used in this work are available on the PHANGS team website.[1]

REFERENCES

Aouad C. J., James P. A., Chilingarian I. V., 2020, MNRAS, 496, 5211
Ballesteros-Paredes J., Gómez G. C., Loinard L., Torres R. M., Pichardo B., 2009, MNRAS, 395, L81
Barnes A. T. et al., 2022, A&A, 662, L6
Beck R., Krause M., 2005, Astron. Nachr., 326, 414
Bronfman L., Cohen R. S., Alvarez H., May J., Thaddeus P., 1988, ApJ, 324, 248
Chevance M. et al., 2022, MNRAS, 509, 272
Chevance M., Krumholz M. R., McLeod A. F., Ostriker E. C., Rosolowsky E. W., Sternberg A., 2023, in Inutsuka S., Aikawa Y., Muto T., Tomida K., Tamura M., eds, ASP Conf. Ser. Vol. 534, Protostars and Planets VII. Astron. Soc. Pac., San Francisco, p. 1
Colombo D. et al., 2014, ApJ, 784, 3
Cox D. P., 2005, ARA&A, 43, 337
Crutcher R. M., 2012, ARA&A, 50, 29
Dale D. A. et al., 2009, ApJ, 693, 1821
Dib S., Walcher C. J., Heyer M., Audit E., Loinard L., 2009, MNRAS, 398, 1201
Dib S., Helou G., Moore T. J. T., Urquhart J. S., Dariush A., 2012, ApJ, 758, 125
Ellsworth-Bowers T. P. et al., 2013, ApJ, 770, 39
Fukui Y., Habe A., Inoue T., Enokiya R., Tachihara K., 2021, PASJ, 73, S1
Ganguly S., Walch S., Seifried D., Clarke S. D., Weis M., 2023, MNRAS, 525, 721
Gong H., Ostriker E. C., 2011, ApJ, 729, 120
He Z.-Z., Li G.-X., Burkert A., 2023, MNRAS, 526, L20
Hughes A. et al., 2013, ApJ, 779, 46
Hunter D. A., Elmegreen B. G., Baker A. L., 1998, ApJ, 493, 595
Ibáñez-Mejía J. C., Mac Low M.-M., Klessen R. S., 2022, ApJ, 925, 196
Jeffreson S. M. R., Kruijssen J. M. D., Keller B. W., Chevance M., Glover S. C. O., 2020, MNRAS, 498, 385
Jeffreson S. M. R., Sun J., Wilson C. D., 2022, MNRAS, 515, 1663
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Kim J. et al., 2022, MNRAS, 516, 3006
Klessen R. S., Hennebelle P., 2010, A&A, 520, A17
Kruijssen J. M. D., Longmore S. N., Elmegreen B. G., Murray N., Bally J., Testi L., Kennicutt R. C., 2014, MNRAS, 440, 3370
Krumholz M. R. et al., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. The University of Arizona Press, Tucson, p. 243
Lang P. et al., 2020, ApJ, 897, 122
Leroy A. K. et al., 2021a, ApJS, 255, 19
Leroy A. K. et al., 2021b, ApJS, 257, 43
Li H.-B., Henning T., 2011, Nature, 479, 499
Liu L., Bureau M., Blitz L., Davis T. A., Onishi K., Smith M., North E., Iguchi S., 2021, MNRAS, 505, 4048
Meidt S. E. et al., 2013, ApJ, 779, 45
Meidt S. E. et al., 2020, ApJ, 892, 73
Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, ApJ, 834, 57
Mulcahy D. D., Beck R., Heald G. H., 2017, A&A, 600, A6
Pessa I. et al., 2022, A&A, 663, A61
Querejeta M. et al., 2021, A&A, 656, A133
Querejeta M. et al., 2024, A&A, 687, A293
Ramírez-Galeano L., Ballesteros-Paredes J., Smith R. J., Camacho V., Zamora-Avilés M., 2022, MNRAS, 515, 2822
Rawat V. et al., 2024, MNRAS, 528, 1460
Sano H. et al., 2021, PASJ, 73, S62
Schinnerer E., Leroy A. K., 2024, ARA&A, 62, 369
Stark A. A., Blitz L., 1978, ApJ, 225, L15
Stark A. A., Lee Y., 2005, ApJ, 619, L159
Sun J. et al., 2022, AJ, 164, 43
Tasker E. J., Tan J. C., 2009, ApJ, 700, 358
Thilliez E., Maddison S. T., Hughes A., Wong T., 2014, Publ. Astron. Soc. Aust., 31, e003
Utreras J. et al., 2020, ApJ, 892, 94
Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt Robert C. J., Thornley M. D., Leroy A., 2008, AJ, 136, 2563
Watkins E. J. et al., 2023, ApJ, 944, L24
Zhou J.-W., Davis T., 2024, Publ. Astron. Soc. Aust., 41, e076
Zhou J.-W. et al., 2023, MNRAS, 519, 2391
Zhou J. W., Dib S., Davis T. A., 2024a, MNRAS, 534, 683
Zhou J. W., Dib S., Juvela M., Sanhueza P., Wyrowski F., Liu T., Menten K. M., 2024b, A&A, 686, A146

APPENDIX: SUPPLEMENTARY MAPS

[1] https://sites.google.com/view/phangs/home.

Downloaded from https://academic.oup.com/mnras/article/537/2/2232/7998935 by guest on 12 February 2025

Original papers of Chapter.4

The original paper of Sec.4.1,

> Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2024, A&A, 682, A128
> *High-resolution APEX/LAsMA 12CO and 13CO (3–2) observation of the G333 giant molecular cloud complex II. Survival and gravitational collapse of dense gas structures under feedback*

The original paper of Sec.4.2,

> Zhou, J. W., Dib, S., Wyrowski, F., et al. 2024, A&A, 682, A173
> *Feedback from protoclusters does not significantly change the kinematic properties of the embedded dense gas structures*

A&A, 682, A128 (2024)
https://doi.org/10.1051/0004-6361/202347377
© The Authors 2024

Astronomy & Astrophysics

High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3–2) observation of the G333 giant molecular cloud complex

II. Survival and gravitational collapse of dense gas structures under feedback

J. W. Zhou[1], F. Wyrowski[1], S. Neupane[1], I. Barlach Christensen[1], K. M. Menten [1], S. H. Li[2], and T. Liu[3]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de
[2] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[3] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, PR China

Received 6 July 2023 / Accepted 7 September 2023

ABSTRACT

Context. Feedback from young massive stars has an important impact on the star formation potential of their parental molecular clouds.
Aims. We investigate the physical properties of gas structures under feedback in the G333 complex using data of the ^{13}CO $J = 3-2$ line observed with the LAsMA heterodyne camera on the APEX telescope.
Methods. We used the Dendrogram algorithm to identify molecular gas structures based on the integrated intensity map of the ^{13}CO (3−2) emission, and extracted the average spectra of all structures to investigate their velocity components and gas kinematics.
Results. We derive the column density ratios between different transitions of the ^{13}CO emission pixel by pixel, and find the peak values $N_{2-1}/N_{1-0} \approx 0.5$, $N_{3-2}/N_{1-0} \approx 0.3$, and $N_{3-2}/N_{2-1} \approx 0.5$. These ratios can also be roughly predicted by the nonlocal thermodynamic equilibrium (NLTE) molecular radiative transfer code RADEX for an average H_2 volume density of $\sim 4.2 \times 10^3$ cm^{-3}. A classical virial analysis does not reflect the true physical state of the identified structures, and we find that external pressure from the ambient cloud plays an important role in confining the observed gas structures. For high-column-density structures, velocity dispersion and density show a clear correlation that is not seen for low-column-density structures, indicating the contribution of gravitational collapse to the velocity dispersion. Branch structures show a more significant correlation between 8 μm surface brightness and velocity dispersion than leaf structures, implying that feedback has a greater impact on large-scale structures. For both leaf and branch structures, $\sigma - N * R$ always has a stronger correlation compared to $\sigma - N$ and $\sigma - R$. The scaling relations are stronger, and have steeper slopes when considering only self-gravitating structures, which are the structures most closely associated with the Heyer relation.
Conclusions. Although the feedback disrupting the molecular clouds will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and changes in gravitational centers, but gravitational collapse is always ongoing.

Key words. ISM: structure – ISM: kinematics and dynamics – stars: formation – ISM: clouds – techniques: image processing

1. Introduction

High-mass stars ($M > 8\ M_\odot$) have a profound impact on the evolution of the interstellar medium (ISM). Throughout their short lifetimes ($\sim 10^6$ yr), radiation-driven stellar winds from high-mass stars create HII regions in the surrounding giant molecular clouds (GMCs; Zinnecker & Yorke 2007; Molinari et al. 2014; Motte et al. 2018). High-mass stars end their lives in the form of supernovae (SNe), whose explosions can release $\sim 10^{51}$ erg of energy near-instantaneously. Shocks from expanding HII regions and supernova remnants (SNRs) can accelerate and heat their surrounding gas, and add turbulence to the gas. In simulations, massive stellar feedback, including ionizing radiation, stellar winds, and SNe (Matzner 2002; Dale et al. 2012, 2014; Rogers & Pittard 2013; Rahner et al. 2017; Smith et al. 2018; Lewis et al. 2023), can suppress the star formation and destroy the natal cloud. However, whether the stellar feedback promotes or suppresses star formation remains controversial.

W49A is one of the most massive and luminous young star-forming regions in the Galaxy. As presented in Rugel et al. (2019), it is more likely that only limited parts of W49A were affected by feedback from the central stellar cluster, while stars in the outer parts of W49A formed independently. Moreover, all feedback models used in Rugel et al. (2019) predict recollapse of the shell after the first star formation event, which means that feedback from the first formed cluster is therefore not strong enough to disperse the cloud. Previous work on the G305 region observed with the Large APEX sub-Millimeter Array (LAsMA) 7 beam receiver on the Atacama Pathfinder Experiment 12 meter submillimeter telescope (APEX) found that strong stellar winds drive turbulence in the G305. GMC in which feedback has triggered star formation through the collect and collapse mechanism (Mazumdar et al. 2021a,b). The dense molecular gas structures inside the cloud serve as star-forming sites and their physical states directly determine the star-formation capability of the molecular cloud under feedback. A basic question is how the

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

dense gas structures survive and maintain star formation activity in a strong feedback environment, which depends on the relative strengths of the gravity and turbulence therein.

The relative importance of turbulence and gravity in massive star-forming regions is a long and widely debated topic, with different views leading to different physical pictures of massive star formation, such as the turbulent-core model (McKee & Tan 2003; Krumholz et al. 2007), the competitive-accretion model (Bonnell et al. 1997, 2001), the inertial-inflow model (Padoan et al. 2020), and the global hierarchical collapse model (Vázquez-Semadeni et al. 2009, 2017, 2019; Ballesteros-Paredes et al. 2011; Hartmann et al. 2012). Larson's laws claim that, in molecular clouds, the velocity dispersion σ scales proportionally to the scale R, and molecular clouds are approximately in virial equilibrium, with a mostly uniform column density. The Larson relation ($\sigma_v \propto R^{0.5}$) is generally used to emphasize the importance of turbulence in molecular clouds, and turbulence acts to sustain the clouds against gravitational collapse (Larson 1981; Solomon et al. 1987; Heyer & Brunt 2004; Mac Low & Klessen 2004; McKee & Ostriker 2007; Hennebelle & Falgarone 2012). Heyer et al. (2009) generalized the Larson relation by extending the Larson ratio $L \equiv \sigma_v/R^{0.5}$ with the surface density Σ of Galactic GMCs, which gives $\sigma_v/R^{0.5} \propto \Sigma^{0.5}$. Subsequently, the Heyer relation was used to emphasize the importance of gravity in molecular clouds (Ballesteros-Paredes et al. 2011, 2018, 2020; Traficante et al. 2018b,a; Vázquez-Semadeni et al. 2019). In particular, in the high-column-density portions of star-forming regions, such as clumps or cores, the Heyer relation always performs better than the Larson relation (Ballesteros-Paredes et al. 2011; Camacho et al. 2016; Traficante et al. 2018a), suggesting strong gravity at relatively small scales in molecular clouds. Ibáñez-Mejía et al. (2016) showed that the Heyer relation cannot be reproduced without self-gravity in simulations of the ISM with SN-driven turbulence. In contrast, purely SN-driven turbulence in the ISM generates the Larson relation.

Regarding the explanation of the Heyer relation, Heyer et al. (2009) claimed that it is consistent with the clouds being in virial equilibrium, as it follows directly from the condition $2E_k = E_g$, where $E_k = M\sigma^2/2$ and $E_g = 3GM^2/5R$. Ballesteros-Paredes et al. (2011) further pointed out that the scaling is also consistent with the clouds undergoing free-fall, in which case $E_k = |E_g|$. However, the differences between the effects of free-fall and virial equilibrium in the $\sigma_v/R^{0.5}$ versus Σ diagram are smaller than the typical uncertainty of the observational data (Ballesteros-Paredes et al. 2011), and are therefore difficult to distinguish. Both explanations involve only gravitational and kinetic energy, which may be a workable approximation for relatively isolated molecular clouds, but is often too simplistic for substructures inside a molecular cloud, especially for a cloud affected by feedback, such as our target G333. When the substructures can self-gravitationally collapse, they may decouple from the surrounding environment (Peretto et al. 2023). If not, the exchange of energy with the surrounding environment will break the conversion between gravitational potential energy, E_g, and kinetic energy, E_k, of the structures, and thus violate the Heyer relation.

In Zhou et al. (2023), we found in the G333 complex that the larger-scale inflow is driven by the larger-scale cloud structure, indicating hierarchical structure in the GMC and gas inflow from large to small scales. The large-scale gas inflow is driven by gravity, implying that the molecular clouds in the G333 complex may be in a state of global gravitational collapse. However, the broken morphology of some very infrared(IR)-bright structures in the G333 complex also indicates that feedback is disrupting star-forming regions. In the present study, we address the question of how the dense molecular structures survive and maintain the gravitational collapse state in a strong feedback environment.

2. Data

2.1. LAsMA data

The observations and data reduction are described in detail in Zhou et al. (2023). We mapped a $3.4° \times 1.2°$ area centered at $(l, b) = (332.33°, -0.29°)$ using the APEX telescope (Güsten et al. 2006)[1]. The 7 pixel Large APEX sub-Millimeter Array (LAsMA) receiver was used to observe the $J = 3-2$ transitions of ^{12}CO ($\nu_{\rm rest} \sim 345.796$ GHz) and ^{13}CO ($\nu_{\rm rest} \sim 330.588$ GHz) simultaneously. The local oscillator frequency was set at 338.190 GHz in order to avoid contamination of the ^{13}CO (3−2) spectra due to bright ^{12}CO (3−2) emission from the image side band. Observations were performed in a position switching on-the-fly (OTF) mode. The data were calibrated using a three-load chopper wheel method, which is an extension of the "standard" method used for millimeter observations (Ulich & Haas 1976) to calibrate the data in the corrected antenna temperature T_A^* scale. The data were reduced using the GILDAS package[2]. The final data cubes have a velocity resolution of 0.25 km s^{-1}, an angular resolution of 19.5″, and a pixel size of 6″. A beam efficiency value $\eta_{\rm mb} = 0.71$ (Mazumdar et al. 2021a) was used to convert intensities from the T_A^* scale to main beam brightness temperatures, $T_{\rm mb}$.

2.2. Archival data

To allow column-density estimates using different ^{13}CO transitions, we also collect ^{13}CO (1−0) data from the Mopra-CO survey (Burton et al. 2013) and ^{13}CO (2−1) from the SEDIGISM survey (Schuller et al. 2021). However, the two surveys only cover Galactic latitudes within ±0.5°, while our LAsMA observations cover the latitude range of approximately $-0.29 \pm 0.6°$. We therefore only consider the overlap region of the three surveys. The data were smoothed to a common angular resolution of ∼35″ and a velocity resolution of ∼0.25 km s^{-1}.

The observed region was covered in the IR range by the Galactic Legacy Infrared Midplane Survey (GLIMPSE, Benjamin et al. 2003). GLIMPS images, obtained with the *Spitzer* Infrared Array Camera (IRAC) at 4.5 and 8.0 μm, were retrieved from the *Spitzer* archive. The angular resolution of the images in the IRAC bands is ∼2″. We also used 870 μm continuum data from the APEX Telescope Survey of the Galaxy (ATLASGAL, Schuller et al. 2009) combined with lower-resolution data from the *Planck* spacecraft, which are sensitive to a wide range of spatial scales at a resolution of ∼21″ (Csengeri et al. 2016). Furthermore, we use Hi-GAL data (Molinari et al. 2010) processed using the PPMAP procedure (Marsh et al. 2015), which provides column density and dust temperature maps with a resolution of ∼12″ (the maps are available online[3], Marsh et al. 2017).

[1] This publication is based on data acquired with the Atacama Pathfinder EXperiment (APEX) while it has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory and the Onsala Space Observatory.
[2] http://www.iram.fr/IRAMFR/GILDAS
[3] http://www.astro.cardiff.ac.uk/research/ViaLactea/

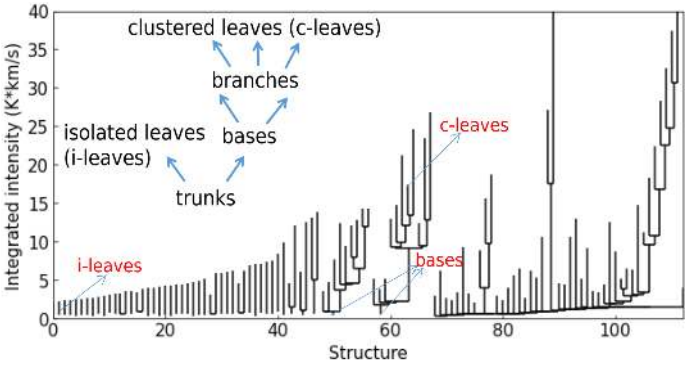


Fig. 1. Hierarchical structures identified by the Dendrogram algorithm. A segment of the Dendrogram tree of subregion S3 in Fig. 2c is used to illustrate the structure types output by the Dendrogram algorithm.

3. Results

3.1. Dendrogram structures

We identify dense gas structures using the Dendrogram algorithm. As described in Rosolowsky et al. (2008), the Dendrogram algorithm decomposes intensity data (a position–position map or a position–position–velocity cube) into hierarchical structures called leaves, branches, and trunks. The relationship between those structures is shown in Fig. 1. Trunks are the largest continuous structures at the bottom of hierarchical structures ("bases"), but by definition they can also be isolated leaves ("i-leaves") without any parent structure. Thus, there are two kinds of "trunks", called "bases" and "i-leaves" in this work. Clustered leaves ("c-leaves") are defined as small-scale, bright structures at the top of the tree that do not decompose into further substructures; they are the smallest structures inside "branches". Branches are the relatively large-scale structures in the tree, and can be broken down into substructures. Between "bases" and "c-leaves", all hierarchical substructures are "branches", and therefore branches can span a wide range of scales. When we treat "bases" as the largest branches, and combine c-leaves and i-leaves, then there are only two kinds of structures (i.e. leaves and branches). However, in some cases, it is necessary to differentiate between i-leaves and c-leaves. In general, c-leaves are concentrated in regions of relatively high column density, while i-leaves are low-column-density structures distributed at the periphery, as shown in Fig. 2. There are no definite limits to the size of the structures at different levels. The size of a leaf structure in a low-column-density region may be larger than a branch structure in a high-column-density region. As shown in Fig. 3a, there is considerable overlap of the scales between leaf and branch structures. In general, branches are larger-scale structures than leaves.

Using the astrodendro package[4], there are three major input parameters for the Dendrogram algorithm: min_value for the minimum value to be considered in the dataset, min_delta for a leaf that can be considered as an independent entity, and min_npix for the minimum area of a structure. From these parameters, we can see that the algorithm does not consider the velocity component or the velocity range of the identified structure carefully. The structure is mainly identified according to the intensity, meaning that the velocity division of a structure is only a result of its intensity division. There is no criterion for a continuous velocity range across a dense structure. However, the velocity range of a structure is crucial for the estimation of its fundamental physical quantities, such as velocity dispersion and mass. Moreover, strict differentiation of the velocity components should be based on the spectral line profiles rather than the intensity thresholds in the algorithm. In this work, instead of identifying structures in the PPV cube, we first identify the intensity peaks on the integrated intensity (Moment 0) map of ^{13}CO (3−2) emission, and then extract the average spectrum of each structure to investigate their velocity components and gas kinematics. For the Moment 0 map, a 5σ threshold has been set, meaning that we therefore only require the smallest area of a structure to be larger than one beam and do not set any other parameters, thereby reducing the dependence of structure identification on the algorithm parameters. In Fig. 4, the structures identified by the Dendrogram algorithm correspond well to the peaks on the integrated intensity maps. In order to retain as many structures as possible, the parameter min_npix was set to one beam, because the hierarchical structures in Dendrogram mean that a leaf structure under strict parameter settings can be a branch structure under loose parameter settings. Moreover, the average spectra fitting described below also further screens the structures, allowing us to exclude structures with poorly defined line profiles.

The algorithm approximates the morphology of each structure as an ellipse, which is used in this work. We do not use the mask output by the algorithm because different parameter settings around the intensity peak will give different masks. In the Dendrogram algorithm, the long and short axes of an ellipse a and b are the rms sizes (second moments) of the intensity distribution along the two spatial dimensions. However, as shown in Fig. 5, a and b will give a smaller ellipse compared to the size of the identified structure. We therefore tried to enlarge the ellipse by 2 and 3 times, and found that multiplying by a factor of 2 is appropriate, similar to the factor of 1.91 suggested by Solomon et al. (1987) and Rosolowsky & Leroy (2006). The effective physical radius of an ellipse is then $R_{\rm eff} = \sqrt{2a \times 2b} * d$, with d = 3.6 kpc for the distance to the G333 complex (Lockman 1979; Bains et al. 2006).

3.2. Velocity components

Based on the Moment 0 map of the ^{13}CO (3−2) emission, 3608 structures are extracted by the Dendrogram algorithm, consisting of 1626 clustered leaves, 1367 branches, and 615 trunks (486 isolated leaves and 129 bases). In the discussion below, we put bases into branches. We extract and fit the averaged spectra of 3608 structures individually using the fully automated Gaussian decomposer GAUSSPY+ (Lindner et al. 2015; Riener et al. 2019). The parameter settings of the decomposition are the same as in Zhou et al. (2023). According to the line profiles, all averaged spectra are divided into three categories:

1. Structures with single velocity components regarded as independent structures (type1, single, 65%, Fig. 6a);
2. Structures with more than one peak, which are separated (type2, separated, 19%, Fig. 6b);
3. Structures with more than one peak, blended together (type3, blended, 16%, Figs. 6c and d).

Spectra averaged across regions that show a single peak in their line profiles probably represent independent structures. From the line profile, we can also determine the complete velocity range of a structure. In order to ensure that other physical quantities (such as column density, temperature) match

[4] https://dendrograms.readthedocs.io/en/stable/index.html

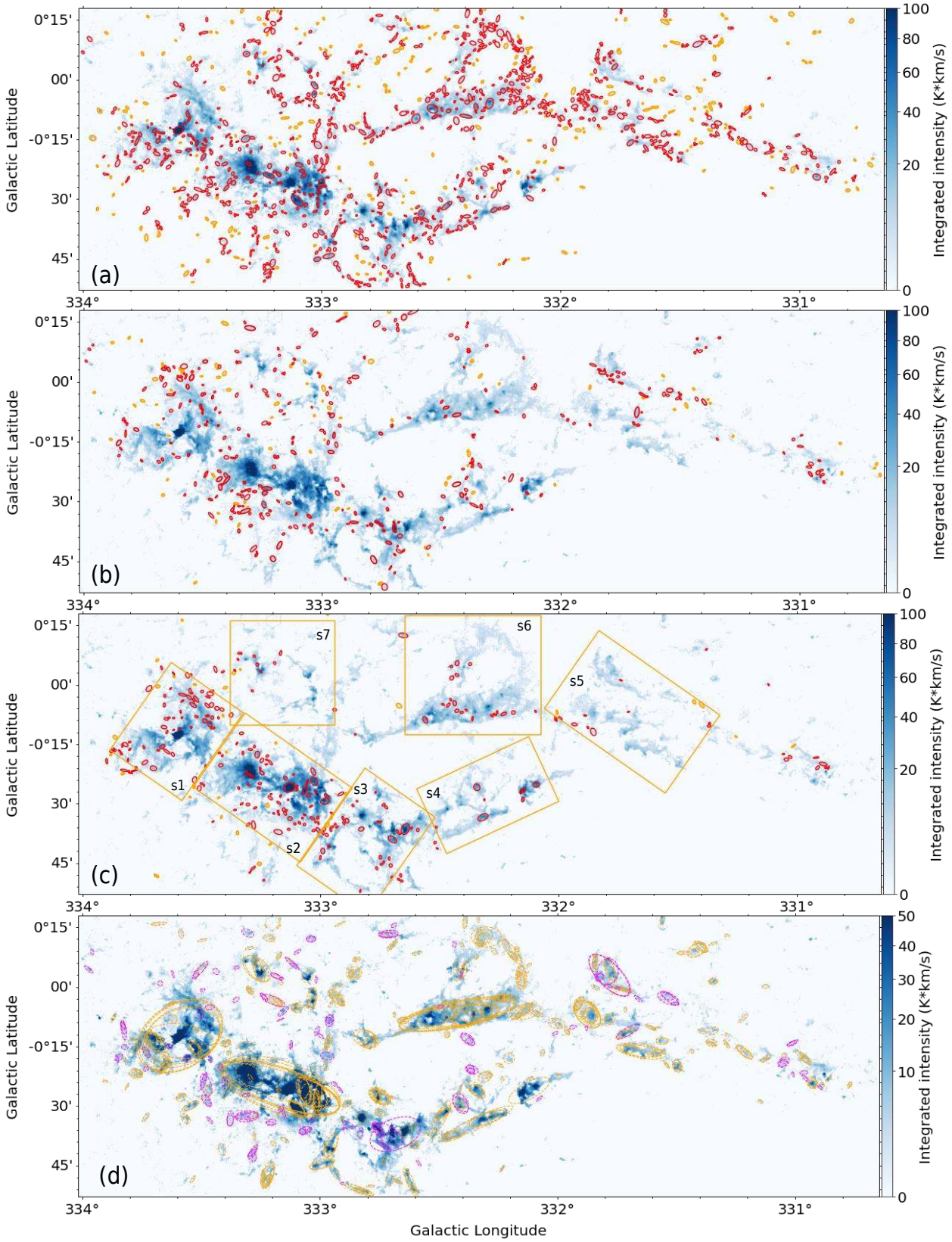


Fig. 2. Different kinds of structures traced by ^{13}CO (3−2) emission classified in Sect. 3.2. (a) Type1 (single velocity component) leaf structures. (b) Type2 (separated velocity components) leaf structures. (c) Type3 (blended velocity components) leaf structures. Orange boxes mark the subregions divided in Zhou et al. (2023). (d) Type1 (orange) and type2 (magenta) branch structures. In panels a–c, orange and red ellipses represent i-leaves and c-leaves, respectively.

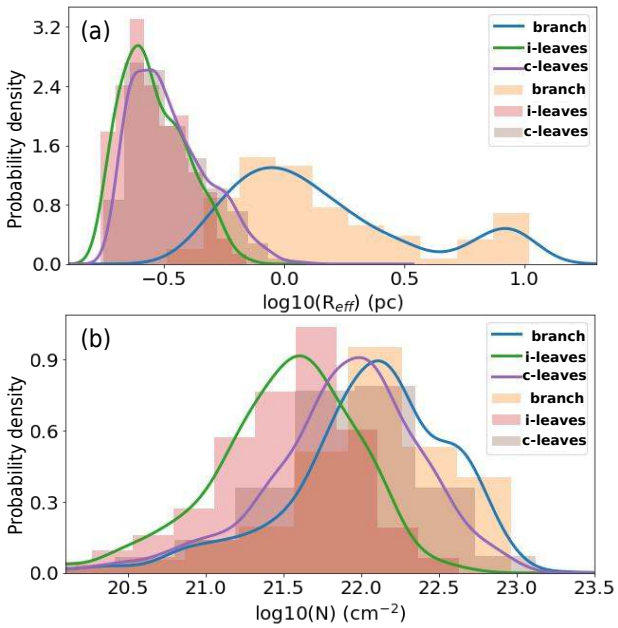


Fig. 3. (a) Effective radius (scale) and (b) column density distribution of Dendrogram structures. Only the type1 and type2 structures (see Fig. 2) are included in the distributions. The probability density is estimated using the kernel density estimation (KDE) method.

the fitted line width, as shown in Fig. 5, we take the velocity range of each structure or each velocity component as [v_c–FWHM, v_c+FWHM], which is necessary to calculate the physical quantities for structures with more than one velocity component.

For type3 structures with significantly overlapping velocity components, complete decomposition cannot easily be obtained, and therefore the decomposition uncertainties directly affect the reliability of the subsequent analysis. In this work, we focus on the structures with independent line profiles (type1 and type2). As shown in Fig. 2, a high-column-density structure does not necessarily imply a complex line profile, and therefore discarding type3 structures will not produce significant sample bias. It is also important to emphasize that for any analysis involving continuum emission without velocity information, only type1 structures will be considered, and subregions 5 and 7 marked in Fig. 2c would also be excluded due to the heavy blending of velocity components described in Zhou et al. (2023).

For a type2 structure, we determine the physical size scales of different velocity components based on their velocity ranges. In Fig. 6b, the total velocity range for deriving the Moment 0 map of the structure is [−60, −35] km s^{-1}, and the area of a type2 structure on Moment 0 map is s and includes n pixels. For two velocity components in a type2 structure, we can also obtain their Moment 0 maps $m01$ and $m02$ in their velocity ranges [v_{c1}–FWHM$_1$, v_{c1}+FWHM$_1$] and [v_{c2}–FWHM$_2$, v_{c2}+FWHM$_2$], respectively. $m01$ and $m02$ contain $n1$ and $n2$ pixels, and their areas are $(n1/n) * s$ and $(n2/n) * s$, which can be used to estimate the physical size scales of the two velocity components.

Generally, the elliptic approximation for the identified structures is good for small-scale leaf structures, but it cannot be satisfactory for some large-scale branch structures because of their complex morphology. We therefore exclude branch structures with complex morphology if the proportion of empty pixels within the effective ellipse of each structure on the Moment 0 map is larger than one-third. Another reason to exclude these morphologically complex structures is that they may not give good effective radius, velocity dispersion, and density estimates. In Fig. 2d, the remaining branch structures correspond well to the background-integrated intensity. For each structure, its velocity range and effective ellipse are used to extract the basic physical quantities based on the column density, temperature, and optical depth cubes derived from the local thermodynamic equilibrium (LTE) analysis in Sect. 3.3.3.

Branch structures are often contained within other branch structures. Some branch structures have similar central coordinates, scales, and morphology, and should therefore be regarded as the same structure to avoid being repeatedly counted. Two branch structures with the area $s1$ and $s2$ ($s1$>$s2$) are considered repetitive if they meet the following conditions: (a) the distance between their central coordinates is less than 1 beam size; and (b) ($s1$–$s2$)/$s2$<1/3. The two clustering conditions can collect the similar branch structures. Each clustering may contain multiple structures, and we keep only one of them in the subsequent analysis. This step will exclude nearly half of branch structures. We should therefore pay more attention to the duplication of branch structures identified by the Dendrogram algorithm before analyzing the identified structures.

3.3. Column density

In this section, we derive the column density of the entire observed field using different methods to find the best estimates for the masses of the identified structures.

3.3.1. Continuum emission

Figures 7a and b present the dust temperature and column density maps, respectively, derived from the Hi-GAL data using the PPMAP procedure (Marsh et al. 2015). As there are some missing values on the PPMAP column density map, we also produced the H_2 column density map using ATLASGAL+*Planck* 870 μm data following the formalism of Kauffmann et al. (2008):

$$N_{H_2} = 2.02\times10^{20}\ \mathrm{cm^{-2}}\left(\mathrm{e}^{1.439(\lambda/\mathrm{mm})^{-1}(T/10\ \mathrm{K})^{-1}}-1\right)\left(\frac{\lambda}{\mathrm{mm}}\right)^3 \cdot\left(\frac{\kappa_\nu}{0.01\ \mathrm{cm^2\ g^{-1}}}\right)^{-1}\left(\frac{F_\nu}{\mathrm{mJy\ beam^{-1}}}\right)\left(\frac{\theta_{\mathrm{HPBW}}}{10\ \mathrm{arcsec}}\right)^{-2}, \quad (1)$$

where F_ν is the flux density, θ_{HPBW} is the beam FWHM, and κ_ν = 0.0185 cm^2 g^{-1} (Csengeri et al. 2016). Assuming a single dust temperature is a crude simplification, and therefore we calculate N_{H_2} pixel by pixel by combining the ATLASGAL+*Planck* 870 μm flux map with *Herschel* dust temperatures derived with the PPMAP procedure. We only use pixels that are above the ∼5σ noise level, ∼0.3 Jy beam^{-1} (Urquhart et al. 2018). From Figs. 7b and c, we can see that the column density derived from ATLASGAL+*Planck* 870 μm data and the *Herschel* multiwavelength data agree with each other, both in terms of their spatial distribution and magnitude.

3.3.2. Molecular line

In this work, we focus on the G333 complex, and limit the velocity range of the ^{13}CO emission to [−60,−35] km s^{-1} (Zhou et al. 2023). To derive the column densities from the ^{13}CO emission, we assume LTE and a beam filling factor of unity. Following the procedures described in Garden et al. (1991) and

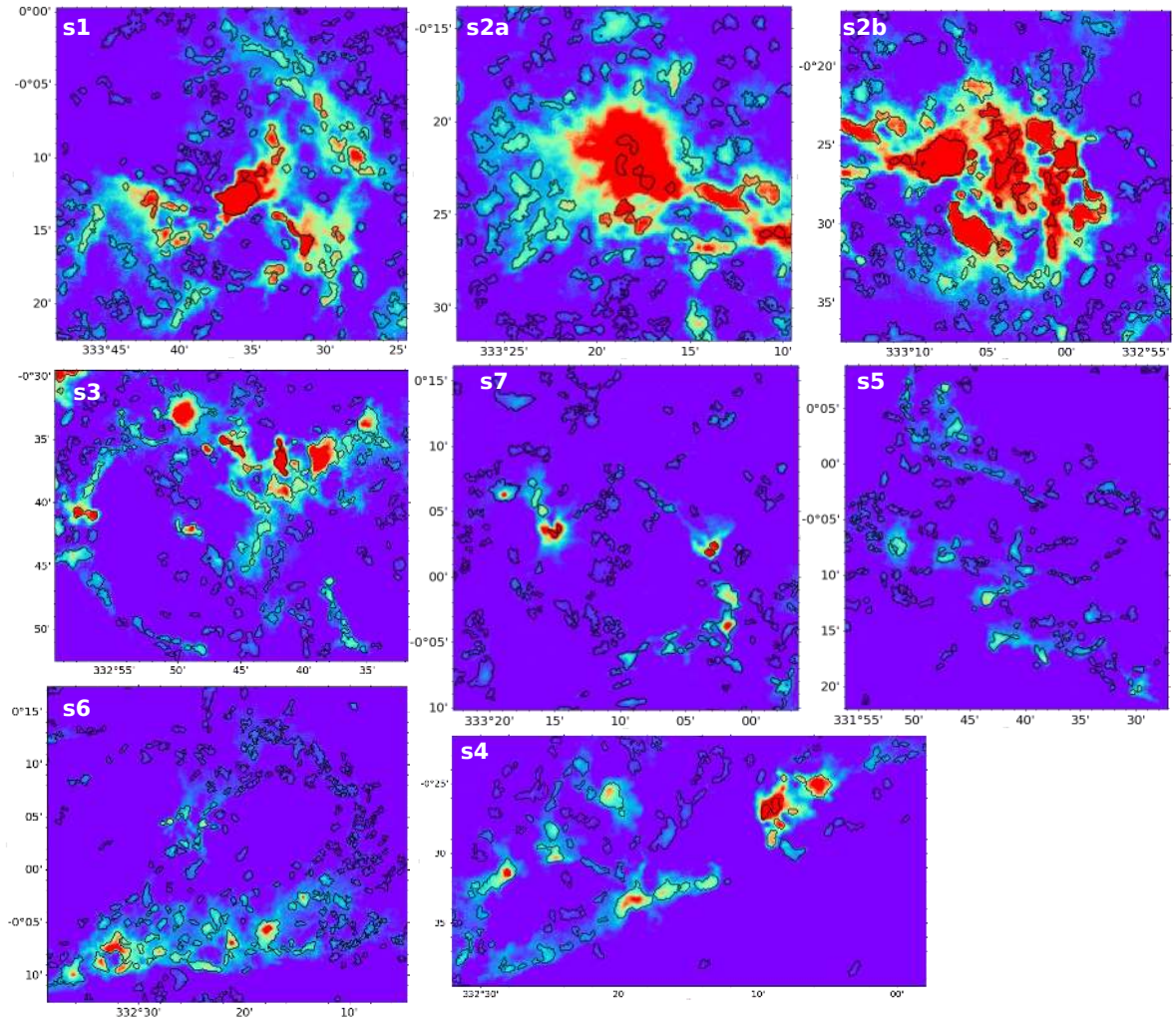


Fig. 4. Masks of leaf structures identified by the Dendrogram algorithm toward the subregions marked in Fig. 2c. Only leaf structures are shown in here.

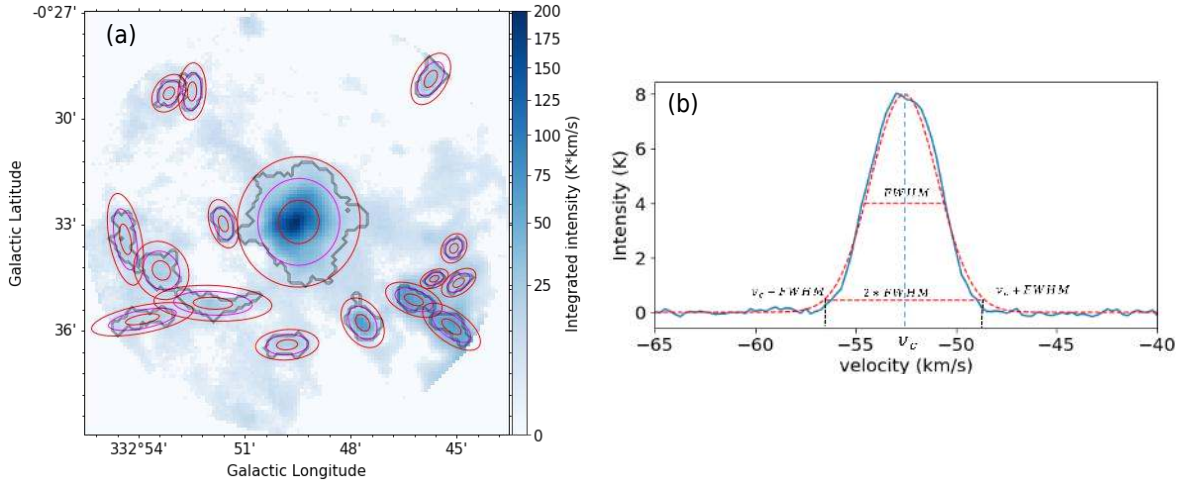


Fig. 5. Part of subregion S3 in Fig. 2c used to illustrate the structures identified by the Dendrogram algorithm. (a) The black contours show the masks of Dendrogram leaves. The long and short axes of the smallest ellipse a and b are the rms sizes (second moments) of the intensity distribution. The ellipses in the second and third layers are enlarged by factors of 2 and 3 in size compared to the smallest one. The middle ellipse visibly corresponds best to the mask. (b) Typical line profile of a leaf structure. The velocity range of the structure is [v_c–FWHM, v_c+FWHM].

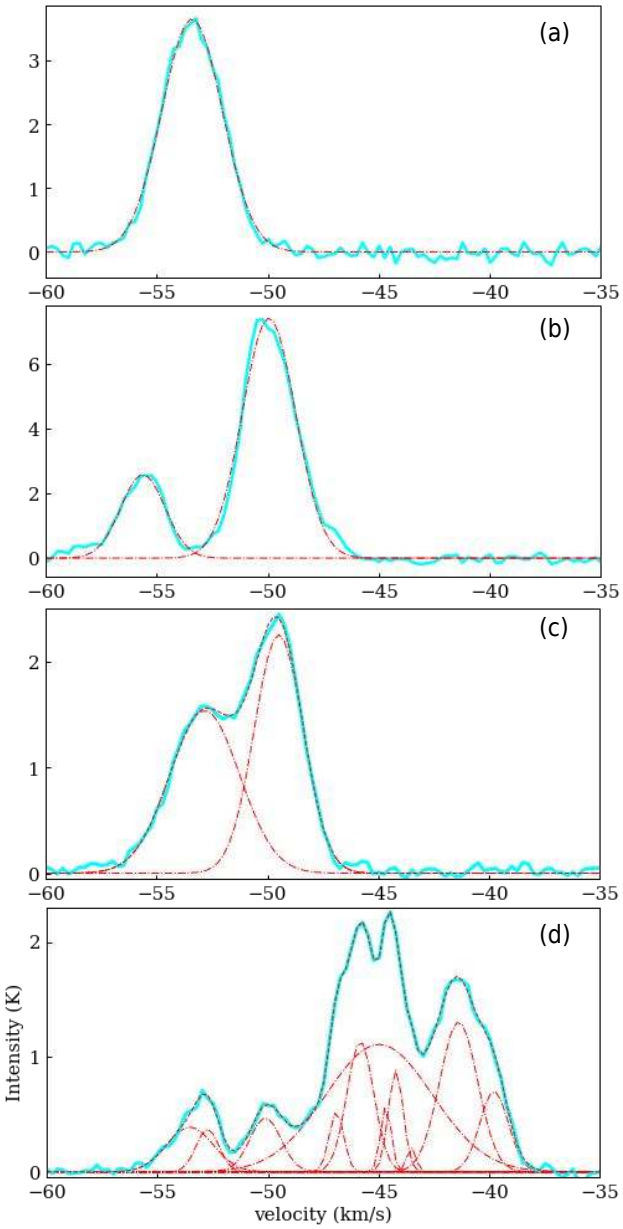


Fig. 6. Typical ^{13}CO (3–2) line profiles of Dendrogram structures.

Mangum & Shirley (2015), for a rotational transition from upper level $J+1$ to lower level J, we can derive the total column density as

$$N_{\rm tot} = \frac{3k}{8\pi^3\mu^2 B(J+1)} \frac{(T_{\rm ex}+hB/3k)\exp[hBJ(J+1)/kT_{\rm ex})]}{1-\exp(-h\nu/kT_{\rm ex})} \times \int \tau {\rm dv}, \quad (2)$$

$$\tau = -\ln[1-\frac{T_{\rm mb}}{J(T_{\rm ex})-J(T_{\rm bg})}], \quad (3)$$

$$\int \tau {\rm dv} = \frac{1}{J(T_{ex})-J(T_{\rm bg})} \frac{\tau}{1-e^{-\tau}} \int T_{\rm mb}{\rm dv}, \quad (4)$$

$$J(T) = \frac{h\nu/k}{e^{h\nu/kT}-1}, \quad (5)$$

where $B = \nu/[2(J+1)]$ is the rotational constant of the molecule and μ is the permanent dipole moment (μ = 0.112 Debye for ^{13}CO). $T_{\rm bg}$ = 2.73 is the background temperature, and $\int T_{\rm mb}{\rm dv}$ represents the integrated intensity. In the above formulas, the correction for high optical depth was applied (Frerking et al. 1982; Goldsmith et al. 2008; Areal et al. 2019). Assuming the ^{12}CO emission to be optically thick, we can estimate the excitation temperature $T_{\rm ex}$ following the formula (Garden et al. 1991; Pineda et al. 2008)

$$T_{\rm ex,3-2} = \frac{16.6{\rm K}}{\ln[1+16.6/(^{12}T_{\rm peak,3-2}+0.038)]}, \quad (6)$$

$$T_{\rm ex,1-0} = \frac{5.53{\rm K}}{\ln[1+5.53/(^{12}T_{\rm peak,1-0}+0.818)]}, \quad (7)$$

where $^{12}T_{\rm peak,3-2}$ and $^{12}T_{\rm peak,1-0}$ are the observed ^{12}CO (3–2) and ^{12}CO (1–0) peak brightness temperature. For the ^{13}CO (2–1) transition, we do not have ^{12}CO (2–1) data and we assume $T_{\rm ex,2-1} = T_{\rm ex,3-2}$.

The distribution of the excitation temperature derived from ^{12}CO (3–2) in Fig. 7d is somewhat similar to the distribution of the dust temperature derived from *Herschel* data shown in Fig. 7a, especially in high-column density regions. We transfer the column densities of ^{13}CO to H_2 column densities by taking the abundance ratio $X_{^{13}CO}$ of H_2 compared with ^{13}CO as $\sim 7.1\times10^5$ (Frerking et al. 1982).

3.3.3. Column density cube

A similar procedure to that presented in Sect. 3.3.2 can be performed for each velocity channel in the ^{13}CO (3–2) cube to obtain a column density cube, which allows us to eliminate the effect of potential overlap of different velocity components on the mass estimation.

3.3.4. Column densities from different ^{13}CO transitions

There are several factors that affect the mass estimate: (1) the overlap of different velocity components; (2) the observed molecular line transition; and (3) the choice between using molecular lines or continuum emission. To address the first factor, we decomposed the velocity components in Zhou et al. (2023) and here we only focus on the peak3 component defined in Zhou et al. (2023) with the velocity range [−60,−35] km s^{-1}. For the second factor, Leroy et al. (2022) measured the low-J ^{12}CO line ratio $R_{21} \equiv {}^{12}$CO (2–1)/^{12}CO (1–0), $R_{32} \equiv {}^{12}$CO (3–2)/^{12}CO (2–1), $R_{31} \equiv {}^{12}$CO (3–2)/^{12}CO (1–0) using whole-disk CO maps of nearby galaxies, and found galaxy-integrated mean values in 16−84% of the emission of R_{21} = 0.65 (0.50−0.83), R_{32} = 0.50 (0.23−0.59), and R_{31} = 0.31 (0.20−0.42). Hence, the 3−2 transition of ^{12}CO resulted in significantly smaller column density estimates compared to the 1−0 transition. To check whether different transitions of ^{13}CO show a similar behavior in a Galactic GMC, we collected ^{13}CO (2−1) and ^{13}CO (1−0) emission of the G333 complex as described in Sect. 2.2.

In Sect. 3.3.2, we derived the column density of different transitions using an LTE analysis. As shown in Fig. 8, the quality of the ^{13}CO J = 1−0 data is not as good as for ^{13}CO J = 2−1 and J = 3−2, and therefore we set a column density threshold of $>10^{21}$ cm^{-2} to exclude the unreliable low-column-density emission from the J = 1−0 transition before the comparison. Figure 9 shows the distribution of pixel-by-pixel column-density

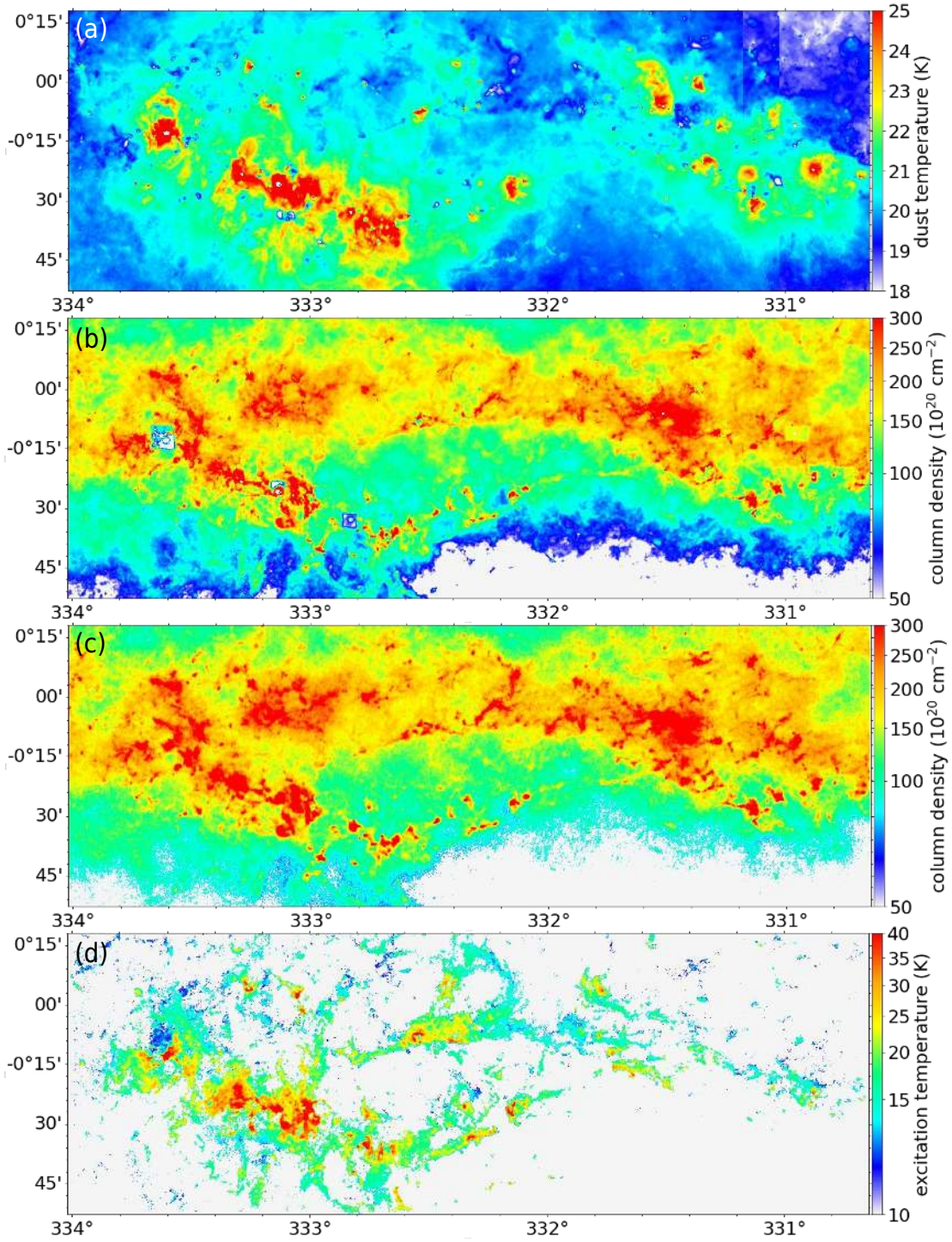


Fig. 7. Temperature and column density maps of the entire field. (a) and (b) Dust temperature and column density distributions in the G333 complex and the G331 GMC derived from Hi-GAL data processed by PPMAP, respectively. (c) Column density distribution in the G333 complex and the G331 GMC derived from ATLASGAL+*Planck* 870 μm data. (d) Excitation temperature distribution in the G333 complex derived from ^{12}CO (3–2) emission by an LTE analysis in the velocity range [−60,−35] km s^{-1}.

ratios between different ^{13}CO transitions. The peak values in the distributions are

$$\frac{N_{2-1}}{N_{1-0}} \approx 0.5, \tag{8}$$

$$\frac{N_{3-2}}{N_{1-0}} \approx 0.3, \tag{9}$$

$$\frac{N_{3-2}}{N_{2-1}} \approx 0.5. \tag{10}$$

Except for the slightly lower ratio between 2−1 and 1−0 transitions, the ratios calculated from different ^{13}CO transitions are comparable with the results derived from ^{12}CO emission in Leroy et al. (2022), although the ratios from ^{12}CO emission are

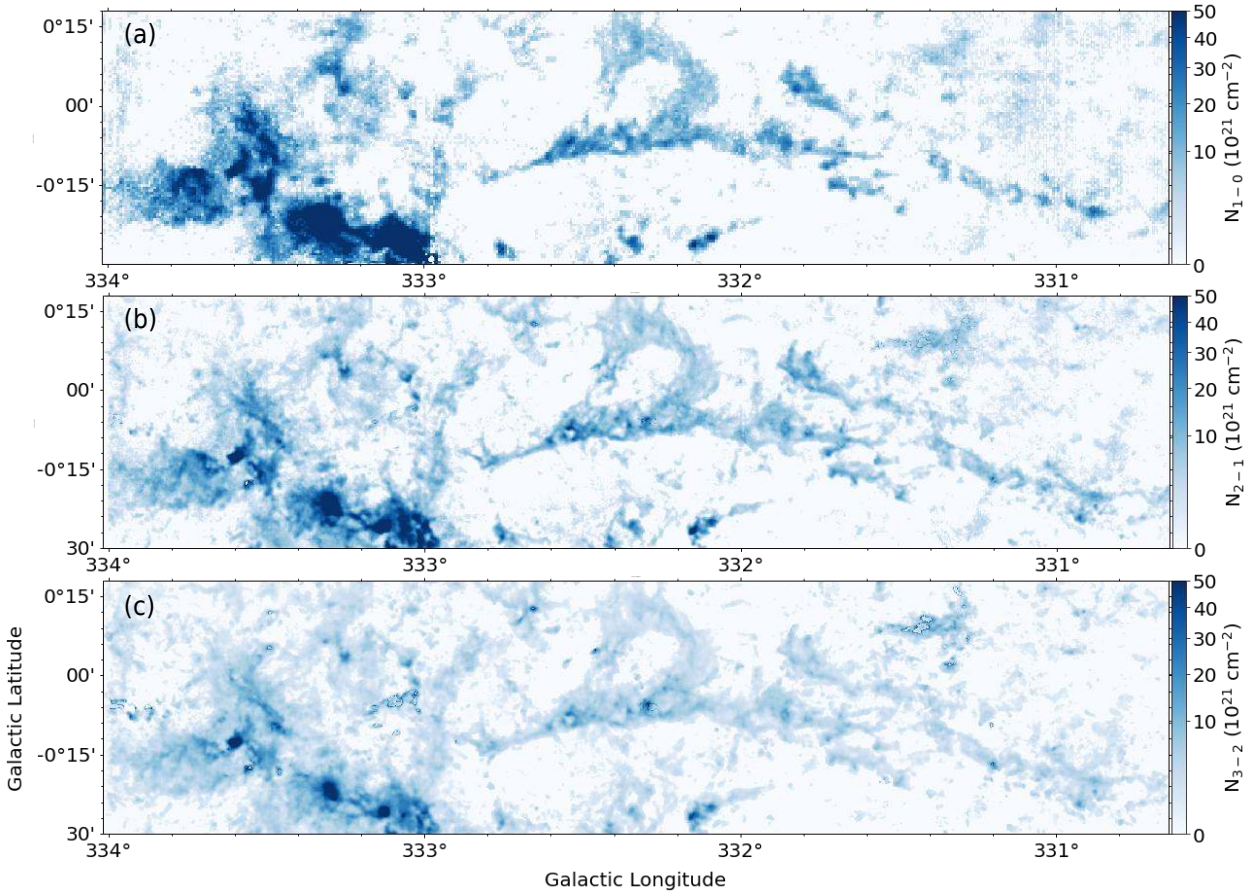


Fig. 8. H_2 column density distribution in the G333 complex derived from ^{13}CO emission. (a) 1–0 transition; (b) 2–1 transition; (c) 3–2 transition.

derived from integrated intensity, rather than from the column density in the case of ^{13}CO.

3.3.5. NLTE estimates

The nonlocal thermodynamic equilibrium (NLTE) molecular radiative transfer algorithm RADEX was used to further test the above results: (1) The column density derived from ^{13}CO $J = 3-2$ transition is significantly lower than $J = 1-0$ transition. (2) The ratios of the column density derived from different transitions of ^{13}CO emission. We use the following input parameters for RADEX: We take $T_{\rm kin}$ = 25 K as the kinematic temperature from Fig. 7a. This is a mean value of the temperature in Fig. 7a for the relatively high-column-density regions in Fig. 7b, which covers the main emission of ^{13}CO J = 3–2. As background temperature, we again use $T_{\rm bg}$ = 2.73 K. A line-width of ~2.5 km s^{-1} is taken from the peak value of all type1 c-leaves, as shown in Fig. 10. For the H_2 volume density $\log_{10}$ ($n_{\rm H_2}$) and ^{13}CO column density $\log_{10}$ ($N_{\rm CO}$), we compute grids in the volume and column density range of [2,6] and [13,17], respectively. We then obtain the intensity $T_{\rm R}$ output from RADEX. Assuming $T_{\rm ex}$= 25 K, using the equations listed in Sect. 3.3.2, for the $J + 1$ to J transition, the column density in the energy level J can be calculated as

$$N_{\rm J} = \frac{2J+1}{Q} N_{\rm tot,CO} \exp[-\frac{hBJ(J+1)}{kT_{\rm ex}}], \quad (11)$$

where the partition function Q is given by $kT_{\rm ex}/hB + 1/3$, B is the rotational constant of the molecule. The rotation temperature $T_{\rm rot}$ can be estimated by the equation

$$\frac{N_{\rm j}}{N_{\rm i}} = \frac{g_{\rm j}}{g_{\rm i}} \exp[-\frac{E_{\rm j} - E_{\rm i}}{kT_{\rm rot}}], \quad (12)$$

where $N_{\rm j}$ and $N_{\rm i}$ are the column densities of any two levels i and j of statistical weights $g_{\rm j}$ and $g_{\rm i}$ and energies $E_{\rm j}$ and $E_{\rm i}$. Using the equations listed in Sect. 3.3.2 again, now the column density $N_{\rm CO,rot}$ can be derived by $T_{\rm rot}$ and $T_{\rm R}$. We also derived the column density $N_{\rm CO,ex}$ by assuming $T_{\rm ex}$= $T_{\rm kin}$ = 25 K. Finally, $N_{\rm CO,rot}$ and $N_{\rm CO,ex}$ are compared with the ^{13}CO column density $N_{\rm CO,radex}$ input in RADEX. As shown in Fig. 11, for CO, ex, using the ^{13}CO (3−2) emission together with the LTE equations indeed gives lower column-density estimates than using the 2−1 and 1−0 emission, which is consistent with our results in Sect. 3.3.4. $N_{\rm CO,J=1-0,ex}$ provides upper limits of the column density derived by different ^{13}CO transitions, and is therefore used to calibrate the column density derived from ^{13}CO (3–2) emission in this work. Generally, for each transition, the column density $N_{\rm CO,rot}$ is higher than $N_{\rm CO,ex}$. The differences of $N_{\rm CO,rot}$ derived from different transitions are also smaller than that of $N_{\rm CO,ex}$. Moreover, $N_{\rm CO,rot}$ is closer to the fiducial $N_{\rm CO,J=1-0,ex}$ than $N_{\rm CO,ex}$. Therefore, using $T_{\rm rot}$ to derive the column density is better than using $T_{\rm ex}$. However, both $J = 2-1$ and $J = 1-0$ data only cover part of the entire observed field, meaning that we cannot obtain the rotational temperature in the full region and therefore do not use it here.

In Fig. 11, we also investigate changes of the column density ratios $N_{\rm CO,ex}$ $J = 3-2/J = 1-0$, $J = 2-1/J = 1-0$, and $J = 3-2/J = 2-1$ with the RADEX input ^{13}CO column density

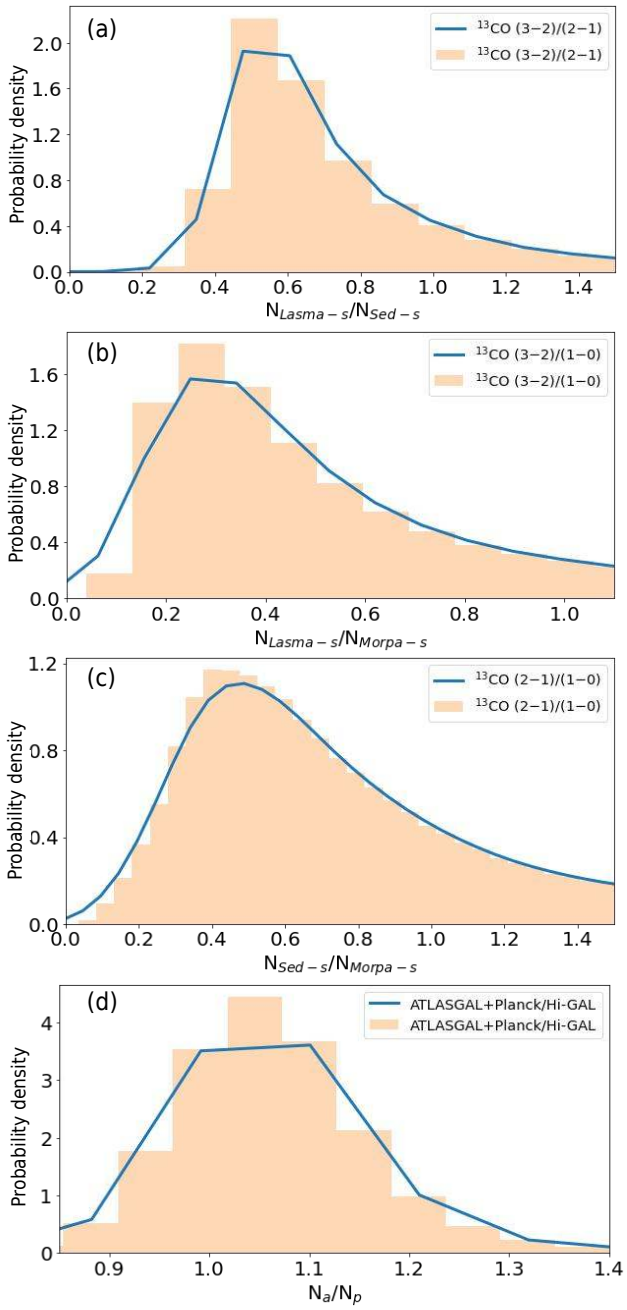


Fig. 9. Distribution of column density ratios. Ratios derived from (a) the ^{13}CO (3–2) and (2–1) emission; (b) the ^{13}CO (3–2) and (1–0); (c) the ^{13}CO (2–1) and (1–0); and (d) the ATLASGAL+*Planck* 870 μm and Hi-GAL data. The probability density is estimated using the KDE method.

$N_{\rm CO,radex}$ for different volume densities $n_{\rm H_2}$. Interestingly, when $n_{\rm H_2}$ is around 4.2×10^3 cm^{-3} (the third column of Fig. 11), the predicted ratios are close to the values in Sect. 3.3.4 for all three ratios. With 4.2×10^3 cm^{-3} in the range of the typically volume densities of H_2 traced by ^{13}CO (Shirley 2015; Liu et al. 2016b; Schuller et al. 2017; Finn et al. 2021), the column density ratios predicted by RADEX are consistent with the ratios derived using LTE from observations of different ^{13}CO transitions. In summary, we divide the column density derived from ^{13}CO $J = 3$–2 emission by a correction factor of 0.3 before converting to H_2 column density.

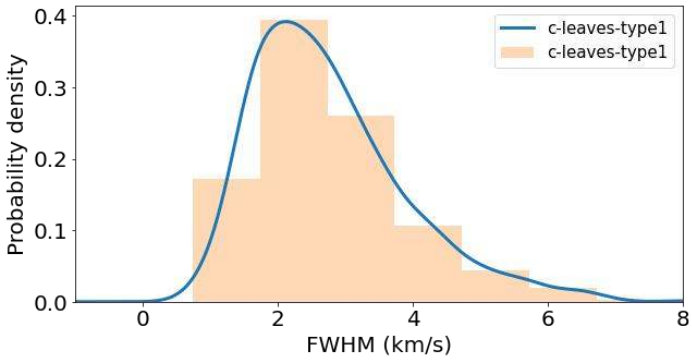


Fig. 10. Line-width distribution of all type1 c-leaves. The probability density is estimated using the KDE method.

3.4. Mass estimation

3.4.1. Mass

The mass of each identified structure is calculated as

$$M = \mu_{\rm H_2} m_{\rm H} \sum N({\rm H_2})(R_{\rm pixel})^2, \quad (13)$$

where $\mu_{\rm H_2} = 2.8$ is the molecular weight per hydrogen molecule, $m_{\rm H}$ is the hydrogen atom mass, $R_{\rm pixel}$ is the size of a pixel. The sum is performed within the elliptical cylinder in the column density cube. As described in Sect. 3.1, the elliptical cylinder has a bottom area of $A = \pi \times 2a \times 2b \times d^2$ and a height range of [v_c–FWHM, v_c+FWHM], where a and b are the long and short axes of the ellipse, and here $d = 3.6$ kpc for the distance to the G333 complex (Lockman 1979; Bains et al. 2006). The average surface density of each structure is then calculated as $\Sigma = M/A$, and the average column density as $N = \Sigma/(\mu_{\rm H_2} m_{\rm H})$.

3.4.2. Molecular line versus continuum emission mass estimates

As described in Zhou et al. (2023), subregions 5 and 7 contain a significant overlap of different velocity components, and should therefore be excluded for column-density estimates based on continuum emission. In Fig. 9d, the column density derived from ATLASGAL+*Planck* 870 μm data is comparable with that estimated from Hi-GAL data processed by the PPMAP procedure. As shown in Figs. 2 and 12a, i-leaves are relatively low-column-density structures distributed at the periphery, and therefore we expect that they will be less massive on average than c-leaves, considering i-leaves and c-leaves structures have similar scales in Fig. 3a. However, in Fig. 12a, i-leaves and c-leaves show similar masses based on the continuum emission. In addition, the continuum mass distribution is relatively narrow, indicating that the contrast between high-column-density and low-column-density structures is not as clear as that derived from molecular line emission due to line-of-sight contamination. In Fig. 12b, we only consider the structures with mean column density greater than 10^{22} cm^{-2}; now the distribution of the masses derived from molecular line emission is similar to that derived from continuum emission after considering the mass correction factor of 0.3. Therefore, in the subsequent analysis, we only adopt the masses estimated from molecular line emission.

3.5. Virial analysis

Having measured the basic physical quantities of the identified structures, we can now investigate their physical properties.

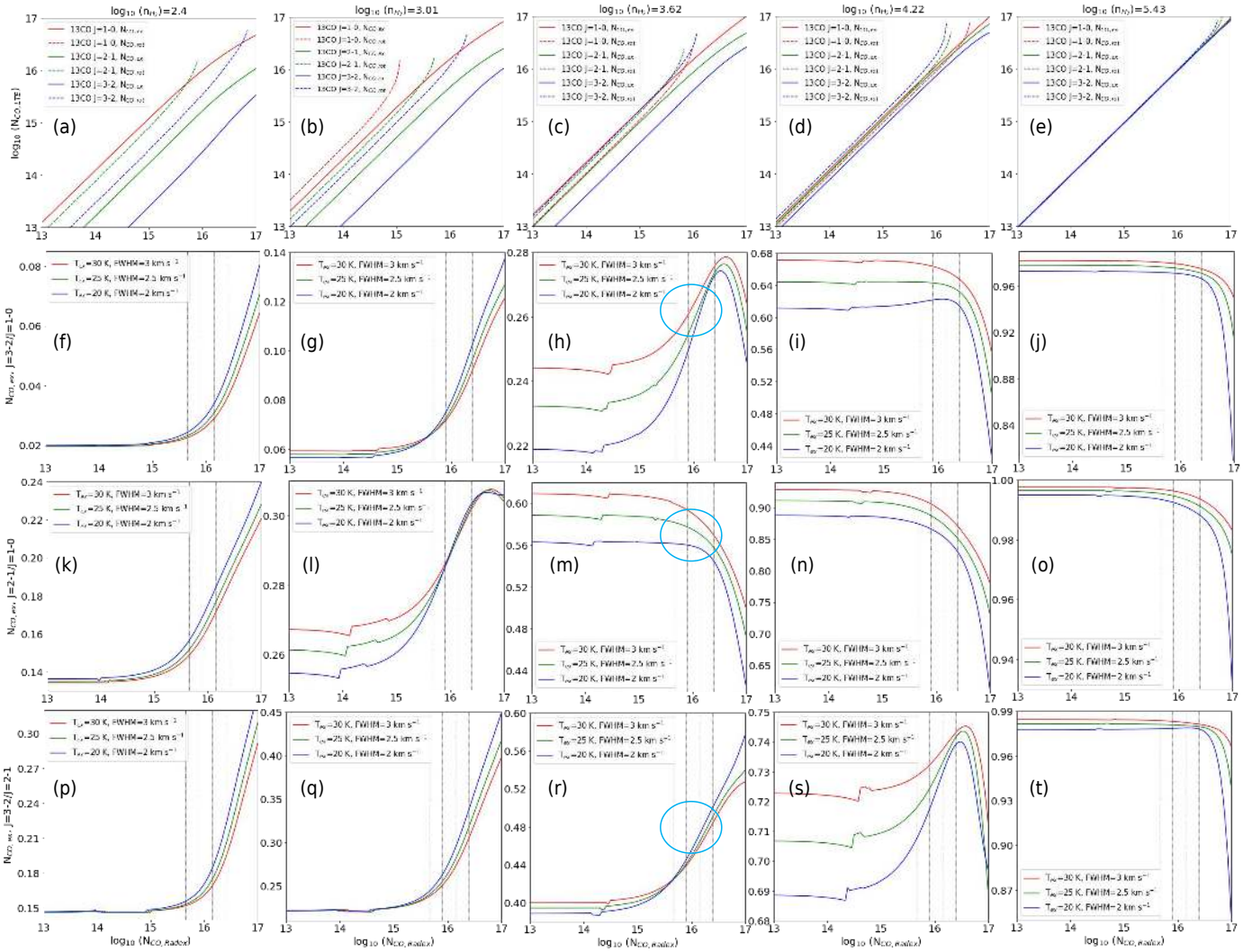


Fig. 11. Calculation results of RADEX. First row: Correlation between RADEX input column densities and column densities of different ^{13}CO transitions derived using LTE equations with $T_{\rm rot}$ and $T_{\rm ex}$ using $T_{\rm R}$ computed by RADEX for different volume densities. Second, third, and fourth rows: Column density ratios of different ^{13}CO transitions derived using LTE equations with $T_{\rm rot}$ and $T_{\rm ex}$ as a function of the RADEX column density input for different volume densities. Vertical lines mark the peak values of the ^{13}CO column density derived from ^{13}CO J = 3–2, J = 2–1, J = 1–0 and ATLASGAL+*Planck* 870 µm emission (from left to right) shown in Figs. 7–9; here the abundance ratio $X_{^{13}CO}$ of H_2 compared with ^{13}CO $\sim 7.1 \times 10^5$ is used. Cyan circles mark the ratios predicted by RADEX when the input ^{13}CO column density takes the peak values of the column density derived by different methods in Sect. 3.3.

3.5.1. Virial parameter

To investigate the energy balance within the extracted structures, we determine the gravitational potential energy and internal kinetic energy to compute the virial parameter (McKee 1989; Bertoldi & McKee 1992):

$$E_g = -\frac{3}{5} a_1 a_2 \frac{GM^2}{R}, \tag{14}$$

$$E_k = \frac{3}{2} M \sigma_{\rm tot}^2. \tag{15}$$

The factor a_1 measures the effects of a nonuniform density distribution and the factor a_2 the effect of the clump's ellipticity. The virial parameter of each decomposed structure is calculated as

$$\alpha_{\rm vir} = 2E_k / \left|E_g\right| = \frac{5}{a_1 a_2} \frac{\sigma_{\rm tot}^2 R}{GM}, \tag{16}$$

with $\sigma_{\rm tot} = \sqrt{\sigma_{\rm nt}^2 + c_s^2}$ being the total velocity dispersion, R the effective radius, and G the gravitational constant, with parameter a_1 equal to $(1 - k/3)/(1 - 2k/5)$ for a power-law-density profile $\rho \propto r^{-k}$, and $a_2 = (\arcsin e)/e$ as the geometry factor. Here, we assume a typical density profile of $k = 1.6$ for all decomposed structures (Butler & Tan 2012; Palau et al. 2014; Li et al. 2019). The eccentricity e is determined by the axis ratio of the dense structure, $e = \sqrt{1 - (b/a)^2}$, and a and b are the long and short axes of the ellipse. Nonmagnetized cores with $\alpha_{\rm vir} < 2$, $\alpha_{\rm vir} \sim 1$, and $\alpha_{\rm vir} < 1$ are considered to be gravitationally bound, in hydrostatic equilibrium, and gravitationally unstable, respectively (Bertoldi & McKee 1992; Kauffmann et al. 2013). Those with $\alpha_{\rm vir} > 2$ could be gravitationally unbound, and are therefore either pressure-confined, or in the process of dispersal.

Figure 13a shows the distribution of virial parameters for all identified structures. We can see that more than half of the leaf structures are gravitationally unbound and only a small fraction are in gravitational collapse. However, in Zhou et al. (2023), we

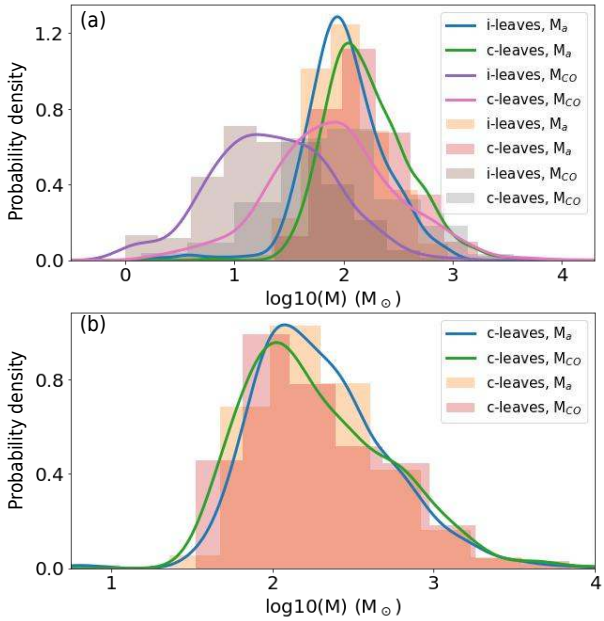


Fig. 12. M_{CO} and M_a represent the mass derived from ^{13}CO (3–2) emission and ATLASGAL+*Planck* 870 μm data, respectively. (a) Mass distribution of all type1 leaf structures. (b) Mass distribution of type1 c-leaves satisfying the density condition >10^{22} cm^{-2}. The probability density is estimated using the KDE method.

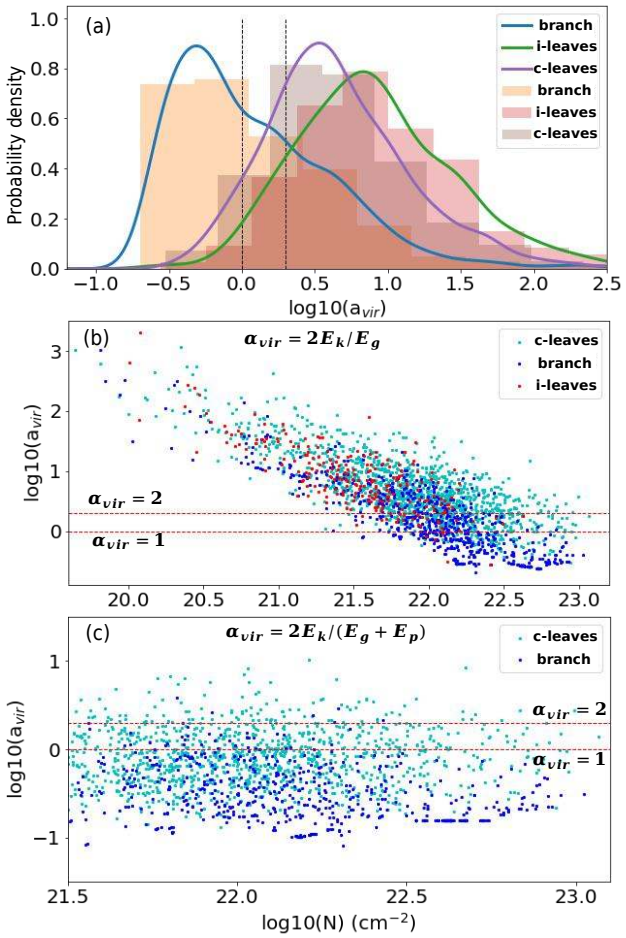


Fig. 13. Virial parameters of dense gas structures. (a) Virial parameters of all type1 and type2 structures. Dashed lines mark the positions $a_{\rm vir} = 1$ and $a_{\rm vir} = 2$. The probability density is estimated using the KDE method. (b) Correlation between virial parameter $a_{\rm vir} = 2E_k/E_g$ and the average column density of the structures. (c) Correlation between virial parameter $a_{\rm vir} = 2E_k/(E_g + E_p)$ and average column density above the threshold of ~3.2 × 10^{21} cm^{-2}.

argue that molecular clouds in the G333 complex are in a state of global gravitational collapse, because the ubiquitous density and velocity fluctuations towards hubs imply the widespread presence of local gravitational collapse. Our previous work provides a more comprehensive approach to study the gas kinematics in the clouds. The dense structures in the clouds are connected to the surrounding environment through filaments, and the gravitational state of the structures can be reflected by the velocity gradients along the filaments, which indicate the converging motions toward gravitational centers (hubs). In Fig. 14, except for the low-column density structures, there is an obvious correlation between velocity dispersion and column density in most of the structures, which indicates a gravitational origin of velocity dispersion, as discussed later in Sect. 4.1. Therefore, most of the structures must be gravitationally bound, or even in a state of gravitational collapse.

The finding that more than half of the structures have virial parameters with value of greater than 2 in Fig. 13 seems reasonable if not considering other forces that can bind the structures. In the ISM, each of the structures is embedded in a larger-scale structure and one can therefore assume that they are confined by various external pressures (Keto & Myers 1986; Lada et al. 2008; Field et al. 2011; Leroy et al. 2015; Kirk et al. 2017; Chen et al. 2019; Li et al. 2020). To reconcile the evidence of gravitational collapse presented by Zhou et al. (2023) with the apparent small fraction of gravitationally unstable gas structures based on the classical virial analysis, we discuss in the following the additional effect of external pressure from ambient cloud structures.

3.5.2. Pressure-confined hydrostatic equilibrium

Previous studies have suggested that the external pressure provided by the larger-scale molecular cloud gas might help to confine dense structures in molecular clouds (Spitzer 1978; McKee 1989; Elmegreen 1989; Ballesteros-Paredes 2006; Kirk et al. 2006, 2017; Lada et al. 2008; Pattle et al. 2015; Li et al. 2020). The external pressure energy can be calculated as

$$E_p = -4\pi P_{\rm cl} R^3, \tag{17}$$

and then the new estimation of the virial parameter is

$$\alpha_{\rm vir} = 2E_k/(\left|E_g\right| + \left|E_p\right|). \tag{18}$$

External pressure can have various origins, such as the turbulent pressure from the HI halo of molecular clouds (Elmegreen 1989), the recoil pressure from photodissociation regions (PDRs) (Field et al. 2011), the infall ram pressure, or other intercloud pressures (Bertoldi & McKee 1992; Lada et al. 2008; Belloche et al. 2011; Camacho et al. 2016). Here, we mainly consider the external pressure from the ambient cloud for each decomposed structure using

$$P_{\rm cl} = \pi G \bar{\Sigma} \Sigma_r, \tag{19}$$

where $P_{\rm cl}$ is the gas pressure, $\bar{\Sigma}$ is the mean surface density of the cloud, and Σ_r is the surface density at the location of each structure (McKee 1989; Kirk et al. 2017). We assumed that

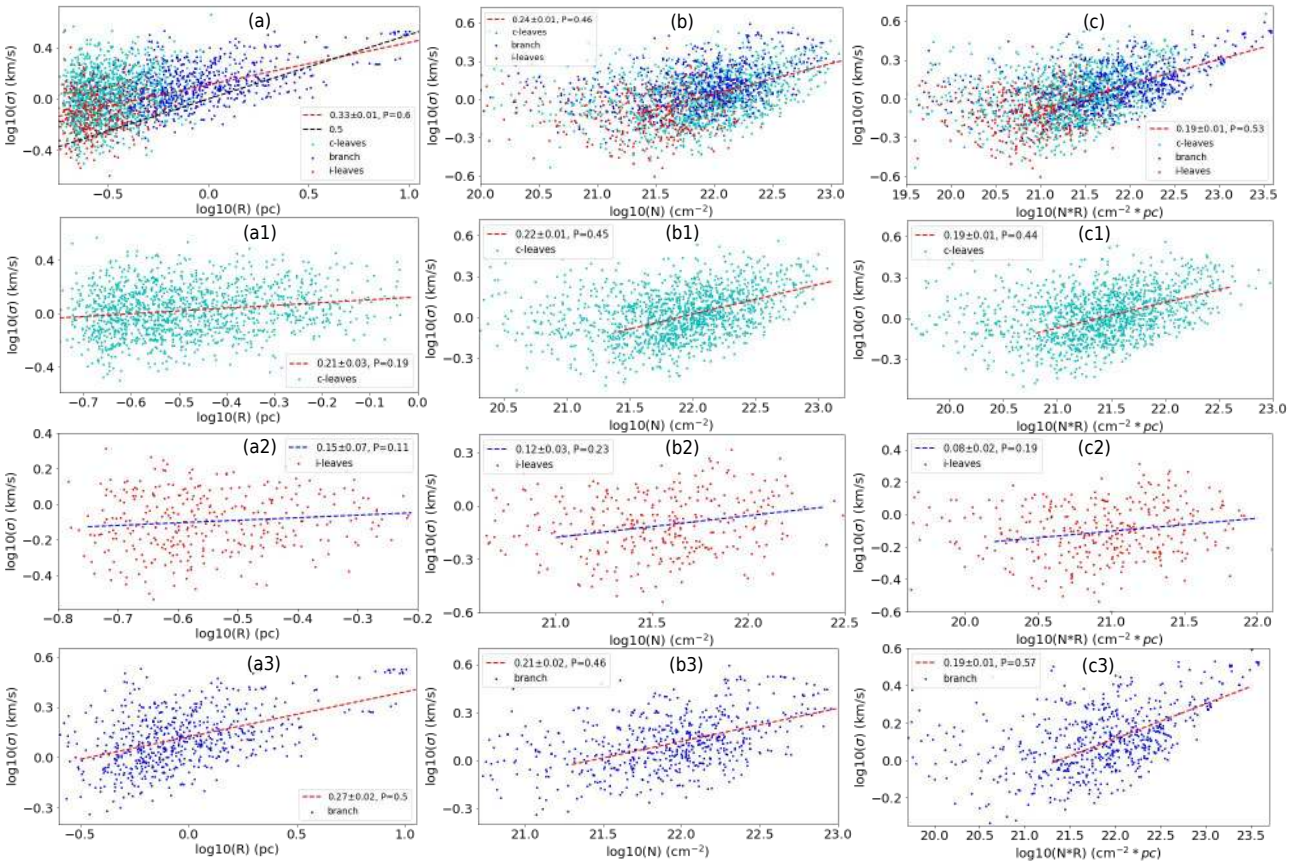


Fig. 14. Scaling relations of leaf and branch structures. (a) $\sigma - R$; (b) $\sigma - N$; (c) $\sigma - N * R$. σ, R and N are the velocity dispersion, effective radius, and column density of each structure, respectively. Here, "P" represents the Pearson coefficient.

Σ_r is equal to half of the observed column density at the footprint of each decomposed structure (Kirk et al. 2017; Li et al. 2020).

The total mass of the G333 complex is calculated with Eq. (13) as $\sim 1.03 \times 10^6\ M_\odot$, which is comparable to the mass of $\sim 1.7 \times 10^6\ M_\odot$ calculated in Miville-Deschênes et al. (2017) using CO (1−0) emission. Using the sum of all nonempty pixels as the total area, the mean surface density of the G333 complex is ~ 0.071 g cm^{-2} (or $\sim 340\ M_\odot$ pc^{-2}), which corresponds to a column density of $\sim 1.5 \times 10^{22}$ cm^{-2}. The G333 complex is located in the molecular ring of the Milky Way, where the mean surface density is $\sim 200\ M_\odot$ pc^{-2} (Heyer & Dame 2015). Given that the G333 complex is the most ATLASGAL-clump-rich giant molecular cloud complex in the southern Milky Way, it should have a higher density than the mean value. The mean surface density of the G333 GMC as calculated by Miville-Deschênes et al. (2017) and Nguyen et al. (2015) is $\sim 120\ M_\odot$ pc^{-2} and $\sim 220\ M_\odot$ pc^{-2}, respectively (see Sect. 4.5 of Zhou et al. 2023 for more details). Adopting a conservative estimate, we take the value of $\bar{\Sigma} \sim 200\ M_\odot$ pc^{-2} as a lower limit, which corresponds to a column density of $\sim 9 \times 10^{21}$ cm^{-2}.

There are also many leaf structures with lower column density; these are usually distributed in the periphery of the clouds. Equation (19) may be not valid for these latter, because Σ_r in the equation is integrated from the cloud surface to depth r of each structure in the cloud (Kirk et al. 2017). We therefore need to set a density threshold for the structures in order to determine whether or not they are eligible to be bound by the external pressure from the ambient cloud. In Fig. 14, high-column-density and low-column-density structures show different behaviors; the turning point corresponds to a column density value of $\sim 3.2 \times 10^{21}$ cm^{-2}, which we use as the threshold.

We treat gravitationally unbound branch structures as c-leaves. Here, we ignore i-leaves structures; they are isolated structures at the cloud periphery and are unlikely to be confined by external pressure from the ambient cloud. For c-leaves and branches, the proportion of the structures above the density threshold ($\sim 3.2 \times 10^{21}$ cm^{-2}) is 82.5%. The proportions of $a_{\rm vir} = 2E_{\rm k}/E_{\rm g} < 2$ and $a_{\rm vir} = 2E_{\rm k}/E_{\rm g} \geq 2$ are 45.3% and 54.7%, respectively. After considering the external pressure ($a_{\rm vir} = 2E_{\rm k}/(E_{\rm g} + E_{\rm p})$), their proportions become 93 and 7%. When accounting for external pressure, the majority of the structures are gravitationally bound, and susceptible to gravitational collapse, as shown in Fig. 13c. However, the peripheral structures are at low column densities and are less bound by external pressure, and are therefore likely to be dispersed by feedback.

Here, we do not consider the HI halo of molecular clouds and also ignore the external pressure exerted by HII regions. The latter should also be important due to the strong feedback in the G333 complex. However, the energy injected into the clouds from HII regions might also destroy the clouds, and therefore we only consider the external pressure from the ambient cloud in binding the structures. The rough estimate in this section shows the important role of the external pressure in confining the observed gas structures.

3.6. Scaling relation

The physical states of the structures can also be reflected by the scaling relations. Figure 14a shows the velocity dispersion–scale

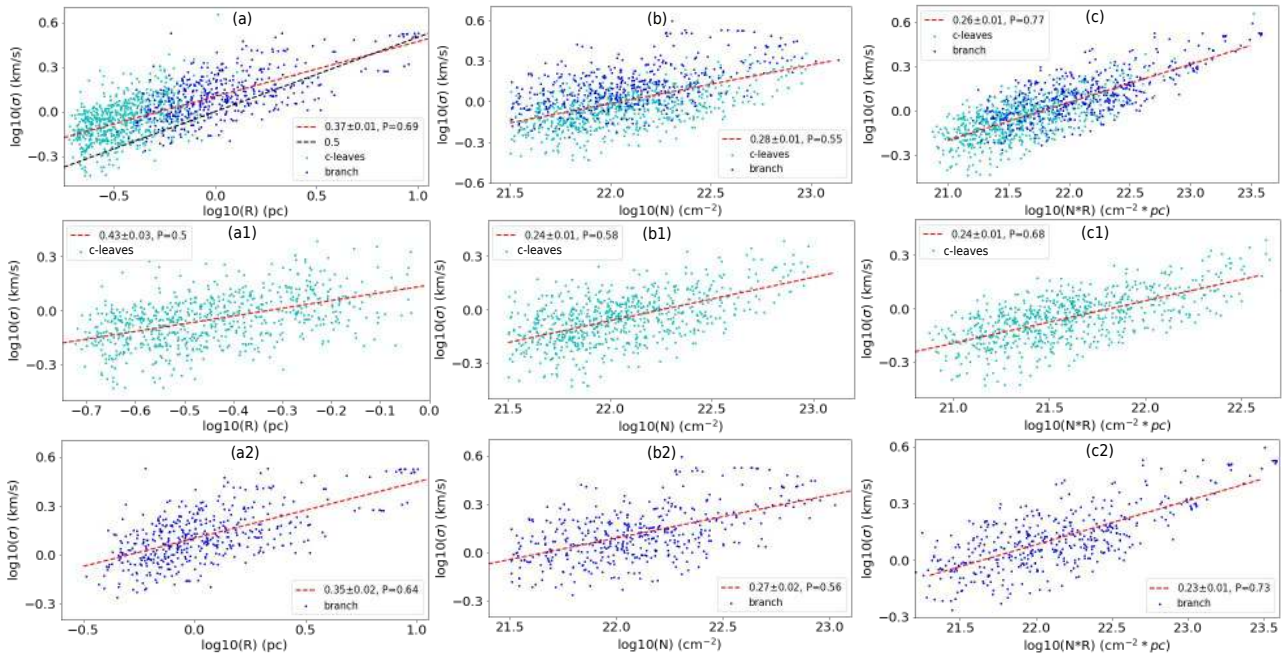


Fig. 15. Scaling relations of c-leaves and branch structures satisfying the conditions $N > 3.2 \times 10^{21}$ cm^{-2} and $a_{\rm vir} = 2E_k/(E_g + E_p) < 1$. (a) $\sigma - R$; (b) $\sigma - N$; (c) $\sigma - N * R$.

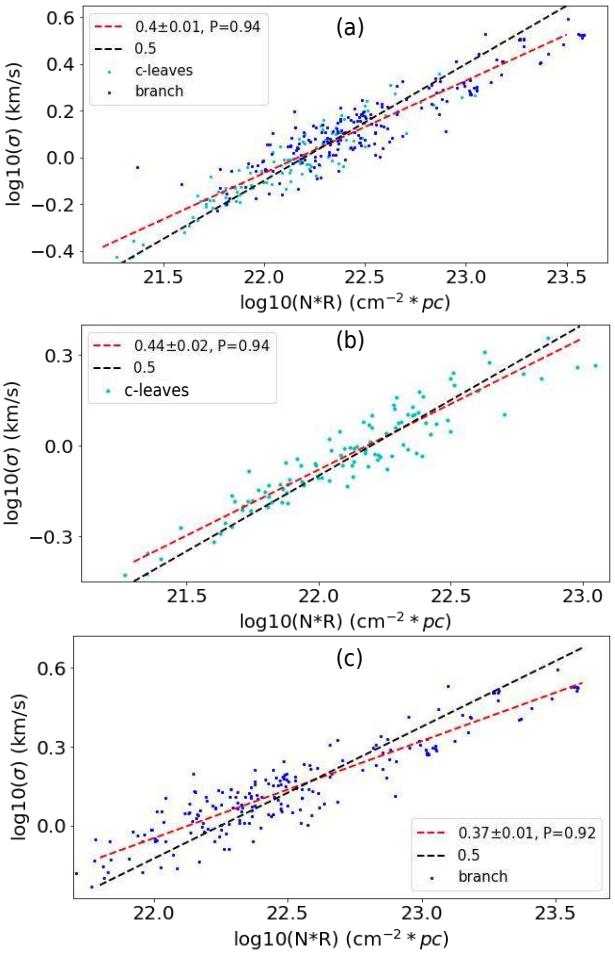


Fig. 16. Scaling relation $\sigma - N * R$ of c-leaves and branch structures satisfying the conditions N>3.2 × 10^{21} cm^{-2} and $a_{\rm vir} = 2E_k/E_g < 1$.

relations of i-leaves, c-leaves, and branch structures. It appears that only branch structures show a clear correlation between velocity dispersion and scale. i-leaves roughly inherit the trend in the $\sigma - R$ relation of branch structures extending from large to small scales, and they can barely be linked behind branch structures in Fig. 14a, although their velocity dispersion shows no significant correlation with scale, which is similar to c-leaves, which deviate more significantly from the Larson relation. The red dashed line, which is fitted to branch and i-leaves structures, has a gradient of 0.33 ± 0.01.

Figure 14b shows the velocity dispersion–column density relation. For c-leaves, there is a moderate correlation between velocity dispersion and column density: the Pearson coefficient is ∼0.45. Interestingly, we can see different behaviors of high-column-density and low-column-density structures. For high-column-density structures, the velocity dispersion and column density show a clear correlation, while low-column-density structures do not. In recent simulations (Ganguly et al. 2024; Weis et al. 2022), dense structures roughly follow the Heyer relation, and less dense structures show no trend with the column density, and populate a low-density tail in the Heyer relation, as shown in Fig. 14c. For a more convenient comparison with $\sigma - R$ and $\sigma - N$ relations, we convert the Heyer relation $\sigma/R^{0.5} \propto N^{0.5}$ to the form $\sigma \propto (R * N)^{0.5}$ (Eq. (3) in Ballesteros-Paredes et al. 2011); both should have a slope of 0.5. From Figs. 14a and b, we conclude that the velocity dispersion of branch structures correlates with both scale and column density, and that the velocity dispersion of c-leaves is only sensitive to column density, while the velocity dispersion of i-leaves shows no significant dependence on either scale or column density.

The structures with column density >3.2 × 10^{21} cm^{-2} can be divided into two types: those that can collapse after adding external pressure ($a_{\rm vir} = 2E_k/(E_g + E_p) < 1$, pressure-assisted collapse), and those that can collapse by self-gravity alone ($a_{\rm vir} = 2E_k/E_g < 1$, self-gravitating collapse). Now we have three structure sets: all identified structures, the structures in

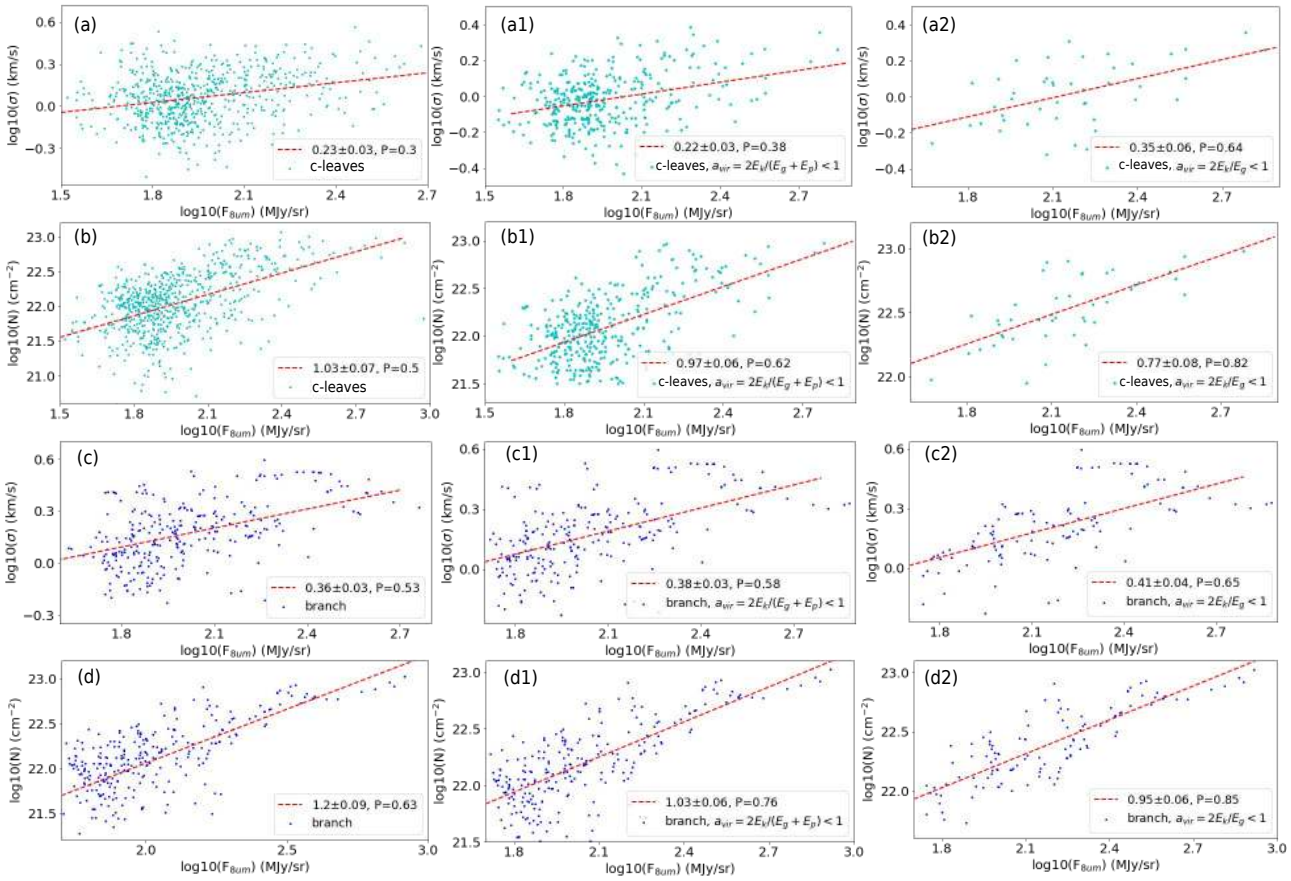


Fig. 17. Correlation between the 8 μm surface brightness and velocity dispersion and column density of the structures. First column: All type1 c-leaves and branch structures. Second column: Type1 c-leaves and branch structures satisfying the condition $a_{\rm vir} = 2E_k/(E_g + E_p) < 1$. Third column: Type1 c-leaves and branch structures satisfying the condition $a_{\rm vir} = 2E_k/E_g < 1$.

pressure-assisted collapse, and the structures in self-gravitating collapse, where the latter are a subset of the former. The scaling relations of these three structure sets are shown in Figs. 14–16, respectively. For both leaf and branch structures, $\sigma - N * R$ always has a stronger correlation than $\sigma - N$ or $\sigma - R$. Moreover, the scaling relations show a stronger correlation and steeper slope when applied to self-gravitating structures, and therefore best follow the Heyer relation.

3.7. Feedback

In the study of the G305 molecular cloud complex, Mazumdar et al. (2021a) argued that the 8 μm emission can be a good indicator of feedback strength. We calculated the average 8 μm surface brightness over each structure to measure the strength of feedback on each structure. In Fig. 17, the 8 μm surface brightness shows a strong positive correlation with column density for both c-leaves and branch structures, which might be an indication for triggering in the G333 complex. However, branch structures show a more obvious correlation between 8 μm surface brightness and velocity dispersion than c-leaves, which is consistent with the results of Mazumdar et al. (2021b), implying that feedback has a greater impact on large-scale structures. The small-scale structures are embedded in large-scale structures, and are therefore less affected by feedback. For large-scale structures, as shown in Fig. 2c, the more evolved subregion 1 and subregion s2b are fragmenting into several pieces, potentially torn apart by the expanding HII regions. These results may explain why the velocity dispersion of branch structures shows a clear correlation with scale, while that of leaf structures does not, as shown in Fig. 14. However, in Fig. 15, after filtering low-column-density and gravitationally unbound structures, the velocity dispersion of c-leaves appears to show a better correlation with scale than in Fig. 14, and the correlation between the 8 μm surface brightness and velocity dispersion of c-leaves is also improved in Fig. 17, which means that leaf structure is also affected by feedback. Here, we should remember that there is considerable overlap in the scale of leaf and branch structures, as described in Sect. 3.1.

Our analysis is based on structural identification, when we say that feedback increases the density and velocity dispersion of the structures, provided that these structures can exist stably. Structures with $a_{\rm vir} = 2E_k/(E_g + E_p) < 1$ and $a_{\rm vir} = 2E_k/E_g < 1$ can be more tenacious in feedback than other structures, and thus exhibit better scaling relations in Figs. 15–17. Dale et al. (2014) used simulations to examine the effects of photoionization and momentum-driven winds from O-stars on molecular clouds, and found that feedback is highly destructive to clouds of lower mass and density, but has little effect on more massive and denser clouds.

4. Discussion

4.1. The origin of velocity dispersion

High-column-density clumps or cores exhibit larger velocity dispersion than low-column-density ones because of gas motions in gravitational collapse (Ballesteros-Paredes et al. 2011, 2018;

Traficante et al. 2018b; Li et al. 2023), as shown in Figs. 14b and 15b, where the positive correlation between velocity dispersion and column density of c-leaves and branch structures indicates the gravitational origin of velocity dispersion. Combined with the discussions in Sects. 3.6 and 3.7, we conclude that both gravitational collapse and feedback contribute significantly to the velocity dispersion of large-scale structures. For small-scale structures, gravitational collapse is an important source of velocity dispersion, but understanding the contribution of feedback therein requires further investigation.

4.2. The Heyer relation in feedback

In Sect. 3.6, self-gravitating structures can better fit the Heyer relation. Considering that global collapse may lag behind the local collapse in the cloud (Heitsch et al. 2008), structures collapsing under self-gravity can be relatively independent of (or "decoupled from") the surrounding environment. Therefore, the explanations of the Heyer relation in Heyer et al. (2009) and Ballesteros-Paredes et al. (2011) can hold. Contrarily, for nonself-gravitating structures, the exchange of energy with the surrounding environment will break the conversion between E_g and E_k, thus breaking the Heyer relation.

Sun et al. (2018) measured cloud-scale molecular gas properties in 15 nearby galaxies, and observed an excess in the velocity dispersion σ at low surface density Σ relative to the expected relation for self-gravity-dominated gas. This behavior leads to a shallower $\sigma - \Sigma$ relation in several galaxies, clearly deviating from the $\sigma - \Sigma^{0.5}$ relation extrapolated from the high-surface-density regime. One of the explanations provided by these latter authors for the deviation is that gas structures in the low-surface-density regime may be more susceptible to external pressure originating from the ambient medium or motions due to the galaxy potential, which is similar to our above explanation.

4.3. Cloud disruption and collapse under feedback

Peretto et al. (2023) performed an analysis of 27 infrared-dark clouds (IRDCs) embedded within 24 molecular clouds. These authors found that the clumps are decoupling from their surrounding cloud and concluded that the observations are best explained by a universal global collapse of dense clumps embedded within stable molecular clouds, thus discovering direct evidence of a transition regime in the dynamical properties of the gas within individual molecular clouds. As discussed in Heitsch et al. (2008) and Ballesteros-Paredes et al. (2011), the decoupling may be due to the global collapse of the molecular cloud lagging behind the local collapse of dense clumps in the cloud. Invoking the notion of the "funnel" structure in PPV space (Zhou et al. 2023), a similar statement is that relatively small-scale hub–filament structures will have a more recognizable "funnel" morphology than large-scale ones due to their strong local gravitational field. Substructure s3a in the G333 complex is a vivid example of the decoupling highlighted in Fig. 1 of Zhou et al. (2023); it is collapsing into a hub–filament structure and is separating from its surrounding environment.

One implication of the work of Peretto et al. (2023) is that star formation is likely to be mostly confined to parsec-scale collapsing clumps, which is also consistent with the results of Zhou et al. (2022, 2023). In our previous works, for both ALMA and LAsMA data, most of the fitted velocity gradients concentrate on scales of ∼1 pc, a scale that is considered to be the characteristic scale of massive clumps (Urquhart et al. 2018). Velocity gradients measured around 1 pc show that the most frequent velocity gradient is $\sim 1.6\,\mathrm{km\,s^{-1}\,pc^{-1}}$. Assuming free-fall, we estimate that the kinematic mass corresponding to 1 pc is $\sim 1190\,M_\odot$, which is also comparable with the typical mass of clumps in the ATLASGAL survey (Urquhart et al. 2018). Therefore, parsec-scale clumps are probably gravity-dominated collapsing objects.

The results in Zhou et al. (2022, 2023) and Peretto et al. (2023) show that the physical properties of parsec-scale clumps in two very different physical environments (infrared dark and infrared bright) are comparable. Therefore, feedback in infrared-bright star-forming regions, such as the G333 complex, will not significantly change the physical properties of parsec-scale clumps, which is also consistent with the finding from surveys that most Galactic parsec-scale massive clumps seem to be gravitationally bound no matter how evolved they are (Liu et al. 2016a; Urquhart et al. 2018; Evans et al. 2021). Although the clumps are exposed to feedback and part of their velocity dispersion is due to feedback, as shown in Sect. 3.7, the clumps are still self-gravitating sufficiently to continue their collapse, even after the lower-density material has been disrupted and is being dispersed. Watkins et al. (2019) found that stellar feedback from O stars does not have a significant impact on the dynamical properties of the dense gas that has already been assembled, but does clearly modify the structure of the larger-scale clouds. The broken morphology of some very infrared-bright structures in the G333 complex also indicates that the feedback is disrupting the molecular clouds.

The effects of feedback in star-forming regions can redistribute, disperse, and enhance preexisting gas structures, and create new structures (Elmegreen & Lada 1977; Dale et al. 2007; Lee & Chen 2007; Nagakura et al. 2009; Krumholz et al. 2014; Fukui et al. 2021). According to the physical picture described in Zhou et al. (2023), the hub–filament structures at different scales may be organized into a hierarchical system, extending up to the largest scales probed, through the coupling of gravitational centers at different scales. Large-scale velocity gradients always involve many intensity peaks, and the larger-scale inflow is driven by the larger-scale structure, implying that the clustering of local small-scale gravitational structures can act as the gravitational center on larger scale. Considering the hierarchical hub–filament structures and the coupling of local gravitational centers in molecular clouds, and feedback do not significantly impact the dynamical properties of the dense gas, therefore although the feedback disrupting the molecular clouds may break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and changes in gravitational centers, but gravitational collapse is always ongoing.

5. Summary

We investigated the kinematics and dynamics of gas structures under feedback in the G333 complex. The main conclusions are as follows:

1. The dense gas structures were identified by the Dendrogram algorithm based on the integrated intensity map of ^{13}CO (3−2). We obtained 3608 structures, and extracted and fitted their averaged spectra one by one. According to the line profiles, all averaged spectra were divided into three categories. The physical quantities of each structure were calculated based on their line profiles.

2. We derived the column density of the entire observed field from ATLASGAL+*Planck* 870 μm data, Hi-GAL data, and different transitions of ^{13}CO (J = 1–0, 2–1 and 3–2). We investigated the column density ratios between them pixel by pixel, and find that the column density derived from ATLASGAL+*Planck* 870 μm data is comparable with that estimated from Hi-GAL data. Molecular line emission gives significantly lower column density estimates than those derived from the continuum emission. The peak values of the column density ratios between different transitions of ^{13}CO emission are $N_{2-1}/N_{1-0} \approx 0.5$, $N_{3-2}/N_{1-0} \approx 0.3$, and $N_{3-2}/N_{2-1} \approx 0.5$. These ratios can be roughly reproduced by the NLTE molecular radiative transfer algorithm RADEX for typical volume densities of $\sim 4.2 \times 10^3$ cm^{-3}. We therefore adopted a correction factor of 0.3 to calibrate the column density derived from ^{13}CO J = 3−2 in order to obtain a value that is more representative of the total column density.

3. Classical virial analysis, which suggests that many structures are unbound, does not reflect the true physical state of the identified structures. After considering external pressure from the ambient cloud, almost all the structures with a higher column density than the threshold $\sim 3.2 \times 10^{21}$ cm^{-2} are gravitationally bound, or are even undergoing gravitational collapse.

4. The positive correlation between velocity dispersion and column density of c-leaves and branch structures reveals the gravitational origin of velocity dispersion.

5. We used the average 8 μm surface brightness as an indicator of feedback strength, which shows a strongly positive correlation with the column density of both c-leaves and branch structures. However, branch structures show a more significant correlation between 8 μm surface brightness and velocity dispersion than c-leaves, implying that feedback has a greater impact on large-scale structures. We conclude that both gravitational collapse and feedback contribute significantly to the velocity dispersion of large-scale structures. For small-scale structures, gravitational collapse is an important source of velocity dispersion, while further investigation is required to understand the contribution of feedback therein.

6. For both leaf and branch structures, $\sigma - N * R$ always has a stronger correlation compared to $\sigma - N$ and $\sigma - R$. The scaling relations are stronger, and have steeper slopes when considering only self-gravitating structures, which are the structures most closely associated with the Heyer relation. However, due to the strong feedback in the G333 complex, only a small fraction of the structures are in a state of self-gravitational collapse.

7. Although the feedback disrupting the molecular clouds will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation as well as changes in gravitational centers, but gravitational collapse is always ongoing.

Acknowledgements. We would like to thank the referee for the detailed comments and suggestions that significantly improve and clarify this work. This publication is based on data acquired with the Atacama Pathfinder Experiment (APEX) under programme ID M-0109.F-9514A-2022. APEX has been a collaboration between the Max-Planck-Institut für Radioastronomie, the European Southern Observatory, and the Onsala Space Observatory. This research made use of astrodendro, a Python package to compute dendrograms of Astronomical data (http://www.dendrograms.org/). J.W.Z. thanks E. Vázquez-Semadeni for the helpful discussions. This work has been supported by the National Key R&D Program of China (No. 2022YFA1603101). T.L. acknowledges the supports by National Natural Science Foundation of China (NSFC) through grants No.12073061 and No.12122307, the international partnership program of Chinese Academy of Sciences through grant No.114231KYSB20200009, and Shanghai Pujiang Program 20PJ1415500.

References

Areal, M. B., Paron, S., Ortega, M. E., & Duvidovich, L. 2019, PASA, 36, e049
Bains, I., Wong, T., Cunningham, M., et al. 2006, MNRAS, 367, 1609
Ballesteros-Paredes, J. 2006, MNRAS, 372, 443
Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F., & Zamora-Avilés, M. A. 2011, MNRAS, 411, 65
Ballesteros-Paredes, J., Vázquez-Semadeni, E., Palau, A., & Klessen, R. S. 2018, MNRAS, 479, 2112
Ballesteros-Paredes, J., André, P., Hennebelle, P., et al. 2020, Space Sci. Rev., 216, 76
Belloche, A., Schuller, F., Parise, B., et al. 2011, A&A, 527, A145
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953
Bertoldi, F., & McKee, C. F. 1992, ApJ, 395, 140
Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 1997, MNRAS, 285, 201
Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 2001, MNRAS, 323, 785
Burton, M. G., Braiding, C., Glueck, C., et al. 2013, PASA, 30, e044
Butler, M. J., & Tan, J. C. 2012, ApJ, 754, 5
Camacho, V., Vázquez-Semadeni, E., Ballesteros-Paredes, J., et al. 2016, ApJ, 833, 113
Chen, H.-R. V., Zhang, Q., Wright, M. C. H., et al. 2019, ApJ, 875, 24
Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, A&A, 585, A104
Dale, J. E., Bonnell, I. A., & Whitworth, A. P. 2007, MNRAS, 375, 1291
Dale, J. E., Ercolano, B., & Bonnell, I. A. 2012, MNRAS, 424, 377
Dale, J. E., Ngoumou, J., Ercolano, B., & Bonnell, I. A. 2014, MNRAS, 442, 694
Elmegreen, B. G. 1989, ApJ, 338, 178
Elmegreen, B. G., & Lada, C. J. 1977, ApJ, 214, 725
Evans, N. J., II., Heyer, M., Miville-Deschênes, M.-A., Nguyen-Luong, Q., & Merello, M. 2021, ApJ, 920, 126
Field, G. B., Blackman, E. G., & Keto, E. R. 2011, MNRAS, 416, 710
Finn, M. K., Indebetouw, R., Johnson, K. E., et al. 2021, ApJ, 917, 106
Frerking, M. A., Langer, W. D., & Wilson, R. W. 1982, ApJ, 262, 590
Fukui, Y., Habe, A., Inoue, T., Enokiya, R., & Tachihara, K. 2021, PASJ, 73, S1
Ganguly, S., Walch, S., Clarke, S. D., & Seifried, D. 2024, MNRAS, in press, https://doi.org/10.1093/mnras/stae032
Garden, R. P., Hayashi, M., Gatley, I., Hasegawa, T., & Kaifu, N. 1991, ApJ, 374, 540
Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428
Güsten, R., Nyman, L. Å., Schilke, P., et al. 2006, A&A, 454, L13
Hartmann, L., Ballesteros-Paredes, J., & Heitsch, F. 2012, MNRAS, 420, 1457
Heitsch, F., Hartmann, L. W., Slyz, A. D., Devriendt, J. E. G., & Burkert, A. 2008, ApJ, 674, 316
Hennebelle, P., & Falgarone, E. 2012, A&ARv, 20, 55
Heyer, M. H., & Brunt, C. M. 2004, ApJ, 615, L45
Heyer, M., & Dame, T. M. 2015, ARA&A, 53, 583
Heyer, M., Krawczyk, C., Duval, J., & Jackson, J. M. 2009, ApJ, 699, 1092
Ibáñez-Mejía, J. C., Mac Low, M.-M., Klessen, R. S., & Baczynski, C. 2016, ApJ, 824, 41
Kauffmann, J., Bertoldi, F., Bourke, T. L., Evans, N. J., I., & Lee, C. W. 2008, A&A, 487, 993
Kauffmann, J., Pillai, T., & Goldsmith, P. F. 2013, ApJ, 779, 185
Keto, E. R., & Myers, P. C. 1986, ApJ, 304, 466
Kirk, H., Johnstone, D., & Di Francesco, J. 2006, ApJ, 646, 1009
Kirk, H., Friesen, R. K., Pineda, J. E., et al. 2017, ApJ, 846, 144
Krumholz, M. R., Klein, R. I., & McKee, C. F. 2007, ApJ, 656, 959
Krumholz, M. R., Bate, M. R., Arce, H. G., et al. 2014, in Protostars and Planets VI, eds. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning (Tucson: University of Arizona Press), 243
Lada, C. J., Muench, A. A., Rathborne, J., Alves, J. F., & Lombardi, M. 2008, ApJ, 672, 410
Larson, R. B. 1981, MNRAS, 194, 809
Lee, H.-T., & Chen, W. P. 2007, ApJ, 657, 884
Leroy, A. K., Bolatto, A. D., Ostriker, E. C., et al. 2015, ApJ, 801, 25
Leroy, A. K., Rosolowsky, E., Usero, A., et al. 2022, ApJ, 927, 149
Lewis, S. C., McMillan, S. L. W., Low, M.-M. M., et al. 2023, ApJ, 944, 211
Li, S., Zhang, Q., Pillai, T., et al. 2019, ApJ, 886, 130
Li, S., Zhang, Q., Liu, H. B., et al. 2020, ApJ, 896, 110
Li, S., Sanhueza, P., Zhang, Q., et al. 2023, ApJ, 949, 109
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Liu, T., Kim, K.-T., Yoo, H., et al. 2016a, ApJ, 829, 59

A&A, 682, A173 (2024)
https://doi.org/10.1051/0004-6361/202348108
© The Authors 2024

Astronomy & Astrophysics

Feedback from protoclusters does not significantly change the kinematic properties of the embedded dense gas structures

J. W. Zhou[1], S. Dib[2], F. Wyrowski[1], T. Liu[3], S. H. Li[2], P. Sanhueza[9], M. Juvela[4], F. W. Xu[5,6], H. L. Liu[7], T. Baug[11], Y. P. Peng[8], K. M. Menten[1], and L. Bronfman[10]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de
[2] Max Planck Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany
[3] Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, PR China
[4] Department of Physics, PO Box 64, 00014 University of Helsinki, Finland
[5] Kavli Institute for Astronomy and Astrophysics, Peking University, Haidian District, Beijing 100871, PR China
[6] Department of Astronomy, School of Physics, Peking University, Beijing 100871, PR China
[7] Department of Astronomy, Yunnan University, Kunming 650091, PR China
[8] Department of Physics, Faculty of Science, Kunming University of Science and Technology, Kunming 650500, PR China
[9] National Astronomical Observatory of Japan, National Institutes of Natural Sciences, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[10] Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile
[11] Satyendra Nath Bose National Centre for Basic Sciences, BlockJD, Sector-III, Salt Lake, Kolkata-700 106, India

Received 28 September 2023 / Accepted 1 December 2023

ABSTRACT

A total of 64 ATOMS sources at different evolutionary stages were selected to investigate the kinematics and dynamics of gas structures under feedback. We identified dense gas structures based on the integrated intensity map of $H^{13}CO^+$ $J = 1$–0 emission, and then extracted the average spectra of all the structures to investigate their velocity components and gas kinematics. For the scaling relations between the velocity dispersion, σ, the effective radius, R, and the column density, N, of all the structures, $\sigma - N * R$ always has a stronger correlation compared to $\sigma - N$ and $\sigma - R$. There are significant correlations between velocity dispersion and column density, which may imply that the velocity dispersion originates in gravitational collapse, also revealed by the velocity gradients. The measured velocity gradients for dense gas structures in early-stage sources and late-stage sources are comparable, indicating gravitational collapse through all evolutionary stages. Late-stage sources do not have large-scale hub-filament structures, but the embedded dense gas structures in late-stage sources show similar kinematic modes to those in early- and middle-stage sources. These results may be explained by the multi-scale hub-filament structures in the clouds. We quantitatively estimated the velocity dispersion generated by the outflows, inflows, ionized gas pressure, and radiation pressure, and found that the ionized gas feedback is stronger than other feedback mechanisms. However, although feedback from HII regions is the strongest, it does not significantly affect the physical properties of the embedded dense gas structures. Combined with the conclusions in our previous work on cloud-clump scales, we suggest that although feedback from cloud to core scales will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and changes in gravitational centers, but gravitational collapse is always ongoing.

Key words. techniques: image processing – stars: formation – evolution – ISM: structure – submillimeter: ISM

1. Introduction

Stellar feedback from high-mass stars ($M > 8\ M_\odot$) can strongly influence their surrounding interstellar medium and regulate star formation. The combined effects of multiple feedback mechanisms may play a major role in determining star formation efficiency and the initial mass function in molecular clouds (Dib et al. 2010; Krumholz et al. 2014). In Krumholz et al. (2014), feedback mechanisms are divided into three categories: (1) momentum feedback, including protostellar outflows and (probably) radiation pressure; (2) explosive feedback, including winds from hot main sequence stars, photoionizing radiation, and supernovae; and (3) thermal feedback, with non-ionizing radiation as the main form. Lopez et al. (2014) assessed the role of stellar feedback at scales ∼10–100 pc toward 32 HII regions in the Large and Small Magellanic Clouds. They employed multi-wavelength data to examine several stellar feedback mechanisms and found that the ionized gas feedback dominates over the other mechanisms in all of the sources. Similar methods were also used in Olivier et al. (2021) but, for 106 deeply embedded HII regions with radii <0.5 pc, they concluded that the dust-processed radiation dominates in 84% of the samples. In simulations, massive stellar feedback, including ionizing radiation, stellar winds, and supernovae (Matzner 2002; Dale et al. 2012, 2014; Rogers & Pittard 2013; Rahner et al. 2017; Smith et al. 2018; Lewis et al. 2023), can suppress star formation and destroy the natal cloud. However, whether stellar feedback promotes or suppresses star formation remains controversial. The simulations of Dale et al. (2014) examined the effects of photoionization and momentum-driven winds from O-stars on molecular clouds, and

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

found that feedback has little effect on denser and more massive clouds, which mainly destroy clouds with a lower density and mass. Dib et al. (2013) showed, using semi-analytical models in which gas-expulsion from protocluster clouds is driven by stellar wind feedback originating from both B and O stars that the gas expulsion timescale depends on the clouds mass and the coupling efficiency of the wind with the surrounding gas. The observations of Pabst et al. (2019), Henshaw et al. (2022) suggest that at early stages energy-winds could dominate, even if this is questioned by recent theoretical works (Lancaster et al. 2021a,b). It is also possible for the feedback to be positive by triggering new star formation. In Deharveng et al. (2005), Zavagno et al. (2006), and Liu et al. (2017), using data at various wavelengths, they identified and studied several objects in which star formation appears to have been triggered at the borders of HII regions. Thompson et al. (2012) studied a large sample of infrared bubbles and estimated that the fraction of massive stars in the Milky Way formed by triggering could be between 14 and 30%. Kendrew et al. (2012) found that approximately ∼22% of massive young stars may have formed as a result of feedback from expanding HII regions.

As presented in Watkins et al. (2019), stellar feedback from O stars clearly modifies the structure of the larger-scale clouds, but without much effect on the dynamical properties of the assembled dense gas. In the G305 GMC, feedback has triggered star formation by the collect and collapse mechanism based on observations using the Large APEX sub-Millimeter Array (LAsMA) seven-beam receiver on the 12-meter Atacama Pathfinder Experiment submillimeter telescope (APEX; Mazumdar et al. 2021a,b). In Rugel et al. (2019), feedback from the first formed cluster in W49A is not strong enough to disperse the cloud, which likely only affects limited parts of W49A. Moreover, all feedback models used in Rugel et al. (2019) predict recollapse of the shell after the first star formation event. The importance of large-scale gravitational collapse in star formation regions has been suggested in many observational and theoretical works (Hartmann & Burkert 2007; Vázquez-Semadeni et al. 2009, 2019; Schneider et al. 2010; Peretto et al. 2013, 2014; Wyrowski et al. 2016; Hacar et al. 2017). In a sample of 11 massive clumps, high-resolution data reveal that protoclusters evolves to tighter by global gravitational collapse (Xu et al. 2024). In the LAsMA observation for the G333 complex under feedback, on large scales, Zhou et al. (2023, 2024) found that the inflow may be driven by the larger-scale structure, consistent with the hierarchical structure in the molecular cloud and gas inflow from large to small scales. The large-scale gas inflow is driven by gravity, implying that the molecular clouds in the G333 complex may be in a state of global gravitational collapse (see also Dib 2023). Ubiquitous density and velocity fluctuations also imply the widespread presence of local gravitational collapse (Henshaw et al. 2020). As for local dense gas structures, most of them exhibit an obvious correlation between velocity dispersion and density, which indicates the gravitational origin of velocity dispersion. For the G333 complex, although the original cloud complex is disrupted by feedback, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation and by changes in gravitational centers, but gravitational collapse is always ongoing. The results in Zhou et al. (2022, 2023, 2024) and Peretto et al. (2023) show that the kinematic properties of parsec-scale clumps in two very different physical environments (infrared dark and infrared bright) are comparable. Thus, feedback in infrared bright star-forming regions, such as the G333 complex, will not change the kinematic properties of parsec-scale clumps, which is also consistent with the survey results that most Galactic parsec-scale massive clumps seem to be gravitationally bound regardless of their evolutionary phases (Liu et al. 2016; Urquhart et al. 2018; Evans et al. 2021).

In this work, we extend the cloud-clump analysis in Zhou et al. (2024) to the clump-core scale using high-resolution ALMA data from the ATOMS (ALMA Three-millimeter Observations of Massive Star-forming regions) survey (Liu et al. 2020). The ATOMS survey has observed 146 active star-forming regions at ALMA Band 3. The data can be used to systematically investigate the spatial distribution of various dense gas tracers in a large sample of Galactic massive clumps. The dense molecular gas structures inside the cloud serve as star-forming sites, their physical states directly determine the star formation capability of the molecular cloud under feedback. A more systematic study is needed to evaluate how massive protostars influence the dense gas distribution and star formation efficiency in their natal clumps. Using ALMA data, we zoomed into the dense gas structures close to or even connecting with protoclusters. The effects of feedback were then inferred through the physical properties of the surrounding dense gas structures. The sample of the ATOMS survey includes 146 active star-forming regions. In Zhou et al. (2022), we find that hub-filament systems are very common within massive protoclusters, and that stellar feedback from HII regions gradually destroys the hub-filament systems as protoclusters evolve. The sources at different evolutionary stages are dominated by different feedback mechanisms. By comparing the physical properties of dense gas structures in the clumps at different evolutionary stages, we will be able to infer the relative importance of various feedback mechanisms.

2. Observations

2.1. ALMA observations

We used ALMA data from the ATOMS survey (Project ID: 2019.1.00685.S; PI: Tie Liu). The ATOMS sources were initially selected from a complete and homogeneous CS J = 2–1 molecular line survey toward IRAS sources with far-infrared colors characteristics of UC HII regions (Bronfman et al. 1996). As discussed in Liu et al. (2020), the ATOMS sources are distributed in very different environments of the Milky Way, and are an unbiased sample of the protoclusters with the strongest CS J = 2–1 line emission ($T_{\rm A} > 2$ K) located in the inner Galactic plane of $-80° < l < 40°$, $|b| < 2°$ (Faúndez et al. 2004).

For the data reduction, the details of the 12 m array and 7 m array ALMA observations were summarized in Liu et al. (2020, 2021). Calibration and imaging were carried out using the CASA software package version 5.6 (McMullin et al. 2007). The 7 m data and 12 m array data were calibrated separately. Then the visibility data from the 7 m and 12 m array configurations were combined and later imaged in CASA. For each source and each spectral window (spw), a line-free frequency range was automatically determined using the ALMA pipeline. This frequency range was used to (a) subtract continuum from line emission in the visibility domain, and (b) make continuum images. Continuum images were made from multifrequency synthesis of data in the line-free frequency ranges in the two 1.875 GHz wide spectral windows, spws 7 and 8, centered on ∼99.4 GHz (or 3 mm). Visibility data from the 12 m and 7 m arrays were jointly cleaned using the task tclean in CASA 5.6. We used natural weighting and a multiscale deconvolver for an optimized sensitivity and image quality. All images were primary beam-corrected. The

continuum image reaches a typical 1 σ rms noise of ~0.2 mJy with a synthesized beam FWHM size of ~2.2″. In this work, we also used $H^{13}CO^+$ J = 1–0 (86.754288 GHz) and HCO^+ J = 1–0 (89.188526 GHz) molecular lines data with spectral resolutions of 0.211 km s^{-1} and 0.103 km s^{-1}. The typical beam FWHM size and channel rms noise level for $H^{13}CO^+$ J = 1–0 line emission are ~2.5″ and 8 mJy beam^{-1}, respectively. For HCO^+ J = 1–0 line emission, the two values are ~2.4″ and 12 mJy beam^{-1}. The typical maximum-recovered angular scale in this survey is about 1 arcmin, which is comparable to the field of view (FOV) in the 12 m array observations (Liu et al. 2020).

2.2. Spitzer infrared data

We also used images at 8.0 μm, obtained by the *Spitzer* Infrared Array Camera (IRAC), as part of the GLIMPSE project (Benjamin et al. 2003). The images of IRAC were retrieved from the *Spitzer* Archive and the angular resolution of the images is about 2″.

3. Results

3.1. Evolutionary stages

Generally, the infrared properties and the presence of H40$_\alpha$ emission can be used to assess the evolutionary stages of the sources (Chambers et al. 2009; Sanhueza et al. 2012). As discussed in Zhou et al. (2022), stellar feedback from HII regions will gradually destroy the hub-filament systems as protoclusters evolve. Thus, the morphology of the hub-filament structure can also characterize the evolutionary stages. In Figs. 2 and 3 of Zhou et al. (2022), for hub-filament systems, the overlaid contours of the 3 mm continuum should be located at the center of filamentary network with the strongest $H^{13}CO^+$ J = 1–0 emission, where is the center of the gravitational potential well and the potential site for high-mass star formation. In the late stage of evolution, the 3 mm continuum emission is almost completely separated from $H^{13}CO^+$ J = 1–0 emission, indicating that the relatively evolved HII regions have dispersed their natal gas (see also Liu et al. 2023). Thus, the separation between 3 mm continuum emission and $H^{13}CO^+$ (1–0) emission can also be used for the classification. Finally, we divided ATOMS sources into three evolutionary stages using the following criteria:

1. The sources at an early stage (early sources): (1) infrared dark and without H40$_\alpha$ emission; (2) well-defined hub-filament structures; and (3) their 3mm continuum emissions concentrate on the centers of filamentary networks with the strongest $H^{13}CO^+$ (1–0) emission.
2. The sources at a middle stage (middle sources) have the same features as the early sources apart from having compact H40$_\alpha$ emission.
3. The sources at a late stage (late sources): (1) obvious triggering signs, such as the bent dense gas structures traced by $H^{13}CO^+$ (1–0) emission; (2) their 3mm emission is extended, and all of them have strong H40$_\alpha$ and 8 μm emission, indicating that 3mm continuum emission is dominated by free–free emission (Keto et al. 2008; Zhang et al. 2023). The two conditions mean that hubs in the late stage have developed into protoclusters. Powerful HII regions driven by protoclusters will destroy the hub-filament systems eventually. We can see the expansion of HII regions is tearing up the surrounding gas structures in Fig. 1c.

The number of early sources, middle sources, and late sources is 30, 11, and 23, respectively. We applied stringent criteria for the selection, and retained only sources that can be

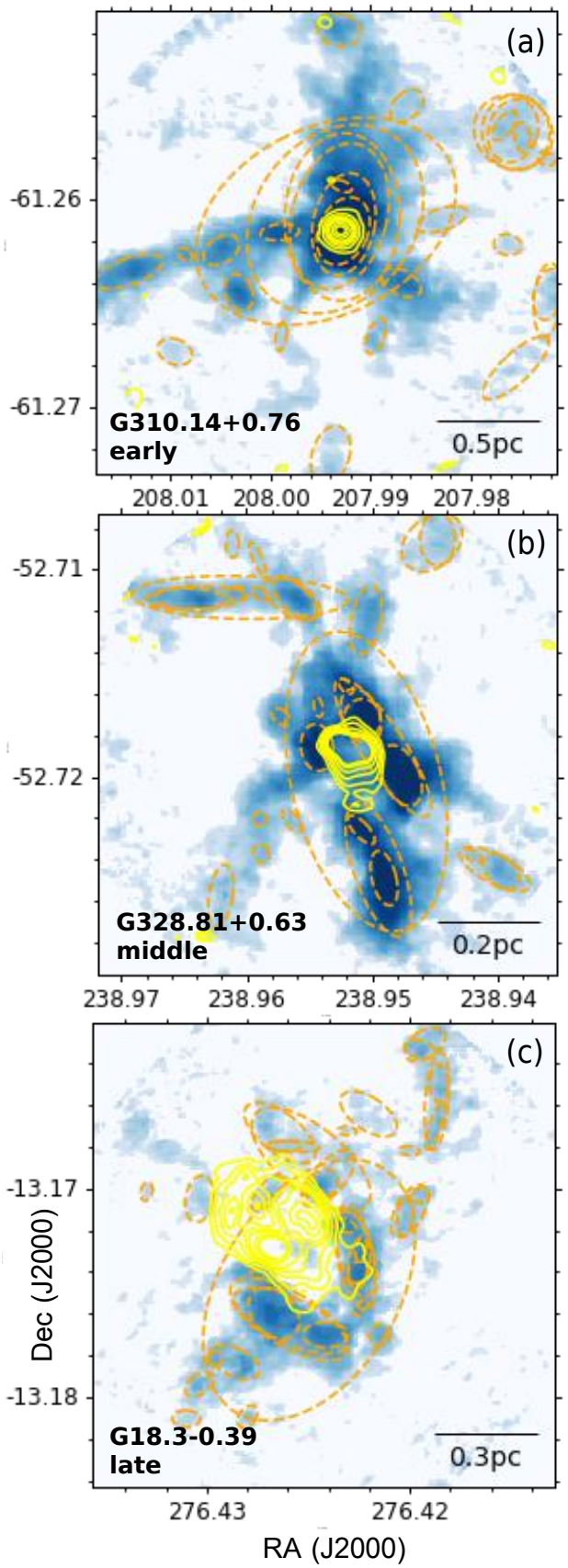


Fig. 1. Moment 0 maps of $H^{13}CO^+$ J = 1–0 for three ATOMS sources at three evolutionary stages. Yellow contours show 3 mm continuum emission with its minimum contour level of 5σ. Orange eclipses represent the retained Dendrogram structures. The maps of all 64 ATOMS sources studied in this work are presented in Appendix B.

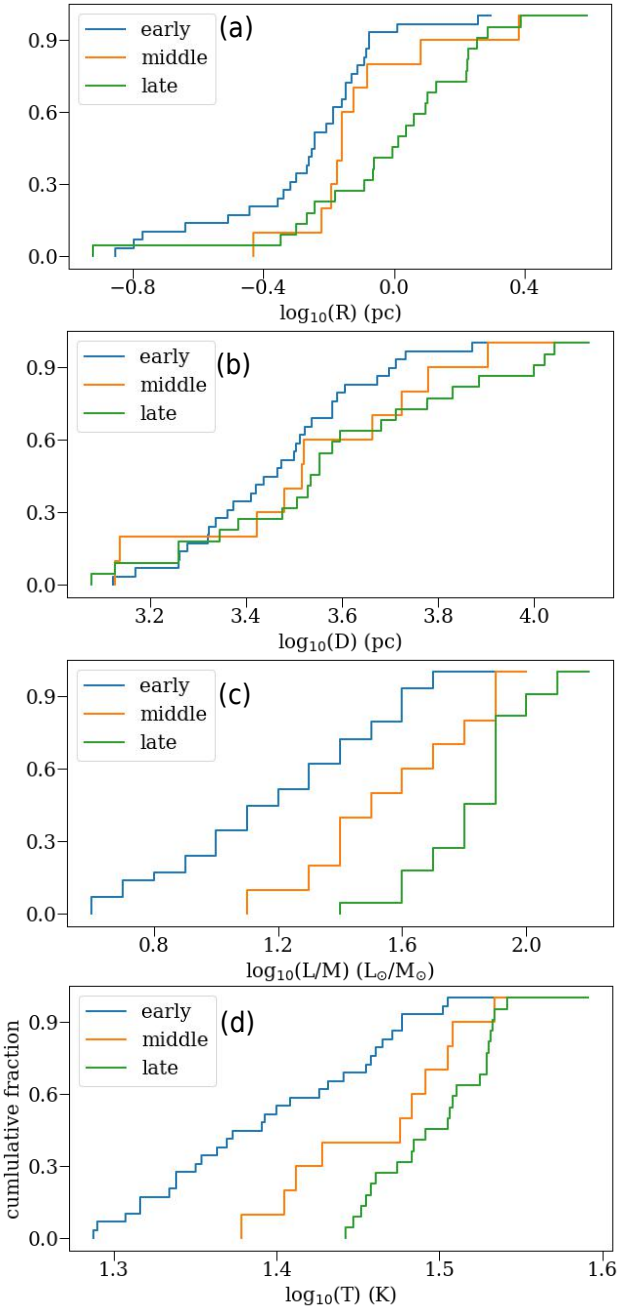


Fig. 2. Physical parameters of 64 ATOMS sources at different evolutionary stages. Panels a, b, c, and d show the distributions of the radius, R, distance, D, luminosity-to-mass ratios, L/M, and dust temperature, T, respectively.

clearly categorized in one out of the three stages. The values of the radius, distance, bolometric luminosity, mass, and dust temperature of ATOMS sources are listed in Table A1 of Liu et al. (2020). The luminosity-to-mass ratios, L/M, and dust temperature, T, of the clumps can be used to trace the evolutionary process of star formation (Saraceno et al. 1996; Krumholz & Tan 2007; Molinari et al. 2008; Liu et al. 2016; Stephens et al. 2016; Urquhart et al. 2018). As shown in Figs. 2c and d, from the early stage through to the middle and late stages, L/M and T of the sources are indeed increasing gradually, confirming that our classification is reasonable.

3.2. Dendrogram structures

The issues of identifying the structures in a PPV cube by the dendrogram algorithm are described in Sects. 3.1 and 3.2 of Zhou et al. (2024), and thus we adopted the same method as in Zhou et al. (2024) to identify dense gas structures. We directly identified hierarchical (sub-)structures based on the 2D integrated intensity (moment 0) map of $H^{13}CO^+$ (1–0) emission, and then extracted the average spectrum of each structure to investigate its velocity components and gas kinematics.

For the moment 0 map, a 5σ threshold was set, and we therefore only required the smallest area of the structure to be larger than 0.5 beam and did not set other parameters to reduce the dependence of the identification on the algorithm parameters. In Fig. 1, the structures identified by the dendrogram correspond well to the integrated intensity maps.

The algorithm approximates the morphology of each structure as an ellipse. In the dendrogram, the long and short axes of an ellipse, a and b, are the rms sizes (second moments) of the intensity distribution along the two spatial dimensions. As described in Zhou et al. (2024), a and b will give a smaller ellipse compared to the size of the identified structure, and thus multiplying by a factor of two to enlarge the ellipse is necessary. Then the effective physical radius of an ellipse is $R_{\rm eff} = \sqrt{2a \times 2b} * \mathrm{D}$, where D is the distance of the source.

As discussed in Zhou et al. (2024), two screening criteria were used to filter the structures: (1) eliminating the repetitive branch structures; and (2) excluding branch structures with complex morphology. Then for 2477 retained structures, according to their averaged spectra, 57 structures with absorption features were eliminated. Then, following the procedure in Zhou et al. (2024), we fitted the averaged spectra of 2420 structures individually using the fully automated Gaussian decomposer `GAUSSPY+` (Lindner et al. 2015; Riener et al. 2019) algorithm. The parameter settings of the decomposition were the same as that in Zhou et al. (2023). According to the line profiles, all the averaged spectra were divided into three categories: type1 (single velocity component, 2038), type2 (separated velocity components, 42), and type3 (blended velocity components, 340). We also refer to Fig. 6 of Zhou et al. (2024) for examples of typical line profiles. In the subsequent analysis, we discarded type2 structures because there are too few of them (1.7%) to simplify the discussion. Furthermore, these structures are likely the product of the overlapping of unrelated gas components.

In this work, we do not classify the structures by the nomenclatures used in *astrodendro*[1], such as leaf, branch, or trunk, because there is always a considerable overlap in the scales between leaf and branch structures, as described in Zhou et al. (2024). Moreover, in Fig. 3, the scales of all structures are presented as a continuous unimodal distribution.

3.3. Column density

To derive the column densities from $H^{13}CO^+$ (1–0) emission, we assumed conditions of local thermodynamic equilibrium (LTE) and a beam filling factor of one. The LTE analysis is described in detail in Appendix A. In Shimajiri et al. (2017), for the Aquila, Ophiuchus, and Orion B clouds, the $H^{13}CO^+$ abundances relative to H_2, $X_{H^{13}CO^+} \equiv N_{H^{13}CO^+}/N_{H_2}$, have mean values in the range $(1.5–5.8) \times 10^{-11}$. These abundance estimates are consistent within a factor of a few with the findings in other regions,

[1] https://dendrograms.readthedocs.io/en/stable/index.html

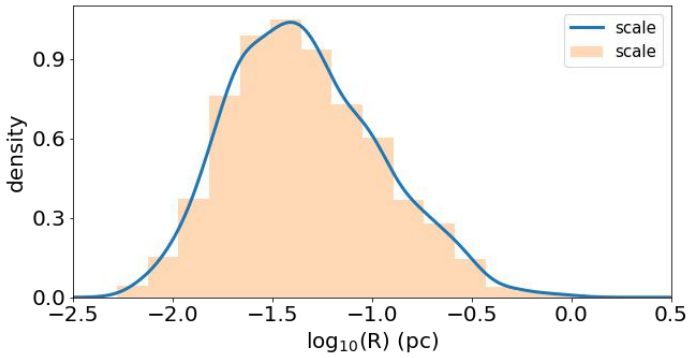


Fig. 3. Scale distribution of all identified structures. The probability density was estimated using the kernel density estimation (KDE) method.

such as $1.1 \pm 0.1 \times 10^{-11}$ in OMC 2-FIR 4 (Shimajiri et al. 2015) and $1.8 \pm 0.4 \times 10^{-11}$ in Sagittarius A (Tsuboi et al. 2011). Peretto et al. (2013) derived a $H^{13}CO^+$ abundance of 5×10^{-11} from Mopra observations towards SDC335 using the 1D non-LTE RADEX radiative transfer code (van der Tak et al. 2007). Here, we take the mean value $X_{H^{13}CO^+} \sim 4.2 \times 10^{-11}$ shown in Table 2 of Shimajiri et al. (2017).

3.4. Velocity components

In Zhou et al. (2024), type3 structures were discarded because of the complex line profiles. In this work, to increase the samples of each source, we needed to consider type3 structures. Possible reasons for the complex profiles were: (1) the optical depth effect; (2) the overlapping of uncorrelated gas components; and (3) the complex gas kinematics inside the structures. Generally, $H^{13}CO^+$ J = 1–0 line emission is optically thin, and thus we did not consider the first factor. The second factor has been discussed in detail in Zhou et al. (2022) and is not an issue in the ATOMS survey, and thus our analysis in the previous work can be done based on the moment maps. The significant overlap for the line profile of a type3 structure may indicate the correlation between different velocity components in the structure. The more severe the overlap between different velocity components, the more likely they belong to the same structure. In the extreme case, they merge into a single-peak line profile. If the velocity differences between different velocity components are small, these velocity components are likely to be different parts of the same structure rather than unrelated overlapping structures, and thus the velocity decomposition is not necessary. For example, in a hub-filament structure, the velocity of gas flow in each filament is different, leading to different velocity components in the same structure. Considering the multi-scale hub-filament scenario described in Zhou et al. (2022), type3 structures may be hub-filament structures at different scales.

In the subsequent analysis, we treat type3 structures as independent structures like type1 structures, but with more complex gas motions. We constantly provide evidence to justify this assumption. If different velocity components in a type3 structure are different parts of the same structure, the velocity dispersion of the type3 structure can be obtained by an intensity-weighted average of the velocity dispersion of all velocity components within this structure.

3.5. Statistics of the basic physical quantities

Below, we discuss the differences in the physical properties of different structure types at different evolutionary stages. For two data sets, the magnitude comparison is based on their median values. Generally, the p value of their Kolmogorov–Smirnov (KS) test is far less than 0.05, if their median values are significantly different, as shown in Fig. 4.

From the first line of Fig. 4, type3 structures have a larger velocity dispersion, higher column density, and lower virial ratios than type1 structures at different evolutionary stages, indicating that type3 structures are mainly concentrated in the hub regions, and thus have a higher column density and present multiple velocity components. In Zhou et al. (2022), from the grid maps of $H^{13}CO$ (1–0) emission, the hub regions always show multiple velocity components that should be attributed to the complex gas motions, not the superposition of uncorrelated foreground or background velocity components.

From the second line of Fig. 4, although late sources have powerful HII regions, their inner gas structures do not have a significantly larger velocity dispersion and virial ratios, compared to that in early and middle sources. In Fig. 4e, the embedded dense structures inside middle sources even have a slightly larger velocity dispersion than those inside late sources. These statistical results are unusual. We need to reconsider the effects of feedback from expanding HII regions (see Sect. 3.8 for more discussion.) Moreover, the higher column density of the embedded dense structures inside middle sources in Fig. 4f is in line with the gravitational collapse of hub-filament systems.

3.6. Scaling relations

The scaling relations between different quantities contain information on the physical conditions of the structures. The scaling relations between the velocity dispersion, σ, effective radius, R, and column density, N, shown in Fig. 5 are similar to those in Zhou et al. (2024). $\sigma - N * R$ always has a stronger correlation compared to $\sigma - N$ and $\sigma - R$[2]. However, the slopes of $\sigma - N * R$ relations are ~0.27, which may indicate the slowing down of a pure free-fall gravitational collapse. There are significant correlations between velocity dispersion and column density, which may imply that the velocity dispersion originates from the gravitational collapse, something which is also revealed by the measured velocity gradients in Sect. 3.7. $\sigma - R$ has a poor correlation, that is, a significant deviation from the Larson-relation (Larson 1981).

Additionally, it is noteworthy that both type1 and type3 structures present similar scaling relations, which further confirms that type3 structures are independent like type1 structures. The structures generated by random overlap should not show good scaling relations.

3.7. Velocity gradient

In Zhou et al. (2022), we used the moment 0 maps (integrated intensity maps) of $H^{13}CO^+$ J = 1–0 to identify filaments in ATOMS sources, and extracted the velocity and intensity along the filament skeletons from the moment 1 and moment 0 maps. Clear velocity and density fluctuations are seen along the filaments, allowing us to fit velocity gradients around the intensity peaks. The filaments like the streamlines, which can trace how gas inflows converge to local dense structures. Local dense structures as gravitational centers will accrete the surrounding

[2] For a more convenient comparison with the $\sigma - R$ and $\sigma - N$ relations, we converted the Heyer relation, $\sigma / R^{0.5} \propto N^{0.5}$, to the form $\sigma \propto (R * N)^{0.5}$ (Eq. (3) in Ballesteros-Paredes et al. 2011). Both should have a slope of 0.5.

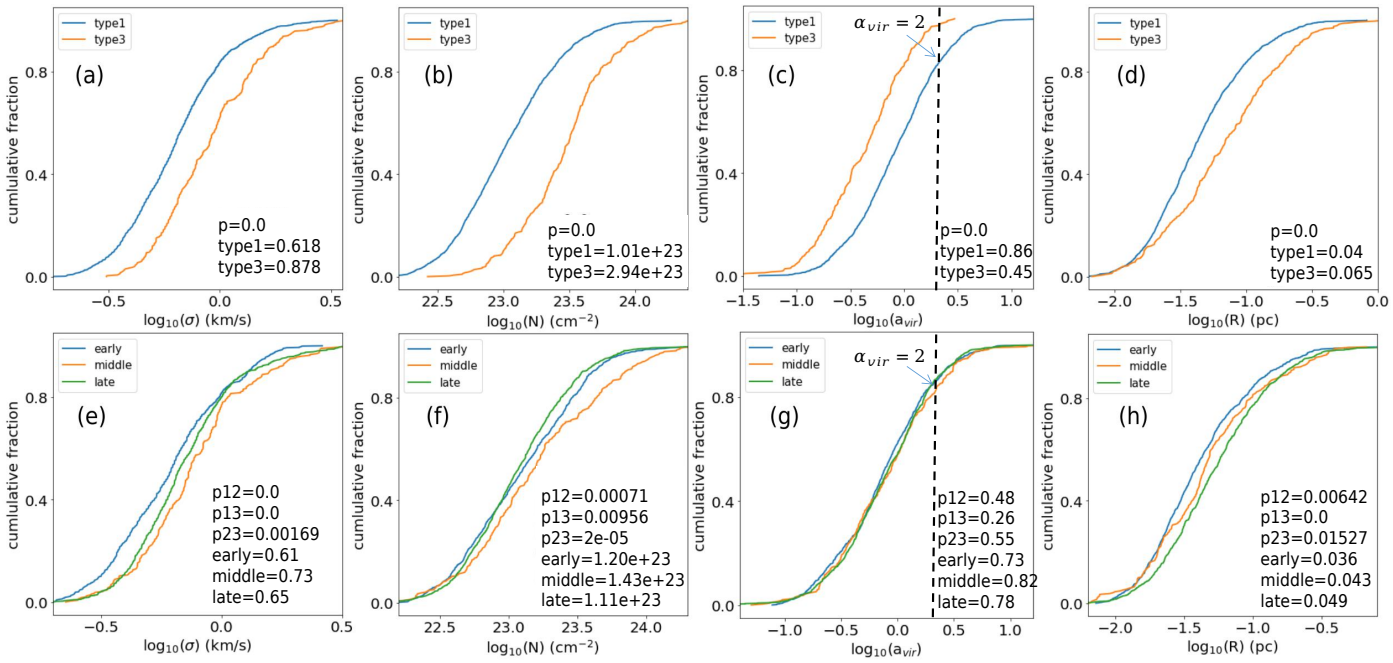


Fig. 4. Physical properties of different structure types at different evolutionary stages. The first row is the comparison between type1 and type3 structures. The second row is the comparison between early, middle and late sources. The dashed lines in panels c and g mark the positions of $\alpha_{\rm vir} = 2$. "p" is the p value of the Kolmogorov–Smirnov (KS) test. "$p12$," "$p13$," and "$p23$" are the p values of KS tests between the structures inside early and middle sources, early and late sources, and middle and late sources, respectively. The median values also show in the legends.

diffuse gases and then form local hub-filament structures. A hub-filament structure can be translated into a funnel structure in the PPV space, as described in Zhou et al. (2023). As illustrated in Fig. 9 of Zhou et al. (2023), the gradient of the funnel profile can reflect the strength of the gravitational field. In Zhou et al. (2022, 2023), the filaments will pass through multiple local hub-filament structures presented as velocity and density fluctuations. Therefore, the local velocity gradients along the filaments can reveal the strength of local gravitational fields. Then, by comparing local velocity gradients, we can assess the differences in the gravitational states of the embedded dense gas structures among different feedback environments.

For all ATOMS sources, the identified filaments, and corresponding extracted velocity and density along the filaments, have been shown in Zhou et al. (2022). We first estimated two overall velocity gradients between the velocity peaks and valleys at the two sides of the center of the gravitational potential well and ignored the local velocity fluctuation, then we also derived local velocity gradients around local intensity peaks of the $H^{13}CO^+$ emission (see Fig. 6 in Zhou et al. 2022). In Fig. 4g, the virial ratios of the embedded dense gas structures in late and early sources are comparable. In Fig. 6, the measured velocity gradients of the structures in early and late sources are also comparable, once again demonstrating that feedback rarely affects the physical properties of the embedded dense gas structures.

3.8. Feedback of early sources versus late ones

In an analysis of the kinematics in early sources, we mainly focused on the outflows and inflows. When investigating the gas kinematics in late sources, we primarily considered the ionized gas pressure. Moreover, the radiation pressure exists in all sources at different evolutionary stages. But late sources must have stronger radiation pressure than early sources, because they embed luminous ionizing (young) stellar objects. We note that inflow is not a feedback mechanism; but as a line-broadening mechanism, it is comparable to other feedback mechanisms. Below, we quantitatively estimate the velocity dispersion generated by the four physical processes, and then compare them with the statistical results in Sect. 3.5.

3.8.1. Outflow

We assumed that the line emission of $H^{13}CO^+$ J = 1–0 and high velocity emission of HCO^+ J = 1–0 are optically thin, and that the excitation temperatures of the $H^{13}CO^+$ J = 1–0 and HCO^+ J = 1–0 lines are the same and constant in both low- and high-velocity outflow gas. Then, according to Eqs. (4) and (5) in Zhou et al. (2021), the mass of the outflows can be estimated by

$$\frac{M_{\rm out}}{M_{\rm gas}} \approx \left[\frac{\rm H^{13}CO^+}{\rm HCO^+}\right] \frac{F_{\rm out,HCO^+}}{F_{\rm gas,H^{13}CO^+}}. \quad (1)$$

The expression in brackets is the abundance ratio of the molecules. Here, we took $\left[\rm H^{13}CO^+/HCO^+\right] \sim 1/60$ (Milam et al. 2005). $F_{\rm out,HCO^+}$ and $F_{\rm gas,H^{13}CO^+}$ are the corresponding total velocity-integrated fluxes (the sum of all pixels in the moment 0 maps) of the outflows and dense gases of the clump traced by the HCO^+ (1–0) and $H^{13}CO^+$ (1–0) emission. $M_{\rm gas}$ was calculated using the column density map derived by LTE analysis in Sect. 3.6. After obtaining the value of $M_{\rm out}$, according to the total velocity-integrated flux of the redshifted and blueshifted lobes of the outflows $F_{\rm out,HCO^+,r}$ and $F_{\rm out,HCO^+,b}$, the masses of the two lobes were

$$M_{\rm out,r} = \frac{F_{\rm out,HCO^+,r}}{F_{\rm out,HCO^+}} \times M_{\rm out} \quad (2)$$

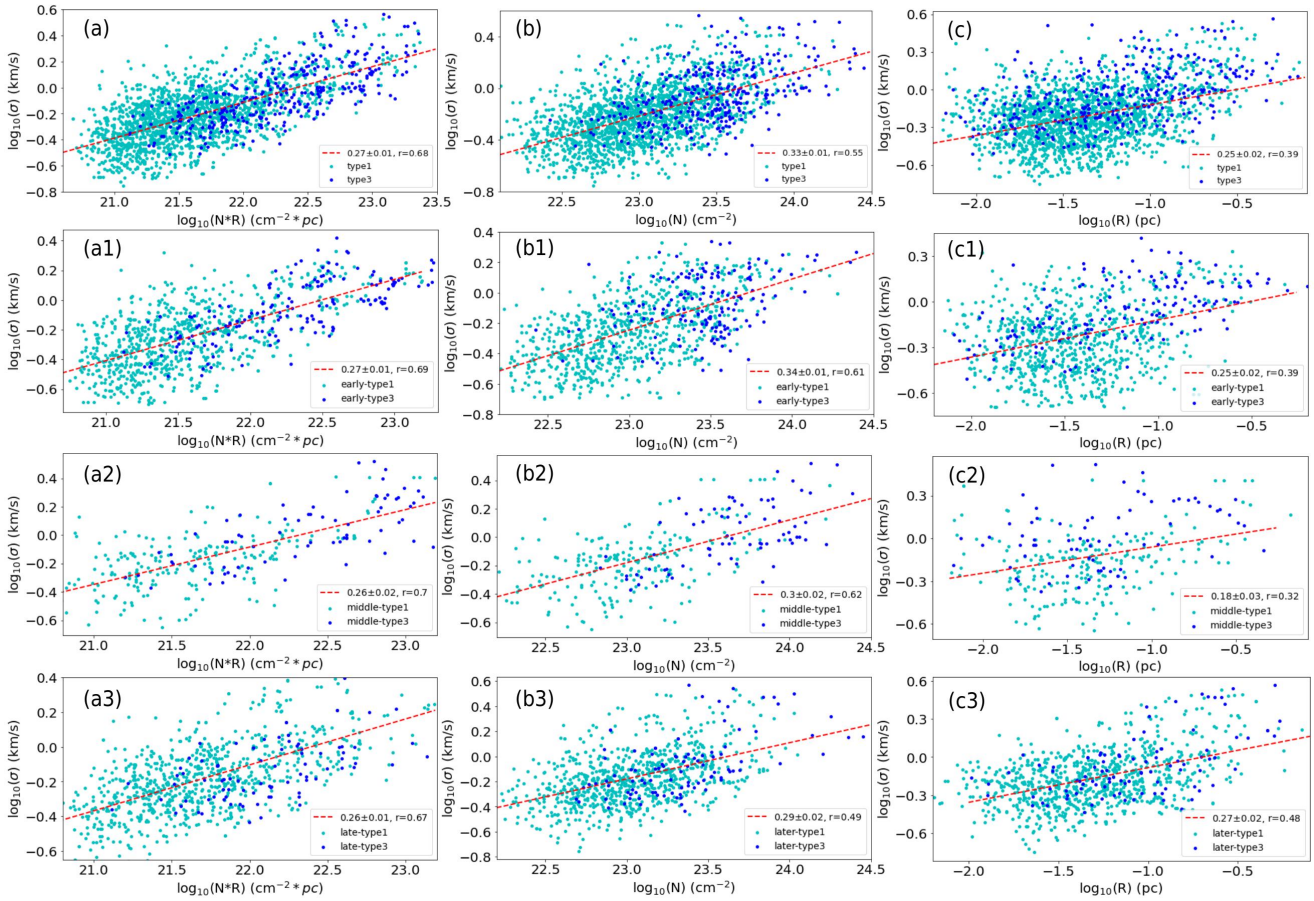


Fig. 5. Scaling relations of type1 and type3 structures. (a) $\sigma - N * R$, (b) $\sigma - N$, and (c) $\sigma - R$. σ, R and N are the velocity dispersion, effective radius, and column density of each structure, respectively. "r" represents the Pearson correlation coefficient. The values in the legends are the exponents of the power laws. The second, third, and fourth rows are the same as the first row, but for early, middle and late sources, respectively.

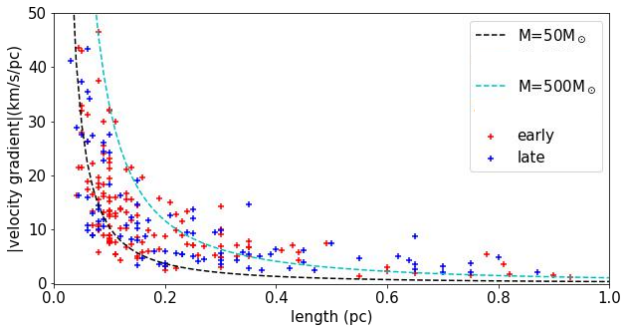


Fig. 6. Velocity gradient versus the length over which the gradient has been fitted. Blue and red crosses represent the velocity gradients fitted in early and late sources, respectively. The dashed lines show free-fall velocity gradients for comparison. For the free-fall model, black and cyan lines denote masses of 50 $M_\odot$ and 500 $M_\odot$, respectively.

and

$$M_{\rm out,b} = \frac{F_{\rm out,HCO^+,b}}{F_{\rm out,HCO^+}} \times M_{\rm out}. \quad (3)$$

Assuming that the momentum of the outflows completely transfers to the surrounding dense gas (i.e., the upper limit), then the corresponding turbulent velocity, $v_{\rm turb,out}$, inside the dense gas caused by the outflows can be estimated from the equation

$$M_{\rm gas} v_{\rm turb,out} = M_{\rm out,r} v_{\rm r} + M_{\rm out,b} v_{\rm b}. \quad (4)$$

According to the moment 1 maps of the two lobes, their average velocities are $v_{\rm r0}$ and $v_{\rm b0}$, the systematic velocity of the clump is v_0, and then the velocities of the two lobes are $v_{\rm r} = v_{\rm r0} - v_0$ and $v_{\rm b} = v_0 - v_{\rm b0}$.

3.8.2. Inflow

Assuming spherical symmetry, the clump has a mass inflow rate of

$$\dot{M} \approx 4\pi r_{\rm in}^2 \rho v_{\rm in}, \quad (5)$$

with

$$\rho = \frac{M_{\rm c}}{4/3\pi R_{\rm c}^3}. \quad (6)$$

$R_{\rm c}$ and $M_{\rm c}$ are the radius and mass of the clump, respectively. $r_{\rm in}$ and $v_{\rm in}$ are the infall radius and velocity. We note that the distance, D, radius, $R_{\rm c}$, dust temperature, $T_{\rm d}$, mass, $M_{\rm c}$, and bolometric luminosity, $L_{\rm bol}$, of the clumps were taken as the values listed in Table A.1 of Liu et al. (2020). For clumps to be hub-filament systems, they should be in a state of global gravitational collapse, and thus we assumed that $R_{\rm c} = r_{\rm in}$.

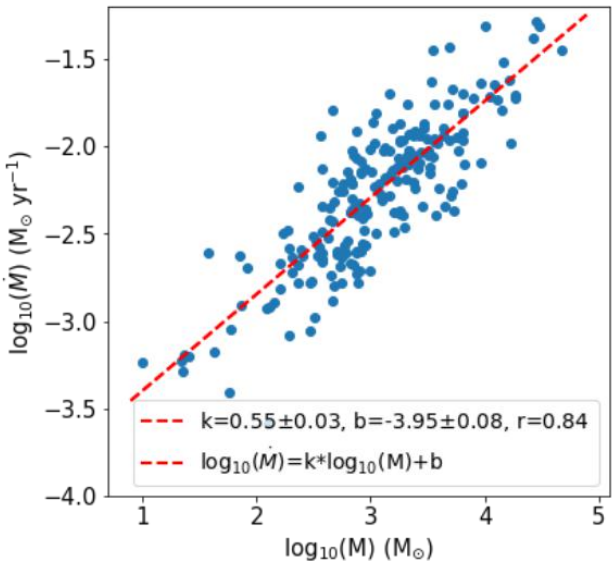


Fig. 7. The relation between the inflow rate and the mass of the clump. Data points are from the tables of He et al. (2015, 2016). "r" is the Pearson coefficient.

For the inflow rate, $\dot{M}$, of the clump, the tables in He et al. (2015, 2016) listed mass infall rates of 231 clumps. Based on the tables, we fitted the quantitative relation between the inflow rate and the mass of the clump, as shown in Fig. 7. Then the infall rates of early sources could be estimated using this relation according to their masses. Finally, the infall velocity could be estimated by

$$v_{\rm in} = \frac{\dot{M} \times R_{\rm c}}{3M_{\rm c}}. \quad (7)$$

3.8.3. Radiation pressure

Following Lopez et al. (2014), the volume-averaged radiation pressure, $P_{\rm rad}$, is

$$P_{\rm rad} = \frac{3L_{\rm bol}}{4\pi R_{\rm c}^2 c}. \quad (8)$$

$L_{\rm bol}$ is the bolometric luminosity of the clump.

3.8.4. Ionized gas pressure

Following Lopez et al. (2014), the ionized gas pressure is given by the ideal gas law, $P_{\rm HII} = (n_{\rm e} + n_{\rm H} + n_{\rm He})kT_{\rm HII}$, where $n_{\rm e}$, $n_{\rm H}$, and $n_{\rm He}$ are the electron, hydrogen, and helium number densities, respectively, and $T_{\rm HII}$ is the temperature of the HII gas, which was assumed to be the same for electrons and ions. If helium is singly ionized, then $n_{\rm e} + n_{\rm H} + n_{\rm He} \approx 2n_{\rm e}$. The electron temperature and density of resolved UC HII regions identified in the ATOMS survey were derived from radio recombination lines, $H_{40}\alpha$, and 3 mm continuum emission in Zhang et al. (2023). The corresponding velocities contributed by $P_{\rm rad}$ and $P_{\rm HII}$ were estimated by $P_{\rm rad} = \rho v_{\rm rad}^2$ and $P_{\rm HII} = \rho v_{\rm HII}^2$.

3.8.5. Relative importance of feedback processes

In Tables 1 and 2, $v_{\rm HII}$ is significantly larger than other velocities, consistent with the conclusions in Lopez et al. (2014) that the ionized gas feedback dominates over the other considered feedback mechanisms. In this work, we did not intend to study which feedback mechanism is the strongest, and thus we ignored some feedback mechanisms, such as dust-reprocessed radiation and stellar wind, which may also play important roles (Pabst et al. 2019; Olivier et al. 2021; Henshaw et al. 2022). Here, we note that the estimated $v_{\rm HII}$ is obviously larger than the observed line widths in $H^{13}CO^+$ J = 1–0. This is not surprising because the energy is deposited into the ionized gas and photodissociation regions. As expected, late sources present significantly stronger radiation pressure than early sources. For early sources, the velocity dispersion caused by the outflows is significantly smaller than that from the infalls and the radiation pressure.

Table 1. Velocity dispersion produced by different physical processes in early sources.

Early-sources	$v_{\rm turb,out}$ (km s^{-1})	$v_{\rm in}$ (km s^{-1})	$v_{\rm rad}$ (km s^{-1})
G261.65-2.09	0.07	0.49 ± 0.16	0.28
G294.51-1.62	0.11	0.59 ± 0.19	0.33
G305.8-0.24	0.10	0.86 ± 0.32	0.6
G305.89+0.02	0.12	0.92 ± 0.33	0.28
G307.76+0.35	0.03	0.81 ± 0.28	0.17
G310.14+0.76	0.15	1.26 ± 0.47	0.91
G313.58+0.32	0.03	0.64 ± 0.21	0.33
G313.77-0.86	0.10	0.95 ± 0.34	0.34
G318.05+0.09	0.13	1.23 ± 0.45	0.61
G326.47+0.7	0.24	0.85 ± 0.3	0.24
G327.12+0.51	0.00	1.28 ± 0.47	0.78
G329.47+0.5	0.26	1.11 ± 0.42	0.29
G329.41-0.46	0.04	1.19 ± 0.5	1.19
G335.59-0.29	0.11	1.14 ± 0.42	0.46
G339.62-0.12	0.09	1.22 ± 0.42	0.41
G339.88-1.26	0.06	0.92 ± 0.33	1.04
G341.13-0.35	0.01	1.31 ± 0.47	0.33
G343.13-0.06	0.32	0.94 ± 0.35	0.74
G352.63-1.07	0.11	0.86 ± 0.28	0.35
G6.8-0.26	0.04	1.08 ± 0.39	0.36
G12.42+0.51	0.06	1.11 ± 0.38	0.42
G12.91-0.26	0.20	1.01 ± 0.4	0.5
G14.33-0.64	0.09	0.98 ± 0.34	0.31
G16.59-0.05	0.04	1.07 ± 0.4	0.47
G19.36-0.03	0.11	1.14 ± 0.41	0.22
G19.88-0.53	0.14	1.1 ± 0.41	0.28
G25.65+1.05	0.13	1.03 ± 0.36	0.37
G28.86+0.07	0.04	1.6 ± 0.65	0.99
G31.58+0.08	0.04	0.98 ± 0.36	0.51
G37.43+1.52	0.05	0.42 ± 0.15	0.48
median	0.10	1.02 ± 0.37	0.39
mean	0.10	1 ± 0.37	0.49

Notes. $v_{\rm turb,out}$, $v_{\rm in}$, and $v_{\rm rad}$ correspond to the outflows, the inflows, and the radiation pressure, respectively.

4. Discussion

4.1. Gravitational collapse under feedback

As discussed in Sect. 3.7, even under strong feedback, the embedded dense structures in late sources can still retain

Table 2. Velocity dispersion produced by different physical processes in late sources.

Late-sources	$v_{\rm HII}$ (km s^{-1})	$v_{\rm rad}$ (km s^{-1})
G268.42-0.85	1.8 ± 0.14	0.61
G305.2+0.03	6 ± 0.5	1.27
G312.11+0.31	4.85 ± 0.43	1.02
G316.14-0.51	4.44 ± 0.36	1.15
G324.2+0.12	8.98 ± 0.71	1.68
G326.45+0.91	5.75 ± 0.54	1.16
G326.72+0.61	6.14 ± 0.52	0.84
G329.47+0.22	6.78 ± 0.44	0.91
G332.65-0.61	5.88 ± 0.42	1.16
G333.31-0.37	7.23 ± 0.61	1.4
G337.12-0.17	13.79± 1.12	2.77
G340.78-1.02	4.08 ± 0.26	1.16
G344.22-0.59	4.41 ± 0.35	0.63
G350.51+0.96	8.9 ± 0.65	1.06
G350.1+0.08	2.98 ± 0.21	1.63
G1.13-0.11	1.44 ± 0.15	1.17
G8.14+0.22	8.32 ± 0.67	1.2
G9.62+0.19	9.5 ± 0.64	1.41
G11.94-0.62	6.68 ± 0.52	0.83
G13.87+0.28	5.09 ± 0.38	1.4
G18.3-0.39	6.8 ± 0.48	0.9
G23.95+0.15	10.33 ± 0.71	1.55
G32.8+0.19	16.13 ± 1.35	1.96
median	6.14 ± 0.5	1.16
mean	6.8 ± 0.53	1.26

Notes. $v_{\rm HII}$ and $v_{\rm rad}$ correspond to the ionized gas pressure and the radiation pressure, respectively.

gravitational collapse. Feedback may redistribute local dense gas structures, as illustrated in Fig. 8, because feedback does not seem to affect the internal dynamics of the embedded dense gas structures. The results are consistent with the conclusion in Zhou et al. (2024): although the feedback disrupting the molecular clouds will break up the original cloud complex, the substructures of the original complex can be reorganized into new gravitationally governed configurations around new gravitational centers. This process is accompanied by structural destruction and generation, and by changes in gravitational centers, but gravitational collapse is always ongoing.

Moreover, in Fig. 6, we can find the same gas kinematic mode presented in Zhou et al. (2023). The variations in the velocity gradients at small and large scales (~0.2 pc as the boundary) are consistent with gravitational free-fall, with central masses of ~50 $M_\odot$ and ~500 $M_\odot$. It means that the velocity gradients on larger scales require larger masses to be maintained, and larger masses imply larger scales; that is to say, the larger-scale inflow is driven by the larger-scale structure, which may be the gravitational clustering of smaller-scale structures, consistent with the hierarchical or multi-scale hub-filament structures in the clouds and the gas inflow from large to small scales. In Zhou et al. (2023), we studied the gas motions of the G333 molecular cloud complex. The longest filament is ~50 pc; however, the largest scale in Fig. 6 is only ~1 pc. The similar gas kinematic modes may indicate the self-similar kinematic structures on cloud-clump and clump-core scales, as illustrated in Fig. 9 of Zhou et al. (2023).

4.2. The role of feedback

In Sect. 3.8, the significantly larger $v_{\rm HII}$ will naturally make us think that late sources should have the strongest feedback effects. However, that is not the case, as shown in Sect. 3.5. In Sect. 3.5, overall, the velocity dispersion and virial ratios of the structures inside early and late sources are comparable. As shown in Sect. 3.7, the gas structures in late sources are also in gravitational collapse like those in early sources. Gravitational collapse may also contribute significantly to the velocity dispersion of the gas structures in late sources. Thus, the proportion of the velocity dispersion of the gas structures contributed by expanding HII regions in late sources decreases further. The clouds exist as gas reservoirs that contain dense gas structures as local gravitational centers. Feedback from HII regions destroys the clouds by mechanical force; in other words, the expanding HII regions push away and reshape the loose cloud complex (see e.g., McLeod et al. 2015; Beuther et al. 2022; Bonne et al. 2023). But the processes do not significantly affect the physical properties of the embedded dense gas structures in the clouds, suggesting that HII regions may not effectively inject momentum into the embedded dense gas structures as star-forming sites. However, the triggering does not necessarily result from momentum being transferred inside the gas structures but by the gas structures being compressed, and so becoming gravitationally unstable.

Although the feedback from HII regions is very strong, it is acting from the outside on the embedded dense gas structures. By comparison, the outflows and inflows can go deep inside of the embedded dense gas structures, and thus the effect of them on the physical properties of the gas structures is not necessarily weaker than the feedback from HII regions. We should not only focus on the strength of a feedback mechanism itself, but also assess the extent to which it can affect the embedded dense gas structures.

In Fig. 4e, the structures in middle sources display a slightly larger velocity dispersion than those in late sources. Notably, middle-stage sources possess larger-scale hub-filament structures than their late-stage counterparts. The complex gas kinematics in these hub regions within middle-stage sources may contribute to the higher velocity dispersion observed. Furthermore, early sources also exhibit large-scale hub-filament structures, but their gas structures present a smaller velocity dispersion compared to those in middle sources. This suggests that feedback from UC-HII regions in middle sources plays a significant role. At least, we can conclude that feedback from protocluster-driven HII regions may not be the most powerful factor affecting the physical properties of local dense gas structures. It is important to note that the above discussion treats each clump in isolation, without considering the potential effects of the surrounding gas environment on the physical properties of the internal structures of the clumps, due to the limited FOV of ALMA observations. Given that the protocluster-driven HII regions may still be bound by the abundant surrounding gas, possible quasi-static expanding processes also indicate the limited effect of feedback from HII regions on the embedded dense gas structures. As shown in Kuhn et al. (2019), star clusters and associations that are still embedded in molecular clouds are less likely to be expanding than those that are partially or fully revealed.

As illustrated in Fig. 8, hierarchical hub-filament structures of star-forming regions mean that star-forming regions exist as loose gravitational complexes. Feedback will reshape the overall structures of star-forming regions. This process is accompanied

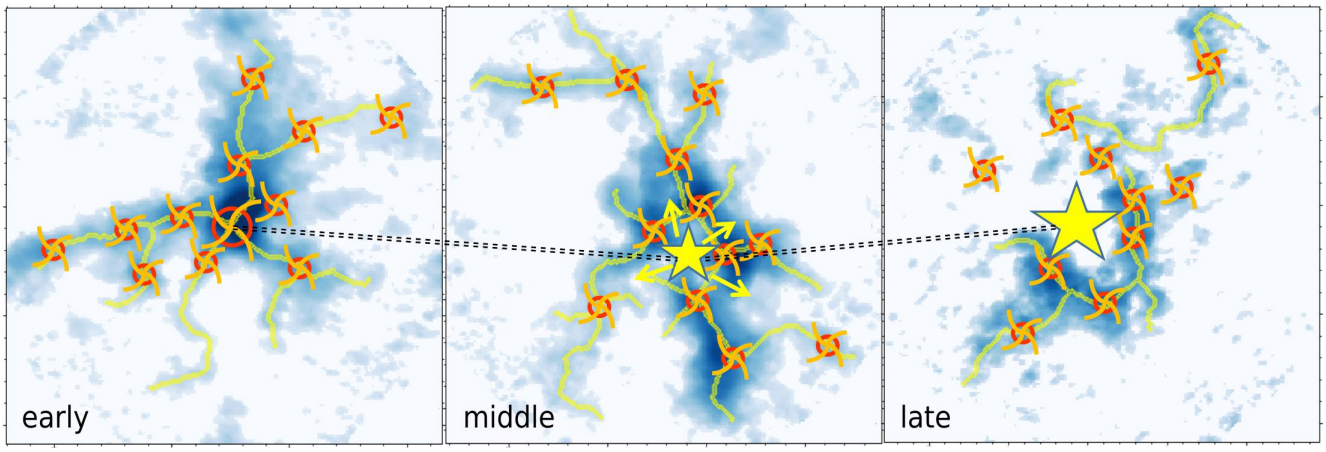


Fig. 8. Schematic diagram of the scenario argued in this work. Evolution of a cloud with multi-scale hub-filament structures with increasing effects of feedback. Here we borrowed three pictures from Fig. 1 as the background, for demonstration only. The yellow lines are the filament skeletons identified in Zhou et al. (2022). Considering the competitive accretion scenario, the precursor of the protocluster is located in the central position of the hub. With the evolution of the protocluster, feedback from HII regions destroys the hub by mechanical force, i.e., the expanding HII regions push away and reshape the hub region, accompanied by the redistribution of small-scale dense structures that may be local hub-filament structures. We note that small-scale hub-filament structures were not directly identified in this work due to the limitation of the resolution.

by the redistribution of the internal structures of the clouds, while the gravitational coupling between the internal structures also reorganizes them into new gravitationally governed configurations around new gravitational centers. Generally, feedback will disperse the clouds. Finally, to some extent, the reorganization is a resistance to the dispersing process, although the resistance will in the end fail. The negative effects of the feedback are that it destroys the gas reservoir inhabited by dense gas structures, inhibiting the further growth of dense structures by accreting the surrounding gas, and thus suppressing the formation of massive stars. On the other hand, if feedback makes the dense gas structures more spread out, and if these dense structures are still in a gas-rich environment, the looser distribution will improve the material accumulation efficiency of the dense structures by weakening the competitive accretion between them, and may eventually form more massive stars.

5. Summary

A total of 64 ATOMS sources at different evolutionary stages were selected to investigate the kinematics and dynamics of gas structures under feedback.

1. We directly identified dense gas structures based on the 2D integrated intensity (moment 0) map of $H^{13}CO^+$ (1–0) emission, and then extracted the average spectrum of each structure to investigate their velocity components and gas kinematics. We discussed the differences in the physical properties of different structure types at different evolutionary stages. (1) Type3 structures with blended velocity components (see Fig. 6 of Zhou et al. 2024) have a larger velocity dispersion, higher column density, and lower virial ratios than type1 structures with a single velocity component at different evolutionary stages, which can be attributed to type3 structures being mainly located in hub regions. (2) Overall, the velocity dispersion and virial ratios of the structures inside early and late sources are comparable. The structures inside middle sources even have a slightly larger velocity dispersion than those inside late sources. These unusual results led us to question the effectiveness of feedback from HII regions in late sources.

2. $\sigma - N * R$ always has a stronger correlation than $\sigma - N$ and $\sigma - R$. There are significant correlations between velocity dispersion and column density, which may imply that the velocity dispersion originates in the gravitational collapse, something also revealed by the measured velocity gradients. Type1 and type3 structures present similar scaling relations, which further confirms that type3 structures are independent structures, not the superposition of uncorrelated foreground or background velocity components.

3. Dense gas structures in late sources are in gravitational collapse like those in early sources. Late sources do not have large-scale hub-filament structures, but the embedded dense gas structures in late sources show similar kinematic modes to those in early and middle sources. These results may be explained by the multi-scale hub-filament structures in the clouds.

4. The differences in the gravitational states of dense gas structures in different feedback environments were investigated by comparing local velocity gradients. The measured velocity gradients of the structures in early and late sources are comparable, indicating that feedback has a limited impact on the kinematic properties of local dense gas structures. The velocity gradients on larger scales require a larger mass to be maintained, and larger masses imply larger scales; that is to say, the larger-scale inflow is driven by the larger-scale structure, which may be the gravitational clustering of smaller-scale structures, consistent with the hierarchical or multi-scale hub-filament structures in the clouds and the gas inflow from large to small scales.

5. We quantitatively estimated the velocity dispersion generated by the outflows, inflows, ionized gas pressure, and radiation pressure, and found that the ionized gas feedback is stronger than other feedback mechanisms. Late sources present significantly stronger radiation pressure than early sources. For early sources, the velocity dispersion caused by the outflows is significantly smaller than that from the infalls and the radiation pressure. However, although feedback from HII regions is the strongest, it does not significantly affect the physical properties of the embedded dense gas structures in the clumps, suggesting that HII regions may not effectively inject momentum into the embedded dense gas structures as star-forming sites.

6. This work on clump-core scales reaches the same conclusion as the study of the G333 molecular cloud complex on cloud-clump scales in Zhou et al. (2024), and thus we suggest that although feedback from cloud to core scales will

break up the original cloud complex, the substructures of the original complex can be reorganized into a new gravitationally governed configuration around new gravitational centers. This process is accompanied by structural destruction and generation, and by changes in gravitational centers, but gravitational collapse is always ongoing.

Acknowledgements. We would like to thank the referee for the detailed and constructive comments and suggestions that significantly improve and clarify this work. We thank J.H. He to share the tables of their studied infall samples. This paper makes use of the following ALMA data: ADS/JAO.ALMA♯2019.1.00685.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

References

Areal, M. B., Paron, S., Ortega, M. E., & Duvidovich, L. 2019, PASA, 36, e049
Ballesteros-Paredes, J., Hartmann, L. W., Vázquez-Semadeni, E., Heitsch, F., & Zamora-Avilés, M. A. 2011, MNRAS, 411, 65
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953
Beuther, H., Schneider, N., Simon, R., et al. 2022, A&A, 659, A77
Bonne, L., Kabanovic, S., Schneider, N., et al. 2023, A&A, 679, L5
Bronfman, L., Nyman, L. A., & May, J. 1996, A&AS, 115, 81
Chambers, E. T., Jackson, J. M., Rathborne, J. M., & Simon, R. 2009, ApJS, 181, 360
Dale, J. E., Ercolano, B., & Bonnell, I. A. 2012, MNRAS, 424, 377
Dale, J. E., Ngoumou, J., Ercolano, B., & Bonnell, I. A. 2014, MNRAS, 442, 694
Deharveng, L., Zavagno, A., & Caplan, J. 2005, A&A, 433, 565
Dib, S. 2023, MNRAS, 524, 1625
Dib, S., Shadmehri, M., Padoan, P., et al. 2010, MNRAS, 405, 401
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Evans, N. J. II, Heyer, M., Miville-Deschênes, M.-A., Nguyen-Luong, Q., & Merello, M. 2021, ApJ, 920, 126
Faúndez, S., Bronfman, L., Garay, G., et al. 2004, A&A, 426, 97
Frerking, M. A., Langer, W. D., & Wilson, R. W. 1982, ApJ, 262, 590
Garden, R. P., Hayashi, M., Gatley, I., Hasegawa, T., & Kaifu, N. 1991, ApJ, 374, 540
Goldsmith, P. F., Heyer, M., Narayanan, G., et al. 2008, ApJ, 680, 428
Hacar, A., Alves, J., Tafalla, M., & Goicoechea, J. R. 2017, A&A, 602, L2
Hartmann, L., & Burkert, A. 2007, ApJ, 654, 988
He, Y.-X., Zhou, J.-J., Esimbek, J., et al. 2015, MNRAS, 450, 1926
He, Y.-X., Zhou, J.-J., Esimbek, J., et al. 2016, MNRAS, 461, 2288
Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nat. Astron., 4, 1064
Henshaw, J. D., Krumholz, M. R., Butterfield, N. O., et al. 2022, MNRAS, 509, 4758
Kendrew, S., Simpson, R., Bressert, E., et al. 2012, ApJ, 755, 71
Keto, E., Zhang, Q., & Kurtz, S. 2008, ApJ, 672, 423
Krumholz, M. R., & Tan, J. C. 2007, ApJ, 654, 304
Krumholz, M. R., Bate, M. R., Arce, H. G., et al. 2014, in Protostars and Planets VI, eds. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 243
Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32
Lancaster, L., Ostriker, E. C., Kim, J.-G., & Kim, C.-G. 2021a, ApJ, 914, 90
Lancaster, L., Ostriker, E. C., Kim, J.-G., & Kim, C.-G. 2021b, ApJ, 922, L3
Larson, R. B. 1981, MNRAS, 194, 809
Lewis, S. C., McMillan, S. L. W., Low, M.-M. M., et al. 2023, ApJ, 944, 211
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Liu, T., Kim, K.-T., Yoo, H., et al. 2016, ApJ, 829, 59
Liu, H.-L., Figueira, M., Zavagno, A., et al. 2017, A&A, 602, A95
Liu, T., Evans, N. J., Kim, K.-T., et al. 2020, MNRAS, 496, 2790
Liu, H.-L., Liu, T., Evans, Neal J., I., et al. 2021, MNRAS, 505, 2801
Liu, H.-L., Tej, A., Liu, T., et al. 2023, MNRAS, 522, 3719
Lopez, L. A., Krumholz, M. R., Bolatto, A. D., et al. 2014, ApJ, 795, 121
Mangum, J. G., & Shirley, Y. L. 2015, PASP, 127, 266
Matzner, C. D. 2002, ApJ, 566, 302
Mazumdar, P., Wyrowski, F., Colombo, D., et al. 2021a, A&A, 650, A164
Mazumdar, P., Wyrowski, F., Urquhart, J. S., et al. 2021b, A&A, 656, A101
McLeod, A. F., Dale, J. E., Ginsburg, A., et al. 2015, MNRAS, 450, 1057
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Data Analysis Software and Systems XVI, eds. R. A. Shaw, F. Hill, & D. J. Bell, Astronomical Society of the Pacific Conference Series, 376, 127
Milam, S. N., Savage, C., Brewster, M. A., Ziurys, L. M., & Wyckoff, S. 2005, ApJ, 634, 1126
Molinari, S., Pezzuto, S., Cesaroni, R., et al. 2008, A&A, 481, 345
Olivier, G. M., Lopez, L. A., Rosen, A. L., et al. 2021, ApJ, 908, 68
Pabst, C., Higgins, R., Goicoechea, J. R., et al. 2019, Nature, 565, 618
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Peretto, N., Fuller, G. A., André, P., et al. 2014, A&A, 561, A83
Peretto, N., Rigby, A. J., Louvet, F., et al. 2023, MNRAS, 525, 2935
Pineda, J. E., Caselli, P., & Goodman, A. A. 2008, ApJ, 679, 481
Rahner, D., Pellegrini, E. W., Glover, S. C. O., & Klessen, R. S. 2017, MNRAS, 470, 4453
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Rogers, H., & Pittard, J. M. 2013, MNRAS, 431, 1337
Rugel, M. R., Rahner, D., Beuther, H., et al. 2019, A&A, 622, A48
Sanhueza, P., Jackson, J. M., Foster, J. B., et al. 2012, ApJ, 756, 60
Saraceno, P., Andre, P., Ceccarelli, C., Griffin, M., & Molinari, S. 1996, A&A, 309, 827
Schneider, N., Csengeri, T., Bontemps, S., et al. 2010, A&A, 520, A49
Shimajiri, Y., Kitamura, Y., Nakamura, F., et al. 2015, ApJS, 217, 7
Shimajiri, Y., André, P., Braine, J., et al. 2017, A&A, 604, A74
Smith, M. C., Sijacki, D., & Shen, S. 2018, MNRAS, 478, 302
Stephens, I. W., Jackson, J. M., Whitaker, J. S., et al. 2016, ApJ, 824, 29
Thompson, M. A., Urquhart, J. S., Moore, T. J. T., & Morgan, L. K. 2012, MNRAS, 421, 408
Tsuboi, M., Tadaki, K.-I., Miyazaki, A., & Handa, T. 2011, PASJ, 63, 763
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627
Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., Ballesteros-Paredes, J., & Klessen, R. S. 2009, ApJ, 707, 1023
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
Watkins, E. J., Peretto, N., Marsh, K., & Fuller, G. A. 2019, A&A, 628, A21
Wyrowski, F., Güsten, R., Menten, K. M., et al. 2016, A&A, 585, A149
Xu, F., Wang, K., Liu, T., et al. 2024, ApJS, 270, 9
Zavagno, A., Deharveng, L., Comerón, F., et al. 2006, A&A, 446, 171
Zhang, C., Zhu, F.-Y., Liu, T., et al. 2023, MNRAS, 520, 3245
Zhou, J.-W., Liu, T., Li, J.-Z., et al. 2021, MNRAS, 508, 4639
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2024, A&A, 682, A128

Original papers of Chapter.5

The original paper of Sec.5.2,

Zhou, J. W., Kroupa, P., & Dib, S. 2024, PASP, 136, 094301
Physical Properties of Embedded Clusters in ATLASGAL Clumps with H ii Regions

The original paper of Sec.5.3,

Zhou, J. W., Dib, S., & Kroupa, P. 2024, PASP, 136, 094302
The Star Formation Histories, Star Formation Efficiencies and Ionizing Sources of ATLASGAL Clumps with H ii Regions

The original paper of Sec.5.4,

Zhou, J.-w., Kroupa, P., & Dib, S. 2024, A&A, 688, L19
Self-similar cluster structures in massive star-forming regions: Isolated evolution from clumps to embedded clusters

The original paper of Sec.5.6,

Zhou, J.-W., Kroupa, P., & Wu, W. 2024, A&A, 691, A204
Evolution of the mass-radius relation of expanding very young star clusters

The original paper of Sec.5.5 and Sec.5.7,

Zhou, J.-w., Dib, S., & Kroupa, P. 2024, A&A, 691, A293
The currently observed clumps cannot be the "direct" precursors of the currently observed open clusters

Zhou, J.-w., Dib, S., & Kroupa, P. 2025, MNRAS, 537, 845
The post-gas expulsion coalescence of embedded clusters as an origin of open clusters

The original paper of Sec.5.8,

Zhou, J.-W., Kroupa, P., under revision
Origin of open clusters revealed by the evolution of the $m_{\rm max} - M_{\rm ecl}$ relation

The original paper of Sec.5.9,

Zhou, J.-W., Jadhav, V., Kroupa, P., submitted
Constraining the gas expulsion modes of embedded clusters by analyzing the expansion speeds of open clusters

The original paper of Sec.5.10,

Zhou, J.-W., under revision
The dependence of black hole formation in open clusters on the cluster formation process

Publications of the Astronomical Society of the Pacific, 136:094301 (10pp), 2024 September https://doi.org/10.1088/1538-3873/ad6f44
© 2024. The Author(s). Published by IOP Publishing Ltd on behalf of the Astronomical Society of the Pacific (ASP).

OPEN ACCESS

Physical Properties of Embedded Clusters in ATLASGAL Clumps with H II Regions

J. W. Zhou[1], Pavel Kroupa[2,3], and Sami Dib[4]
[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany; jwzhou@mpifr-bonn.mpg.de
[2] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14-16, 53115 Bonn, Germany; pkroupa@uni-bonn.de
[3] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic
[4] Max Planck Institute für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany; dib@mpia.de
Received 2024 March 4; accepted 2024 August 14; published 2024 September 3

Abstract

Using the optimal sampling model, we synthesized the embedded clusters of ATLASGAL clumps with H II regions (H II-clumps). The 0.1 Myr isochrone was used to estimate the bolometric luminosity of each star in an embedded cluster, we also added the accretion luminosity of each star in the embedded cluster. The total bolometric luminosity of synthetic embedded clusters can well fit the observed bolometric luminosity of H II-clumps. More realistically, we considered the age spread in the young star and protostar populations in embedded clusters of H II-clumps by modeling both constant and time-varying star formation histories (SFHs). According to the age distribution of the stellar population, we distributed the appropriate isochrones to each star, and sorted out the fraction of stellar objects that are still protostars (Class 0 and Class I phases) to properly add their accretion luminosities. Compared to a constant SFH, burst-like and time-dependent SFHs can better fit the observational data. We found that as long as 20% of the stars within the embedded cluster are still accreting, the contribution of accretion luminosity will be significant to the total bolometric luminosity of low-mass H II-clumps with mass $\log_{10}(M_{\rm cl}/M_{\odot}) < 2.8$. Variations in the accretion rate, the SFE and the initial mass function and more physical processes like the external heating from H II regions and the flaring from pre-main sequence stars and protostars need to be investigated to further explain the excess luminosity of low-mass H II-clumps.

Unified Astronomy Thesaurus concepts: Compact H II region (286); Dust continuum emission (412); Molecular gas (1073); Protostars (1302); Protoclusters (1297); Initial mass function (796)

1. Introduction

Until a decade ago, the fields of stellar dynamics and chemical abundances on the one hand and star cluster formation on the other were operating almost independently with the stellar kinematics and abundances community focusing on the long-term dynamical evolution of stellar clusters (Aarseth 1974; Heggie et al. 2006), while the star formation community has directed its attention toward the dynamics of interstellar gas and the impacts of star formation and stellar feedback (Bate 1998; Dib et al. 2008; Smith et al. 2009; Dale 2017; Zhou et al. 2024). As highlighted in Krumholz et al. (2019), maintaining this division is no longer tenable. Advances in observational capabilities and large surveys of the populations of young and less young clusters in the Milky Way now allow us to trace the origins of field stars back to their birthplaces, or at least reconstruct some of the now-dissolved structures in which they originated (Wright et al. 2024; Cantat-Gaudin et al. 2024). To guide such reconstructions, input from both theoretical and observational perspectives within the star formation community is essential.

However, our understanding of massive stars and star clusters formation remains incomplete in the present day. Most and perhaps all the observed stars were formed in embedded clusters (Kroupa 1995a, 1995b, 2005; Lada & Lada 2003; Megeath et al. 2016; Dinnbier et al. 2022), introducing considerable complexity in the identification and analysis of individual young stellar objects (YSOs). Unlike their lower-mass counterparts, massive stars progress to the main sequence while still deeply embedded in their natal clumps, rendering all initial stages invisible even when observed at mid-infrared wavelengths (not all). Adding to the challenge, massive stars are both rare and undergo rapid evolution. Typically, the sites where they form are situated at much greater distances compared to their low-mass counterparts, further complicating observational efforts (Urquhart et al. 2022; Wells et al. 2022; Reyes-Reyes et al. 2024).

Fortunately, our understanding of massive star and cluster formation has seen significant advancements, thanks to the

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

combination of Galactic plane surveys and high-angular resolution images obtained through submillimeter facilities (Motte et al. 2018). Notably, surveys like HiGAL (Molinari et al. 2010) and ATLASGAL (Schuller et al. 2009) have played a crucial role by providing unbiased samples of dense clumps, identified by their dust emission. Dense clumps serve as the initial stages of massive star and cluster formation. Investigations utilizing extensive and statistically representative clump samples can offer valuable perspectives on the effectiveness of transforming molecular gas into stellar clusters. This approach enables us to evaluate how the physical environment influences this intricate process, providing a deeper understanding of the kinematics and dynamics involved (Urquhart et al. 2022). In Urquhart et al. (2022), 5007 ATLASGAL clumps have been classified into four evolutionary stages, i.e., quiescent, protostellar, YSOs and H II regions. In this work, we focus on the clumps at the most evolved stage, called H II-clumps. H II-clumps represent the final stage of embedded cluster formation in the clumps, which are critical to build the bridge between the stellar kinematics and abundances community and the star formation community, because protoclusters or embedded clusters in H II-clumps are the first stage of gas-free star clusters. Currently, it is difficult to construct complete star catalogs for early-stage and distant star clusters from observations (Feigelson et al. 2013; Romine et al. 2016; Kuhn et al. 2017). If there is a continuous transition from protoclusters to embedded clusters, and assuming the stellar initial mass function (IMF) is a universal optimal distribution function rather than a probability density (Kroupa et al. 2013; Yan et al. 2023),[5] we can directly infer the complete census of stars in H II-clumps. In this work we will investigate the physical properties of embedded clusters in H II-clumps and estimate the final star formation efficiency of the clumps.

2. Samples

The physical parameters of 1246 H II-clumps have been calculated and listed in Table 4 of Urquhart et al. (2022). Clump sizes are determined based on the pixel count within the Full Width at Half Maximum (FWHM) contour, i.e., above 50% of the peak of the ATLASGAL dust continuum emission. The clump's mass, denoted as $M_{\rm cl}$ and measured at FWHM, is determined by integrating the flux density at 870 μm within the 50% contour, as outlined in Urquhart et al. (2022). The bolometric luminosity ($L_{\rm bol,obs}$) is derived from the spectral energy distributions of the clumps, which are constructed using continuum data spanning from mid-infrared to sub-millimeter wavelengths (Konig et al. 2017; Urquhart et al. 2022). There are 1226 H II-clumps with mass $M_{\rm cl}$ and bolometric luminosity $L_{\rm bol,obs}$ estimates, Figure 1 shows their $L_{\rm bol,obs}$–$M_{\rm cl}$ relation.

[5] See however Marks et al. (2012), Dib (2023) for how the slope of the IMF in the intermediate-to high-mass end can vary with the environment.

3. Results

3.1. Synthetic Embedded Clusters

Adopting the optimal sampling algorithm, we generate a stellar population for an embedded cluster with a mass denoted as $M_{\rm ecl}$ using the publicly accessible GalIMF code. A detailed description of the optimal sampling and of the GalIMF code that optimally populates clusters and galaxies with stars can be found in Schulz et al. (2015) and Yan et al. (2017). There are 3 input parameters to the code: the stellar mass $M_{\rm ecl}$ of the embedded cluster, the initial metallicity and the age of the embedded cluster. The metallicity we adopt here is the solar metallicity ($[{\rm Fe/H}] = 0$). For the age, Urquhart et al. (2022) calculated the freefall times of all clumps, which provides a lower limit to the star formation timescale. The freefall timescales of the clumps are a few times 10^5 yr, as shown in Figure 2(a). Considering the embedded cluster in a H II-clump should be a mixture of pre-main-sequence (PMS) stars and protostars with spanning ages. The age in the code was taken to be 1 Myr.

The H II-clump mass $M_{\rm cl}$ and its embedded star cluster mass $M_{\rm ecl}$ satisfy $M_{\rm ecl} = {\rm SFE} \ ^* \ M_{\rm cl}$, where SFE is the final star formation efficiency of the clump. In this work, we take ${\rm SFE} = 0.33$. This value has been widely used in previous simulations and has been proven to be effective in reproducing the observational properties of stellar clusters (Banerjee & Kroupa 2013, 2014, 2015, 2017; Kroupa et al. 2001; Banerjee et al. 2012; Oh et al. 2015; Oh & Kroupa 2016, 2018; Brinkmann et al. 2017; Pavlik et al. 2019; Wang et al. 2019, 2020; Dinnbier et al. 2022). Such a SFE is also consistent with the value obtained from hydrodynamic calculations including self-regulation (Machida & Matsumoto 2012; Bate et al. 2014) and as well with observations of embedded systems in the solar neighborhood (Lada & Lada 2003; Megeath et al. 2016). In Figure 1, the mass range is $\log_{10}(M_{\rm cl}/M_\odot) \in [1, 5]$, and we take the values uniformly at a step size of 0.1 in this range. Then according to the corresponding $M_{\rm ecl}$, a population of stars is synthesized using the GalIMF code, which will be used to fit the data points in Figure 1.

3.2. Isochrone Model

In Yan et al. (2023), assuming that the PMS stars have a bolometric luminosity equal to the zero age main sequence stars, the bolometric luminosity of each optimally sampled star in a embedded star cluster was calculated by the empirical massluminosity relations summarized in Equation (12) of Yan et al. (2023). Then, the luminosity of all stars were added together to get the luminosity of the embedded cluster. Yan et al. (2023) found that the optimal sampling model qualitatively agrees with the observation in Urquhart et al. (2022). However, the observational data systematically become more luminous than the optimal sampling model for low-mass

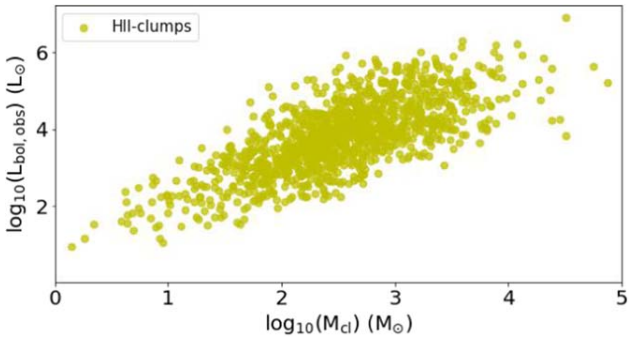


Figure 1. The $L_{\rm bol,obs}$–$M_{\rm cl}$ relation of H II-clumps.

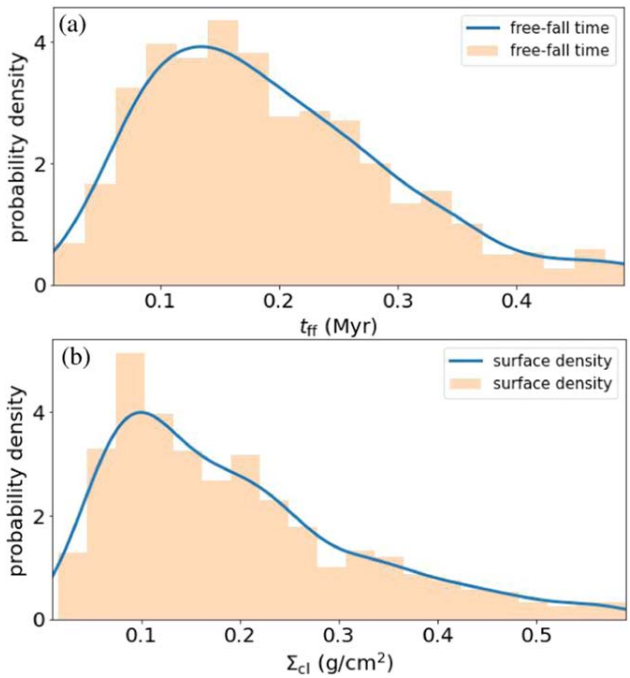


Figure 2. Physical properties of H II-clumps. (a) Distribution of the freefall time; (b) Distribution of the surface density.

H II-clumps, as shown in Figure 3(a). Low-mass embedded clusters in low-mass H II-clumps are predominantly populated by low-mass stars. As these low-mass stars become observable within H II-clumps, they are still in the process of contracting onto the main sequence with significantly larger luminosities. Consequently, it is essential to incorporate their PMS evolutionary tracks into the optimal sampling model. To achieve this goal, we use PMS isochrones to estimate the bolometric luminosity for each star rather than using the empirical massluminosity relation. For a synthetic embedded cluster, the mass of each star is known. Using the isochrones

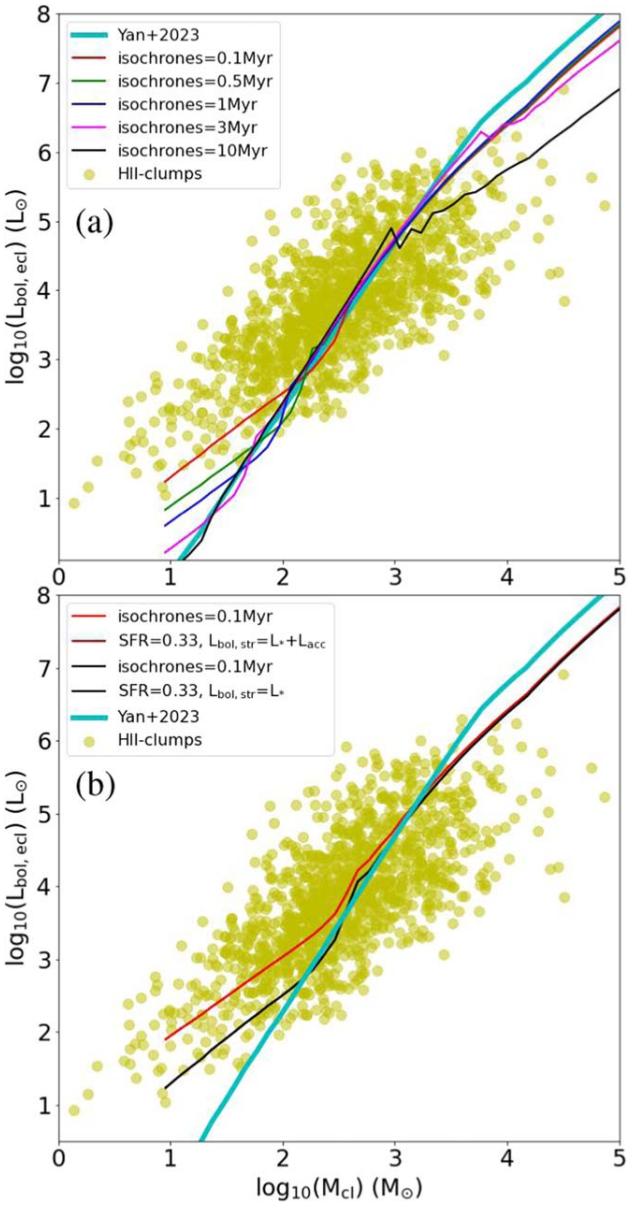


Figure 3. Fitting the observed H II-clumps in Urquhart et al. (2022) using the optimal sampling model. (a) Assuming $L_{\rm bol,str} = L_*$, L_* were estimated by different isochrones; (b) Assuming $L_{\rm bol,str} = L_* + L_{\rm acc}$, L_* were estimated by the 0.1 Myr isochrone.

that are displayed in Figure 4, the bolometric luminosity of a star can be estimated from its mass.

We use the isochrones from the Modules for Experiments in Stellar Astrophysics (MESA) Isochrones and Stellar Tracks (MIST) project. The input physics, general overview of the models, and comparisons with observations are described extensively in Choi et al. (2016). The MIST stellar evolutionary tracks are computed with the MESA code (Paxton et al.

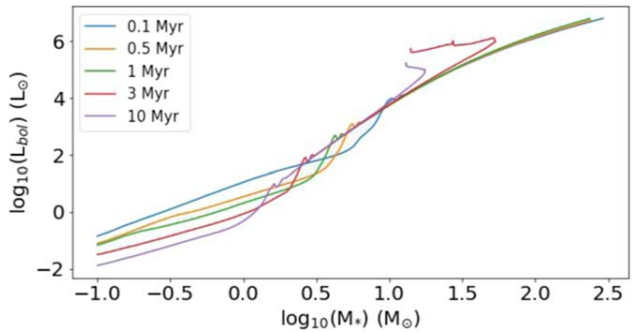


Figure 4. The isochrones from the MESA Isochrones and Stellar Tracks (MIST) project.

2011, 2013, 2015). The isochrone construction procedure is explained in detail in Dotter (2016). The default set of ages in the isochrones ranges are from $\log_{10}(\mathrm{Age/yr}) = 5.0$ to $\log_{10}(\mathrm{Age/yr}) = 10.3$ with a step of 0.05. Figure 4 displays the isochrones of 0.1, 0.5, 1, 3 and 10 Myr. As shown in Figure 3(a), using the 0.1 Myr isochrone appears to best fit the observed data points. Although it is unlikely that all the stars in a H II-clump are that young.

3.3. Accretion Luminosity

In Figure 3(a), after using the isochrones, the fitting to the data is better than that in Yan et al. (2023) who only considered the zero age main sequence luminosities of all stars in the embedded cluster. However, the models are still systematically lower than the observations for low-mass H II-clumps, even using the 0.1 Myr isochrone. For H II-clumps, feedback from the H II region may cause part of the clump mass to be dispersed. Therefore, the observed data points will shift to the left, thus deviating from the models. In order to check this possibility, we compared the mass–size relations of the ATLASGAL clumps at four different evolutionary stages in Figure 5. The clumps in the three earlier evolutionary stages should not experience significant mass dispersal, maintaining a nearly identical mass–size relation. Being more evolved, H II clumps would exhibit a notable deviation from the initial mass–size relation if they had undergone significant mass dispersal. However, as shown in Figure 5, they only show a minor deviation, suggesting that they have not experienced significant mass dispersal. Actually, as discussed in Urquhart et al. (2022), this minor deviation may be attributed to observational biases.

Given that H II-clumps have ongoing star formation, the accretion luminosity $L_{\rm acc}$ from protostars may contribute to the total bolometric luminosity, and it is not explicitly included in the PMS evolutionary tracks. For each accreting protostar in a clump, its total bolometric luminosity $L_{\rm bol,str}$ is the sum of the accretion luminosity $L_{\rm acc}$ and the luminosity of the protostar

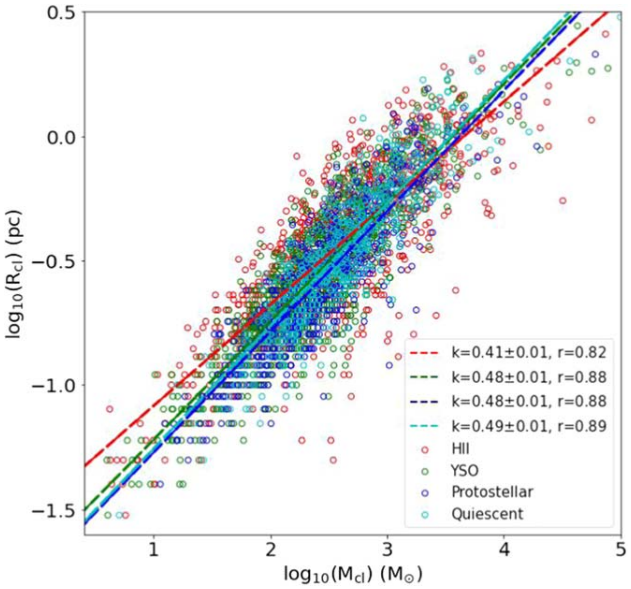


Figure 5. The mass–size relations of the ATLASGAL clumps at four different evolutionary stages. "k" is the slope of the linear fitting. "r" represents the Pearson correlation coefficient.

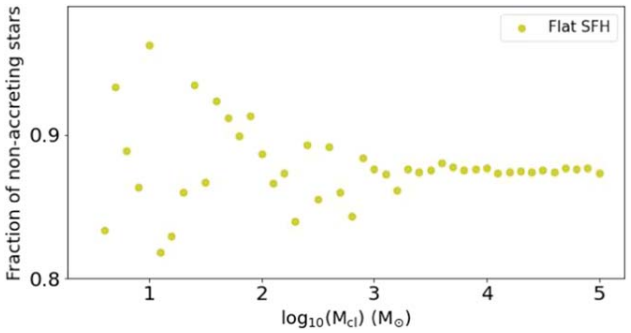


Figure 6. Fraction of non-accreting stars in the case of a flat SFH.

L_*. As shown in Figure 4 of Hosokawa & Omukai (2009), when $M_* < 10\,M_\odot$, $L_{\rm acc}$ is much larger than L_*.

L_* can be calculated using the isochrones. For the accretion, Davies et al. (2011) compared different accretion modes, and found that the prescription of accretion in McKee & Tan (2003) excellently fits the observations. Thus, we compute $L_{\rm acc}$ using Equations (41) and (73) in McKee & Tan (2003)

$$L_{\rm acc} = 4685\left(\frac{f_{\rm acc}}{0.5}\right)\left(\frac{M_*}{30M_\odot}\right)\left(\frac{\dot{M}_*}{10^{-4}M_\odot\ \mathrm{yr}^{-1}}\right)\left(\frac{10R_\odot}{r_*}\right)L_\odot, \quad (1)$$

$$\dot{M}_* = 4.6\times10^{-4}\left(\frac{M_{*\rm f}}{30M_\odot}\right)^{3/4}\Sigma_{\rm cl}^{3/4}\left(\frac{M_*}{M_{*\rm f}}\right)^{0.5} M_\odot\ \mathrm{yr}^{-1}, \quad (2)$$

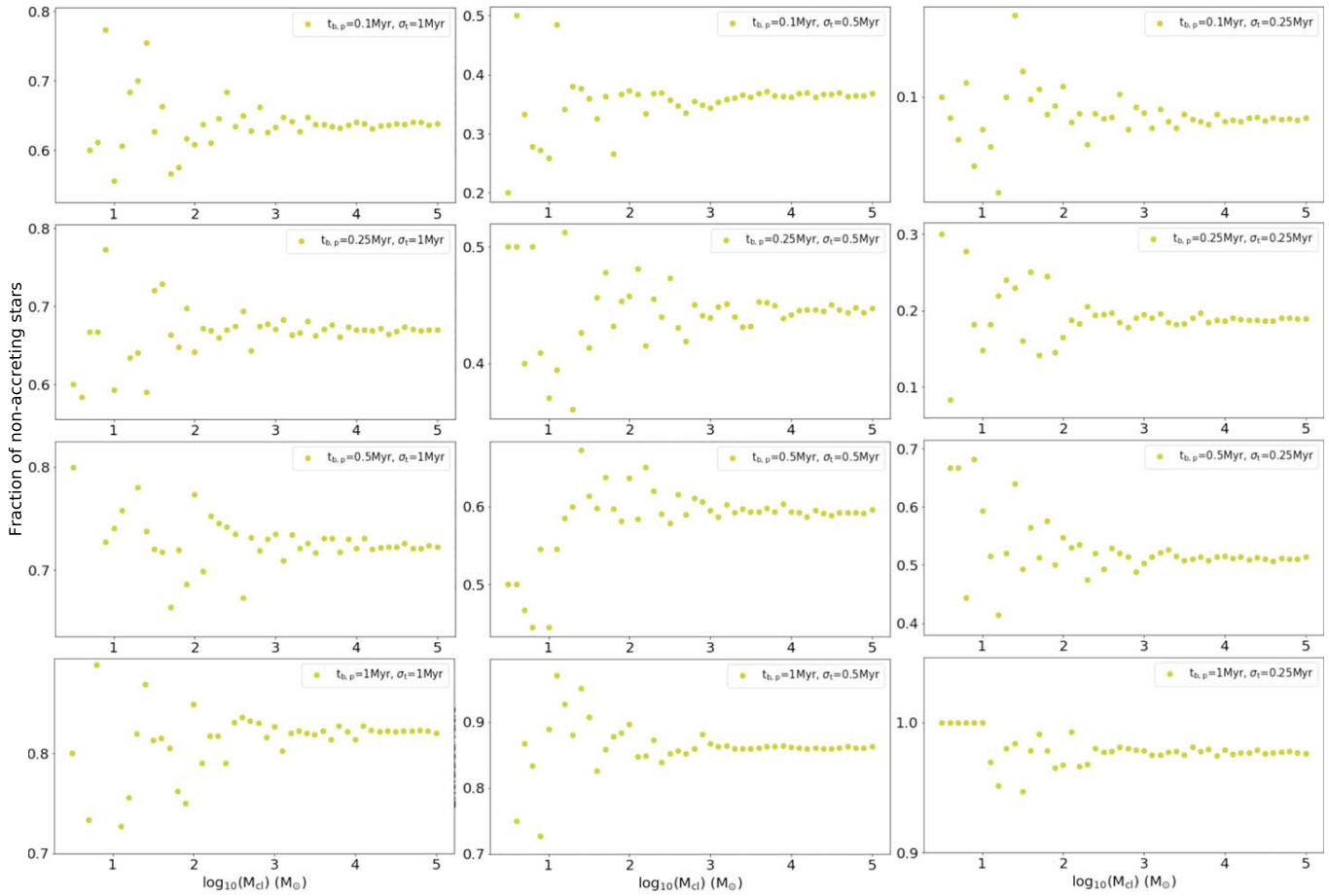


Figure 7. Fraction of non-accreting stars in the case of a Gaussian SFH. From left to right, they are $\sigma_t = 1$, 0.5 and 0.25 Myr, respectively; From top to bottom, they are $t_{b,p} = 0.1$, 0.25, 0.5 and 1 Myr, respectively.

where f_{acc} is the fraction of the protostar's gravitational energy radiated by the star+disk, the fiducial value is 0.5. M_{*f} is the final mass of the star. M_* is at any time the mass of a star, r_* is its corresponding radius. Σ_{cl} is the surface density of the clump. The values of M_* and r_* can be estimated from the isochrones. From Figure 2(b), we take the surface density $\Sigma_{cl} \approx 0.15\ \mathrm{g\ cm^{-2}}$ as the fiducial value in our calculation. We also make a simplified assumption, i.e., $M_{*f} \approx M_*$. As shown in Figure 3(b), after adding the accretion luminosity for each star, the model can fit the observed data.

3.4. Age Distribution

In the above analysis, we only considered the 0.1 Myr isochrone. Actually, many stars in H II-clumps should be older than 0.1 Myr. There is certainly an age spread in the young star and protostar populations in embedded clusters of H II-clumps. Thus, we should choose the appropriate isochrones based on the ages of the stars. Moreover, not all stars in H II-clumps are in accretion, which also depends on their ages. In order to address these issues, we use the method presented in Dib et al. (2024) to create an age distribution of the stellar population in each synthetic cluster to properly distribute the isochrone to each star. Then, we sorted out the fraction of stellar objects that are still protostars to properly add their accretion luminosities.

We examine scenarios where the star formation history (SFH) remains constant and others where it varies over time. In the cases of a constant SFH, we randomly sample the age of stars, t_b, in the 0−4 Myrs time range. For a time-varying SFH, we employ a Gaussian function, allowing us to adjust both the peak position and the distribution width in the 0−4 Myrs time range. For this choice of the SFH, the star age, t_b, is randomly drawn from the Gaussian distribution which is given by:

$$P(t_b) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{t_b - t_{b,p}}{\sigma_t}\right)^2\right), \quad (3)$$

where $t_{b,p}$ is the position of the peak and σ_t the standard deviation.

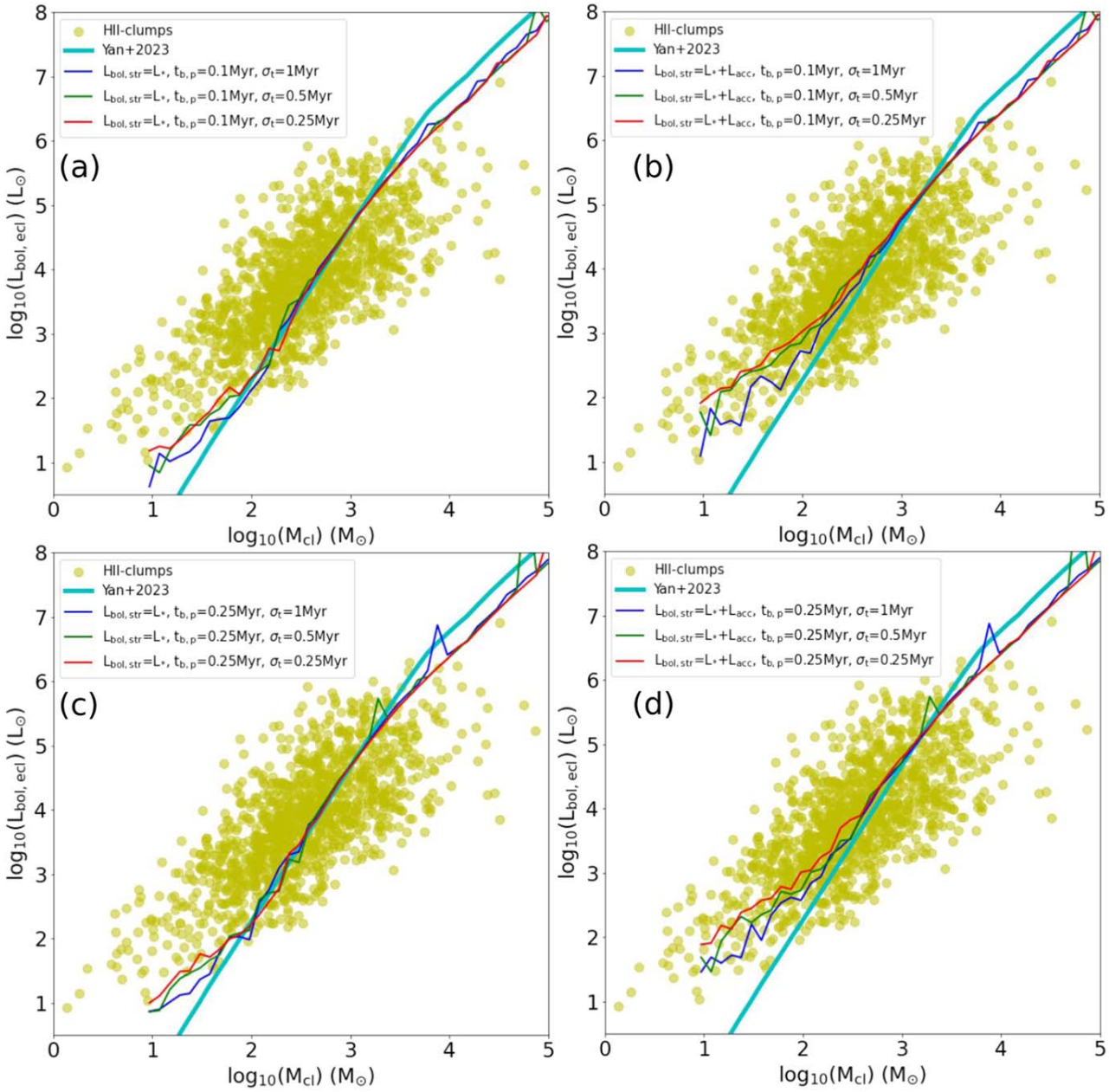


Figure 8. Similar to Figure 3(b), but taking into account the burst-like and time-dependent star formation history (SFH) of the embedded clusters and the age distribution of their stellar populations. (a) $t_{b,p} = 0.1$ Myr and $L_{bol,str} = L_*$; (b) $t_{b,p} = 0.1$ Myr and $L_{bol,str} = L_*+L_{acc}$; (c) $t_{b,p} = 0.25$ Myr and $L_{bol,str} = L_*$; (d) $t_{b,p} = 0.25$ Myr and $L_{bol,str} = L_*+L_{acc}$. $L_{bol,str}$, L_* and L_{acc} are the total bolometric luminosity, the stellar luminosity and the accretion luminosity of the protostar, respectively.

In this scenario of burst-like, time-dependent SFH, the t_b values of stars are randomly chosen based on the probability distribution function $P(t_b)$ as defined in Equation (3). In Figure 3(a), the 0.1 Myr isochrone can best fit the observed data points, indicating that most of the stars should be very young, <1 Myr. Thus, $t_{b,p}$ should be less than 1 Myr. Young clusters continue to undergo star formation for several million years until they either deplete their original molecular gas reservoirs or until gas is expelled by feedback processes, ultimately leading to the cessation of star formation. As revealed in Figure 5, H II-clumps have not significantly dispersed their natal molecular gas. Thus, they should be very young. The ratio of PMS stars with disks to protostars (also called the Class II/Class I ratio) for the cluster cores in

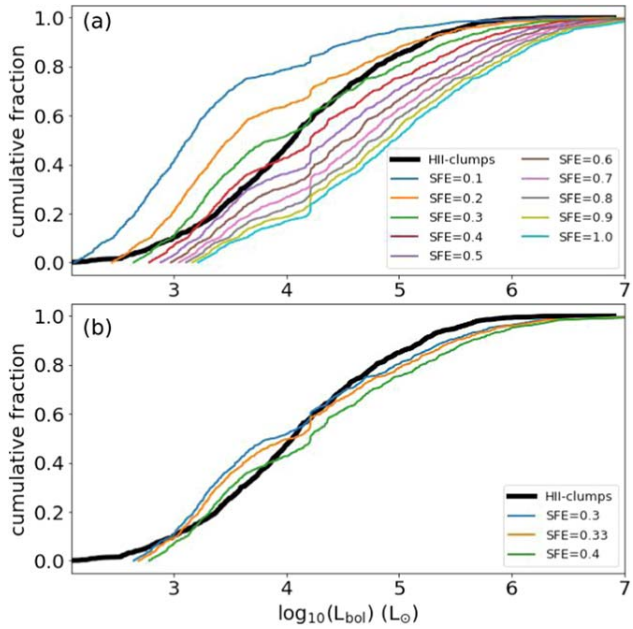


Figure 9. Cumulative distributions of $\log_{10}(L_{\rm bol,obs})$ of H II-clumps (black thick lines) and $\log_{10}(L_{\rm bol,ecl})$ of synthetic embedded clusters under different SFEs.

Gutermuth et al. (2009) requires a period of sustained star formation that lasts 2–3.5 Myr (Megeath et al. 2022). The ages of the Orion Nebula Cluster stars may in actuality span 1–4 Myr (Da Rio et al. 2010; Reggiani et al. 2011; Getman et al. 2014; Beccari et al. 2017; Jerabkova et al. 2019). Therefore, we should consider different values of $\sigma_{\rm t}$ to cover possible time spans of star formation. Finally, we took $t_{\rm b,p} = 0.1$, 0.25, 0.5, 1 Myr and $\sigma_{\rm t} = 0.25$, 0.5 and 1 Myr. There are a total of 12 cases here.

The typical values of the young protostellar stage is 0.5 Myr (≈0.1 Myr for the Class 0 phase and ≈0.44 Myr for the Class I phase) and that of the Class II phases is ≈2 Myr (Evans et al. 2009; Hsieh & Lai 2013; Pokhrel et al. 2020; Megeath et al. 2022; Dib et al. 2024). In this work, only the protostars with ages less than 0.5 Myr are used to calculate the accretion luminosity.

In Figure 6, for the case of a flat SFH, nearly 90% stars are excluded. Thus, the case with a flat SFH can not account for difference in luminosity. In the cases with time-dependent SFHs, as shown in Figures 7 and 8, as long as 20% of the stars within the embedded cluster are still accreting, the contribution of accretion luminosity will be significant for low-mass clumps. Currently, we have assigned corresponding isochrones based on the age of each star, rather than using the 0.1 Myr isochrone for all stars in Figure 3(b). As shown in Figure 8, the fitting lines remain comparable to Figure 3(b). For low-mass H II-clumps, the contribution of the stellar luminosity to the total luminosity is minimal. For high-mass H II-clumps, isochrones of different ages yield consistent results. Therefore, in the above analysis, the choice of isochrones is not important.

Table 1
The Median Values of $\log_{10}(L_{\rm bol,ecl})$ for the Synthetic Embedded Clusters Under Different SFEs

SFE	p	median
0.32	0.00003	4.004
0.33	0.00009	4.056
0.34	0.00038	4.103
0.35	0.00164	4.145
0.36	0.00195	4.178
0.37	0.00055	4.202
0.38	0.00038	4.212
0.39	0.00007	4.213
0.4	0.00003	4.215
0.41	0.00002	4.218

Note. "p" is the p-value of the Kolmogorov–Smirnov (KS) test between $\log_{10}(L_{\rm bol,obs})$ and $\log_{10}(L_{\rm bol,ecl})$.

3.5. The Final Star Formation Efficiency of the Clump

In the above discussion, we adopted a constant SFE = 0.33 for all H II-clumps following Kroupa et al. (2001), Megeath et al. (2016), Yan et al. (2017, 2023), Banerjee & Kroupa (2018), although the masses of all H II-clumps can span more than 5 orders of magnitude. In Section 3.3, by adding the accretion luminosity, the model is able to fit the observations. However, one could argue that the deviation between the model without the accretion luminosity and the observation in Figure 3(a) is due to assuming a constant SFE = 0.33. As described in Dib et al. (2013), Wells et al. (2022), the SFE may change as a function of clump mass. If the SFE decreases with increasing mass of the clump, low-mass clumps with high SFE can also decrease the deviation between the model and the observed data in Figure 3(a). Thus, there is a critical question, for all H II-clumps with significantly different masses, whether they have similar final star formation efficiency? The good fit in Figure 3(b) supports the case of a SFE ≈ 0.33, but the scatter at any fixed cluster mass could be explained by a certain degree of variation of the SFE.

In order to evaluate the effect of the SFE for H II-clumps, we consider values of the SFE that fall in the range [0.1, 1] with a step of 0.01. Therefore, each H II-clump has 91 alternative SFEs, which correspond to 91 synthetic embedded clusters using the method in Section 3.1. In Yan et al. (2017), the assumed lower limit of embedded cluster mass is $M_{\rm ecl,min} = 5\,M_{\odot}$. When SFE = 0.1, 924 H II-clumps in the sample in Section 2 can meet this requirement, i.e., $M_{\rm cl}{}^{*}0.1 > 5M_{\odot}$. The median value of the observed bolometric luminosity $\log_{10}(L_{\rm bol,obs})$ for 924 H II-clumps is 4.057. Table 1 displays the median values of the total bolometric luminosity

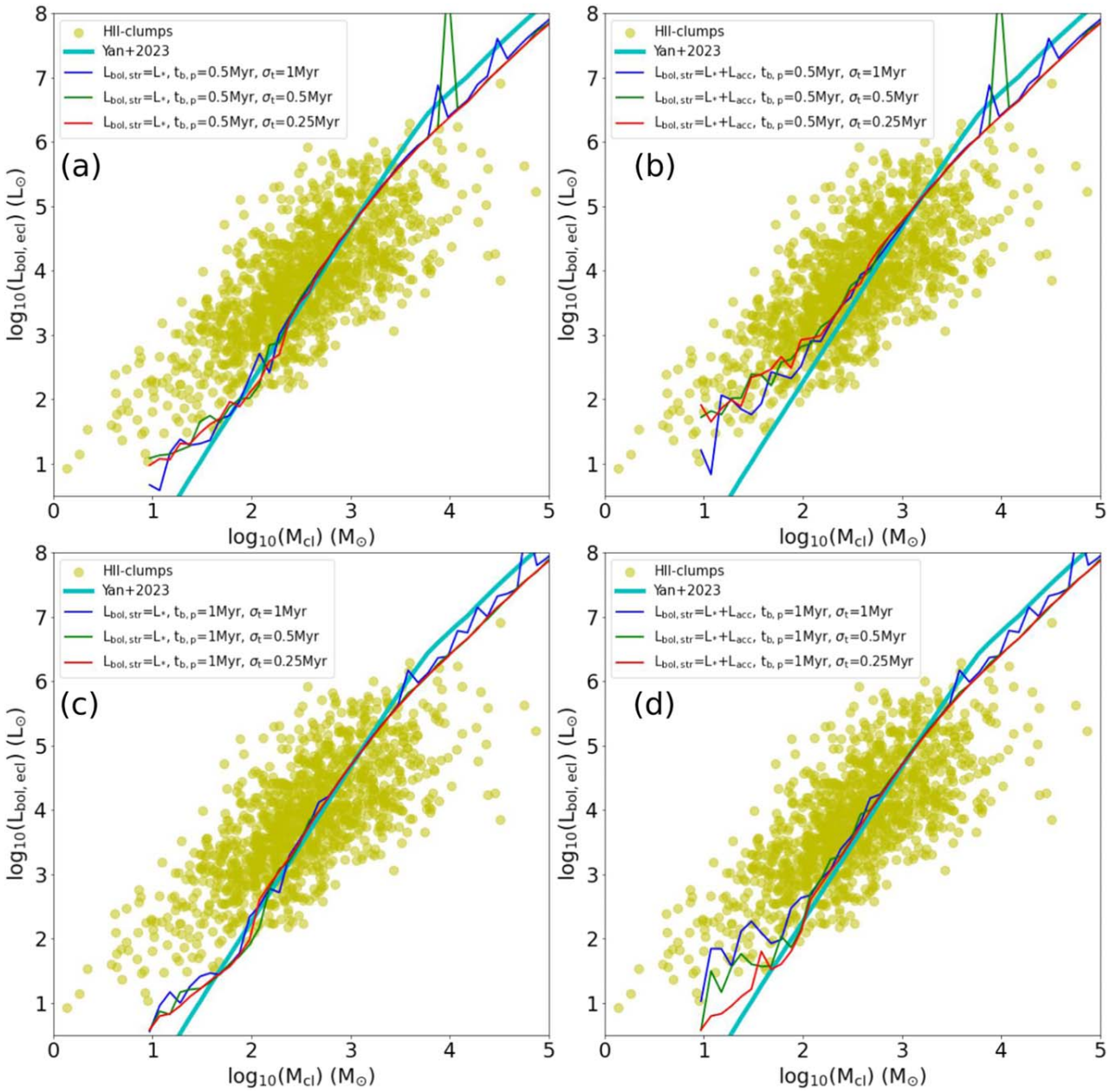


Figure 10. Same as Figure 8, but for $t_{\rm b,p} = 0.5$ Myr and $t_{\rm b,p} = 1$ Myr.

$\log_{10}(L_{\rm bol,ecl})$ for the synthetic embedded clusters under different SFEs. When SFE = 0.36, the p-value of the Kolmogorov–Smirnov (KS) test between $\log_{10}(L_{\rm bol,obs})$ and $\log_{10}(L_{\rm bol,ecl})$ has the highest value, although this value is smaller than 0.05. The corresponding median value of $\log_{10}(L_{\rm bol,ecl})$ is 4.178, comparable with that of $\log_{10}(L_{\rm bol,obs})$. Interestingly, when SFE = 0.33, the median value of $\log_{10}(L_{\rm bol,ecl})$ is almost the same with that of $\log_{10}(L_{\rm bol,obs})$. In Figures 9(a) and (b), when SFE $\approx$ 0.33, the cumulative distributions of $\log_{10}(L_{\rm bol,obs})$ and $\log_{10}(L_{\rm bol,ecl})$ are also comparable. In summary, SFE = 0.33 can be a typical value for all H II-clumps.

However, the analysis presented here follows the procedure in Section 3.3. We do not consider the SFH and still assume all stars are in accretion. Thus, we should be cautious about the results. Low-mass H II-clumps with high SFE may lead to the excess luminosity shown in Figure 3. The current analysis can not exclude this possibility.

4. Discussion and Conclusions

In Figure 4 of Hosokawa & Omukai (2009), when $M_* > 10\,M_\odot$, L_* begins to exceed $L_{\rm acc}$. Thus, for high-mass stars, the contribution from L_* to $L_{\rm bol,str}$ is more important than $L_{\rm acc}$. For the synthetic embedded clusters, when $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.2$, stars with masses larger than $10\,M_\odot$ begin to appear. In Figure 3(b), we can see the black line indeed begins to approach the red line around $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.2$. But now, the number of low-mass stars is still much larger than high-mass stars in the synthetic embedded cluster, the total accretion luminosity of low-mass stars remains dominant. When $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.8$, the luminosity of high-mass stars becomes dominant and $L_{\rm acc}$ can be ignored. Therefore, there is a mass limit of $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.8$, when the H II-clump mass is less than this limit, one must consider the contribution of the accretion luminosity from protostars to the observed bolometric luminosity $L_{\rm bol,obs}$ of the H II-clump.

In the above analysis, we employed classical PMS contraction tracks, but it needs to be kept in mind that the true path an accreting PMS star takes in the Hertzsprung–Russell diagram (HRD) is likely to differ substantially as its internal structure, mass and current radius will reflect the accretion history (Wuchterl & Tscharnuter 2003). The accretion rates onto protostars are time-variable (McKee & Tan 2003; Wuchterl & Tscharnuter 2003; Schmeja & Klessen 2004; Dib et al. 2010; Fischer et al. 2017) and this may lead to a dispersion of the data points in Figure 3. Moreover, although our results seem to indicate that a constant SFE = 0.33 is quite plausible, variations in the SFE at any given clump mass, along with potential variations of the IMF, could also help explain the scatter in the $L_{\rm bol,obs}$–$M_{\rm cl}$ relation. Moreover, Low-mass H II-clumps with high SFE may also explain the excess luminosity. Compared to the very limited analysis in Section 3.5, a comprehensive discussion on the impact of SFE must take into account the SFH. However, the SFH of H II-clumps is quite uncertain, which will make the estimation of SFE quite uncertain as well. Furthermore, SFE is just one of many factors influencing the total luminosity of H II-clumps. We cannot simply explain the observational data by increasing SFE alone. Various factors must be considered together, which is certainly complex.

In this work, we demonstrate that the accretion process has important contribution to the total luminosity of low-mass H II-clumps. However, many H II-clumps exhibit significantly higher luminosity than the fitting lines in Figures 8 and 10, which indicates that the accretion should not be the only important factor. Other physical processes, such as the external heating from H II regions and the flaring from PMS stars and protostars (Flaccomio et al. 2018; Zhang et al. 2020; Getman & Feigelson 2021), may also contribute to the total luminosity.

Acknowledgments

We would like to thank the referee for sharing detailed comments and suggestions that helped improve and clarify several aspects of this work.

ORCID iDs

J. W. Zhou https://orcid.org/0009-0004-2927-239X
Sami Dib https://orcid.org/0000-0002-8697-9808

References

Aarseth, S. J. 1974, A&A, 35, 237
Banerjee, S., & Kroupa, P. 2013, ApJ, 764, 29
Banerjee, S., & Kroupa, P. 2014, ApJ, 787, 158
Banerjee, S., & Kroupa, P. 2015, MNRAS, 447, 728
Banerjee, S., & Kroupa, P. 2017, A&A, 597, A28
Banerjee, S., & Kroupa, P. 2018, in The Birth of Star Clusters, ed. S. Stahler (Cham: Springer), 143
Banerjee, S., Kroupa, P., & Oh, S. 2012, ApJ, 746, 15
Bate, M. R. 1998, ApJL, 508, L95
Bate, M. R., Tricco, T. S., & Price, D. J. 2014, MNRAS, 437, 77
Beccari, G., Petr-Gotzens, M. G., Boffin, H. M. J., et al. 2017, A&A, 604, A22
Brinkmann, N., Banerjee, S., Motwani, B., & Kroupa, P. 2017, A&A, 600, A49
Cantat-Gaudin, T., Fouesneau, M., Rix, H.-W., et al. 2024, A&A, 683, A128
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Da Rio, N., Robberto, M., Soderblom, D. R., et al. 2010, ApJ, 722, 1092
Dale, J. E. 2017, MNRAS, 467, 1067
Davies, B., Hoare, M. G., Lumsden, S. L., et al. 2011, MNRAS, 416, 972
Dib, S. 2023, ApJ, 959, 88
Dib, S., Brandenburg, A., Kim, J., Gopinathan, M., & André, P. 2008, ApJL, 678, L105
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Dib, S., Shadmehri, M., Padoan, P., et al. 2010, MNRAS, 405, 401
Dib, S., Zhou, J. W., Comerón, S., et al. 2024, arXiv:2405.00095
Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61
Dotter, A. 2016, ApJS, 222, 8
Evans, N. J. I., Dunham, M. M., Jørgensen, J. K., et al. 2009, ApJS, 181, 321
Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, ApJS, 209, 26
Fischer, W. J., Megeath, S. T., Furlan, E., et al. 2017, ApJ, 840, 69
Flaccomio, E., Micela, G., Sciortino, S., et al. 2018, A&A, 620, A55
Getman, K. V., & Feigelson, E. D. 2021, ApJ, 916, 32
Getman, K. V., Feigelson, E. D., & Kuhn, M. A. 2014, ApJ, 787, 109
Gutermuth, R. A., Megeath, S. T., Myers, P. C., et al. 2009, ApJS, 184, 18
Heggie, D. C., Trenti, M., & Hut, P. 2006, MNRAS, 368, 677
Hosokawa, T., & Omukai, K. 2009, ApJ, 691, 823
Hsieh, T.-H., & Lai, S.-P. 2013, ApJS, 205, 5
Jerabkova, T., Beccari, G., Boffin, H. M. J., et al. 2019, A&A, 627, A57
Konig, C., Urquhart, J. S., Csengeri, T., et al. 2017, A&A, 599, A139
Kroupa, P. 1995a, MNRAS, 277, 1491
Kroupa, P. 1995b, MNRAS, 277, 1507
Kroupa, P. 2005, in Proc. of the Symp. 576, The Three-Dimensional Universe with Gaia, ed. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 629
Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699
Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, Vol. 5 (Dordrecht: Springer), 115
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Kuhn, M. A., Medina, N., Getman, K. V., et al. 2017, AJ, 154, 87
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Machida, M. N., & Matsumoto, T. 2012, MNRAS, 421, 588

Publications of the Astronomical Society of the Pacific, 136:094302 (11pp), 2024 September https://doi.org/10.1088/1538-3873/ad77f4
© 2024. The Author(s). Published by IOP Publishing Ltd on behalf of the Astronomical Society of the Pacific (ASP).

OPEN ACCESS

The Star Formation Histories, Star Formation Efficiencies and Ionizing Sources of ATLASGAL Clumps with H II Regions

J. W. Zhou[1], Sami Dib[2], and Pavel Kroupa[3,4]
[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany; jwzhou@mpifr-bonn.mpg.de
[2] Max Planck Institute für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany; dib@mpia.de
[3] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14-16, 53115 Bonn, Germany; pkroupa@uni-bonn.de
[4] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic
Received 2024 July 22; accepted 2024 September 6; published 2024 September 26

Abstract

1226 ATLASGAL clumps with H II regions (H II-clumps) were matched with radio sources in the CORNISH-North/South surveys, and 392 of them have corresponding radio sources. We determined the stellar luminosity $L_{*,\mathrm{T84}}$ according to the Lyman continuum flux N_{Ly}. When the bolometric luminosity of H II-clumps is less than $\log_{10}(L_{\mathrm{bol,obs}}/L_{\odot})\approx 3.7$, corresponding to a clump mass $\log_{10}(M_{\mathrm{cl}}/M_{\odot})\approx 2.55$, the values of $L_{*,\mathrm{T84}}$ derived from N_{Ly} overestimate the actual stellar luminosities, because the accretion onto the protostars contributes significantly to the radio emission. After subtracting the accretion luminosity from $L_{*,\mathrm{T84}}$, we obtained reasonable estimates of the stellar luminosity. Using the 0.5 Myr isochrone, we calculated the stellar masses according to the stellar luminosities, and found that they roughly follow the $m_{\mathrm{max}}-M_{\mathrm{ecl}}$ relation of embedded clusters, consistent with the ionizing sources representing the most massive stars in the embedded clusters of H II-clumps. We also studied the contribution of the possible flaring activity to the observed stellar luminosity and found that they can be neglected. We further studied the change of SFE with the clump mass. According to the derived mass of the most massive star in each HII-clump, using the theoretical $m_{\mathrm{max}}-M_{\mathrm{ecl}}$ relation, we calculated the mass of the corresponding embedded cluster and then the SFE of the clump. The SFE decreases with increasing clump mass, with a median value of ≈0.3. We also independently derived the SFE for each H II-clump based on the model developed in our previous work. The SFEs of H II-clumps derived from the observation and the model are in good agreement. Concerning the star formation histories of the ATLASGAL clumps, low-mass clumps may reach the peak of star formation earlier than high-mass clumps, consistent with the shorter free-fall time of low-mass clumps.

Unified Astronomy Thesaurus concepts: Accretion (14); Compact H II region (286); Young stellar objects (1834); Radio source catalogs (1356); Star forming regions (1356); Protoclusters (1297)

1. Introduction

Understanding the contribution of stars in different mass ranges to the total energetics and dynamics of star clusters is crucial. A majority, if not all, of the stars we observe were born within embedded clusters (Kroupa 1995a, 1995b; Lada & Lada 2003; Kroupa 2005; Megeath et al. 2016; Dinnbier et al. 2022), which adds significant intricacy to pinpointing and analyzing individual protostellar objects. Unlike their lower-mass counterparts, massive stars evolve towards the main sequence while still deeply embedded within their parental clumps. This renders their initial stages invisible, even when observed using mid-infrared wavelengths. Nowadays, our comprehension of massive star and cluster formation has significantly advanced, owing to the combination of Galactic plane surveys and high-angular resolution images obtained through submillimeter facilities (Motte et al. 2018). Surveys such as HiGAL (Molinari et al. 2010) and ATLASGAL (Schuller et al. 2009) have played pivotal roles by providing unbiased compilations of dense clumps that track the early phases of massive star and cluster formation. Explorations employing representative clump samples can provide valuable insights into the efficiency of converting molecular gas into stellar clusters (Elia et al. 2021; Urquhart et al. 2022; Wells et al. 2022).

In Urquhart et al. (2022), 5007 ATLASGAL clumps have been categorized into four distinct evolutionary stages, with the most advanced stage being referred to as H II-clumps, i.e., ATLASGAL clumps with H II regions. In Zhou et al. (2024a) (paper I), we synthesized the embedded clusters within H II-clumps under the assumption that the stellar initial mass function (IMF) follows a universal optimal distribution function rather than a probability density (Kroupa et al. 2013; Yan et al. 2023). We utilized the 0.1 Myr isochrone to estimate the bolometric luminosity of individual stars within an embedded cluster,

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

augmenting this with the accretion luminosity of each star in the cluster. The cumulative bolometric luminosity of the synthetic embedded clusters aligns closely with the observed bolometric luminosity of H II-clumps, validating the effectiveness of the method. As a follow up study based on paper I, in this work, we focus on the ionizing sources in H II-clumps. They may represent the most massive stars in the embedded clusters. We explore their physical properties by comparing the observations with numerical simulations.

2. Sample

The physical parameters of 1246 H II-clumps have been calculated and listed in Urquhart et al. (2022). 1226 H II-clumps with mass ($M_{\rm cl}$) and bolometric luminosity ($L_{\rm bol,obs}$) measurements were matched with radio sources in the catalogs of the CORNISH-North (Purcell et al. 2013) and the CORNISH-South (Irabor et al. 2023) surveys. The separations between the central coordinates of a radio source and a H II-clump was required to be smaller than the radius of the clump. Generally, one H II-clump matches with one radio source. If one H II-clump includes more than one radio source, we only consider the radio source with the strongest radio emission, because we are only interested in the most massive star in a H II-clump. Finally, the numbers of radio sources matched with HII-clumps are 244 and 148 in the CORNISH-South and the CORNISH-North survey catalogs, respectively.

Figure 1(a) displays the samples of H II clumps in the $L_{\rm bol,obs}$–$M_{\rm cl}$ diagram. We can see that the samples are mainly concentrated at the upper end, similar to Figure 5 in Urquhart et al. (2022), where the H II-clumps were divided into two populations, i.e., radio-loud and radio-quiet. Our samples are radio-loud, which are more luminous than the radio-quiet H II-clumps. This indicates that they are hosting high-mass stars. In Figure 1(b), the radii of radio sources are systematically smaller than the radii of the H II-clumps. These radio sources should be excited by the most massive stars in the embedded clusters of H II-clumps. From the radio emission, we can infer some physical parameters of the massive stars.

3. Results and Discussion

3.1. Lyman Continuum Flux

As discussed in Kalcheva et al. (2018), assuming optically thin radio sources at 5 GHz (CORNISH-North) and 5.5 GHz (CORNISH-South) is justified for our samples. We calculated the physical parameters of the matched radio sources following the formalism of Mezger & Henderson (1967) and Rubin (1968) using the equations collected in Sánchez-Monge et al. (2013). The emission measure (EM) and the electron density

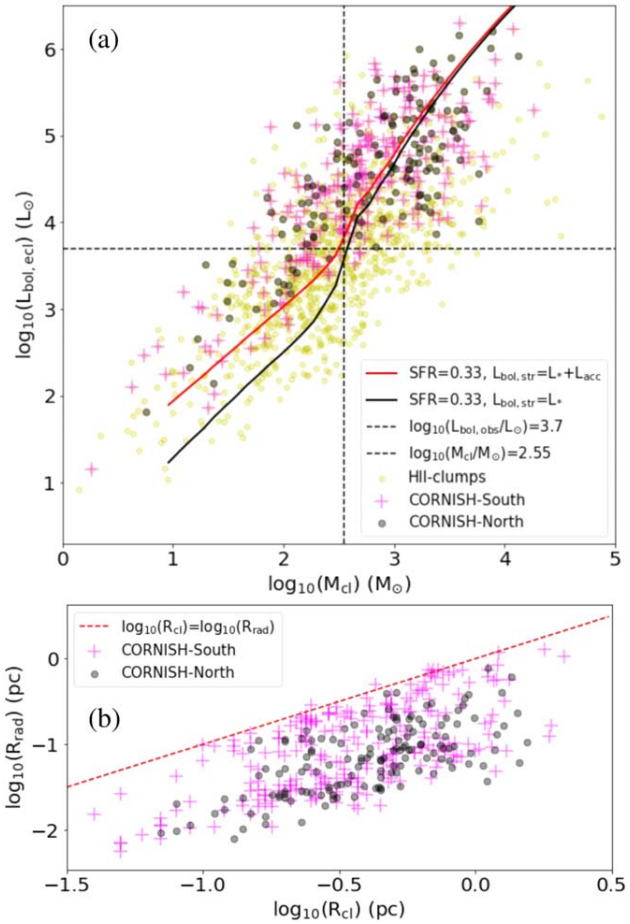


Figure 1. (a) The $L_{\rm bol,obs}$–$0M_{\rm cl}$ relation of H II-clumps. Black dots and magenta pluses mark the H II-clumps matched with radio sources in the CORNISH-North and CORNISH-South surveys. The red and black lines show the fitting to the observed HII-clumps using the optimal sampling model in Zhou et al. (2024a). The red line considers the accretion luminosity of each star in an embedded cluster, the black line does not. The vertical and horizontal lines mark the turning points of the bolometric luminosity and mass of H II-clumps discussed in Section 3.2; (b) The radii of radio sources ($R_{\rm rad}$) vs. the radii of the corresponding H II-clumps ($R_{\rm cl}$).

($n_{\rm e}$) were calculated using:

$$\left[\frac{\rm EM}{\rm cm^{-6}\,pc}\right] = 1.7\times10^{7}\left[\frac{S_\nu}{\rm Jy}\right]\left[\frac{\nu}{\rm GHz}\right]^{0.1}\left[\frac{T_{\rm e}}{\rm K}\right]^{0.35}\left[\frac{\theta_{\rm s}}{''}\right]^{-2}, \quad (1)$$

and

$$\left[\frac{n_{\rm e}}{\rm cm^{-3}}\right] = 2.3\times10^{6}\left[\frac{S_\nu}{\rm Jy}\right]^{0.5}\left[\frac{\nu}{\rm GHz}\right]^{0.05}\left[\frac{T_{\rm e}}{\rm K}\right]^{0.175} \times\left[\frac{d}{\rm pc}\right]^{-0.5}\left[\frac{\theta_{\rm s}}{''}\right]^{-1.5}, \quad (2)$$

Table 1
Explanation of the Physical Quantities

$L_{\rm bol,obs}$	The bolometric luminosity of a H II-clump
$M_{\rm cl}$	The mass of a H II-clump
$M_{\rm ecl}$	The mass in stars of an embedded cluster in a H II-clump
$N_{\rm Ly}$	The Lyman continuum flux
$L_{*,\rm T84}$	The luminosity of a star calculated by interpolating the star's luminosity given as a function of the Lyman continuum flux in Table 1 of Thompson (1984)
$m_{*,\rm iso}$	The mass of a star estimated by the 0.5 Myr isochrone according to $L_{*,\rm T84}$
$m_{\rm max}$	The mass of the most massive star in a synthetic embedded cluster generated by the GalIMF code
$L_{\rm acc}$	The accretion luminosity of the most massive star in a synthetic embedded cluster
L_{*}	The cleaned luminosity of a star after subtracting its accretion luminosity from $L_{*,\rm T84}$
$L_{\rm x}$ and $L_{\rm opt}$	The luminosities of a flare in X-ray and optical bands

where S_ν is the integrated radio flux density, $T_{\rm e}$ is the electron temperature assumed to be 10^4 K, $\theta_{\rm s}$ is the angular diameter of the source, d is the distance, and ν is the observing frequency. S_ν and $\theta_{\rm s}$ are assigned the values in the catalogs of Purcell et al. (2013) and Irabor et al. (2023). The values of d are the distances of the corresponding H II-clumps listed in Urquhart et al. (2022). The number of Lyman-continuum photons per second ($N_{\rm Ly}$; hereafter Lyman continuum flux) is calculated from the flux density and distance as:

$$\left[\frac{N_{\rm Ly}}{\rm photons\ s^{-1}}\right] = 8.9\times10^{40}\left[\frac{S_\nu}{\rm Jy}\right]\left[\frac{\nu}{\rm GHz}\right]^{0.1} \times\left[\frac{T_{\rm e}}{10^4\ \rm K}\right]^{-0.45}\left[\frac{d}{\rm pc}\right]^{2}. \quad (3)$$

Table 1 of Davies et al. (2011) gives the relationship between the stellar parameters and Lyman continuum flux. For stars with masses above 60 $M_\odot$, they made the assumption that Lyman continuum flux per unit area is constant, which is likely to slightly underestimate the Lyman flux of very massive stars. In Figure 2(a), we compared the data in Table 1 of Davies et al. (2011) with Table 1 of Thompson (1984). It seems that Table 1 of Thompson (1984) preforms better at the end with high $N_{\rm Ly}$ and stellar luminosity, L_*. The relationship between stellar mass and luminosity in Davies et al. (2011) was taken from Meynet & Maeder (2000), which is consistent with the isochrones of 0.5 Myr or 1 Myr from the MESA Isochrones and Stellar Tracks (MIST) project (Choi et al. 2016), as shown in Figure 2(b).

In this work, we determined the stellar luminosity $L_{*,\rm T84}$ corresponding to the calculated $N_{\rm Ly}$ by interpolating the stellar luminosity given as a function of the Lyman continuum flux in Table 1 of Thompson (1984), shown in Figure 2(a).

3.2. The Turning Points

Table 1 summarizes some key physical quantities in this work and their meanings. The stellar luminosity $L_{*,\rm T84}$ of the brightest star should be smaller than the bolometric luminosity $L_{\rm bol,obs}$ of the corresponding H II-clump. However, in Figure 3(a), for low-luminosity H II-clumps, $L_{*,\rm T84}$ is significantly larger than $L_{\rm bol,obs}$, which is abnormal. In Equation (3), $N_{\rm Ly}\propto S_\nu$, and in Figure 2, $N_{\rm Ly}\propto L_*$. If the bolometric luminosity $L_{\rm bol,obs}$ of an H II-clump is mainly from the stars, $L_{\rm bol,obs}\propto L_*$. Finally, $L_{\rm bol,obs}\propto S_\nu$. However, in Figure 3(b), there is a turning point around $\log_{10}(L_{\rm bol,obs}/L_\odot)\approx 3.7$. For H II-clumps with luminosity lower than this value, there is no clear correlation between $L_{\rm bol,obs}$ and radio flux density S_ν. The turning point $\log_{10}(L_{\rm bol,obs}/L_\odot)\approx 3.7$ corresponds to the HII-clump mass $\log_{10}(M_{\rm cl}/M_\odot)\approx 2.55$ in Figure 1(a).

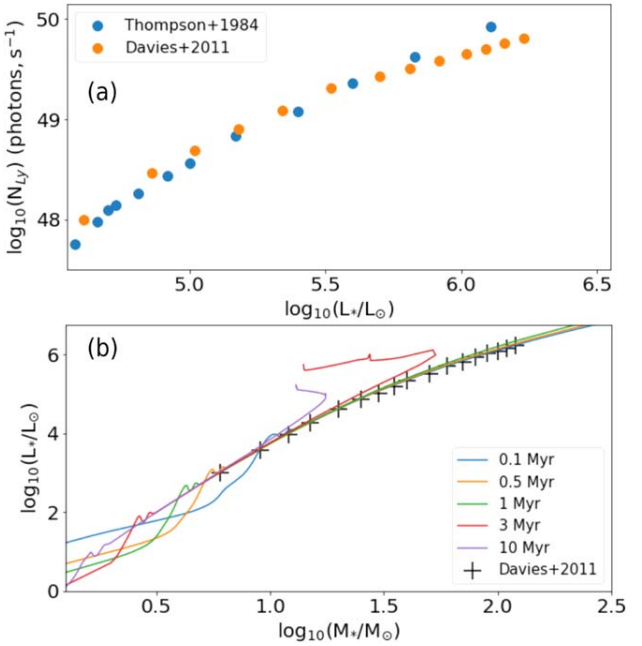


Figure 2. (a) The relationships between stellar luminosity L_* of individual stars and Lyman continuum flux $N_{\rm Ly}$ in Table 1 of Davies et al. (2011) (orange dots) and Table 1 of Thompson (1984) (blue dots), respectively; (b) The relationship between stellar mass and luminosity in Table 1 of Davies et al. (2011). The color-coded lines are the MESA isochrones, see text for details.

Using the 0.5 Myr isochrone, we estimated the stellar mass $m_{*,\rm iso}$ according to the stellar luminosity $L_{*,\rm T84}$. The ionizing sources represent the most massive stars in H II-clumps, i.e.,

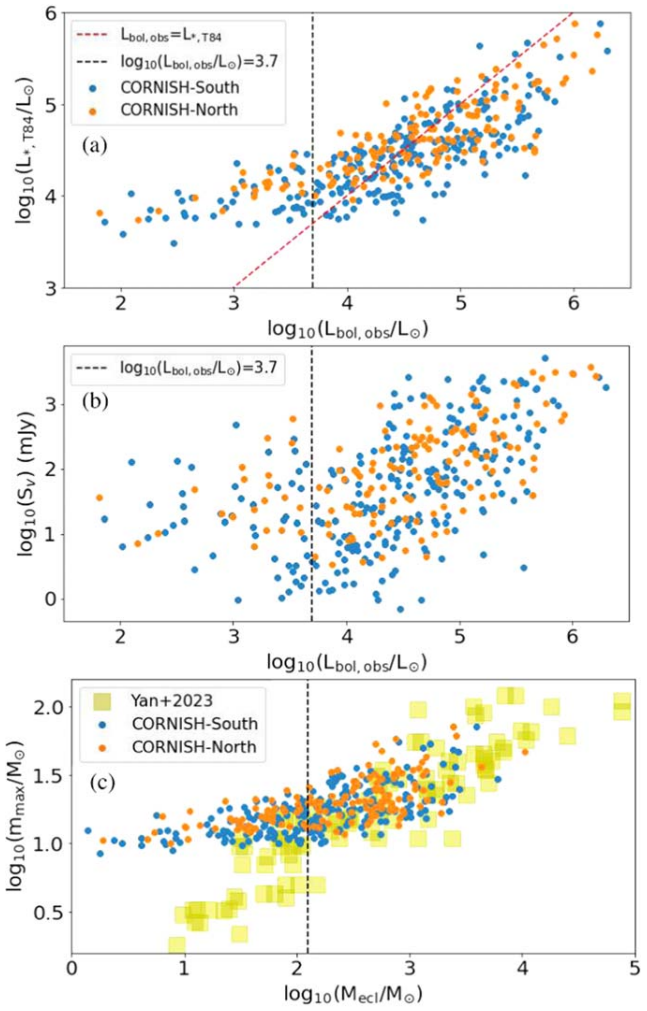


Figure 3. (a) The correlation between $L_{*,\mathrm{T84}}$ and $L_{\mathrm{bol,obs}}$; (b) The correlation between S_ν and $L_{\mathrm{bol,obs}}$; (c) The $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation. Yellow squares represent the star clusters from the table of Yan et al. (2023). These vertical dashed lines mark the turning points, see text for the details.

$m_{*,\mathrm{iso}} \approx m_{\mathrm{max}}$. Considering the $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation (Weidner & Kroupa 2006; Weidner et al. 2013; Yan et al. 2023), a correlation between the embedded star cluster mass M_{ecl} and the mass of the most massive star m_{max}, there should be a correlation between $m_{*,\mathrm{iso}}$ and the mass of the corresponding embedded cluster. The clump mass M_{cl} and its embedded cluster mass M_{ecl} satisfy $M_{\mathrm{ecl}} = \mathrm{SFE} \times M_{\mathrm{cl}}$, where SFE is the final star formation efficiency of the clump, and which we assumed to be SFE = 0.33 (Lada & Lada 2003; Machida & Matsumoto 2012; Bate et al. 2014; Megeath et al. 2016; Banerjee & Kroupa 2018; Zhou et al. 2024a, 2024b), see Section 3.6 for more discussion. For 100 embedded clusters with age < 5 Myr, the table in Yan et al. (2023) gives the values of their m_{max} and M_{ecl}. This sample fits the $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation well. In Figure 3(c), we compared our sample with the

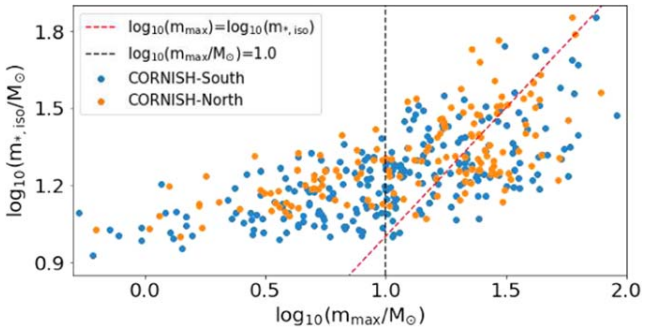


Figure 4. The correlation between the most massive star with mass m_{max} in each synthetic embedded cluster and the corresponding $m_{*,\mathrm{iso}}$.

sample from Yan et al. (2023). When $\log_{10}(M_{\mathrm{ecl}}/M_\odot) > 2.1$, the $m_{*,\mathrm{iso}}$–M_{ecl} relation follows the $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation. Assuming SFE = 0.33, from the turning point $\log_{10}(M_{\mathrm{ecl}}/M_\odot) \approx 2.1$, we have $\log_{10}(M_{\mathrm{cl}}/M_\odot) \approx 2.55$, consistent with the findings above.

3.3. Synthetic Embedded Clusters

We adopted the optimal sampling algorithm (Kroupa et al. 2013) to produce a population of stars for an embedded cluster with the mass M_{ecl} using the publicly available GalIMF code. Detailed descriptions of the optimal sampling and the publicly available GalIMF code that optimally populates clusters and galaxies with stars are available in Schulz et al. (2015) and Yan et al. (2017). The code has three input parameters: the mass of the embedded cluster, M_{ecl}, its age, and a metallicity associated with this adopted age. In this work, we assume that the metallicity of the clumps and their associated embedded clusters is the solar metallicity ([Fe/H] = 0). For the age, following Zhou et al. (2024a), it is taken to be 1 Myr.

In Figure 4, we picked out the most massive star with mass m_{max} in each synthetic embedded cluster and compared it with the corresponding $m_{*,\mathrm{iso}}$. When $\log_{10}(m_{\mathrm{max}}/M_\odot) > 1$, m_{max} and $m_{*,\mathrm{iso}}$ are roughly comparable. If the mass of the most massive star in a synthetic embedded cluster is $\log_{10}(m_{\mathrm{max}}/M_\odot) \approx 1$, which corresponds to $\log_{10}(M_{\mathrm{cl}}/M_\odot) \approx 2.55$ or $\log_{10}(M_{\mathrm{ecl}}/M_\odot) \approx 2.1$, consistent with the results in Section 3.2. Moreover, using the fitted $m_{\mathrm{max}} - M_{\mathrm{ecl}}$ relation, i.e., Equation 1 in Weidner et al. (2013), when $\log_{10}(M_{\mathrm{ecl}}/M_\odot) \approx 2.1$, we can also obtain $\log_{10}(m_{\mathrm{max}}/M_\odot) \approx 1$.

3.4. Accretion

The turning point $\log_{10}(M_{\mathrm{cl}}/M_\odot) \approx 2.55$ is used to judge the relative importance of accretion luminosity in Zhou et al. (2024a), also presented in Figure 1(a). Therefore, the abnormal tail in Figures 3 and 4 on the left of the vertical dashed lines means that if the accretion luminosity is dominant, the formalism in Section 3.1 cannot be used, because the main Lyman continuum flux or radio emission are not predominantly

from the stars. The high radio flux density of low-luminosity and low-mass H II-clumps in Figure 3(b) indicate that accretion is an effective way to produce radio emission. Therefore, the Lyman continuum flux $N_{\rm Ly}$ derived from radio flux density S_ν includes a significant contribution from the accretion process. The previous studies of H II regions in Lumsden et al. (2013), Sánchez-Monge et al. (2013), Urquhart et al. (2013) have noted that there is an excess of Lyman photons over what would be expected from the ionizing stars compared to the bolometric luminosities. The inclusion of the luminosity contribution from accretion may provide an explanation.

In Section 3.1, we obtained the stellar luminosity $L_{*,\rm T84}$ from the Lyman continuum flux $N_{\rm Ly}$. Actually, $L_{*,\rm T84}$ is not only the stellar luminosity, which includes the real stellar luminosity L_*, but the sum of L_* and the accretion luminosity $L_{\rm acc}$, i.e., $L_{*,\rm T84} = L_* + L_{\rm acc}$. Using the method described in Zhou et al. (2024a), we calculated the accretion luminosity $L_{\rm acc}$ of the most massive stars of the synthetic embedded clusters, and subtracted $L_{\rm acc}$ from the corresponding $L_{*,\rm T84}$. Then we repeated the analysis in Section 3.2 using L_* rather than $L_{*,\rm T84}$. As can be seen in Figure 5(a), the abnormal tail disappears.

3.5. Flares

The value of L_* still remains somewhat higher in Figure 5(a). The abnormal tail in Figure 5(b) is not completely eliminated, although it shows a substantial improvement over the case without considering accretion, presented in Figure 4. Therefore, the stellar luminosity L_* may still be overestimated. There are likely other physical processes contributing to the luminosity that have not been considered, similar to the accretion luminosity, which also needs to be subtracted.

Flaring represents a common occurrence of magnetic activity in low-mass stars, including our Sun. While flares, originating from both the Sun and other stars, are predominantly observed in the soft X-ray band, the majority of the emitted energy is released at optical/UV wavelengths (Flaccomio et al. 2018; Kuznetsov & Kolotkov 2021). For a sample of pre-main-sequence (PMS) stars in the NGC 2264 star-forming region, Flaccomio et al. (2018) detected seventy-eight X-ray flares with optical and/or mid-infrared counterparts. The optical emission of flares, encompassing both emitted energy and peak flux, demonstrates a strong correlation with, and notably surpasses, the X-ray emission. The luminosities in X-ray ($L_{\rm x}$) and optical ($L_{\rm opt}$) bands satisfy $L_{\rm opt} = a \times L_{\rm x}^{b}$ with a ranging from 5.2 to 14.7 and b between 0.6 and 1.0. Here we take the average values, i.e., $L_{\rm opt} \approx 10 \times L_{\rm x}^{0.8}$.

Getman & Feigelson (2021) identified a sample of 1086 X-ray superflares and megaflares. These events are generated by young stars across all masses throughout evolutionary stages spanning from protostars to stars without disks. Only the sample of Getman & Feigelson (2021) has mass estimates of the PMS stars. We converted the X-ray luminosity of the flare

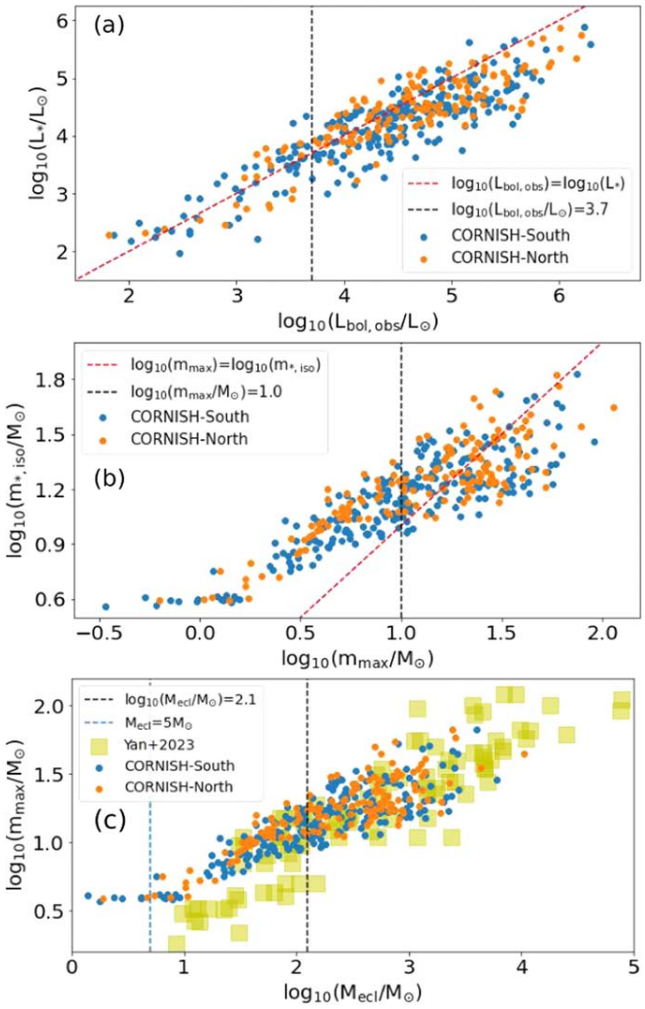


Figure 5. (a) The correlation between L_* and $L_{\rm bol,obs}$; (b) Comparison between $m_{\rm max}$ from the synthetic clusters and $m_{*,\rm iso}$ from the observation; (c) The $m_{\rm max} - M_{\rm ecl}$ relation. Yellow squares represent the star clusters from the table of Yan et al. (2023).

measured in Getman & Feigelson (2021) into optical luminosity using the relation fitted in Flaccomio et al. (2018). In Figure 6(a), the distribution of $L_{\rm opt}$ is highly scattered, and it has almost no correlation with stellar mass. The ratio of the optical and bolometric luminosities ($L_{\rm opt}/L_{\rm bol}$) of the PMS stars decreases rapidly as mass increases. Even for low-mass stars, the ratio is always less than 1. Thus, the luminosity from the flares can be neglected, and does not make a significant contribution to the abnormal tail in Section 3.2.

3.6. Star Formation Efficiency

3.6.1. Observations

In Figure 5(c), for low-mass clusters, there is a mismatch between our sample and the sample in Yan et al. (2023). As

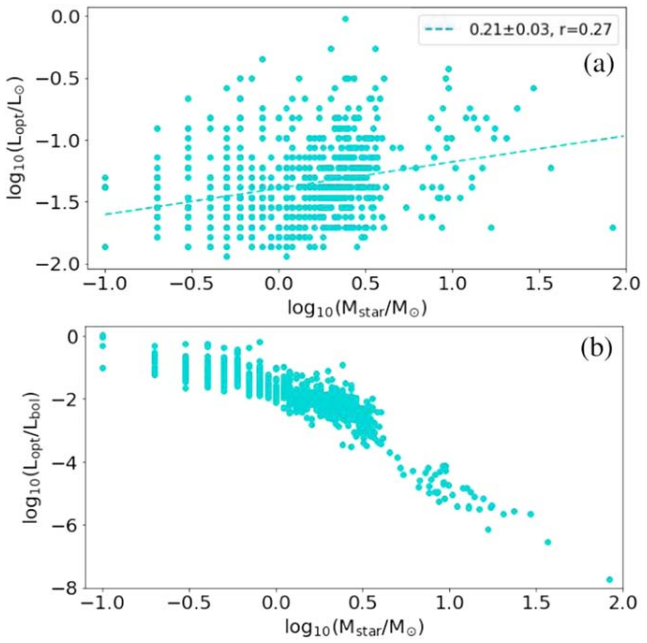


Figure 6. (a) Correlation between the optical luminosity and the stellar mass. "r" is the Pearson correlation coefficient of the linear fitting; (b) Correlation between the ratio of optical and bolometric luminosities of a flare and the stellar mass.

discussed in Section 3.5, we may be overestimating the stellar mass. Another possibility is that the low-mass clumps have higher SFEs > 0.33. Assuming all measured maximum stellar masses follow the $m_{\rm max} - M_{\rm ecl}$ relation, we can calculate the mass of the embedded cluster and then the star formation efficiency for each clump. The third-order polynomial fit to the observed $m_{\rm max} - M_{\rm ecl}$ relation presented in Weidner et al. (2013) has large errors. As an alternative, we derived a theoretical $m_{\rm max} - M_{\rm ecl}$ relation directly from the IMF. The details are described in Section A and the embedded cluster mass, $M_{\rm ecl}$, is given by

$$M_{\rm ecl} \approx \frac{5.37}{\frac{0.77}{m_{\rm max}^{1.3}} - 0.001} - \frac{3.33}{m_{\rm max}^{0.3}\left(\frac{0.77}{m_{\rm max}^{1.3}} - 0.001\right)}, \quad (4)$$

where we use the $m_{\rm max}$ values displayed in Figure 5(c) to derive the corresponding value of $M_{\rm ecl}$ for each embedded cluster.

The abnormal tail in Figure 5(b) is visible in Figure 7(a) again, presented as the SFEs close to 1. We excluded this tail in the subsequent analysis by requiring the clump mass larger than $10^{1.5}$ $M_{\odot}$. Figure 7(b) displays a strong correlation between the SFE and the clump mass. The SFE decreases with increasing clump mass, with an upper limit of ≈0.65, This is also visible in Figure 5 of Wells et al. (2022). We fit the relation between $M_{\rm cl}$ and the SFE with a power law of the form SFE $\propto M_{\rm cl}^{k}$ and find a value of $k = -0.44 \pm 0.02$. This anti-correlation between the mass of the clumps and their SFE has been predicted theoretically by Dib et al. (2013). For the simple case of uniform density clumps, Dib et al. (2013) predicted the clump mass and the SFE to follow SFE $\propto M_{\rm cl}^{-0.6}$ and to be the result of an increasingly stronger effect of stellar feedback in clumps with higher masses which leads to a faster gas expulsion and limits the value of the SFE. Interestingly, the median value of the SFEs is ≈0.3, consistent with our previous work (Zhou et al. 2024a, 2024b). An SFE ≈ 0.33 is also consistent with the value obtained from hydrodynamic calculations including self-regulation (Machida & Matsumoto 2012; Bate et al. 2014) and also with observations of embedded systems in the solar neighborhood (Lada & Lada 2003; Megeath et al. 2016).

3.6.2. Models

In order to assess the uncertainty in the observations, now we independently study the clump's SFE using theoretical models. In Figure 7(b), the upper limit of the SFE is ≈0.65. In Figure 4 of Wells et al. (2022), a SFE larger than 0.5 is necessary to fit the low-mass H II-clumps. Therefore, we consider the SFEs that fall in the range [0.05, 0.65] with a step of 0.05. Then each H II-clump has 13 alternative SFEs. We consider both the stellar luminosity and accretion luminosity of each star in the embedded cluster of a H II-clump, and also apply the PMS evolutionary track for each star. There is certainly an age spread of the stellar populations in embedded clusters within H II-clumps. Therefore, selecting the appropriate isochrones according to the stars' ages is essential. Additionally, not all stars in H II-clumps are in the accretion phase, as this too is age-dependent. We should identify the fraction of stellar objects that remain protostars to properly add their accretion luminosities. In order to address these issues, we apply the method described in Zhou et al. (2024a) to generate an age distribution for the stellar population in each synthetic cluster. The age distribution of stars within a H II-clump depends on the star formation history (SFH) of the clump. As discussed in Zhou et al. (2024a), compared to a constant SFH, burst-like and time-dependent SFHs can better fit the observational data. Thus, we only consider the time-dependent SFHs in this work. As described in Zhou et al. (2024a), there are 12 different SFHs with different age peaks ($t_{\rm p}$) and age standard deviations ($\sigma_{\rm t}$), i.e., $t_{\rm p} = 0.1$, 0.25, 0.5, 1 Myr and $\sigma_{\rm t} = 0.25$, 0.5, 1 Myr. Finally, each clump has 156 different combinations of SFEs and SFHs, resulting in 156 distinct bolometric luminosities. The combination of a SFE and a SFH that produces the bolometric luminosity of the synthetic embedded cluster closest to the observed bolometric luminosity of the corresponding H II-clump is considered as the optimal one. The optimal SFE and SFH distributions of all H II-clumps are

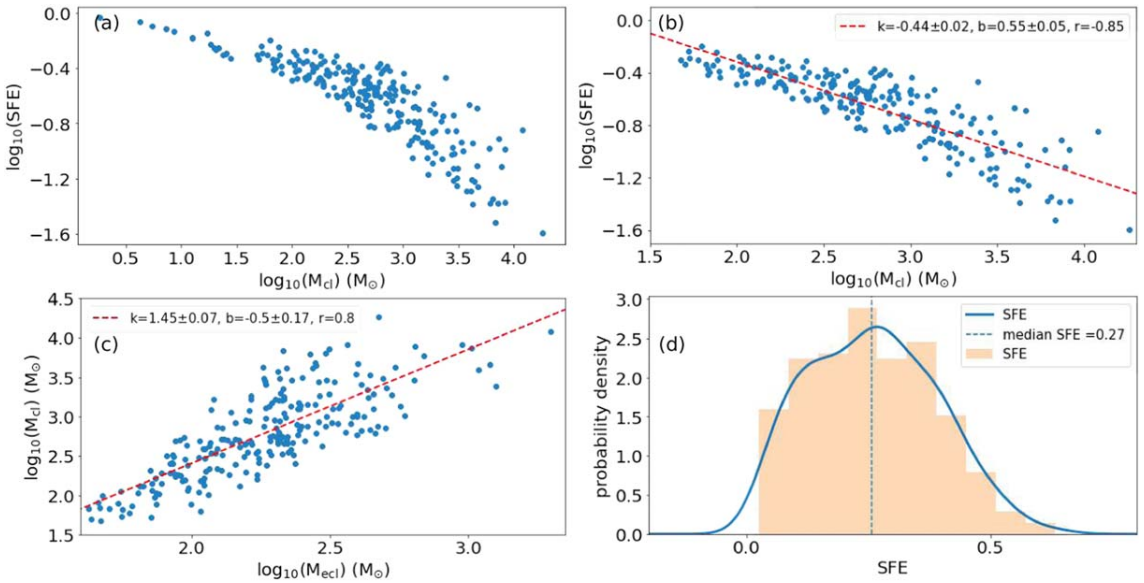


Figure 7. Star formation efficiency derived from the observations. (a) The correlation between the star formation efficiency and the clump mass for all the data points presented in Figure 5(b); (b) Eliminated the abnormal tail in panel (a); (c) The correlation between the embedded cluster mass and the corresponding H II-clump mass; (d) The distribution of the SFEs. Red dashed line is the linear fitting to the data points. "k" and "b" are the slope and the intercept, "r" is the Pearson correlation coefficient.

shown in Figure 8. The optimal combination has the smallest difference from the observed bolometric luminosity. To test the robustness of the model, we also selected the second optimal combination, which has the second smallest difference from the observed bolometric luminosity. As shown in Figure B1 and B2, the optimal and the second optimal combinations give almost the same results.

Compared with Figure 7, overall, the SFEs of H II-clumps derived from the observation and the model are comparable. From Figures 8(b) and (c), we have

$$\log_{10}(\mathrm{SFE}) = (-0.37 \pm 0.01) \times \log_{10}(M_{\mathrm{cl}}) + (0.42 \pm 0.04)\,, \quad (5)$$

and

$$\log_{10}(M_{\mathrm{cl}}) = (1.02 \pm 0.02) \times \log_{10}(M_{\mathrm{ecl}}) + (0.52 \pm 0.05). \quad (6)$$

Then we use Equation (5) to calculate the SFE or the mass of the embedded cluster for each H II-clump. As shown in Figure 9, after the modification of the SFE, the mismatch in Figure 5(c) indeed disappears.

3.7. Star Formation History

Figure 10 displays the distributions of the optimal t_{p} and σ_{t} for H II-clumps with different masses. For each t_{p} or σ_{t}, we calculated the median and mean masses of the corresponding H II-clump groups. The median and mean clump masses are comparable for different σ_{t}. However, there is a weak trend

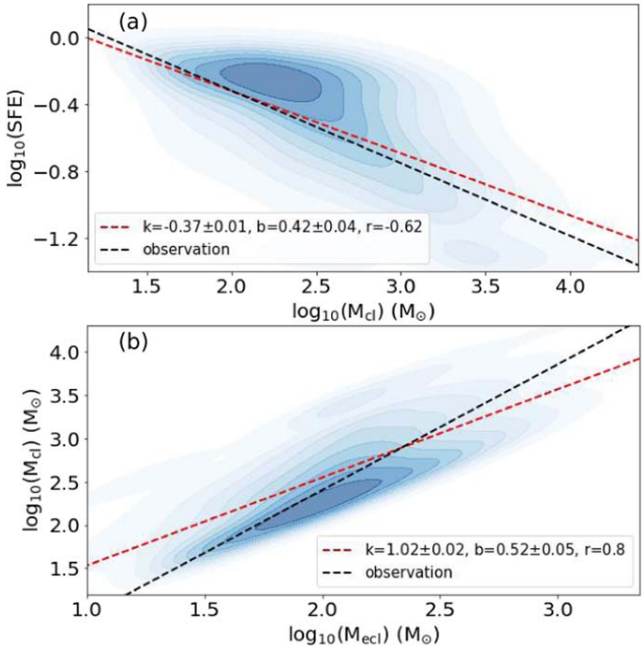


Figure 8. Star formation efficiency derived from the model. (a) The correlation between the star formation efficiency and the clump mass for all H II-clumps with a mass larger than $10^{1.5}$ $M_{\odot}$; (b) The correlation between the embedded cluster mass and the corresponding H II-clump mass. Red dashed line is the linear fitting to the data points. Black dashed lines come from Figure 7. "k" and "b" are the slope and the intercept, "r" is the Pearson correlation coefficient.

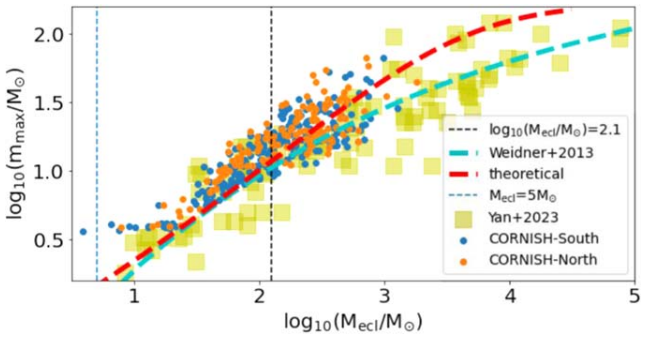


Figure 9. Same as Figure 5(c), but after modifying the SFE using Equation (5). Cyan dashed line is the $m_{\rm max} - M_{\rm ecl}$ relation fitted in Weidner et al. (2013) based on the observation. Red dashed line is the $m_{\rm max} - M_{\rm ecl}$ relation derived from the IMF theory, i.e., Equation (4).

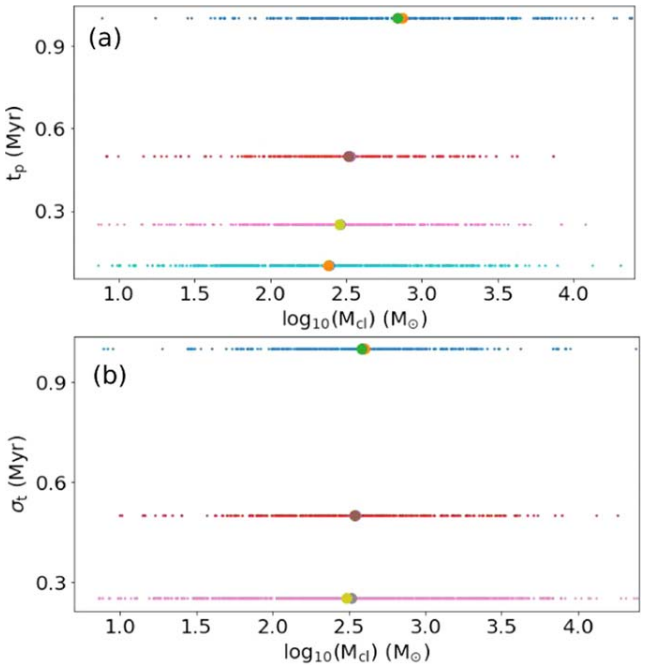


Figure 10. The distributions of the optimal $t_{\rm p}$ and $\sigma_{\rm t}$ for all H II-clumps with mass estimation. For each $t_{\rm p}$ or $\sigma_{\rm t}$, we calculated the median and mean masses of the corresponding HII-clump groups, marked by big dots.

that clumps with larger masses have larger $t_{\rm p}$, indicating that low-mass clumps reach the peak of star formation earlier than high-mass clumps, possibly because star formation occurs more rapidly in low-mass clumps. The clump free-fall time ($\tau_{\rm ff}$) is useful in providing a lower limit to the star formation timescales. In Figure 11, we compared the free-fall time of ATLASGAL clumps in the earliest ("quiescent") and the latest ("H II region") stages (Urquhart et al. 2022). Low-mass clumps indeed have a shorter free-fall time.

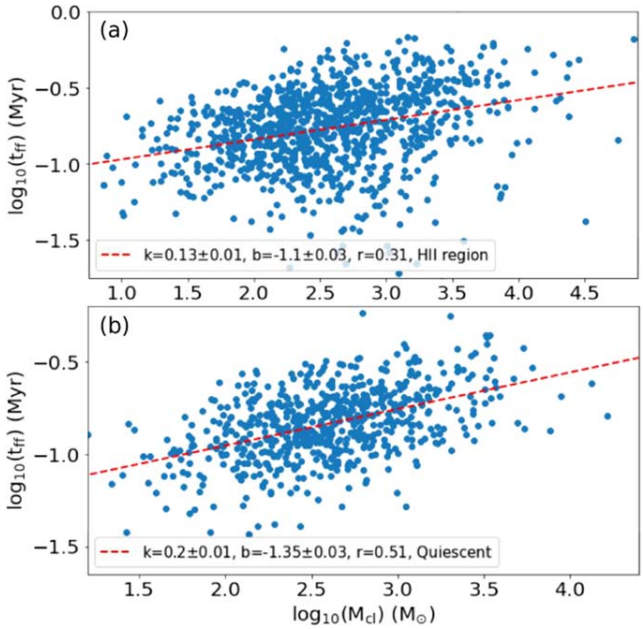


Figure 11. The correlation between the free-fall time and the clump mass in the earliest ("quiescent") and the latest ("H II region") stages.

4. Conclusion

In this work, 1226 H II-clumps from the ATLASGAL survey with mass $M_{\rm cl}$ and bolometric luminosity $L_{\rm bol,obs}$ estimates were matched with radio sources in the catalogs of the CORNISH-North and the CORNISH-South surveys. 392 H II-clumps have corresponding radio sources in the CORNISH-North/South surveys. These H II-clumps are mainly concentrated at the upper end in the $L_{\rm bol,obs}$–$M_{\rm cl}$ diagram, which are more luminous than the remaining H II-clumps, indicating they host high-mass stars. Radio sources are excited by the most massive stars in the embedded clusters of H II-clumps. We calculated the EM, the electron density ($n_{\rm e}$) and the Lyman continuum flux ($N_{\rm Ly}$) for each radio source. Using the relationship between the stellar parameters and the Lyman continuum flux in Table 1 of Thompson (1984), we determined the stellar luminosity $L_{*,\rm T84}$ according to the calculated $N_{\rm Ly}$. We found a turning point around $\log_{10}(L_{\rm bol,obs}/L_{\odot}) \approx 3.7$, which corresponds to a H II-clump mass $\log_{10}(M_{\rm cl}/M_{\odot}) \approx 2.55$. When the bolometric luminosity of H II-clumps is less than this value, $L_{*,\rm T84}$ is significantly larger than $L_{\rm bol,obs}$. Thus, there is an abnormal tail in the $L_{\rm bol,obs}$-$L_{*,\rm T84}$ diagram. Using the 0.5 Myr isochrone, we estimated the stellar mass $m_{*,\rm iso}$ according to the stellar luminosity $L_{*,\rm T84}$.

To investigate the physical origin of this turning point, we adopted the optimal sampling algorithm to produce a population of stars for an embedded cluster using the publicly available GalIMF code. Then, we picked out the most massive star with mass $m_{\rm max}$ in each synthetic embedded cluster and

compared it with the corresponding $m_{*,\rm iso}$. When $\log_{10}(m_{\rm max}/M_\odot) > 1$, $m_{\rm max}$ and $m_{*,\rm iso}$ are roughly comparable. Actually, the turning point at $\log_{10}(M_{\rm cl}/M_\odot) \approx 2.55$ exactly corresponds to $\log_{10}(m_{\rm max}/M_\odot) \approx 1$, which is also used to judge the relative importance of accretion luminosity in Zhou et al. (2024a). The abnormal tail means that if the accretion luminosity is dominant, we can not directly estimate the stellar parameters from the radio emission, and have to subtract the contribution of accretion from the radio flux density. To address this issue, we calculated the accretion luminosity $L_{\rm acc}$ of the most massive star in each synthetic embedded cluster, and subtracted $L_{\rm acc}$ from the corresponding $L_{*,\rm T84}$. However, the cleaned stellar luminosity L_* ($=L_{*,\rm T84}$–$L_{\rm acc}$) may be still overestimated. There are likely other physical processes contributing to the luminosity apart from the accretion. For example, the flaring activity represents a common occurrence of magnetic activity in low-mass stars. We found that the ratio of the optical and bolometric luminosities ($L_{\rm opt}/L_{\rm bol}$) of PMS stars decreases rapidly as mass increases. Even for low-mass stars, the ratio is always less than 1. Thus, the luminosity from the flaring activity can be neglected. Other physical processes need to be further investigated.

In the above analysis, we assumed a constant SFE = 0.33. We further studied the change of SFE with the clump mass. According to the derived mass of the most massive star in each H II-clump, using the theoretical $m_{\rm max} - M_{\rm ecl}$ relation, we calculated the mass of the corresponding embedded cluster and then the SFE of the clump. We find a strong anti-correlation between the SFE and the clump mass. The SFE decreases with increasing clump mass in agreement with theoretical expectations (Dib et al. 2013), with a median value of ≈0.3. We also independently derived the SFE for each H II-clump based on the model developed in Zhou et al. (2024a). We consider the SFEs that fall in the range [0.05, 0.65] with a step of 0.05. We employ 12 different star formation histories for each H II-clump to distribute the ages to the stars in the clump. According to the stars' age, we consider both the stellar luminosity and accretion luminosity of each star in the embedded cluster of a H II-clump, and apply the PMS evolutionary track for each star. Finally, each clump has 156 different combinations of SFEs and SFHs, resulting in 156 distinct bolometric luminosities. The optimal one has the smallest difference from the observed bolometric luminosity of the corresponding H II-clump. Overall, the SFEs of H II-clumps derived from the observations and the models are comparable. Using the quantitative relation between the SFE and the clump mass, we calculated the SFE and the mass of the embedded cluster for each H II-clump. As shown in Figure 9, after the modification of the SFE, the mismatch in Figure 5(c) disappears.

For the optimal $t_{\rm p}$ or $\sigma_{\rm t}$, we estimated the median and mean masses of the corresponding H II-clump groups. There is a weak trend that clumps with larger masses have larger $t_{\rm p}$, indicating that low-mass clumps reach the peak of star formation earlier than high-mass clumps, possibly because star formation occurs more rapidly in low-mass clumps, consistent with the shorter free-fall time of low-mass clumps.

Acknowledgments

We would like to thank the referee for the comments and suggestions that helped improve and clarify this work.

Appendix A
The Theoretical $m_{\rm max} - M_{\rm ecl}$ Relation

The stellar IMF, $\xi_\star(m)$, is canonical (Kroupa 2001):

$$\xi_\star(m, M) = \begin{cases} 0, & m < 0.08\ M_\odot, \\ 2k_\star m^{-1.3}, & 0.08\ M_\odot \leqslant m < 0.5\ M_\odot, \\ k_\star m^{-2.3}, & 0.5\ M_\odot \leqslant m < 1 M_\odot, \\ k_\star m^{-\alpha_3}, & 1\ M_\odot \leqslant m < m_{\rm max}(M), \\ 0, & m_{\rm max}(M) \leqslant m, \end{cases} \quad \text{(A1)}$$

where $\alpha_3 = 2.3$ is the constant Salpeter-Massey index for the invariant canonical IMF but will change for larger $\rho_{\rm cl}$ (the density of star-forming clump) to account for IMF variation under star-burst conditions (Elmegreen & Shadmehri 2003; Dib et al. 2007; Dabringhausen et al. 2009; Marks et al. 2012). 0.08 $M_\odot$ in Equation (A1) is about the lower mass limit of stars (Thies et al. 2015). The mass conservation of the embedded cluster gives

$$M_{\rm ecl} = \int_{0.08\ M_\odot}^{m_{\rm max}} m\ \xi_\star(m) dm. \quad \text{(A2)}$$

The optimal sampling normalization condition is

$$1 = \int_{m_{\rm max}}^{150\ M_\odot} \xi_\star(m) dm, \quad \text{(A3)}$$

where 150 $M_\odot$ is the adopted stellar upper mass limit (Weidner & Kroupa 2004; Figer 2005; Oey & Clarke 2005; Koen 2006; Maíz Apellániz et al. 2007). When $m_{\rm max} > 1 M_\odot$, the optimal sampling normalization condition becomes

$$1 = \int_{m_{\rm max}}^{150\ M_\odot} k_\star m^{-\alpha_3}\ dm. \quad \text{(A4)}$$

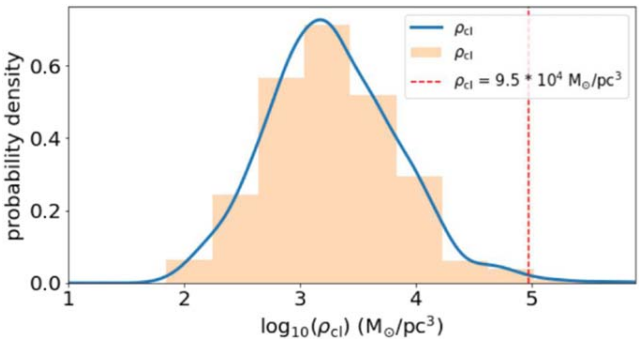


Figure A1. The volume density distribution of ATLASGAL clumps with H II regions.

For larger $\rho_{\rm cl}$, $\xi_\star(m)$ becomes top-heavy where a $\alpha_3(\rho_{\rm cl})$ relation is adopted from Marks et al. (2012):

$$\alpha_3 = \begin{cases} 2.3, & \rho_{\rm cl} < 9.5 \times 10^4, \\ 1.86 - 0.43\log_{10}(\rho_{\rm cl}/10^6), & \rho_{\rm cl} \geqslant 9.5 \times 10^4. \end{cases} \quad \text{(A5)}$$

Here

$$\rho_{\rm cl} = 3M_{\rm cl}/4\pi r_{\rm h}^3 \quad \text{(A6)}$$

in the unit of $[M_\odot\,{\rm pc}^{-3}]$ is the clump density when the embedded cluster is forming. For $M_{\rm cl}$ and $r_{\rm h}$ in the equation, we take the clump's mass and radius calculated in Urquhart et al. (2022). As shown in Figure A1, for ATLASGAL clumps with H II regions (H II-clumps), generally, they have $\rho_{\rm cl} < 9.5 \times 10^4$. Thus, we take $\alpha_3 = 2.3$ in this work.

Setting $x \equiv m_{\rm max}$ and solving the above equations, we obtain

$$\begin{aligned} M_{\rm ecl} &= \int_{0.08\,M_\odot}^{x} m\,\xi_\star(m)dm \\ &= \int_{0.08\,M_\odot}^{0.5\,M_\odot} m\,2k_\star m^{-1.3}\,dm + \int_{0.5\,M_\odot}^{x} m\,k_\star m^{-2.3}\,dm \\ &\approx \frac{5.37}{\frac{0.77}{x^{1.3}} - 0.001} - \frac{3.33}{x^{0.3}\left(\frac{0.77}{x^{1.3}} - 0.001\right)} \end{aligned} \quad \text{(A7)}$$

Appendix B
Robust of the Model

To test the robustness of the model, we also selected the second optimal combination of SFE and SFH, rather than the optimal combination in Section 3.6.2, which has the second smallest difference from the observed bolometric luminosity. The analysis here is the same as in Section 3.6.2.

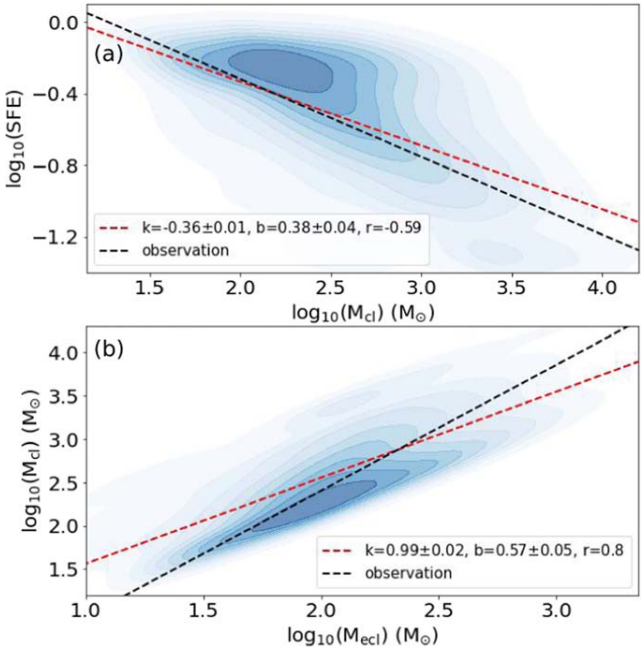


Figure B1. Same as Figure 8. But the star formation efficiency is derived from the case of the second optimal combination, as described in Section 3.6.

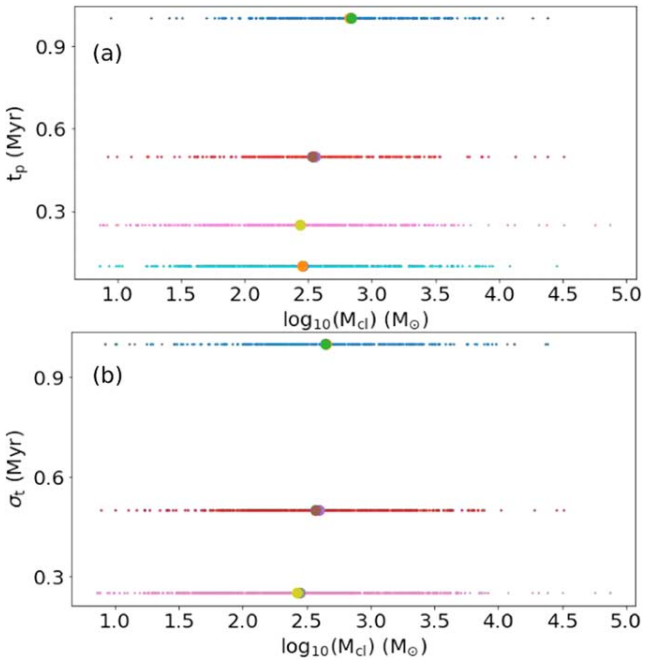


Figure B2. Same as Figure 10, but for the case of the second optimal combination, as described in Section 3.6.

ORCID iDs

J. W. Zhou https://orcid.org/0009-0004-2927-239X
Sami Dib https://orcid.org/0000-0002-8697-9808

References

Banerjee, S., & Kroupa, P. 2018, in The Birth of Star Clusters, ed. S. Stahler, Vol. 424 (Cham.: Springer), 143
Bate, M. R., Tricco, T. S., & Price, D. J. 2014, MNRAS, 437, 77
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Dabringhausen, J., Kroupa, P., & Baumgardt, H. 2009, MNRAS, 394, 1529
Davies, B., Hoare, M. G., Lumsden, S. L., et al. 2011, MNRAS, 416, 972
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Dib, S., Kim, J., & Shadmehri, M. 2007, ApJ, 381, L40
Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61
Elia, D., Merello, M., Molinari, S., et al. 2021, MNRAS, 504, 2742
Elmegreen, B. G., & Shadmehri, M. 2003, MNRAS, 338, 817
Figer, D. F. 2005, Natur, 434, 192
Flaccomio, E., Micela, G., Sciortino, S., et al. 2018, A&A, 620, A55
Getman, K. V., & Feigelson, E. D. 2021, ApJ, 916, 32
Irabor, T., Hoare, M. G., Burton, M., et al. 2023, MNRAS, 520, 1073
Kalcheva, I. E., Hoare, M. G., Urquhart, J. S., et al. 2018, A&A, 615, A103
Koen, C. 2006, MNRAS, 365, 590
Kroupa, P. 1995a, MNRAS, 277, 1491
Kroupa, P. 1995b, MNRAS, 277, 1507
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2005, in Proc. of the Gaia Symp. "The Three-Dimensional Universe with Gaia" (ESA SP-576), ed. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 629

A&A, 688, L19 (2024)
https://doi.org/10.1051/0004-6361/202450412
© The Authors 2024

Astronomy & Astrophysics

LETTER TO THE EDITOR

Self-similar cluster structures in massive star-forming regions: Isolated evolution from clumps to embedded clusters

Jian-wen Zhou[1,⋆], Pavel Kroupa[2,3,⋆], and Sami Dib[4,⋆]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[2] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany
[3] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, 180 00 Praha 8, Czech Republic
[4] Max Planck Institute für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Received 16 April 2024 / Accepted 25 July 2024

ABSTRACT

We used the dendrogram algorithm to decompose the surface density distributions of stars into hierarchical structures. These structures were tied to the multiscale structures of star clusters. A similar power-law for the mass-size relation of star clusters measured at different scales suggests a self-similar structure of star clusters. We used the minimum spanning tree method to measure the separations between clusters and gas clumps in each massive star-forming region. The separations between clusters, between clumps, and between clusters and clumps were comparable, which indicates that the evolution from clump to embedded cluster proceeds in isolation and locally, and does not affect the surrounding objects significantly. By comparing the mass functions of the ATLASGAL clumps and the identified embedded clusters, we confirm that a constant star formation efficiency of ≈0.33 can be a typical value for the ATLASGAL clumps.

Key words. stars: formation – ISM: clouds – ISM: structure – local insterstellar matter – galaxies: star clusters: general

1. Introduction

Massive star-forming regions (MSFRs) represent a major mode of star formation within the Galaxy (Lada & Lada 2003; Zinnecker & Yorke 2007; Motte et al. 2018; Dib 2023). Young stellar clusters within these regions can serve as progenitors for open clusters. It has been suggested that most if not all observed stars originate from embedded clusters (Kroupa 1995a,b, 2005; Lada & Lada 2003; Megeath et al. 2016; Dinnbier et al. 2022). Measurements of the internal structure and dynamics of young clusters and star-forming regions are necessary to fully understand the process of their formation and dynamical evolution.

The project called Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX; Feigelson et al. 2013) examines 20 nearby ($d < 3.6$ kpc) MSFRs using a combination of archival *Chandra* X-ray imaging, 2MASS+UKIDSS near-IR (NIR), and *Spitzer* mid-IR (MIR) survey data. The MYStIX regions include the Orion nebula, the Flame nebula, W 40, RCW 36, NGC 2264, the Rosette nebula, the Lagoon nebula, NGC 2362, DR 21, RCW 38, NGC 6334, NGC 6357, the Eagle nebula, M 17, W 3, W 4, the Carina nebula, the Trifid nebula, NGC 3576, and NGC 1893, from which a sample of 31 784 MYStIX probable complex members (MPCMs; Broos et al. 2013) was obtained. This catalog of 31 784 young stars includes both high- and low-mass stars and disk-bearing and disk-free stars (Broos et al. 2013). Although the MYStIX samples are not complete, the young stars were identified in a uniform way for the different regions. This allowed a comparative analysis of the stellar populations in various regions. A comparison of the properties of a large cluster sample spanning a range of environments and ages enables inferences about processes such as gas removal, cluster expansion, subcluster mergers, dynamical boundedness and relaxation, and cluster dispersal, which all affect the spatial structure of clusters (Kuhn et al. 2014, 2015a,b).

In 17 of the MYStIX MSFRs, Kuhn et al. (2014) identified 142 subclusters of young stars using finite-mixture models. Their physical parameters were summarized in Kuhn et al. (2015b). Kuhn et al. (2015a) presented a census of the total intrinsic populations (estimated total numbers of OB and pre-main-sequence stars down to 0.1 $M_\odot$) and the surface density distributions for these 17 MSFRs using catalogs of young stars. The observed surface densities for X-ray selected MPCMs were calculated using an adaptive-smoothing method and were then corrected to the intrinsic populations by dividing the observed surface densities by the detection fraction maps. These stellar surface density maps can thus be directly compared with each other. However, the structures observed in the surface density maps represent the projected distributions of stars within star-forming regions, leading to possible overlaps of physically separate star groups on the map. Additionally, some of the small dense subclusters might not be visible in the surface density maps. More details about these aspects are given in Kuhn et al. (2015a).

2. Method

Observations of MSFRs at various stages of their star-forming lifespan unveil diverse stellar cluster structures. Notably, a

⋆ Corresponding authors; jwzhou@mpifr-bonn.mpg.de; pkroupa@uni-bonn.de; sami.dib@gmail.com

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

striking feature of the spatial distributions of young stars is their organization into clusters and subclusters. In addition to the main clusters within these regions, smaller stellar groups can be observed outside the main clusters, or even as subclusters within the main clusters (Lada & Lada 2003; Portegies Zwart et al. 2010; Krumholz et al. 2019). Stars form from hierarchically collapsing molecular clouds, which often results in clustered star formation that mirrors the intricate structure of the original molecular cloud. In a similar way to what Zhou et al. (2024a,b) reported for molecular clouds and clumps, we used the dendrogram algorithm to decompose the surface density distributions of stars into hierarchical structures. This revealed the multiscale structures of the star clusters.

3. Results and discussion

3.1. Dendrogram structures

We identified local embedded star clusters using the dendrogram algorithm according to the surface density distributions of stars derived in Kuhn et al. (2015a). As described in Rosolowsky et al. (2008), the dendrogram algorithm can decompose the density maps into hierarchical structures called branches and leaves (here, clusters and their subclusters). The *astrodendro* package[1] includes three main input parameters for the dendrogram algorithm: *min_value* for the minimum value to be considered in the dataset, *min_delta* for a leaf that can be considered as an independent entity, and *min_npix* for the minimum area of a structure.

In some observations, the cluster size was measured by the length scale over which the density profile exceeds twice the standard deviation of the surface density in the surrounding field (Lada et al. 1991; Carpenter 2000; Kumar et al. 2006). According to the surface density map, we estimated the mean value of the low-density regions located at the edge as the background density, that is, $\Sigma_{\rm rms}$, then we set the parameters $min_value = 2{*}\Sigma_{\rm rms}$ and $min_delta = 2{*}\Sigma_{\rm rms}$. The identification is not sensitive to these two parameters. Taking $5{*}\Sigma_{\rm rms}$ does not change the results significantly. The number of the extracted structures in the two cases are very similar because the stellar surface density of local subclusters is much higher than the background stellar surface density ($\Sigma_{\rm rms}$). The mass of the smallest stellar group observed in Kuhn et al. (2015b) is $\approx 5\,M_\odot$, according to the initial mass–radius relation of embedded clusters deduced by Marks & Kroupa (2012), that is,

$$\frac{r_{\rm h}}{\rm pc} = 0.10^{+0.07}_{-0.04} \times \left(\frac{M_{\rm ecl}}{M_\odot}\right)^{0.13\pm0.04}, \quad (1)$$

where $r_{\rm h}$ and $M_{\rm ecl}$ are the half-mass radius and the mass in stars of the embedded cluster, and the corresponding minimum radius $r_{\rm min}$ of an embedded cluster is $\approx$0.12 pc. The minimum area can be calculated by $\pi r^2_{\rm min}$, which is used to fix the parameter *min_npix*. Figure 1 displays an example that shows the hierarchical structures of the clusters in NGC 6334, decomposed by the dendrogram algorithm.

3.2. Cluster mass

The *astrodendro* package can output the mask of each structure, which can be directly used to calculate the physical quantities of the identified structures. The area of each pixel is A_0 (unit: pixel^2), and the number of pixels included in a mask is n. The mass of a cluster is then

$$M = \sum_{i=1}^{n} \Sigma_{\rm i} A_0 * 0.5, \quad (2)$$

where we assumed that the average mass of stars in a cluster is 0.5 $M_\odot$. $\Sigma_{\rm i}$ is the surface density of stars in a pixel (unit: n_s/pixel^2), with n_s the number of stars per pixel. The effective radius of the cluster is given by $R_{\rm eff} = n * A_0/\pi$.

The algorithm can also approximate the morphology of each structure as an ellipse. In the dendrogram algorithm, the long and short axes of an ellipse, a and b, are the rms sizes in radians (second moments) of the density distribution along the two spatial dimensions. However, as shown in Fig. 1, a and b give an ellipse that is much smaller than the size of the identified structure, but multiplying by a factor of 2.5 is appropriate to cover the structure (Zhou et al. 2024a,b). The effective physical radius of an ellipse is then $R_{\rm eff} = \sqrt{2.5a \times 2.5b} * d$, with d being the distance to each star-forming region. The elliptical mask can also be used to calculate the mass using Eq. (2).

Fig. 2 shows that similar calculation results are obtained with the original or the elliptical mask. Moreover, the radii and masses of leaf clusters are systematically smaller and lower than those of the subclusters in Table 1 of Kuhn et al. (2015b). Thus, they systematically move down to the lower left side. In Fig. 1(a), the

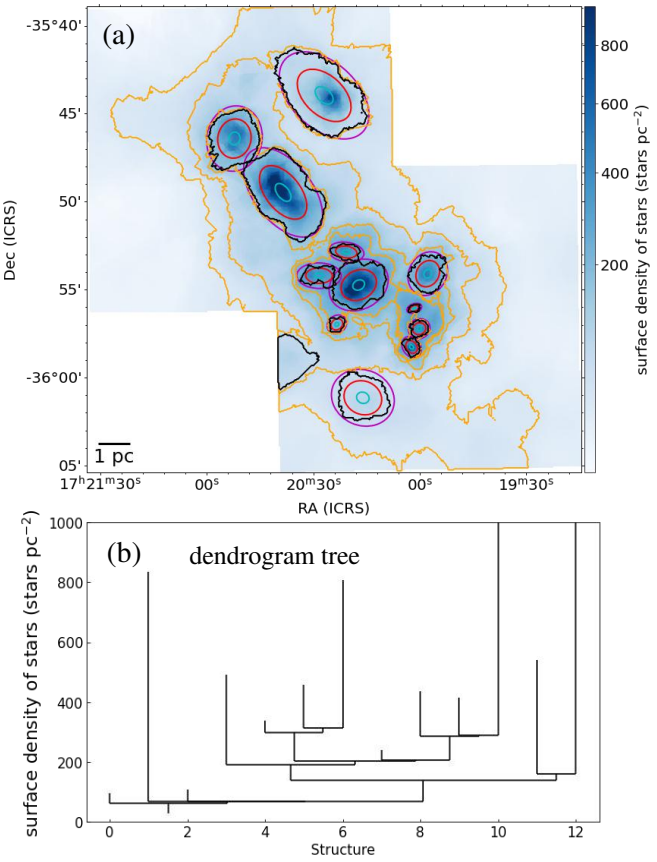


Fig. 1. Hierarchical structures of the clusters in NGC 6334. (a) The background is the surface density of stars. The black and orange contours show the masks of leaf and branch clusters, respectively. The magenta, red, and cyan cycles are the approximated ellipses with effective radii $R_{\rm eff}$, $R_{\rm eff,1}$, and $R_{\rm eff,2}$ defined in Sect. 3.3, respectively. (b) The dendrogram tree shows the hierarchical cluster structures.

[1] https://dendrograms.readthedocs.io/en/stable/index.html

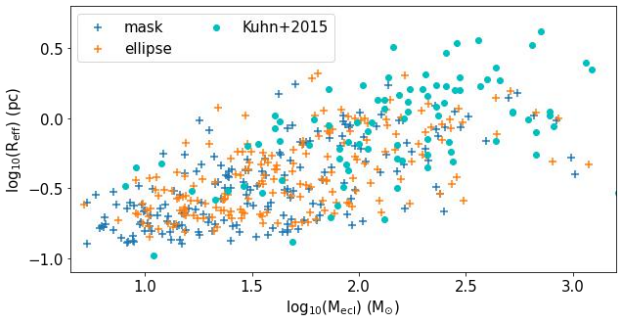


Fig. 2. Leaf clusters with the radii and masses calculated by the original mask and the elliptical mask (blue and orange pluses). The cyan dots show the subclusters in Table 1 of Kuhn et al. (2015b).

branch clusters present a complex morphology, and thus, only the original mask was used to do the calculations.

3.3. Self-similar mass-size relations

Fig. 3 shows that after the observed radii are unified to the half-mass radii (Zhou et al. 2024c), the embedded clusters in the catalogs of Lada et al. (1991), Carpenter et al. (1993), Lada & Lada (2003), Carpenter (2000), Kumar et al. (2006), Faustini et al. (2009) are quite comparable with the leaf clusters identified in this work. Their mass-size relations have almost the same slope, ≈0.33.

As marked in Fig. 1(a) by colored ellipses, we calculated the physical parameters of leaf clusters by taking the effective radius of the ellipse as $R_{\rm eff} = \sqrt{2.5a \times 2.5b} * d$, $R_{\rm eff,1} = \sqrt{1.5a \times 1.5b} * d$, and $R_{\rm eff,2} = \sqrt{0.5a \times 0.5b} * d$. As shown in Fig. 4(a), for these three cases, the data points mainly move down to the lower left side along the line of the mass-size relation with a constant slope of ≈0.3. Therefore, the slope of the mass-size relation is independent of the cluster radius and mass measurements, implying self-similar cluster structures. Interestingly, when the three different radii and mass measurements are mixed together, a steeper mass-size relation is obtained, similar to the mass-size relation found by Pfalzner et al. (2016).

Branch clusters are the complexes of leaf clusters. As shown in Fig. 4(b), they follow a similar trend as those shown in Fig. 4(a). When leaf and branch clusters were mixed atogether, we again obtained the steeper mass-size relation. The data points in Table 1 of Kuhn et al. (2015b) also follow this trend, although they were identified and measured with completely different methods. The radii and masses of the subclusters in Kuhn et al. (2015b) are quite comparable with the middle-scale branch clusters in this work, and they have similar slopes for the mass-size relations in Fig. 4(b).

In conclusion, different measurements for embedded clusters at different scales and ages give comparable mass-size relations. Here and in Kuhn et al. (2015b), a unified measurement of the samples results in a shallower slope for the mass-size relation. In contrast, when different measurements at different scales are mixed, the mass-size relation is always steeper. Although the mixing may not make sense, it is unavoidable when different catalogs from different observations are collected together. Thus, in Zhou et al. (2024c), we unified the observed radii from different observations to the half-mass radii using the templates obtained from direct N-body simulations. In this work, the unified measurement of the samples from the MYStIX project and the unified identification of the clusters effectively avoids this issue.

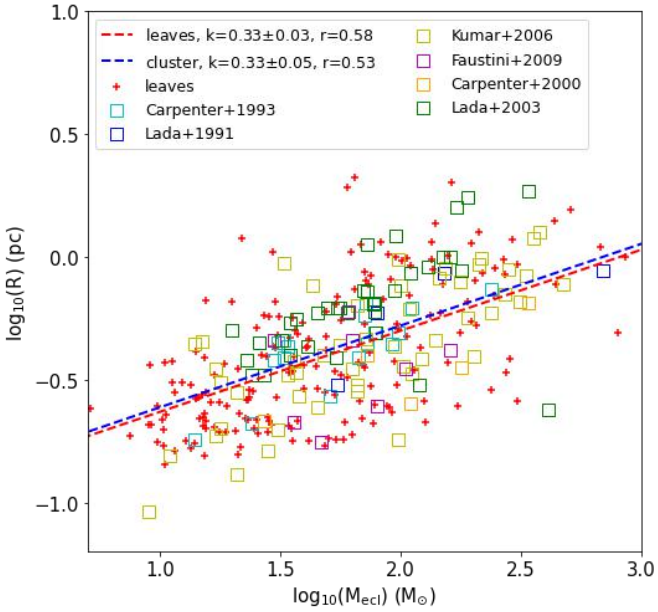


Fig. 3. Mass-size relations of leaf clusters (red plus) identified in this work fit by the dashed red line, and embedded clusters (colored squares) in the catalogs of Lada et al. (1991), Carpenter et al. (1993), Lada & Lada (2003), Carpenter (2000), Kumar et al. (2006), Faustini et al. (2009) fit by the dashed black line. The slope for all is ≈0.33.

The exponent $k \approx 0.33$ in Figs. 3 and 4(a) indicates a uniform distribution of the mass through the scale if the structures were to be roundish. This is consistent with the suggestions made in Zhou et al. (2024c), namely that all subclusters initially have a Plummer profile. The Plummer radius serves as a measure of the core radius, which is the region where the density remains relatively constant. For a Plummer profile with the Plummer radius $r_{\rm pl}$, the mass within the radius r is given by

$$M(r) = M_{\rm ecl} \frac{\left(\frac{r}{r_{\rm pl}}\right)^3}{\left[1 + \left(\frac{r}{r_{\rm pl}}\right)^2\right]^{\frac{3}{2}}}. \quad (3)$$

Thus, different radii in different observations can be unified to the half-mass radii. This operation does not change the slope of the mass-size relation, which only shifts the line of the mass-size relation, as shown in Fig. 4(a).

3.4. Isolated evolution

It remains controversial whether massive clusters are formed by the merger of clusters with lower masses. If there is a considerable merger between embedded clusters, we expect the separations between embedded clusters to be significantly smaller than their precursors, that is, clumps. In order to test this hypothesis, we used the clumps from the ATLASGAL survey, which provides the most well-characterized currently available sample of high-mass star-forming clumps (Urquhart et al. 2022). Seven of the 17 MSFRs are covered in the ATLASGAL survey, which are the Lagoon nebula, NGC 6334, NGC 6357, the Eagle nebula, M 17, the Carina nebula, and the Trifid nebula. We only considered the clumps contained in the fields of view of the MYStIX project. We also filtered out clumps whose kinematic distances differed significantly from the distance of the corresponding

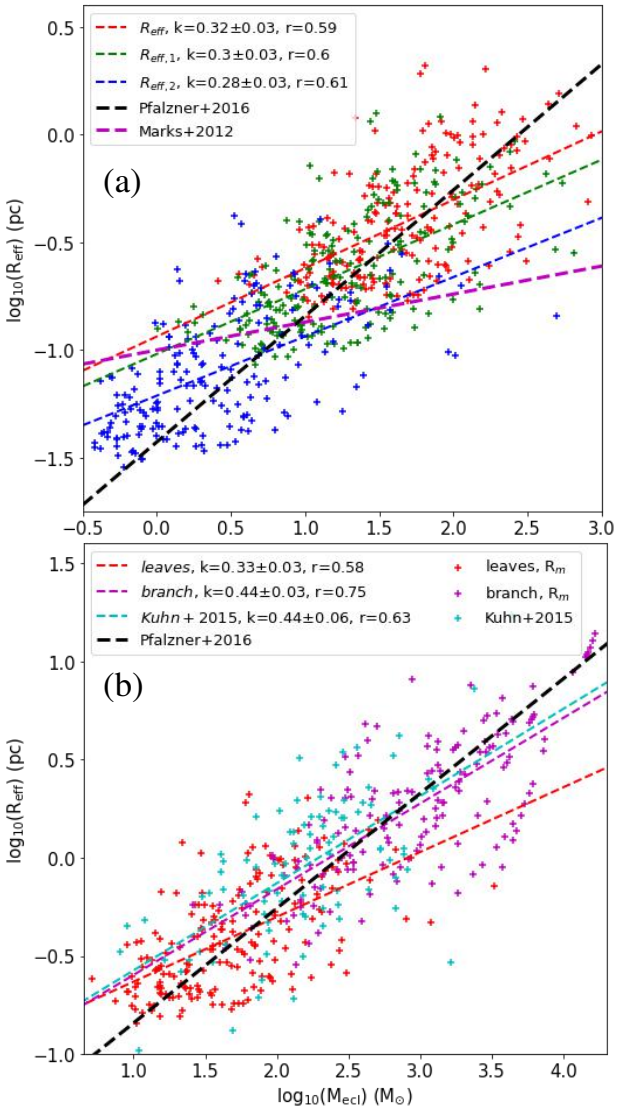


Fig. 4. Mass-size relation of star clusters measured at different scales. (a) Red, green, and blue pluses represent leaf clusters with effective radii $R_{\rm eff}$, $R_{\rm eff,1}$, and $R_{\rm eff,2}$ defined in Sect. 3.3 and marked by colored ellipses in Fig. 1(a), respectively. (b) Magenta, cyan, and red pluses represent branch clusters, which are the subclusters in Table 1 of Kuhn et al. (2015b), and leaf clusters, respectively. k and r are the slope of the mass-size relation and the Pearson correlation coefficient, respectively.

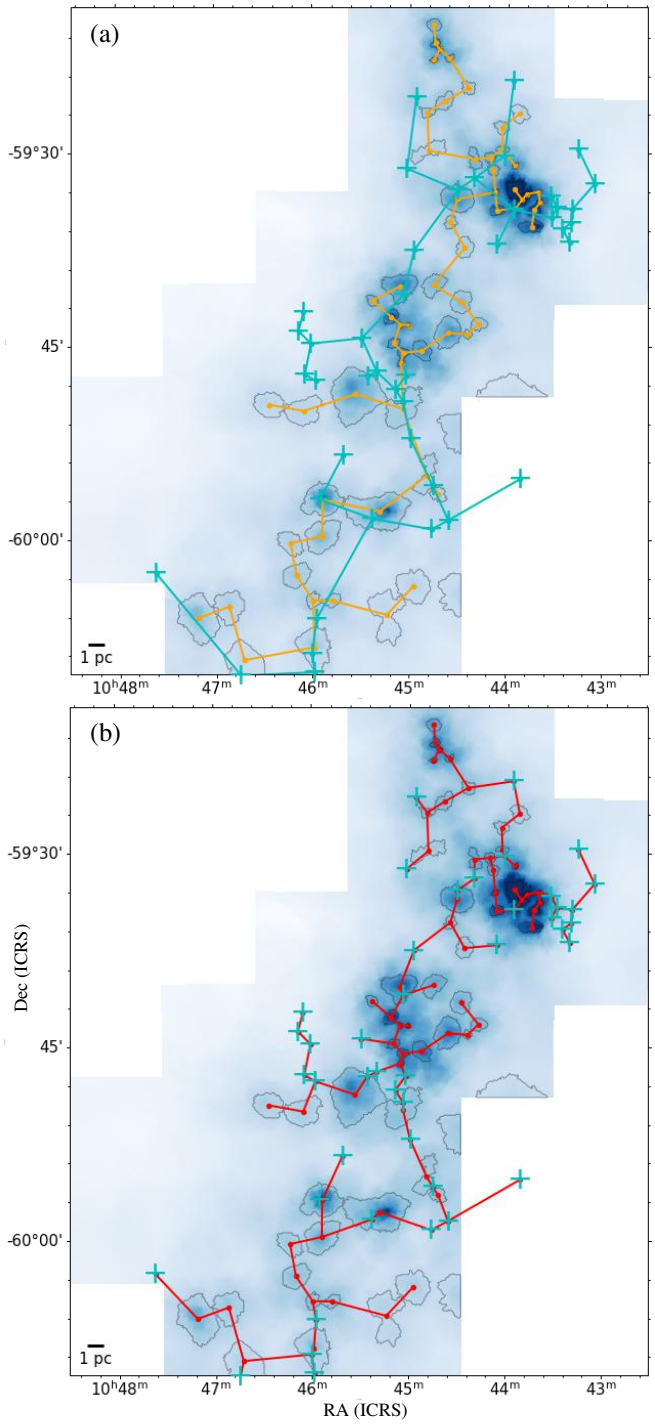


Fig. 5. Carina nebula illustrates as an example the minimum spanning tree (MST) method. The cyan plus marks the positions of ATLASGAL clumps. The black contours show the masks of leaf clusters. (a) Orange and cyan lines show the minimum spanning trees between clusters and between clumps, respectively. (b) MST between clusters and clumps.

massive star-forming region. In Fig. 5, the minimum spanning tree (MST) method (Cartwright & Whitworth 2004; Dib et al. 2018; Dib & Henning 2019) was used to measure the separation between each two neighboring objects (between clusters, between clumps, and between clusters and clumps). Interestingly, in Table 1 and Fig. 6, the peak values of the separation distributions in three cases are comparable. This indicates the following:

1. The evolution from clump to embedded cluster is fixed at the initial position of the clump. There is no significant positional change between the embedded cluster and its precursor, as illustrated in Fig. 7.
2. There is no significant merger between embedded clusters for the samples in the MYStIX project.
3. Feedback from embedded clusters does not redistribute the surrounding clumps, consistent with the conclusions in Zhou et al. (2024a,b): Feedback from embedded clusters does not significantly change the physical properties of the surrounding dense gas structures.

In short, the evolution from a clump to an embedded cluster proceeds in isolation and locally, with no significant effect on the surrounding objects. Neither does the well-established $m_{\rm max} - M_{\rm ecl}$ relation (Larson 2003; Bonnell et al. 2004; Weidner & Kroupa 2006; Weidner et al. 2013; Stephens et al. 2017; Yan et al. 2023),

Table 1. Median spacing between clusters, between clumps, and between clusters and clumps within each massive star-forming region.

Regions	Cluster (pc)	Clump (pc)	Cluster & Clump (pc)
Lagoon	0.84	1.04	0.82
NGC 6334	1.34	0.82	0.75
NGC 6357	1.25	0.78	0.76
Eagle	0.76	1.41	0.72
M 17	0.66	1.02	0.66
Carina	1.86	2.36	1.42
Trifid	1.67	2.19	1.67
Mean value	1.20	1.37	1.00

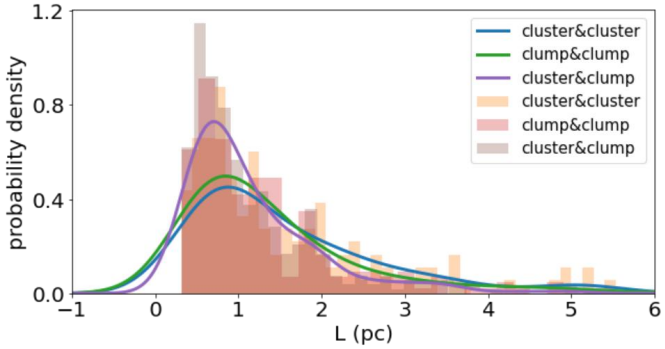


Fig. 6. Separation distributions of each two neighboring objects in three cases. The peak values for all are similar and close to 1 pc.

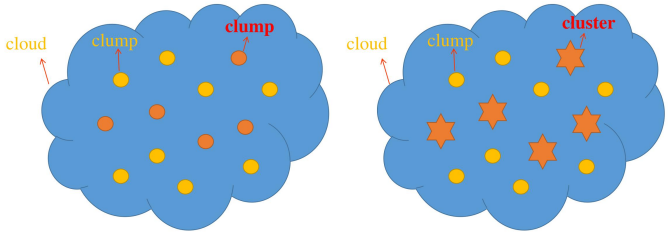


Fig. 7. Evolution from clump to embedded cluster fixed at the initial position of the clump.

a correlation between the embedded star cluster mass $M_{\rm ecl}$ and the mass of the most massive star $m_{\rm max}$, support the hypothesis that mergers during the formation of embedded clusters are common. Using Gaia DR2 data, Kuhn et al. (2019) investigated the kinematics of some subclusters identified in Kuhn et al. (2014), They found no evidence that these groups are merging. We note that the imbalanced populations of star clusters and clumps in terms of numbers may impact the statistics of the spacing. More samples are necessary for further validation.

3.5. Star formation efficiency

The massive star-forming regions in the MYStIX project analyzed in Kuhn et al. (2014) are located at distances <3.6 kpc. We also adopted this distance criterion to select the ATLASGAL clumps in Urquhart et al. (2022), which were used for a comparison with the identified embedded clusters. This distance limitation can also effectively eliminate a possible distance bias (Urquhart et al. 2022).

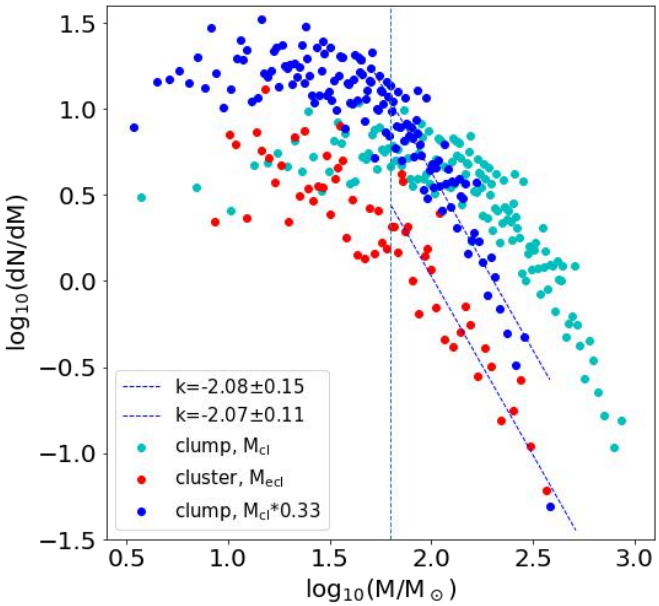


Fig. 8. Cluster and clump mass functions. The embedded clusters are taken from Fig. 3. The vertical dashed line marks the start point of the linear fit. k is the slope of the fitting.

Clumps are the precursors of embedded clusters. Urquhart et al. (2022) classified 5007 ATLASGAL clumps into four evolutionary stages, that is, quiescent, protostellar, young stellar objects (YSOs), and HII regions. The ATLASGAL clumps with HII regions (HII-clumps) represent the final stages of embedded cluster formation in the clumps. When we consider the hub-filament structure of the clump in high-resolution ALMA observations (Zhou et al. 2022, 2024d), the hub is the true formation site of an embedded cluster within the clump. Thus, the size of the clump should be significantly larger than the initial radius of the embedded cluster generated in the clump, which is also indicated by the initial mass–radius relation of embedded star clusters of Marks & Kroupa (2012). Therefore, the mass-size relation of the clumps is not comparable with that of the embedded clusters. The star formation efficiency (SFE) cannot be inferred from the differences between the mass-size relations of the clumps and embedded clusters, as was done in Pfalzner et al. (2016). However, before strong feedback occurs, the total mass of a clump at different evolutionary stages remains almost constant. The SFE can be deduced from the mass functions of the clumps and embedded clusters.

A constant star formation efficiency (SFE) of ≈0.33 is revealed in Fig. 8. This value has been widely used in previous simulations and has been proven to be effective (Kroupa et al. 2001; Banerjee et al. 2012; Banerjee & Kroupa 2013, 2014, 2015, 2017; Oh et al. 2015; Oh & Kroupa 2016, 2018; Brinkmann et al. 2017; Wang et al. 2019, 2020; Pavlik et al. 2019; Dinnbier et al. 2022). This SFE is also consistent with the value obtained from hydrodynamic calculations including self-regulation (Machida & Matsumoto 2012; Bate et al. 2014) and also with observations of embedded systems in the solar neighborhood (Lada & Lada 2003; Megeath et al. 2016).

4. Conclusion

We used the dendrogram algorithm to decompose the surface density distributions of stars from the MYStIX project into

hierarchical structures. The algorithm presents multiscale structures of star clusters. Our main conclusions are listed below.

1. Different measurements of embedded clusters at different scales and ages give comparable mass-size relations. The simple explanation is that all embedded clusters possess a Plummer profile.
2. In a unified measurement of the same sample, the mass-size relation has a shallower slope. Instead, mixing different measurements at different scales together always yields a steeper mass-size relation. This issue cannot be avoided when different catalogs from different observations are collected together. The observed radii from different observations must be unified before an analysis.
3. We used the minimum spanning tree (MST) method to measure the separations between clusters and clumps in each massive star-forming region. The separations between cluster-cluster, clump-clump, and cluster-clump are comparable, which indicates that the evolution from a clump to an embedded cluster is fixed at the initial position of the clump. There is no significant positional change between the embedded cluster and its precursor. There is no significant merger between embedded clusters for the samples in the MYStIX project. Feedback from embedded clusters does not redistribute the surrounding clumps. However, more samples are necessary for a further validation.
4. The mass-size relation of the clumps is not comparable with that of the embedded clusters. However, by comparing the mass functions of the ATLASGAL clumps and the identified embedded clusters, we confirmed that a constant star formation efficiency of ≈0.33 can be a typical value for the ATLASGAL clumps.

Acknowledgements. We would like to thank the referee for the detailed comments and suggestions that helped improve and clarify this work.

References

Banerjee, S., & Kroupa, P. 2013, ApJ, 764, 29
Banerjee, S., & Kroupa, P. 2014, ApJ, 787, 158
Banerjee, S., & Kroupa, P. 2015, MNRAS, 447, 728
Banerjee, S., & Kroupa, P. 2017, A&A, 597, A28
Banerjee, S., Kroupa, P., & Oh, S. 2012, ApJ, 746, 15
Bate, M. R., Tricco, T. S., & Price, D. J. 2014, MNRAS, 437, 77
Bonnell, I. A., Vine, S. G., & Bate, M. R. 2004, MNRAS, 349, 735
Brinkmann, N., Banerjee, S., Motwani, B., & Kroupa, P. 2017, A&A, 600, A49
Broos, P. S., Getman, K. V., Povich, M. S., et al. 2013, ApJS, 209, 32
Carpenter, J. M. 2000, AJ, 120, 3139
Carpenter, J. M., Snell, R. L., Schloerb, F. P., & Skrutskie, M. F. 1993, ApJ, 407, 657
Cartwright, A., & Whitworth, A. P. 2004, MNRAS, 348, 589
Dib, S. 2023, MNRAS, 524, 1625
Dib, S., & Henning, T. 2019, A&A, 629, A135
Dib, S., Schmeja, S., & Parker, R. J. 2018, MNRAS, 473, 849
Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61
Faustini, F., Molinari, S., Testi, L., & Brand, J. 2009, A&A, 503, 801
Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, ApJS, 209, 26
Kroupa, P. 1995a, MNRAS, 277, 1491
Kroupa, P. 1995b, MNRAS, 277, 1507
Kroupa, P. 2005, in The Three-Dimensional Universe with Gaia, eds. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, ESA SP, 576, 629
Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2014, ApJ, 787, 107
Kuhn, M. A., Getman, K. V., & Feigelson, E. D. 2015a, ApJ, 802, 60
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2015b, ApJ, 812, 131
Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32
Kumar, M. S. N., Keto, E., & Clerkin, E. 2006, A&A, 449, 1033
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lada, E. A., Depoy, D. L., Evans, N. J., II, & Gatley, I. 1991, ApJ, 371, 171
Larson, R. B. 2003, ASP Conf. Ser., 287, 65
Machida, M. N., & Matsumoto, T. 2012, MNRAS, 421, 588
Marks, M., & Kroupa, P. 2012, A&A, 543, A8
Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, AJ, 151, 5
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Oh, S., & Kroupa, P. 2016, A&A, 590, A107
Oh, S., & Kroupa, P. 2018, MNRAS, 481, 153
Oh, S., Kroupa, P., & Pflamm-Altenburg, J. 2015, ApJ, 805, 92
Pavlik, V., Kroupa, P., & Šubr, L. 2019, A&A, 626, A79
Pfalzner, S., Kirk, H., Sills, A., et al. 2016, A&A, 586, A68
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, ApJ, 679, 1338
Stephens, I. W., Gouliermis, D., Looney, L. W., et al. 2017, ApJ, 834, 94
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
Wang, L., Kroupa, P., & Jerabkova, T. 2019, MNRAS, 484, 1843
Wang, L., Kroupa, P., Takahashi, K., & Jerabkova, T. 2020, MNRAS, 491, 440
Weidner, C., & Kroupa, P. 2006, MNRAS, 365, 1333
Weidner, C., Kroupa, P., & Pflamm-Altenburg, J. 2013, MNRAS, 434, 84
Yan, Z., Jerabkova, T., & Kroupa, P. 2023, A&A, 670, A151
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Dib, S., Wyrowski, F., et al. 2024a, A&A, 682, A173
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2024b, A&A, 682, A128
Zhou, J. W., Kroupa, P., & Wu, W. J. 2024c, MNRAS, submitted
Zhou, J. W., Dib, S., Juvela, M., et al. 2024d, A&A, 686, A146
Zinnecker, H., & Yorke, H. W. 2007, ARA&A, 45, 481

A&A, 691, A204 (2024)
https://doi.org/10.1051/0004-6361/202452064
© The Authors 2024

Astronomy & Astrophysics

Evolution of the mass–radius relation of expanding very young star clusters

Jian-Wen Zhou[1,⋆], Pavel Kroupa[2,3,⋆], and Wenjie Wu[2]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[2] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany
[3] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, 180 00 Praha 8, Czech Republic

Received 30 August 2024 / Accepted 30 September 2024

ABSTRACT

The initial mass–radius relation of embedded star clusters is an essential boundary condition for understanding the evolution of embedded clusters in which stars form to their release into the galactic field via an open star cluster phase. The initial mass–radius relation of embedded clusters deduced by Marks & Kroupa (2012, A&A, 543, A8) is significantly different from the relation suggested by Pfalzner et al. (2016, A&A, 586, A68). Here, we use direct *N*-body simulations to model the early expansion of embedded clusters after the expulsion of their residual gas. The observationally deduced radii of clusters up to a few million years old, compiled from various sources, are well fitted by *N*-body models, implying that these observed very young clusters are most likely in an expanding state. We show that the mass–radius relation of Pfalzner et al. (2016) reflects the expansion of embedded clusters following the initial mass–radius relation of Marks & Kroupa (2012). We also suggest that even the embedded clusters in ATLASGAL clumps with HII regions are probably already in expansion. All the clusters collected here from different observations show a mass–radius relation with a similar slope, which may indicate that all clusters were born with a profile resembling that of the Plummer phase-space distribution function.

Key words. stars: evolution – galaxies: star clusters: general – galaxies: star formation

1. Introduction

The formation of stars within molecular clouds and the early stages of stellar evolution are active research topics. Most, and perhaps even all, of the observed stars were formed in embedded clusters (Kroupa 1995a,b; Lada & Lada 2003; Kroupa 2005; Megeath et al. 2016; Dinnbier et al. 2022). The central hub of a hub-filament system in molecular cloud could serve as the precursor of an embedded cluster. In hub-filament systems, converging flows are funneling matter into the hub through the filaments, such that cores embedded in the hub can prolong the accretion time for growing massive stars due to the sustained supply of matter (Myers 2009; Schneider et al. 2012; Peretto et al. 2013; Longmore et al. 2014; Motte et al. 2018; Vázquez-Semadeni et al. 2019; Kumar et al. 2020; Zhou et al. 2022). Ultimately the feedback energy from the protostars and pre-main sequence stars expels the remaining gas, and the embedded cluster expands. Therefore, measurements of the internal dynamics of young clusters and star-forming regions are necessary to fully understand the process of their formation and dynamical evolution.

Increasing observational evidence suggests that early expansion plays a fundamental role in the dynamical evolution of young star clusters. A useful visualization of the expansion of real star clusters is provided by the collation in Fig. 4 of Brandner (2008). *Gaia* DR2 and DR3 data have opened a new window onto the internal kinematics of young star clusters and large star-forming complexes. Several case studies based on *Gaia* DR2 and DR3 data and multi-object spectroscopy have revealed the expansion of very young clusters in star-forming regions (Wright et al. 2019; Cantat-Gaudin et al. 2019; Kuhn et al. 2020; Lim et al. 2020; Swiggum et al. 2021; Lim et al. 2022; Mužić et al. 2022; Das et al. 2023; Flaccomio et al. 2023; Jadhav et al. 2024). Statistically, in Kuhn et al. (2019), for 28 clusters and associations with ages of about 1–5 Myr, proper motions from *Gaia* DR2 reveal that at least 75% of these systems are expanding. Della Croce et al. (2023) presented a comprehensive kinematic analysis of almost all known young ($t < 300$ Myr) Galactic clusters based on the improved astrometric quality of the *Gaia* DR3 data, finding that a remarkable fraction (up to 80%) of clusters younger than ≈30 Myr are currently experiencing significant expansion. Wright et al. (2023) presented a large-scale 3D kinematic study of ≈2000 spectroscopically confirmed young stars ($t < 20$ Myr) in 18 star clusters and OB associations (groups) from the combination of *Gaia* astrometry and *Gaia*-ESO Survey spectroscopy, showing that nearly all of the groups studied are in the process of expansion.

Bastian & Goodwin (2006) compared the luminosity profiles of young massive clusters, such as M82-F, NGC 1569-A, and NGC 1705-1, with *N*-body simulations of clusters that had experienced rapid gas expulsion, resulting in the loss of a substantial portion of their mass. They found the observed profiles to closely agree with the simulations. Goodwin & Bastian (2006) analyzed the differences between luminosity-based and dynamical mass estimates for young massive stellar clusters. Their results revealed significant discrepancies, which they attributed to the possibility that many young clusters are not in virial equilibrium due to undergoing violent relaxation following gas expulsion. Additionally, these authors noted that the increasing core radii observed in young clusters within the Large Magellanic Cloud and Small Magellanic Cloud can be well explained as an effect

⋆ Corresponding authors; jwzhou@mpifr-bonn.mpg.de; pkroupa@uni-bonn.de

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

of rapid gas loss. In Kroupa et al. (2001) and Banerjee & Kroupa (2013, 2014, 2015), direct N-body modeling of realistic star clusters with a monolithically formed structure and undergoing residual gas expulsion consistently reproduced the characteristics of several well-observed very young star clusters, namely the Orion Nebula Cluster (ONC) and the clusters R136 and NGC 3603, thereby shedding light on the birth conditions of massive star clusters. Similar simulations were also carried out in Banerjee & Kroupa (2017) to explain how young massive star clusters attain their current shapes and sizes from an initially compact embedded morphology. These authors found that young massive clusters cannot expand to their current sizes if only two-body relaxation, stellar mass loss and dynamical heating through initial binaries are considered. They also confirmed that a substantial residual gas expulsion with ≈30% star formation efficiency can adequately swell the newborn embedded clusters. The simulations of Banerjee & Kroupa (2013, 2014, 2015, 2017) start from an initially compact configuration of the computed cluster with the initial mass–radius relation of embedded clusters deduced by Marks & Kroupa (2012):

$$\frac{r_{\rm h}}{\rm pc} = 0.10^{+0.07}_{-0.04} \times \left(\frac{M_{\rm ecl}}{M_\odot}\right)^{0.13\pm0.04}. \quad (1)$$

This relation was derived from the radii and densities of open star clusters and the binding energy distributions of the surviving binary stars in them as these constrain the birth, or the initial half-mass radii, of the precursor embedded star clusters. Based on the universality of the early cluster expansion, in the present work, we collected various cluster samples to characterize the early cluster expansion using the numerical simulation recipes in Kroupa et al. (2001), Banerjee & Kroupa (2013, 2014, 2015) and Dinnbier et al. (2022), with our main goal being to study the evolution of the initial mass–radius relation of embedded clusters.

2. *N*-body simulations

2.1. Basic parameters

The parameters for the simulation were summarized from previous work, i.e. Kroupa et al. (2001); Baumgardt & Kroupa (2007); Banerjee et al. (2012); Banerjee & Kroupa (2013, 2014, 2015); Oh et al. (2015); Oh & Kroupa (2016); Banerjee & Kroupa (2017); Brinkmann et al. (2017); Oh & Kroupa (2018); Wang et al. (2019); Pavlik et al. (2019); Wang et al. (2020b); Dinnbier et al. (2022). The previous simulations have already demonstrated the effectiveness and rationality of these parameter settings (see below). The influence of different parameter settings on simulation results and the multidimensional parameter space are discussed in the aforementioned literature and their references. In the present work, we mainly use the mature simulation recipe to interpret observational data. With this contribution, we are not aiming to provide all possible solutions, as this would entail searching for solutions in a high-dimensional space (cluster mass, age, age spread, star formation efficiency, time of onset of gas expulsion, initial spatial configuration of the embedded cluster – spherical or not, and fractal or not). This is beyond the scope of a reasonable project. Instead, we test whether the initial conditions previously and independently deduced from other data and work provide reasonable physical representations of the stellar populations and their spatial distribution.

We computed five clusters with total stellar masses of 100 $M_\odot$, 300 $M_\odot$, 1000 $M_\odot$, and 3000 $M_\odot$ in the first 4 Myr. The time step is 0.125 Myr. The initial density profile of all clusters is the Plummer profile (Aarseth et al. 1974; Heggie & Hut 2003; Kroupa 2008). The half-mass radius r_h of the cluster is given by the $r_{\rm h} - M_{\rm ecl}$ relation (equation (1)). All clusters are fully mass segregated (S=1), have no fractalization, and are in the state of virial equilibrium (Q=0.5). S and Q are the degree of mass segregation and the virial ratio of the cluster, respectively. More details can be found in Küpper et al. (2011) and the user manual for the `McLuster` code. The initial segregated state is detected for young clusters and star-froming clumps and clouds (Littlefair et al. 2003; Chen et al. 2007; Portegies Zwart et al. 2010; Kirk & Myers 2011; Getman et al. 2014; Lane et al. 2016; Alfaro & Román-Zúñiga 2018; Plunkett et al. 2018; Pavlik et al. 2019; Nony et al. 2021; Zhang et al. 2022; Xu et al. 2024), but the degree of mass segregation is not clear. In simulations of the very young massive clusters R136 and NGC 3603 with gas expulsion by Banerjee & Kroupa (2013), mass segregation does not influence the results. In Fig. A.3, we compare S=1 (fully mass segregated) and S=0.5 (partly mass segregated), and find the results are similar. We also discuss settings with and without fractalization in Fig. A.2; the results of both are also consistent.

Subvirial initial conditions of very young clusters have also been studied (Adams et al. 2006; Proszkow & Adams 2009; Proszkow et al. 2009). An embedded cluster forms from a gas clump on a timescale of about 1 Myr (Kroupa 2005, 2008; Zhou et al. 2024a), such that the subvirial bulk state cannot persist for longer than this while the individual protostar takes about 0.1 Myr to reach most of its mass (Wuchterl & Tscharnuter 2003; Duarte-Cabral et al. 2013). Protostars thus decouple from the hydrodynamical flow and become ballistic particles (Stutz & Gould 2016) that orbit within the evolving overall potential while also having close encounters with other protostars. At the time of greatest compactness, when the embedded cluster has reached the mass–radius relation of embedded clusters (Marks & Kroupa 2012), the embedded cluster can thus be assumed to be close to virial equilibrium, since most of the protostars have orbited a few times within the contracting gas clump. An initial fractal or other subclustering can be set up in models, but such structures may be unrealistic, because the protostars emerge from the gas over the crossing timescale such that there is never a fractal initial state with all stars present because it washes out within the crossing timescale. Using the observational data on NGC 3603 and R 136, Banerjee & Kroupa (2015) and Banerjee & Kroupa (2018) show that any initial subclustering needs to be so compact that it becomes almost equal to monolithically formed clusters. In summary, the virialized state at the time when gas expulsion occurs is a reasonable assumption for the idealized initial conditions applied here.

The IMFs of the clusters are chosen to be canonical (Kroupa 2001) with the most massive star following the $m_{\rm max}$–$M_{\rm ecl}$ relation of Weidner et al. (2013) and Yan et al. (2023). All stars are initially in binaries, that is, $f_{\rm b}$=1 (Kroupa 1995a,b; Belloni et al. 2017), where $f_{\rm b}$ is the primordial binary fraction. We assume the clusters to be at solar metallicity, that is, Z = 0.02 (von Steiger & Zurbuchen 2016). The clusters traverse circular orbits within the Galaxy, positioned at a Galactocentric distance of 8.5 kpc, moving at a speed of 220 km s^{-1}.

The models here assume the canonical Kroupa or Belloni binary fraction of 100% at birth (Kroupa 1995a,b; Belloni et al. 2017). The binary-star distribution functions underlying this assumption are the theoretical values that idealized star

formation tends to if stars form in isolation. In a real embedded cluster, binary systems with a separation of greater than the distances between the protostars will not form and, for example, young protostars in Orion are observed to have a companion fraction of only 44% down to reasonably close separations (≈20 AU, see Tobin et al. 2022). The canonical Kroupa or Belloni distribution functions are however consistent with this because the wide binaries are immediately disrupted even prior to the first integration step in the N-body model (see Fig. 12 in Kroupa 2008). As these wide binaries contain, in their sum, a negligible amount of binding energy compared to that of the embedded cluster, their disruption does not affect the further evolution noticeably. The kinematical cooling of an embedded cluster similar to the ONC through the disruption of the wide binary population was measured for the first time by Kroupa et al. (1999), but the effect is negligible concerning the dynamical evolution of the embedded cluster. The disruption of the birth binary population to a distribution of binary systems that is in equilibrium with the embedded cluster takes about one cluster crossing time (<1 Myr). For the birth density of the ONC, $\approx 10^5\ M_\odot/\mathrm{pc}^3$, the binary fraction is about 40 per cent (see Fig.10 in Marks et al. 2011), which is in good agreement with the above-mentioned observational result of Tobin et al. (2022). In any case, in Fig. A.4, we compare f_b=1 and f_b=0.44, and find similar results.

Following previous simulations, we consider the essential dynamical effects of the gas-expulsion process by applying a diluting, spherically symmetric external gravitational potential to a model cluster, as in Kroupa et al. (2001). Specifically, we use the potential of the spherically symmetric, time-varying mass distribution:

$$\begin{aligned} M_g(t) &= M_g(0) \qquad t \le \tau_d, \\ M_g(t) &= M_g(0)\exp\left(-\tfrac{(t-\tau_d)}{\tau_g}\right) t > \tau_d. \end{aligned} \tag{2}$$

Here, $M_g(t)$ is the total mass in gas; this gas is spatially distributed with the Plummer density distribution (Kroupa 2008) and starts depleting after a delay of τ_d, and is totally depleted within a timescale of τ_g. The Plummer radius of the gas distribution is kept time-invariant (Kroupa et al. 2001); this assumption is an approximate model of the effective gas reduction within the cluster in the situation where gas is blown out while new gas is also accreting onto the cluster along filaments such that the gas mass reduces over time while the radius of the gas distribution remains roughly constant. As discussed in Urquhart et al. (2022), the mass and radius distributions of the ATLASGAL clumps at different evolutionary stages are quite comparable. We use an average velocity of the outflowing gas of $v_g \approx 10$ km s^{-1}, which is the typical speed of sound in an HII region. This gives $\tau_g = r_h(0)/v_g$, where $r_h(0)$ is the initial half-mass radius of the cluster. We note that comparable outflows are detected in the about 0.1 Myr-old Treasure Chest cluster (Smith et al. 2005) and in the few-million-years-old star-burst clusters in the Antannae galaxies (Whitmore et al. 1999; Zhang et al. 2001). Regarding the delay-time, we take the representative value of $\tau_d \approx 0.6$ Myr (Kroupa et al. 2001; this being about the lifetime of the ultracompact HII region), and assume a star formation efficiency (SFE) ≈0.33, which means $M_g(0) = 2M_\mathrm{ecl}(0)$.

Embedded clusters form in clumps. More massive clumps can produce more massive clusters, leading to stronger feedback and larger gas expulsion velocity (Dib et al. 2013). There should be correlations between the feedback strength, the clump (or cluster) mass, and the gas expulsion velocity (v_g). Low-mass clusters should have a slower gas-expulsion process compared with high-mass clusters. As shown in Pang et al. (2021), low-mass clusters tend to agree with the simulations of slow gas expulsion. This work primarily focuses on low-mass star clusters, whose masses are far below that of the young massive clusters studied in Bastian & Goodwin (2006), Goodwin & Bastian (2006) and Banerjee & Kroupa (2013, 2014, 2017). Therefore, a gas-expulsion velocity of $v_g \approx 10$ km s^{-1} is an upper limit. Yang et al. (2022) conducted the largest outflow survey so far toward ATLASGAL clumps, which mainly produced low-mass embedded clusters (Zhou et al. 2024c,a). An upper limit on the outflow velocity of ≈10 km s^{-1} is clearly shown in their Fig. 5.

In Fig. A.5, we compare SFE = 0.33 and SFE = 0.2. When SFE = 0.2, $M_g(0) = 4M_\mathrm{ecl}(0)$, which is twice that at SFE = 0.33. Hence, we also doubled the gas expulsion timescale τ_g. Moreover, a large amount of gas can also affect the physical state of the cluster, causing it not to be in virial equilibrium but instead to collapse due to excessive gravitational binding. Therefore, we set Q=0.2 for SFE = 0.2. In Fig. A.5, clusters with SFE = 0.2 expand faster than with SFE = 0.33. A smaller SFE implies a greater amount of gas, which affects the gas expulsion process, leading to differences in the results. Further discussion of the details is not the focus of this work. In any case, considering the measurement errors in cluster ages, simulations with SFE = 0.2 can still roughly explain the observations, as shown in the following text. In this work, we assume a star formation efficiency (SFE) of 0.33. This value has been widely used in the previous simulations listed above and has been proven to be effective. Such a SFE is also consistent with those obtained from hydrodynamic calculations including self-regulation (Machida & Matsumoto 2012; Bate et al. 2014) and with observations of embedded systems in the solar neighborhood (Lada & Lada 2003; Gutermuth et al. 2009; Megeath et al. 2016). In Zhou et al. (2024c), by comparing the mass functions of the ATLASGAL clumps and the embedded clusters, we found that a SFE of ≈ 0.33 could be a typical value for the ATLASGAL clumps. In Zhou et al. (2024b), assuming SFE = 0.33, it was shown that the total bolometric luminosity of synthetic embedded clusters can well fit the observed bolometric luminosity of ATLASGAL clumps with HII regions. In Zhou et al. (2024a), we directly calculated the SFE of ATLASGAL clumps with HII regions, and found a median value of ≈0.3.

2.2. Procedure

The `McLuster` program (Küpper et al. 2011) was used to set the initial configurations. The dynamical evolution of the model clusters was computed using the state-of-the-art `PeTar` code (Wang et al. 2020a). `PeTar` employs the well-tested analytical single and binary stellar evolution recipes (SSE and BSE) (Hurley et al. 2000, 2002; Giacobbo et al. 2018; Tanikawa et al. 2020; Banerjee et al. 2020).

2.3. Evolution of cluster radii

The computed model parameters are listed in Table 1. We studied the variations of the half-mass radius r_h (the radius within which 50% of the mass in stars is found) and the full radius r_f (the radius within which 90% of the mass in stars is found) of the cluster in the first 4 Myr. The $r_\mathrm{h} - M_\mathrm{ecl}$ relation (equation (1)) predicts the initial half-mass radius of the cluster at t = 0 Myr. In Fig. 1, we fitted the $r_\mathrm{h} - M_\mathrm{ecl}$ relations at different time nodes, namely t = 1 Myr, 2 Myr, 3 Myr, and 4 Myr. In the subsequent analysis, these are referred to as the cluster's 1 Myr,

Table 1. Computed model parameters.

Cluster	$M_{\rm ecl}(M_\odot)$	$r_{\rm h}(pc)$	$m_{\rm max}(M_\odot)$	$\tau_{\rm g}$(Myr)
1	30	0.16	5	0.012
2	100	0.18	10	0.014
3	300	0.21	17	0.016
4	1000	0.25	30	0.019
5	3000	0.28	43	0.022

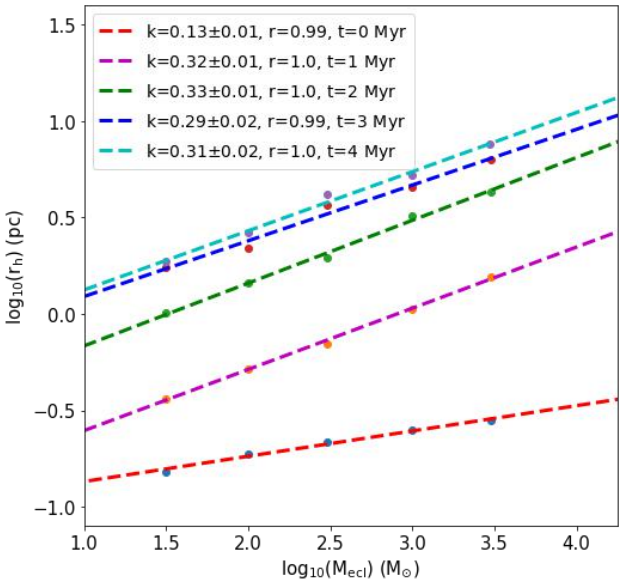


Fig. 1. Fits to the expanding cluster in simulations by least-squares. The dots are from different time nodes; i.e., t = 0 Myr, 1 Myr, 2 Myr, 3 Myr, and 4 Myr. Here, "k" is the slope of the linear fit with the unit pc/$M_\odot$, and "r" represents the Pearson correlation coefficient.

2 Myr, 3 Myr, and 4 Myr expanding lines. In the first 4 Myr, the cluster mass in stars does not change significantly, and we therefore assume stellar mass conservation, although the stars are allowed to evolve. To determine $r_{\rm h}(t)$ at a given mass of the embedded cluster in stars, $M_{\rm ecl}$, the distance from the star to the cluster center is recorded, within which half of the cluster's mass is contained. At time t, the data $r_{\rm h}(M_{\rm ecl};t)$ are fitted by a linear relation using least-squares (Fig. 1).

3. Cluster radii in observations

In observations, different methods are used to estimate the size of a cluster, and the measured radius can be the half-mass radius, $r_{\rm h}$, or the full radius, $r_{\rm f}$. We can trace the evolution of $r_{\rm h}$ and $r_{\rm f}$ at the same time in simulations, which can be the benchmarks to classify the observed radii of very young star clusters (<4 Myr).

3.1. Half-mass radius

For the clusters in Lada et al. (1991), they determined cluster boundaries using the magnitude limit of $m_{\rm K} < 13$, and also required the lowest contour level to correspond to a surface density of two times the background density. In Carpenter et al. (1993), the magnitude limit is $m_{\rm H} = 15.5$. For the samples in these two papers, the boundaries of the clusters based on the identification criteria are probably not representative of the total extent of the cluster, and therefore the measured radius is likely to be smaller than the full radius $r_{\rm f}$ of the cluster. As shown in Fig. 2a, the samples can be well fitted by the 1 Myr expanding line. Considering that the ages of these clusters are indeed around 1 Myr (Weidner et al. 2013), the simulations well explain the observed radii, and the measured radii in these observations are close to the half-mass radii of the clusters.

For the embedded cluster catalog in Lada & Lada (2003), as shown in Fig. 2b, the samples move systematically upwards, indicating older ages or larger radius measurements. The catalog gives the absolute magnitude limits $m_{\rm K}$ of the corresponding imaging observations. Figure 2c shows the distribution of the magnitude limits. If we take $m_{\rm K} < 15.5$ and $m_{\rm K} > 15.5$ as high and low standards of the magnitude limits, the median radii of the clusters in the two groups are ≈0.61 pc and 0.76 pc, respectively. As expected, low standards indeed give slightly larger radius estimates due to the larger star counts. According to the selection criterion in Lada & Lada (2003), most of the clusters should have an age of around 1 Myr. We note that part of the samples of Carpenter et al. (1993) and Lada et al. (1991) are also found in the catalog of Lada & Lada (2003). Except for those samples with significant deviations, the almost systematic upshifting of most of the sample in Lada & Lada (2003) in Fig. 2b may be attributed to the low standards, which cause slightly larger radius estimates. However, given that accurate age estimation of embedded clusters is always very challenging in observations, the differences are just as likely to be caused by the different ages of the cluster samples.

3.2. Full radius

In some observations, the cluster radius was measured by: 1. The surface density profile decreases until it reaches a constant value (Faustini et al. 2009); 2. The density profile exceeds twice the standard deviation of the surface density in the surrounding field (Carpenter 2000; Kumar et al. 2006). It seems these observations measured the full cluster radius. In Fig. 3, we used the $r_{\rm h} - M_{\rm ecl}$ and $r_{\rm f} - M_{\rm ecl}$ relations to compare with the observations. The two relations give the age estimates of ≈2 Myr and ≈1 Myr, respectively. For the samples in Kumar et al. (2006), they are the youngest stellar clusters associated with massive protostellar candidates, thus, their ages should be ≤1 Myr. Therefore, the $r_{\rm f}$–$M_{\rm ecl}$ relation gives a better fit than the $r_{\rm h}$–$M_{\rm ecl}$ relation, which means the above observations roughly measured the full radii of the samples. The ages of the clusters in Carpenter (2000) and Faustini et al. (2009) are <2 Myr and ≤1 Myr, respectively. All of them are well fitted in Fig. 3.

3.3. Recent observations

Readers may consider the observational sample presented above to be dated and therefore question its reliability. Below, we further validate the above results using more recent observations.

The MYStIX (Massive Young Star-Forming Complex Study in Infrared and X-ray) project (Feigelson et al. 2013) examined 20 nearby (d<3.6 kpc) massive star-forming regions (MSFRs) using a combination of archival *Chandra* X-ray imaging, *2MASS+UKIDSS* near-infrared (NIR), and *Spitzer* mid-infrared (MIR) survey data. The catalog of 31 784 young stars from the project includes both high-mass and low-mass stars, and both disk-bearing and disk-free stars (Broos et al. 2013). Although the MYStIX samples are not "complete", the identification of young stars was performed in a uniform way for the different regions, enabling a comparative analysis of the stellar populations within

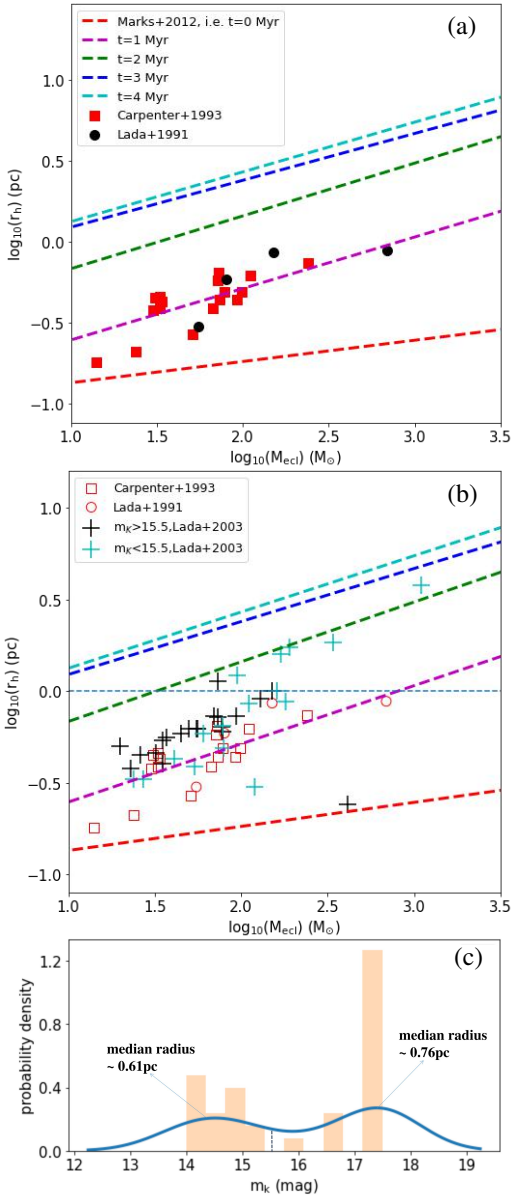


Fig. 2. Fitting the mass–radius relation of embedded clusters in observations using N-body simulations. (a) Dashed lines are the $r_{\rm h} - M_{\rm ecl}$ relations at different time nodes (t = 1 Myr, 2 Myr, 3 Myr and 4 Myr), i.e. the 1 Myr, 2 Myr, 3 Myr and 4 Myr expanding lines. Red squares and black dots represent the clusters from Carpenter et al. (1993) and Lada et al. (1991), respectively. (b) Plus symbols show the clusters from Lada & Lada (2003), while the black and red plus symbols show the clusters selected by the magnitude limits $m_{\rm K}$<15.5 and $m_{\rm K}$>15.5, respectively. The horizontal line marks the position of the radius r = 1 pc. (c) Distribution of the magnitude limits $m_{\rm K}$ of the clusters in Lada & Lada (2003).

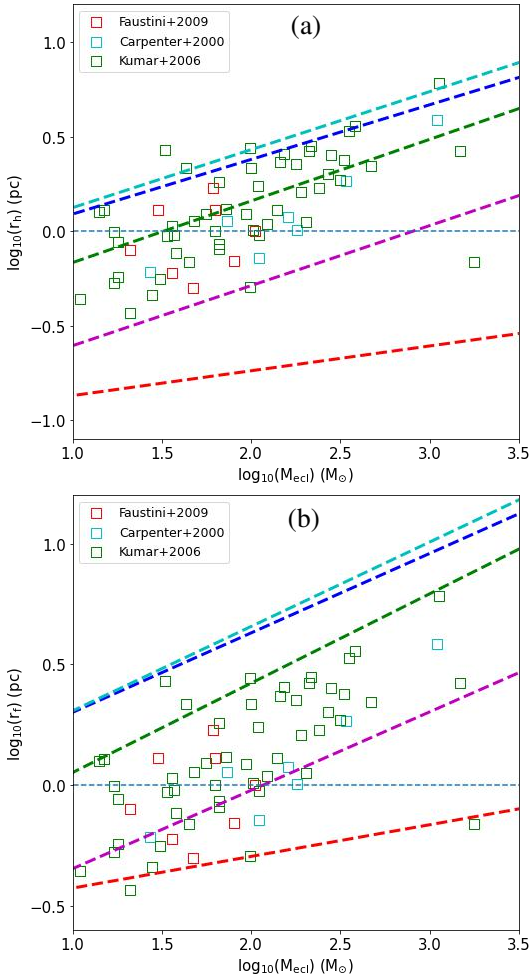


Fig. 3. Fitting the mass–radius relation of embedded clusters in observations using N-body simulations. Colored squares represent the clusters from Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009). The horizontal line marks the position of the radius r = 1 pc. The dashed lines are color-coded according to age as in Figs. 1 and 2.

these diverse regions. Across 17 massive star-forming regions in the MYStIX project, Kuhn et al. (2014) identified 142 subclusters (their "subclusters" being synonymous to the embedded clusters here) of young stars using finite mixture models. Their physical parameters were listed in Kuhn et al. (2015), such as the radius, the age, and the total number of stars. Getman et al. (2014) provided age estimates for over 100 subclusters using the novel Age_{JX} method, which encompasses a wide age range between 0.5 and 5 Myr, showing that individual regions often present spatial segregation.

The radii of the subclusters in Table 1 of Kuhn et al. (2015) are four times the core radius based on their model and are considered by these authors to be roughly the projected half-mass radii. As shown in Fig. 4a, the $r_{\rm h}$–$M_{\rm ecl}$ relations can well fit the observations. However, the ages assigned to individual young stars may exhibit considerable uncertainty due to statistical errors in luminosities, uncertainties in dereddening, and the inherent variability in the X-ray-luminosity–mass relation. Additionally, different evolutionary models might introduce systematic shifts in ages (Getman et al. 2014; Kuhn et al. 2015). To further complicate this discussion, the identification of subclusters in Kuhn et al. (2014) is not based on the Plummer profile as used in the present simulations. Furthermore, the table only provides the total number of stars in each subcluster, and to derive mass we assume an average stellar mass of 0.5 $M_\odot$. This may lead to underestimation of the mass of the subcluster. Using the samples from Lada & Lada (2003), in Fig. A.1, the actual correlation between the total number of stars and the mass of an embedded cluster systematically shifts upward compared to the 0.5 $M_\odot$ assumption. Despite many uncertainties, the good fittings in Fig. 4a remain impressive. Moreover, the expansion of subclusters has been confirmed in Kuhn et al. (2015) and Kuhn et al. (2019) from observations.

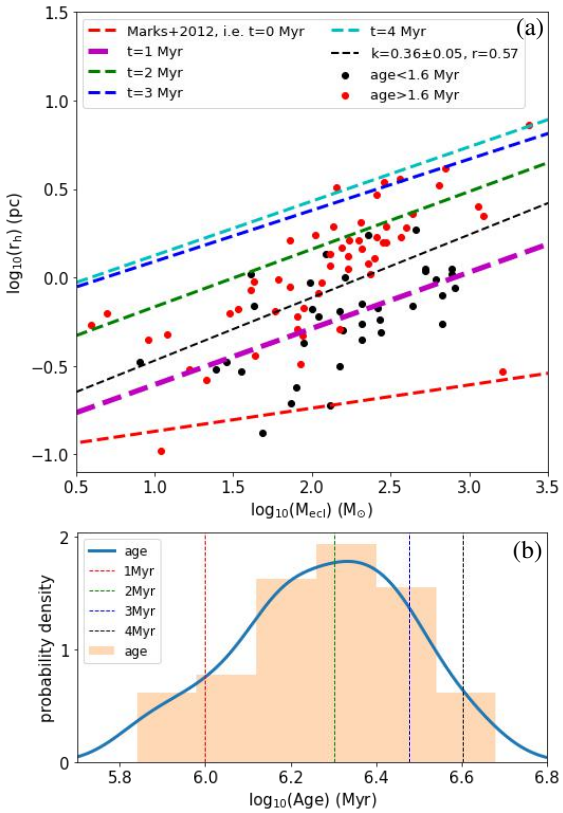


Fig. 4. Fitting the mass–radius relation of embedded clusters in observations using N-body simulations. (a) Dashed lines are the r_h–M_{ecl} relations at different time nodes. Red and black dots represent the subclusters from Kuhn et al. (2014), which have ages of >1.6 Myr and <1.6 Myr, respectively. The dashed black line represents the fitted mass–radius relation. Here, "k" is the slope of the linear fitting with units of pc/M$_\odot$, and "r" is the Pearson correlation coefficient. (b) Age distribution of the subclusters in Table 1 of Kuhn et al. (2015).

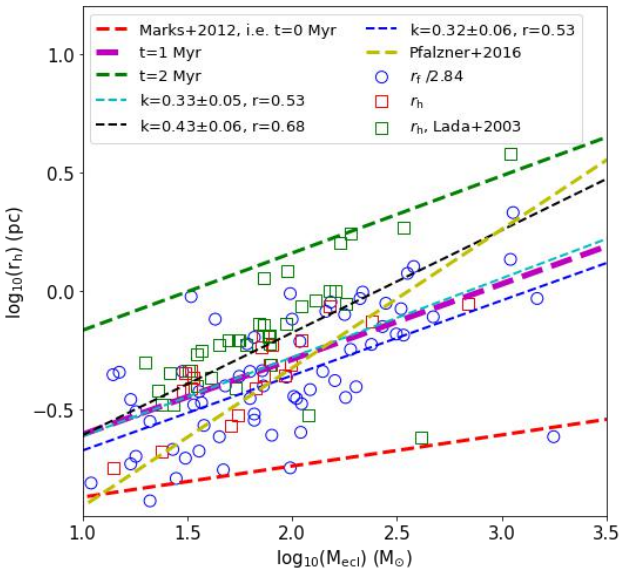


Fig. 5. Fitting the mass–radius relation of embedded clusters in observations using N-body simulations. For the samples in Fig. 3, their radii were divided by 2.84, and then merged with the samples in Fig. 2. The radii of all samples were unified to the half-mass radius. The yellow dashed line is the mass–radius relation of Pfalzner et al. (2016). Dashed cyan, black, and blue lines represent the mass–radius relations of all samples, the samples in Lada et al. (1991); Carpenter et al. (1993); Lada & Lada (2003) and the samples in Carpenter (2000); Kumar et al. (2006); Faustini et al. (2009), respectively. Here, "k" is the slope of the fitting with the unit pc/M$_\odot$, and "r" represents the Pearson correlation coefficient.

4. Mass–radius relation

Assuming all embedded clusters have the Plummer profile with the Plummer radius r_{pl}, the mass within radius r is

$$M(r) = M_{ecl}\frac{\left(\frac{r}{r_{pl}}\right)^3}{\left[1+\left(\frac{r}{r_{pl}}\right)^2\right]^{\frac{3}{2}}}. \quad (3)$$

The full radius (0.9 mass fraction) and the half-mass radius (0.5 mass fraction) satisfy $r_f \approx 3.707 r_{pl}$ and $r_h \approx 1.305 r_{pl}$, and therefore $r_f \approx 2.84 r_h$ (Kroupa 2008). For the samples discussed above, we can unify their radii to the half-mass radii. The radii of the samples in Lada et al. (1991), Carpenter et al. (1993) and Lada & Lada (2003) can be approximated to the half-mass radii. For the samples in Carpenter (2000), Kumar et al. (2006) and Faustini et al. (2009), their radii need to be divided by 2.84. As shown in Fig. 5, different samples can indeed merge together after the unification. We fitted the mass–radius relation for all samples, which gives a slope of ≈0.33 (dashed cyan line in Fig. 5). If we fit the samples in Lada et al. (1991); Carpenter et al. (1993); Lada & Lada (2003) and Carpenter (2000); Kumar et al. (2006); Faustini et al. (2009) separately, the slopes are ≈0.43 (dashed black line in Fig. 5) and ≈0.32 (dashed blue line in Fig. 5), respectively. Only using the samples in Kuhn et al. (2015) gives a slope of ≈0.36 in Fig. 4a.

The equation (4) of Pfalzner et al. (2016) has a steeper slope of ≈0.58 (dashed yellow line in Fig. 5). Pfalzner et al. (2016) also found that all collected data in the K-band show a mass–radius relation with a similar slope. These authors therefore argued that clusters might "grow" from low-mass clusters to their final mass without changing their size, and their size would be determined by the extent of the clump. However, as shown in this work, the unique mass–radius relation of Pfalzner et al. (2016) actually reflects the expansion of embedded clusters. There is no self-similar mass–radius development during the evolution of embedded clusters, because the expanding lines in Fig. 2 are not parallel at the beginning. The real initial mass–radius relation of embedded clusters (equation (1)) is significantly different from the relation of Pfalzner et al. (2016). Any observable cluster is already in a state of expansion to varying degrees. Thus, the current mass–radius relation of clusters cannot reveal their initial formation conditions. In Pfalzner et al. (2016), there is no cluster containing a thousand stars and with a half mass radius of <0.4 pc in their samples because all the samples are expanding and thus detach from the initial compact configuration predicted by the initial mass–radius relation of embedded clusters (equation (1)). As confirmed in the following section, even the embedded clusters in ATLASGAL clumps with HII regions (HII-clumps) are already in expansion.

In Pfalzner et al. (2016), all collected samples show a mass–radius relation with a similar slope, also shown in Figs. 2, 3, 4a and 5 in this work, although different methods were used to determine the radius and mass in different observations. As explained in Pfalzner et al. (2016), assuming a unique mass–radius relation and a self-similar mass–radius development throughout, the different mass and radius measurements simply move the data points along the line of the mass–radius

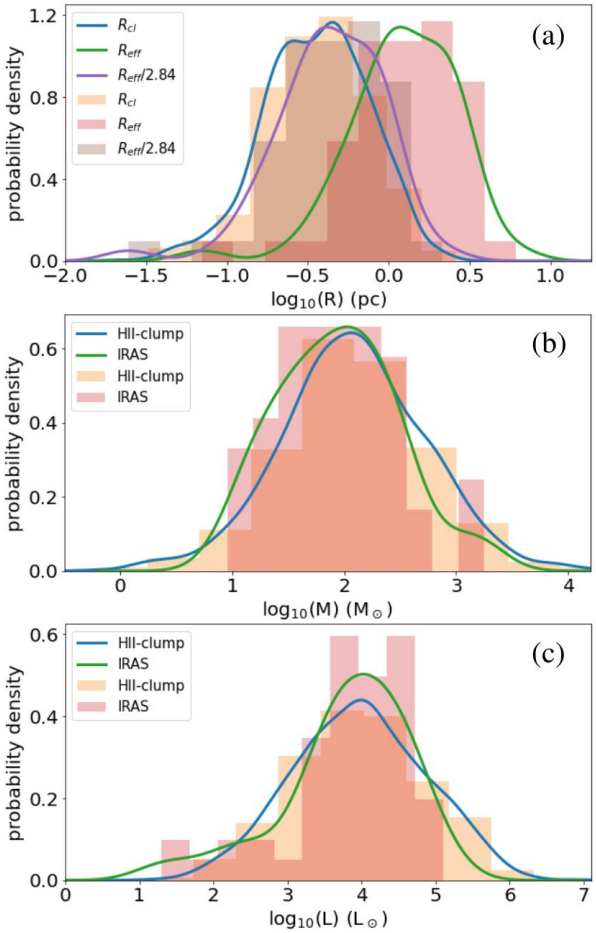


Fig. 6. Comparison of the radius, mass, and bolometric luminosity of embedded clusters in HII clumps of Urquhart et al. (2022) and the clusters (IRAS sources) in Kumar et al. (2006).

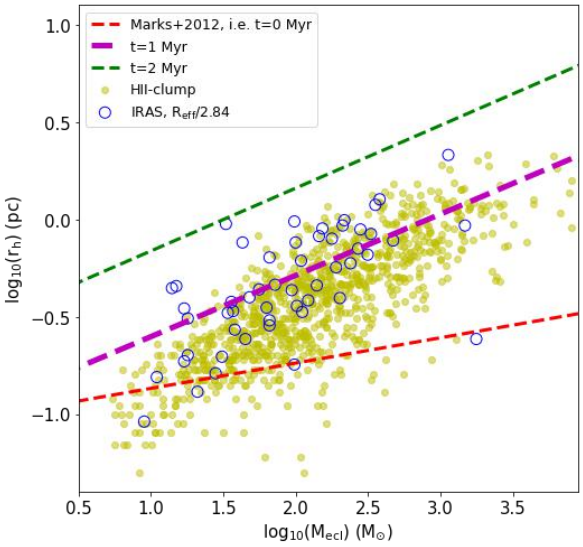


Fig. 7. Fitting the masses and sizes of embedded clusters in HII clumps (yellow dots) using the expanding lines in Figs. 1 and 2. Blue circles represent the clusters in Kumar et al. (2006).

relation. In the discussion above, we have excluded this interpretation. Another simple explanation is that all clusters initially have the Plummer profile and then expand through the expulsion of residual gas, as modeled in Sect. 2.1. As done above, different radii in different observations can be unified to the half-mass radii. This operation will not change the slope of the mass–radius relation, which only shifts the line of the mass–radius relation, as shown in Pfalzner et al. (2016) and this work, also see the discussion in Zhou et al. (2024c).

5. Expansion of embedded clusters in HII-clumps

The sample in Kumar et al. (2006) consists of the youngest stellar clusters associated with massive protostellar candidates (IRAS sources). As such, they are comparable to embedded clusters in HII-clumps, as shown in Figs. 6 and 7. The cluster radius in Kumar et al. (2006) is the effective radius $R_{\rm eff}$, which is approximated to the full radius of the cluster, as described in Sect. 3.2. The radii of HII-clumps, $R_{\rm cl}$, in Urquhart et al. (2022) are determined from the number of pixels within the FWHM contour, that is, those above 50 per cent of the peak of the ATLASGAL dust continuum emission. As shown in Fig. 6a, $R_{\rm eff}$ is significantly larger than $R_{\rm cl}$. However, $R_{\rm eff}/2.84$ is comparable with $R_{\rm cl}$. $R_{\rm eff}/2.84$ is approximated to the half-mass radius of the cluster in Kumar et al. (2006), as discussed in Sect. 4. Therefore, $R_{\rm cl}$ measured in Urquhart et al. (2022) is close to the half-mass radius of the embedded cluster in a HII-clump. A HII-clump mass $M_{\rm cl}$ and its embedded star cluster mass $M_{\rm ecl}$ satisfy $M_{\rm ecl} = {\rm SFE} * M_{\rm cl}$, where SFE is the final star formation efficiency of the clump. As discussed in Sect. 2.1, we take SFE = 0.33. In Fig. 7, the masses and radii of the embedded clusters in HII-clumps can be well fitted on the $r_{\rm h}$–$M_{\rm ecl}$ plane. Generally, the samples are between the 0 and 1 Myr expanding lines, which is consistent with their ages ($\leq$1 Myr). However, some samples are located under the initial 0 Myr line. This may be due to the fact that these clumps reach a higher gas density during gas collapse than implied by the $r_{\rm h}$–$M_{\rm ecl}$ relation for embedded clusters, or it could be that the measurement of their radii is biased or affected by resolution. As shown in Urquhart et al. (2022), evolutionary sequences still exist within HII-clumps. For the samples under the initial 0 Myr line, they may be so young that the inner cluster formation is not yet complete.

6. Conclusion

We used standard N-body simulations as in previous works. The comparison between the previously deduced initial conditions of the simulations with the data used here has not been performed up to date, and this work constitutes a first step in an attempt to understand the data in terms of a realistic physical model that was developed previously on different data of young clusters and stellar populations.

In this work, we traced the evolution of the $r_{\rm h}$–$M_{\rm ecl}$ and $r_{\rm f}$–$M_{\rm ecl}$ relations in the first 4 Myr using direct N-body simulations. We used the expanding lines at different time nodes as benchmarks to classify the observed radii of very young star clusters. The good fit of the expanding lines to the samples compellingly suggests that the observed very young clusters are in expansion. After unifying the radii of all samples to the half-mass radii, we fitted the mass–radius relations for different categories, finding that they are comparable with the mass–radius relation documented by Pfalzner et al. (2016). Thus, the mass–radius relation of Pfalzner et al. (2016) actually reflects the expansion of embedded clusters starting with the initial mass–radius relation of Marks & Kroupa (2012). We also suggest that even the embedded clusters in ATLASGAL clumps with HII regions

(HII-clumps) are already in expansion. The collected samples all show a mass–radius relation with a similar slope, which may indicate that embedded clusters are born with a profile resembling that of the Plummer phase-space density distribution function (e.g., Aarseth et al. 1974; Heggie & Hut 2003). This is interesting, because the Plummer phase space density distribution function is the simplest analytical (two-parameter) distribution function that solves the time-independent collisionless Boltzmann equation (e.g., Kroupa 2008), and because the molecular cloud filaments in which stars have been found to have Plummer-like cross sections (André et al. 2022), as well as both open (Röser et al. 2011; Röser & Schilbach 2019) and globular star clusters (Plummer 1911) are also well described by Plummer models.

References

Aarseth, S. J., Henon, M., & Wielen, R. 1974, A&A, 37, 183
Adams, F. C., Proszkow, E. M., Fatuzzo, M., & Myers, P. C. 2006, ApJ, 641, 504
Alfaro, E. J., & Román-Zúñiga, C. G. 2018, MNRAS, 478, L110
André, P. J., Palmeirim, P., & Arzoumanian, D. 2022, A&A, 667, L1
Banerjee, S., & Kroupa, P. 2013, ApJ, 764, 29
Banerjee, S., & Kroupa, P. 2014, ApJ, 787, 158
Banerjee, S., & Kroupa, P. 2015, MNRAS, 447, 728
Banerjee, S., & Kroupa, P. 2017, A&A, 597, A28
Banerjee, S., & Kroupa, P. 2018, in Astrophysics and Space Science Library, 424, The Birth of Star Clusters, ed. S. Stahler, 143
Banerjee, S., Kroupa, P., & Oh, S. 2012, ApJ, 746, 15
Banerjee, S., Belczynski, K., Fryer, C. L., et al. 2020, A&A, 639, A41
Bastian, N., & Goodwin, S. P. 2006, MNRAS, 369, L9
Bate, M. R., Tricco, T. S., & Price, D. J. 2014, MNRAS, 437, 77
Baumgardt, H., & Kroupa, P. 2007, MNRAS, 380, 1589
Belloni, D., Askar, A., Giersz, M., Kroupa, P., & Rocha-Pinto, H. J. 2017, MNRAS, 471, 2812
Brandner, W. 2008, in Astronomical Society of the Pacific Conference Series, 387, Massive Star Formation: Observations Confront Theory, eds. H. Beuther, H. Linz, & T. Henning, 369
Brinkmann, N., Banerjee, S., Motwani, B., & Kroupa, P. 2017, A&A, 600, A49
Broos, P. S., Getman, K. V., Povich, M. S., et al. 2013, ApJS, 209, 32
Cantat-Gaudin, T., Jordi, C., Wright, N. J., et al. 2019, A&A, 626, A17
Carpenter, J. M. 2000, AJ, 120, 3139
Carpenter, J. M., Snell, R. L., Schloerb, F. P., & Skrutskie, M. F. 1993, ApJ, 407, 657
Chen, L., de Grijs, R., & Zhao, J. L. 2007, AJ, 134, 1368
Das, S. R., Gupta, S., Prakash, P., Samal, M., & Jose, J. 2023, ApJ, 948, 7
Della Croce, A., Dalessandro, E., Livernois, A. R., & Vesperini, E. 2024, A&A, 683, A10
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61
Duarte-Cabral, A., Bontemps, S., Motte, F., et al. 2013, A&A, 558, A125
Faustini, F., Molinari, S., Testi, L., & Brand, J. 2009, A&A, 503, 801
Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, ApJS, 209, 26
Flaccomio, E., Micela, G., Peres, G., et al. 2023, A&A, 670, A37
Getman, K. V., Feigelson, E. D., Kuhn, M. A., et al. 2014, ApJ, 787, 108
Giacobbo, N., Mapelli, M., & Spera, M. 2018, MNRAS, 474, 2959
Goodwin, S. P., & Bastian, N. 2006, MNRAS, 373, 752
Goodwin, S. P., & Whitworth, A. P. 2004, A&A, 413, 929
Gutermuth, R. A., Megeath, S. T., Myers, P. C., et al. 2009, ApJS, 184, 18
Heggie, D., & Hut, P. 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics (Cambridge University Press)
Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, MNRAS, 315, 543
Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, MNRAS, 329, 897
Jadhav, V. V., Kroupa, P., Wu, W., Pflamm-Altenburg, J., & Thies, I. 2024, A&A, 687, A89
Kirk, H., & Myers, P. C. 2011, ApJ, 727, 64
Kroupa, P. 1995a, MNRAS, 277, 1491
Kroupa, P. 1995b, MNRAS, 277, 1507
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2005, in ESA Special Publication, Vol. 576, The Three-Dimensional Universe with Gaia, eds. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 629
Kroupa, P. 2008, in The Cambridge *N*-Body Lectures, 760, ed. S. J. Aarseth, C. A. Tout, & R. A. Mardling, 181
Kroupa, P., Petr, M. G., & McCaughrean, M. J. 1999, New A, 4, 495
Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2014, ApJ, 787, 107
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2015, ApJ, 812, 131
Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32
Kuhn, M. A., Hillenbrand, L. A., Carpenter, J. M., & Avelar Menendez, A. R. 2020, ApJ, 899, 128
Kumar, M. S. N., Keto, E., & Clerkin, E. 2006, A&A, 449, 1033
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87
Küpper, A. H. W., Maschberger, T., Kroupa, P., & Baumgardt, H. 2011, MNRAS, 417, 2300
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lada, E. A., Depoy, D. L., Evans, Neal J., I., & Gatley, I. 1991, ApJ, 371, 171
Lane, J., Kirk, H., Johnstone, D., et al. 2016, ApJ, 833, 44
Lim, B., Hong, J., Yun, H.-S., et al. 2020, ApJ, 899, 121
Lim, B., Nazé, Y., Hong, J., et al. 2022, AJ, 163, 266
Littlefair, S. P., Naylor, T., Jeffries, R. D., Devey, C. R., & Vine, S. 2003, MNRAS, 345, 1205
Longmore, S. N., Kruijssen, J. M. D., Bastian, N., et al. 2014, in Protostars and Planets VI, eds. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 291
Machida, M. N., & Matsumoto, T. 2012, MNRAS, 421, 588
Marks, M., & Kroupa, P. 2012, A&A, 543, A8
Marks, M., Kroupa, P., & Oh, S. 2011, MNRAS, 417, 1684
Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, AJ, 151, 5
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Mužić, K., Almendros-Abad, V., Bouy, H., et al. 2022, A&A, 668, A19
Myers, P. C. 2009, ApJ, 700, 1609
Nony, T., Robitaille, J. F., Motte, F., et al. 2021, A&A, 645, A94
Oh, S., & Kroupa, P. 2016, A&A, 590, A107
Oh, S., & Kroupa, P. 2018, MNRAS, 481, 153
Oh, S., Kroupa, P., & Pflamm-Altenburg, J. 2015, ApJ, 805, 92
Pang, X., Li, Y., Yu, Z., et al. 2021, ApJ, 912, 162
Pavlik, V., Kroupa, P., & Šubr, L. 2019, A&A, 626, A79
Peretto, N., Fuller, G. A., Duarte-Cabral, A., et al. 2013, A&A, 555, A112
Pfalzner, S., Kirk, H., Sills, A., et al. 2016, A&A, 586, A68
Plummer, H. C. 1911, MNRAS, 71, 460
Plunkett, A. L., Fernández-López, M., Arce, H. G., et al. 2018, A&A, 615, A9
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Proszkow, E.-M., & Adams, F. C. 2009, ApJS, 185, 486
Proszkow, E.-M., Adams, F. C., Hartmann, L. W., & Tobin, J. J. 2009, ApJ, 697, 1020
Röser, S., & Schilbach, E. 2019, A&A, 627, A4
Röser, S., Schilbach, E., Piskunov, A. E., Kharchenko, N. V., & Scholz, R. D. 2011, A&A, 531, A92
Schneider, N., Csengeri, T., Hennemann, M., et al. 2012, A&A, 540, L11
Smith, N., Stassun, K. G., & Bally, J. 2005, AJ, 129, 888
Stutz, A. M., & Gould, A. 2016, A&A, 590, A2
Swiggum, C., D'Onghia, E., Alves, J., et al. 2021, ApJ, 917, 21
Tanikawa, A., Yoshida, T., Kinugawa, T., Takahashi, K., & Umeda, H. 2020, MNRAS, 495, 4170
Tobin, J. J., Offner, S. S. R., Kratter, K. M., et al. 2022, ApJ, 925, 39
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061
von Steiger, R., & Zurbuchen, T. H. 2016, ApJ, 816, 13
Wang, L., Kroupa, P., & Jerabkova, T. 2019, MNRAS, 484, 1843
Wang, L., Iwasawa, M., Nitadori, K., & Makino, J. 2020a, MNRAS, 497, 536
Wang, L., Kroupa, P., Takahashi, K., & Jerabkova, T. 2020b, MNRAS, 491, 440
Weidner, C., Kroupa, P., & Pflamm-Altenburg, J. 2013, MNRAS, 434, 84
Whitmore, B. C., Zhang, Q., Leitherer, C., et al. 1999, AJ, 118, 1551
Wright, N. J., Jeffries, R. D., Jackson, R. J., et al. 2019, MNRAS, 486, 2477
Wright, N. J., Jeffries, R. D., Jackson, R. J., et al. 2023, arXiv e-prints [arXiv:2311.08358]
Wuchterl, G., & Tscharnuter, W. M. 2003, A&A, 398, 1081
Xu, F., Wang, K., Liu, T., et al. 2024, ApJS, 270, 9
Yan, Z., Jerabkova, T., & Kroupa, P. 2023, A&A, 670, A151
Yang, W. J., Menten, K. M., Yang, A. Y., et al. 2022, A&A, 658, A192
Zhang, Q., Fall, S. M., & Whitmore, B. C. 2001, ApJ, 561, 727
Zhang, Y., Tanaka, K. E. I., Tan, J. C., et al. 2022, ApJ, 936, 68
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Dib, S., & Kroupa, P. 2024a, PASP, 136, 094302
Zhou, J. W., Kroupa, P., & Dib, S. 2024b, PASP, 136, 094301
Zhou, J.-w., Kroupa, P., & Dib, S. 2024c, A&A, 688, L19

A&A, 691, A293 (2024)
https://doi.org/10.1051/0004-6361/202451728
© The Authors 2024

Astronomy & Astrophysics

The currently observed clumps cannot be the "direct" precursors of the currently observed open clusters

Jian-wen Zhou[1,⋆], Sami Dib[2,⋆], and Pavel Kroupa[3,4,⋆]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[2] Max Planck Institute für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[3] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany
[4] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, 180 00 Praha 8, Czech Republic

Received 31 July 2024 / Accepted 1 October 2024

ABSTRACT

We categorized clumps, embedded clusters, and open clusters and conducted a comparative analysis of their physical properties. Overall, the radii of open clusters are significantly larger than those of embedded clusters and clumps. The radii of embedded clusters are larger than those of clumps, which may be due to the expansion of embedded clusters. The open clusters have significantly higher masses than embedded clusters, by about one order of magnitude. Given the current mass distribution of clumps in the Milky Way, the evolutionary sequence from a single clump evolving into an embedded cluster and subsequently into an open cluster cannot account for the observed open clusters with old ages and high masses, which is also supported by *N*-body simulations of individual embedded clusters. To explain the mass and radius distributions of the observed open clusters, initial embedded clusters with masses higher than 3000 $M_\odot$ are necessary. However, the upper limit of the embedded cluster sample is less than 1000 $M_\odot$, and only a few ATLASGAL clumps have a mass higher than 3000 $M_\odot$. Thus, the currently observed clumps cannot be the "direct" precursors of the currently observed open clusters. If the Milky Way has a burst-like and time-dependent star formation history, the currently observed open clusters with old ages and high masses may come from massive clumps in the past. There is also a very real possibility that these open clusters originate from post-gas expulsion coalescence of multiple embedded clusters. We compared the separation of open clusters and the typical size of molecular clouds, and find that most molecular clouds may only form one open cluster, which supports the scenario of post-gas expulsion coalescence. Further study is necessary to distinguish between the different scenarios.

Key words. stars: evolution – ISM: clouds – open clusters and associations: general – galaxies: clusters: general – galaxies: star formation

1. Introduction

Massive star-forming regions (MSFRs) constitute a primary mode of star formation within the Galaxy (Lada & Lada 2003; Motte et al. 2018). Young stellar clusters within these regions can potentially evolve into open clusters (OCs). It has been proposed that the majority of observed stars, if not all, originate from embedded clusters (Kroupa 1995a,b; Lada & Lada 2003; Kroupa 2005; Megeath et al. 2016; Dinnbier et al. 2022). Despite their significance, the characteristics of embedded clusters and their connection to OCs remain unclear. The physical processes responsible for the fragmentation of cluster-forming gas and the subsequent dynamics of the resulting stars are not well understood (Megeath et al. 2016). While it is probable that embedded clusters are precursors to OCs, Lada & Lada (2003) found that only 7% of embedded clusters survive the gas dispersal phase.

Understanding the structure and dynamical evolution of embedded clusters represents a crucial step toward unraveling the formation processes of OCs. Recent observational advancements indicate that early expansion plays a pivotal role in shaping the development of young star clusters. Utilizing *Gaia* Data Release (DR) 2 and DR3 data alongside multi-object spectroscopy, studies have unveiled the expansion phenomena in very young clusters within star-forming regions through both case studies (Wright et al. 2019; Cantat-Gaudin et al. 2019; Kuhn et al. 2020; Lim et al. 2020; Swiggum et al. 2021; Lim et al. 2022; Mužić et al. 2022; Das et al. 2023; Flaccomio et al. 2023) and statistical works (Kuhn et al. 2019; Della Croce et al. 2024; Wright et al. 2023). *N*-body simulations of massive star clusters with monolithic structures and residual gas expulsion (Kroupa et al. 2001; Banerjee & Kroupa 2013, 2014, 2015; Haghi et al. 2020) as well as semi-analytical models (Dib et al. 2007, 2010; Dib 2023) have consistently replicated the characteristics of well-observed very young massive clusters such as the Orion Nebula Cluster, R136, NGC 3603, and the Arches cluster, offering insights into their initial formation conditions. For lower-mass embedded clusters, Zhou et al. (2024e) compiled various cluster samples to examine the early expansion dynamics using numerical simulation recipes from Kroupa et al. (2001), Banerjee & Kroupa (2013, 2014, 2015), and Dinnbier et al. (2022). A special focus of that work was the evolution of the initial mass-radius relation of embedded clusters.

In Zhou et al. (2024e), direct *N*-body simulations were employed to model the early expansion of embedded clusters after the expulsion of their residual gas. The models effectively fit the observational data on cluster radii from various sources, indicating that observed very young clusters are likely in an

⋆ Corresponding authors; jwzhou@mpifr-bonn.mpg.de; sami.dib@gmail.com; pkroupa@uni-bonn.de

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

expanding phase. They also suggest that even embedded clusters within ATLASGAL clumps with HII regions may already be undergoing expansion. In Zhou et al. (2024d), the dendrogram algorithm was employed to analyze surface density maps of stars derived from the Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX) project (Feigelson et al. 2013; Kuhn et al. 2015b). This approach decomposes the maps into hierarchical structures, revealing the multi-scale features of star clusters. Additionally, the minimum spanning tree (MST) method was utilized to quantify the distances between clusters and clumps within each MSFR. Notably, the distances between clusters, between clumps, and between clusters and clumps were found to be comparable. This suggests that the evolution from clumps to embedded clusters occurs locally and independently, with minimal influence on surrounding objects. In this work we further investigated the evolution from embedded clusters to OCs by analyzing observational data and conducting *N*-body simulations.

2. Sample

2.1. Embedded clusters

In Zhou et al. (2024d), we identified a sample of embedded clusters based on the surface density maps of stars for 17 MSFRs with distances less than 3.6 kpc from the MYStIX project. To this we added the embedded cluster catalogs of Lada et al. (1991), Carpenter et al. (1993), Lada & Lada (2003), Carpenter (2000), Kumar et al. (2006), and Faustini et al. (2009) and required the distance of the clusters to be less than 3.6 kpc.

2.2. Clumps

We selected clumps from the ATLASGAL survey (Urquhart et al. 2022) and limited the kinematic distance of the clumps to be less than 3.6 kpc. In addition to being consistent with the distance restriction of the embedded cluster sample, this distance limitation can also effectively eliminate a possible distance bias (Urquhart et al. 2022).

2.3. Open clusters

Hunt & Reffert (2023) conducted a blind, all-sky search for OCs using 729 million sources from *Gaia* DR3 down to magnitude G ∼20, creating a homogeneous catalog of 7167 clusters. The Hunt & Reffert (2024) catalog (see their Table 3) includes 5647 OCs and 1309 moving groups, of which 3530 OCs and 539 moving groups are of high quality. For the 3530 high-quality OCs, most of them (3103) have a distance of less than 3.6 kpc.

3. Results

3.1. Mass and radius

For the ATLASGAL clumps in Urquhart et al. (2022), the radius was determined from the number of pixels within the full width at half maximum (FWHM) contour (i.e., above 50% of the peak of the ATLASGAL dust continuum emission). The clump's mass, also measured at the FWHM, was determined by integrating the flux density at 870 μm within the 50% contour. Furthermore, the radii of ATLASGAL clumps with HII regions (HII clumps) were confirmed to be equal to the half-mass radii of embedded clusters in the clumps in Zhou et al. (2024e). In Hunt & Reffert (2024), the radius of OCs was determined by

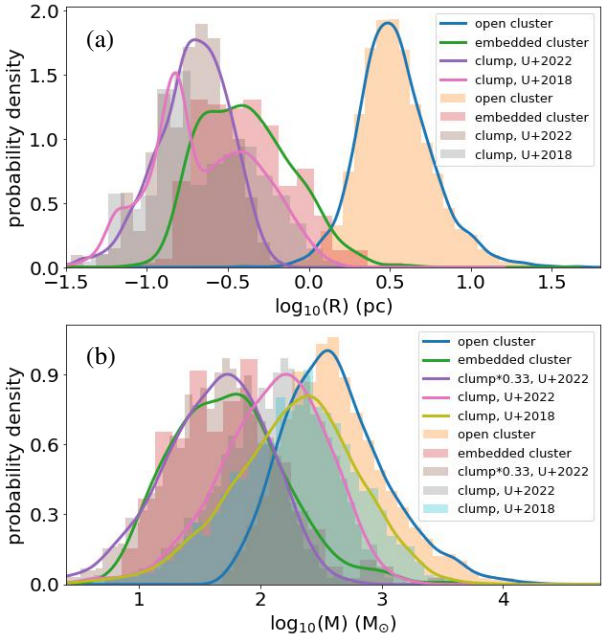


Fig. 1. Physical properties of Galactic clumps, embedded clusters, and open clusters. (a) Radius distribution; (b) mass distribution. The lines are from the histograms.

including half of the members (r_{50}) that is close to the half-mass radius. For embedded clusters, their radii were unified to the half-mass radii, as done in Zhou et al. (2024e). Therefore, the radii of clumps, embedded clusters, and OCs can be directly compared. Generally, the radii of OCs are significantly larger than those of embedded clusters and clumps. The radii of embedded clusters are larger than those of clumps, which may be due to the expansion of embedded clusters, as presented in Zhou et al. (2024e) and clearly shown in Appendix A.

In Fig. 1b, the mass distribution decreases from OCs to clumps to embedded clusters. Interestingly, the mass distribution of clumps, when multiplied by 0.33, is comparable to that of embedded clusters, implying that 0.33 could be a typical star formation efficiency (SFE) for clumps in the Milky Way. The mass distribution of the clumps is significantly smaller than that of the OCs. The mass of ATLASGAL clumps in Urquhart et al. (2018) is the full mass rather than the FWHM mass in Urquhart et al. (2022), which is still significantly lower than the mass of OCs.

These offsets raise questions about the origin of OCs with masses significantly higher than those of their progenitors (clumps or embedded clusters). From clumps to embedded clusters, the SFE is ≈0.3 (Lada & Lada 2003; Megeath et al. 2016; Zhou et al. 2024d,c,b), and there is considerable mass loss during the evolution. Therefore, for a single clump to evolve into a single embedded cluster and then evolve into a single OC, with the current clump mass distribution in the Milky Way, this evolutionary sequence cannot produce the observed OCs' mass distribution, as verified by the direct *N*-body simulations in Sect. 3.3.

If an OC originates from an embedded cluster, the initial mass of the embedded cluster should be much higher than that of the OC due to the considerable mass loss during the evolution, as shown in Sect. 3.4. However, the opposite is seen in Fig. 1b, where the mass distribution of OCs is significantly larger than that of embedded clusters (by about one order of magnitude), suggesting that most OCs do not originate from single embedded clusters.

Table 1. Computed model parameters.

$M_{\rm ecl}$($M_\odot$)	$r_{\rm h}$(pc)	$m_{\rm max}$($M_\odot$)	$\tau_{\rm g}$(Myr)	$\tau_{\rm d}$(Myr)	-s (Myr)	-r (pc)	–r-search-min(pc)	–r-bin(pc)
300	0.21	17	0.016	0.6	1.22e-04	3.08e-03	3.55e-03	2.46e-04
1000	0.25	30	0.019	0.6	1.22e-04	3.21e-03	3.93e-03	2.57e-04
3000	0.28	43	0.022	0.6	6.10e-05	2.12e-03	2.75e-03	1.70e-04

Notes. $M_{\rm ecl}$ and $r_{\rm h}$ are the mass and the half-mass radius of the cluster. $m_{\rm max}$ is the mass of the most massive star in the cluster. $\tau_{\rm g}$ and $\tau_{\rm d}$ are the decay time and the delay time during the gas expulsion process. The tree time step (-s), the changeover radius (-r), the neighbor searching radius (–r-search-min), and the multiple group radius (–r-bin) are the important parameters in `PeTar`.

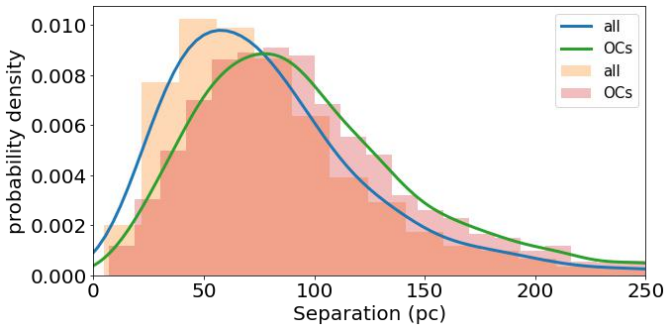


Fig. 2. Separation distribution of each of the two neighbouring objects. The "all" signifies all 5858 OCs and moving groups, and "OCs" high-quality OCs.

3.2. Separations of open clusters

We used the minimum spanning tree (MST) analysis to measure the inter-cluster spacing based on the 3D coordinates of OCs. The MST connects all clusters in a network with the least amount of total distance, without creating any loops. The length of the line connecting two clusters represents the separation between them. In Fig. 2, the typical inter-cluster separation of 3103 high-quality OCs is ∼80 pc. Considering all 5858 OCs and moving groups (<3.6 kpc), the lower limit of the separation is ∼60 pc. However, the most probable size for molecular clouds in the entire Milky Way disk is ∼30 pc (Miville-Deschênes et al. 2017). This implies that most molecular clouds may only form one OC. Each molecular cloud typically includes many embedded clusters, as presented in Appendix A. If each individual embedded cluster could evolve into an OC, the spacing between OCs should not be so large; it should at least be smaller than the size of the parental molecular cloud.

In Zhou et al. (2024d), for the MSFRs in the MYStIX project, the separations between clusters, between clumps, and between clusters and clumps are quite comparable. The typical separation is ∼1 pc. Therefore, as the embedded clusters in each MSFR expand, they are able to undergo merger (post-gas expulsion coalescence). Significant differences in separation for embedded clusters and OCs indicate that there are considerable mergers between embedded clusters as they evolve into OCs. We will discuss this picture in a forthcoming work.

3.3. Single embedded cluster simulation

3.3.1. Fast gas expulsion

The parameters for the simulation are summarized from previous works (i.e., Kroupa et al. 2001; Baumgardt & Kroupa 2007; Banerjee et al. 2012; Banerjee & Kroupa 2013, 2014, 2015, 2017; Oh et al. 2015; Oh & Kroupa 2016; Brinkmann et al. 2017; Oh & Kroupa 2018; Wang et al. 2019; Pavlik et al. 2019; Dinnbier et al. 2022; Zhou et al. 2024e). The previous simulations have already demonstrated the effectiveness and rationality of the parameter settings (see below). The influence of different parameter settings on simulation results and the discussion of the multidimensional parameter space can also be found in the works cited above. In this work, we mainly utilized the mature simulation recipe to interpret observational data. The model parameters of the simulations are listed in Table. 1.

We computed three clusters with star masses of 300 $M_\odot$, 1000 $M_\odot$, and 3000 $M_\odot$ in the first 100 Myr. The initial density profile of all clusters was the Plummer profile (Aarseth et al. 1974; Heggie & Hut 2003; Kroupa 2008), an appropriate choice since the molecular clouds' filaments in which stars form have been found to have Plummer-like cross sections (Malinen et al. 2012; André et al. 2022), and open star clusters can also be described by the Plummer model (Röser et al. 2011; Röser & Schilbach 2019). Moreover, such a specific initial profile does not significantly affect the overall expansion rate of a cluster, as discussed in Banerjee & Kroupa (2017), which is primarily governed by the total stellar mass loss and the dynamical interactions occurring within the inner part of the cluster. The half-mass radius, r_h, of the cluster is given by the $r_{\rm h}$–$M_{\rm ecl}$ relation (Marks & Kroupa 2012):

$$\frac{r_{\rm h}}{\rm pc} = 0.10^{+0.07}_{-0.04} \times \left(\frac{M_{\rm ecl}}{M_\odot}\right)^{0.13\pm0.04}. \quad (1)$$

All clusters are fully mass segregated (S=1), with no fractalization, and in a state of virial equilibrium (Q=0.5). S and Q are the degree of mass segregation and the virial ratio of the cluster, respectively. More details can be found in Küpper et al. (2011) and the user manual for the `McLuster` code. The initial segregated state is detected for young clusters and star-forming clumps/clouds (Littlefair et al. 2003; Chen et al. 2007; Portegies Zwart et al. 2010; Kirk & Myers 2011; Getman et al. 2014; Lane et al. 2016; Alfaro & Román-Zúñiga 2018; Plunkett et al. 2018; Pavlik et al. 2019; Nony et al. 2021; Zhang et al. 2022; Xu et al. 2024), but the degree of mass segregation is not clear. In simulations of the very young massive clusters R136 and NGC 3603 with gas expulsion by Banerjee & Kroupa (2013), mass segregation does not influence the results. In Zhou et al. (2024e), we compared S=1 (fully mass segregated) and S=0.5 (partly mass segregated) and found similar results. We also discussed settings with and without fractalization in Zhou et al. (2024e); the results of the two are also consistent. The initial mass functions of the clusters were chosen to be canonical (Kroupa 2001), with the most massive star following the $m_{\rm max} - M_{\rm ecl}$ relation of Weidner et al. (2013):

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \quad (2)$$

where $y = \log_{10}(m_{\max}/M_\odot)$, $x = \log_{10}(M_{\rm ecl}/M_\odot)$, $a_0 = -0.66$, $a_1 = 1.08$, $a_2 = -0.150$, and $a_3 = 0.0084$. We assumed the clusters to be at solar metallicity (i.e., $Z = 0.02$; von Steiger & Zurbuchen 2016). The clusters travel along circular orbits within the Galaxy, positioned at a galactocentric distance of 8.5 kpc, and are moving at a speed of 220 km s^{-1}.

The initial binary setup follows the method described in Wang et al. (2019). All stars are initially in binaries, that is to say, $f_{\rm b}$=1, where $f_{\rm b}$ is the primordial binary fraction. Kroupa (1995a,b) propose that stars with masses below a few $M_\odot$ are initially formed with a universal binary distribution function and that star clusters start with a 100% binary fraction. Inverse dynamical population synthesis was employed to derive the initial distributions of binary periods and mass ratios. Belloni et al. (2017) introduce an updated model of pre-main-sequence eigenevolution, originally developed by Kroupa (1995b), to account for the observed correlations between the mass ratio, period, and eccentricity in short-period systems. For low-mass binaries, we adopted the formalism developed in Kroupa (1995a,b) and Belloni et al. (2017) to characterize the period, mass ratio, and eccentricity distributions. For high-mass binaries (OB stars with masses >5 $M_\odot$), we utilized the Sana et al. (2012) distribution, which is derived from O stars in OCs. The distribution functions of the period, mass ratio, and eccentricity are presented in Oh et al. (2015) and Belloni et al. (2017).

Accurately modeling gas removal from embedded clusters is challenging due to the complexity of radiation hydrodynamical processes, which involve uncertain and intricate physical mechanisms. To simplify the approach, we simulated the key dynamical effects of gas expulsion by applying a diluting, spherically symmetric external gravitational potential to a model cluster, following the method presented in Kroupa et al. (2001) and Banerjee & Kroupa (2013). This analytical approach is partially validated by Geyer & Burkert (2001), who conducted comparison simulations using the smoothed particle hydrodynamics method to treat the gas. The hydrodynamics+N-body simulations in Farias et al. (2024) also confirm that the exponential decay model presented in Eq. (3) generally provides a good description of gas removal driven by radiation and wind feedback. Specifically, we used the potential of the spherically symmetric, time-varying mass distribution:

$$\begin{aligned} M_g(t) &= M_g(0) \qquad t \le \tau_d, \\ M_g(t) &= M_g(0)\exp\left(-\tfrac{(t-\tau_d)}{\tau_g}\right) t > \tau_d. \end{aligned} \tag{3}$$

Here, $M_g(t)$ is the total mass of the gas; it is spatially distributed with the Plummer density distribution (Kroupa 2008) and starts depleting after a delay of τ_d, and is totally depleted within a timescale of τ_g. The Plummer radius of the gas distribution is kept time-invariant (Kroupa et al. 2001). This assumption is an approximate model of the effective gas reduction within the cluster in the situation that gas is blown out while new gas is also accreting into the cluster along filaments such that the gas mass ends up being reduced with time but the radius of the gas distribution remains roughly constant. As discussed in Urquhart et al. (2022), the mass and radius distributions of the ATLASGAL clumps at different evolutionary stages are quite comparable. We used an average velocity of the outflowing gas of $v_g \approx 10$ km s^{-1}, which is the typical sound speed in an HII region. This gives $\tau_g = r_h(0)/v_g$, where $r_h(0)$ is the initial half-mass radius of the cluster. As for the delay time, we took the representative value of $\tau_d \approx 0.6$ Myr (Kroupa et al. 2001), this being about the lifetime of the ultracompact HII region. As shown in Banerjee & Kroupa (2013), varying the delay time, τ_d, primarily results in a temporal shift in the cluster's rapid expansion phase, without significantly impacting its subsequent evolution for times greater than τ_d. Protoclusters typically form in hub-filament systems (Motte et al. 2018; Vázquez-Semadeni et al. 2019; Kumar et al. 2020; Zhou et al. 2022), which are located in hub regions. Compared to the surrounding filamentary gas structures, the hub region, as the center of gravitational collapse, is usually more regular, as shown in Zhou et al. (2022, 2024a). Thus, modeling a spherically symmetric mass distribution is appropriate (see Sect. 3.3.3 for more discussion).

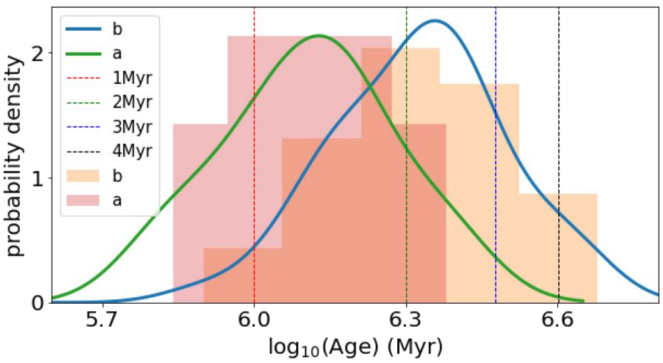


Fig. 3. Age distribution of embedded clusters in the seven MSFRs displayed in Appendix A. The ages are taken from Table 1 of Kuhn et al. (2015a). The "a" represents NGC 6334 and M 17, and "b" represents the Lagoon nebula, NGC 6357, the Eagle nebula, the Carina nebula, and the Trifid nebula.

In this work, we assumed a SFE ≈0.33 as a benchmark (i.e., $M_g(0) = 2M_{\rm ecl}(0)$). This value has been widely used in the simulations cited above and has been proven effective in reproducing the observational properties of stellar clusters. Such a SFE is also consistent with the value obtained from hydrodynamical calculations that include self-regulation (Machida & Matsumoto 2012; Bate et al. 2014) and as well with observations of embedded systems in the solar neighborhood (Lada & Lada 2003; Megeath et al. 2016). In Zhou et al. (2024d), by comparing the mass functions of the ATLASGAL clumps and the identified embedded clusters, we found that a SFE of ≈0.33 is typical for the ATLASGAL clumps, which are also presented in Fig. 1b here. In Zhou et al. (2024c), assuming SFE = 0.33, it was shown that the total bolometric luminosity of synthetic embedded clusters can fit the observed bolometric luminosity of ATLASGAL clumps with HII regions. In Zhou et al. (2024b), we directly calculated the SFE of ATLASGAL clumps with HII regions and found a median value of ≈0.3.

3.3.2. Timescale of gas free

Seven of the 17 MSFRs investigated in the MYStIX project are covered in the ATLASGAL survey (Schuller et al. 2009): the Lagoon nebula, NGC 6334, NGC 6357, the Eagle nebula, M 17, the Carina nebula, and the Trifid nebula. As shown in Appendix A, the ATLASGAL+*Planck* or only ATLASGAL 870 μm data (Csengeri et al. 2016) were used to trace the gas distribution around embedded or very young clusters. On these maps, clusters and clumps are well separated. However, NGC 6334 and M 17 still have abundant gas, which means they are younger than the other MSFRs. As shown in Fig. 3, the clusters in NGC 6334 and M 17 ("a") are indeed younger (the median value is ≈1.4 Myr) than those in other MSFRs ("b"; the median

value is ≈2.2 Myr). If embedded clusters in NGC 6334 and M 17 just finished the gas expulsion, then this implies an upper limit on the gas expulsion time of around 1.4 Myr.

3.3.3. Slow and moderate gas expulsions

Embedded clusters form in clumps. More massive clumps can produce more massive clusters, leading to stronger feedback and a higher gas expulsion velocity (Dib et al. 2013). There should be correlations between the feedback strength, the clump (or cluster) mass, and the gas expulsion velocity (v_g). And low-mass clusters should have a slower gas expulsion process compared with high-mass clusters. As shown in Pang et al. (2021), low-mass clusters tend to agree with the simulations of slow gas expulsion. Except for the feedback strength, the SFE determines the total amount of the remaining gas, which also influences the timescale of gas expulsion. The gas expulsion process is driven by feedback, and the effectiveness of the feedback will depend on the geometric shape of the gas shell surrounding the embedded cluster (Wünsch et al. 2010; Rahner et al. 2017). Therefore, a complex or nonspherical gas distribution would also change the timescale of gas expulsion. In short, these parameter uncertainties can ultimately be incorporated into the timescale of gas expulsion. Thus, apart from the fast gas expulsion with the gas decay time, τ_g, described above, we also simulated slow and moderate gas expulsions. As described in Sect. 3.3.2, the upper limit on the gas expulsion timescale may be ≈1.4 Myr. Therefore, for the slow gas expulsion, the gas decay time was set to $10\tau_g$. The moderate gas expulsion is between the fast and slow gas expulsions. Considering the large parameter space between τ_g and $10\tau_g$, we simulated two kinds of moderate gas expulsions: $2\tau_g$ ("moderate2") and $5\tau_g$ ("moderate1") . Figure 4 shows the evolution of gas mass under different gas expulsion modes over time.

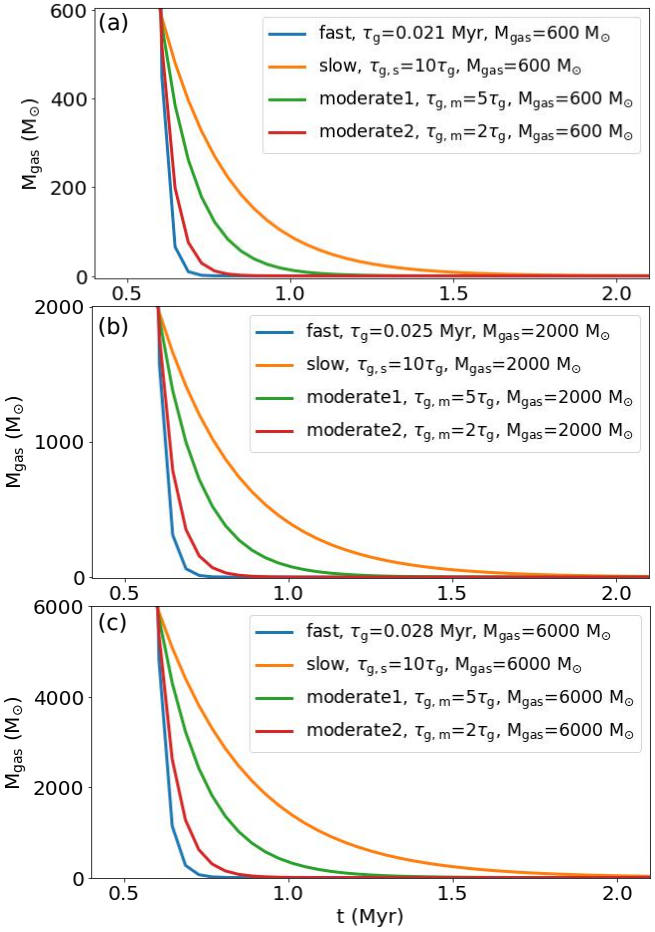


Fig. 4. Evolution of the gas mass over time under different gas expulsion modes.

3.3.4. Star formation efficiency

The total amount of the residual gas not only affects the gas expulsion timescale, but it also significantly influences the strength of the gas potential. The strength of the external gas potential may have a considerable impact on the evolution of embedded star clusters. The total amount of residual gas is determined by the SFE of the clump. In the simulations described above, we assumed a SFE = 0.33 for all star clusters. In Zhou et al. (2024b), we calculated the SFE of ATLASGAL clumps with HII regions and found a strong anticorrelation between the SFE and the clump mass. Here, we directly derived the residual gas mass using Eq. (6) in Zhou et al. (2024b), which encodes the variation in SFE with clump mass:

$$\log_{10}(M_{\rm cl}) = (1.02 \pm 0.02) \times \log_{10}(M_{\rm ecl}) + (0.52 \pm 0.05), \quad (4)$$

and

$$M_g(0) = M_{\rm cl} - M_{\rm ecl}, \quad (5)$$

where $M_{\rm cl}$ is the clump mass.

3.3.5. Procedure

The `McLuster` program (Küpper et al. 2011) was used to set the initial configurations. The dynamical evolution of the model clusters was computed using the state-of-the-art `PeTar` code (Wang et al. 2020). `PeTar` employs well-tested analytical single and binary stellar evolution recipes (SSE/BSE) (Hurley et al. 2000, 2002; Giacobbo et al. 2018; Tanikawa et al. 2020; Banerjee et al. 2020). In `PeTar`, there are some important parameters: the tree time step (-s), the changeover radius (-r), the neighbor searching radius (–r-search-min), and the multiple group radius (–r-bin). The tree time step is a fixed time step used to calculate the long-range (particle-tree) force. The changeover region is the overlap shell between the long-range and short-range interaction. In particular, -s and -r significantly affect the accuracy and performance of the simulations. If -s and -r are not specified, `PeTar` determines them automatically, assuming that the input model is a spherical symmetric star cluster with a King- or a Plummer-like density profile. The neighbor searching radius is defined for each particle to identify neighboring candidates expected to be within the changeover region during the subsequent tree time step. Upon determining the changeover radius (-r), `PeTar` automatically computes the neighbor searching radius. The multiple group radius (–r-bin) is used as the criterion for selecting group members, which is automatically determined based on the changeover inner radius. In this work, all these parameters were automatically determined by `PeTar` (see Table 1).

3.4. Simulation results analysis

In Fig. 5, as expected, the mass loss of the clusters strongly depends on the gas expulsion modes. In the scenario of fast gas expulsion, gas is removed almost instantaneously, which

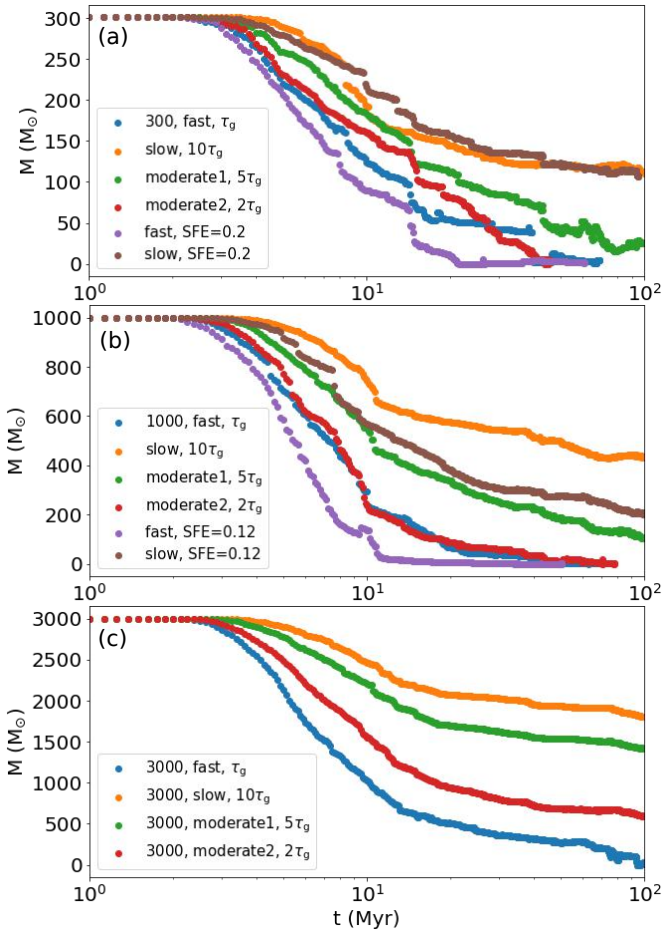


Fig. 5. Evolution of the cluster mass within the tidal radius over time under different gas expulsion modes and with different SFEs.

strongly impacts the clusters. In contrast, with slow gas expulsion, the gas disperses over a timescale significantly longer than the cluster's crossing (dynamical) time, enabling the cluster to adapt to the changing gravitational potential and expand while maintaining near dynamical equilibrium, as discussed in Baumgardt & Kroupa (2007). This quasi-static (or adiabatic) evolution results in substantially less stellar mass loss, as shown in Fig. 5. Therefore, it best fits the mass distribution of the observed OCs. However, it fails to match the radius distribution, as shown in Fig. 6d1. For the fast gas expulsion, the clusters cannot survive more than 100 Myr. Similarly, in Kroupa (2001) and Brinkmann et al. (2017), the massive cluster containing many O stars most likely survives as an OC. Low-mass embedded clusters or groups dissolve quickly, in part due to the loss of their residual gas. A lower SFE means more residual gas, leading to more intense expansion and mass loss of the cluster after gas expulsion, as shown in Fig. 5b. A lower SFE is equivalent to a shorter gas expulsion timescale.

For high-mass clusters, such as 3000 $M_\odot$, τ_g and $2\tau_g$ will give significantly different results. Thus, the simulations are very sensitive to the gas decay time. We need to carefully constrain this parameter from observations in future work. Since $\tau_g = r_h(0)/v_g$, with $r_h(0)$ being constrained by Eq. (1), we can focus on the measurement of the gas expulsion velocity (v_g) in observations, especially the correlation between v_g and the clump (or cluster) mass, as discussed in Sect. 3.3.3.

In Fig. 6, fast gas expulsion is required to explain the radius distribution of the observed OCs, but to account for the mass distribution, slow gas expulsion is necessary. This contradiction makes both fast and slow gas expulsions inappropriate. To reconcile this contradiction, the gas expulsion should be somewhere between fast and slow. In terms of both mass and radius, "moderate2" can best fit the observed OCs, and the main contribution is from the high-mass cluster (3000 $M_\odot$). To explain the mass and radius distributions of the observed OCs, initial embedded clusters with masses higher than 3000 $M_\odot$ are necessary. However, in Fig. 1, the upper limit of the embedded cluster sample is less than 1000 $M_\odot$, and only a few clumps have a mass higher than 3000 $M_\odot$. For an OC, its progenitor should have a significantly higher mass. Thus, the currently observed clumps cannot be the "direct" precursors of the currently observed OCs.

4. Discussion

There may be two evolutionary tracks from embedded clusters to OCs: a single embedded cluster evolution scenario and a multiple embedded clusters merger (post-gas expulsion coalescence) scenario. In Zhou et al. (2024d), we used the MST method to measure the separations between clusters and gas clumps in each MSFR shown in Appendix A. The separations between clusters, between clumps, and between clusters and clumps are comparable, which indicates that the evolution from clump to embedded cluster proceeds in isolation and locally, and does not significantly affect the surrounding objects. Using *Gaia* DR2 data, Kuhn et al. (2019) investigated the kinematics of subclusters identified in Kuhn et al. (2014) for the same MSFRs and found no evidence that these groups are merging. Therefore, in this work, we only focused on the first scenario. However, as presented above, the masses of current clumps or embedded clusters in the Milky Way are too low to have evolved into the observed OCs.

For the analysis in this work, we assumed a constant star formation history for the Milky Way. If its star formation history is instead burst-like and time-dependent, the currently observed OCs may be remnants of massive clumps from the past. The epoch of peak star formation rate density is thought to lie between redshifts of 2 and 3 (Madau et al. 1998; Hopkins & Beacom 2006; Bouwens et al. 2011; Moster et al. 2013). *Gaia* DR2 reveals a star formation burst in the Galactic disk 2–3 Gyr ago (Mor et al. 2019). As an inference, the peak period of star formation in the Milky Way has passed, and most of the massive clumps have evolved into OCs or older clusters. At present, only relatively low-mass clumps remain. This possibility is worth exploring further in future work.

Another explanation is that these massive and relatively old OCs come from post-gas expulsion coalescence of multiple embedded clusters. As shown in Appendix A, the embedded clusters or the very young clusters are very near to each other in each MSFR. There is a very real possibility that they will merge after expansion. One molecular cloud can produce multiple embedded clusters, which can merge into larger clusters. This scenario is consistent with the finding in Sect. 3.2, that most molecular clouds can only form one OC. In a forthcoming work (Zhou et. al. in preparation), we will simulate the post-gas expulsion coalescence of subclusters in each MSFR, with the initial conditions of the simulations directly derived from the observation. There is now extensive literature arguing that star clusters form from mergers between embedded clusters, both from simulations (Fujii et al. 2012; Howard et al. 2018; Sills et al. 2018;

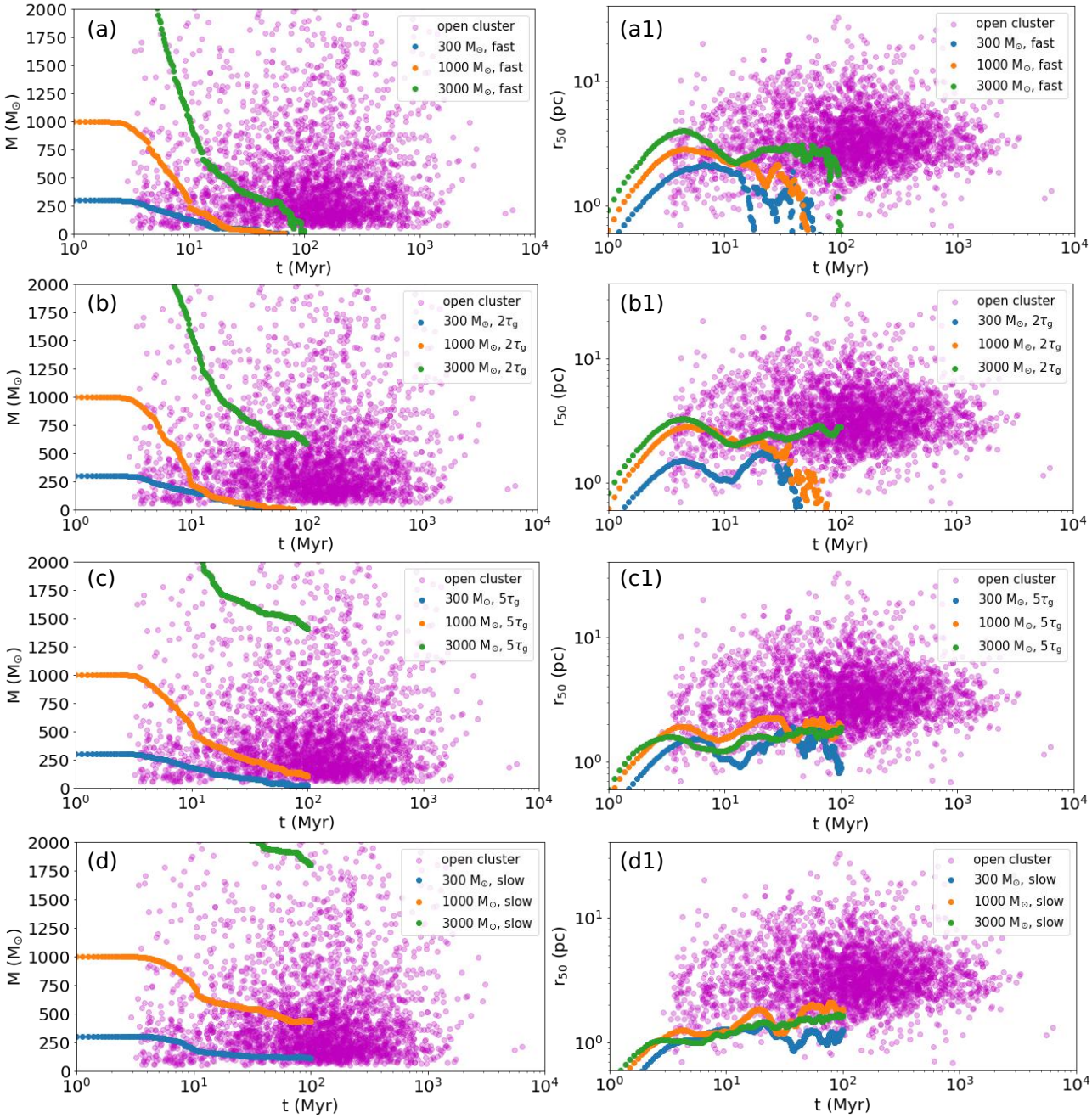


Fig. 6. Fitting the physical parameters of OCs in Hunt & Reffert (2024) using single embedded cluster simulations.

Fujii et al. 2022; Dobbs et al. 2022; Cournoyer-Cloutier et al. 2023; Reina-Campos et al. 2024) and observations (Sabbi et al. 2012; Pang et al. 2022; Della Croce et al. 2023; Fahrion & De Marchi 2024). A comparison with these works will also be presented in the forthcoming work.

In the current work, the initial density profile of all clusters is the Plummer profile. However, observations of embedded clusters show clusters with high ellipticities (Kuhn et al. 2014; Getman et al. 2018). Due to the limitations of the method used here, we are currently unable to discuss the impact of the cluster's ellipticity on its subsequent evolution. For now, we have only performed N-body simulations. In future work, it is necessary to compare them with more comprehensive hydrodynamics+ N-body simulations, such as those presented in Fujii et al. (2021), Rieder et al. (2021), and Polak et al. (2024), to further test and constrain the initial conditions used in the N-body simulations presented in this work.

5. Conclusion

We collected samples of Galactic clumps, embedded clusters, and OCs to compare their physical properties. Overall, the radii of OCs are significantly larger than those of embedded clusters and clumps. The radii of embedded clusters are larger than those of clumps, which may be due to the expansion of embedded clusters. The mass distribution of OCs covers a significantly larger mass range than that of embedded clusters, by about one order of magnitude. Given the current mass distribution of clumps in the Milky Way, it is unlikely that a single clump evolving into a single embedded cluster, and subsequently into a single OC, could

account for the observed OCs with old ages and high masses, which is also supported by *N*-body simulations of individual embedded clusters.

We used standard *N*-body simulations as in previous works to interpret the physical parameters of the observed OCs. The mass loss of the simulated clusters strongly depends on the gas expulsion modes. For the fast gas expulsion, the clusters ($\lesssim$3000 $M_\odot$) cannot survive more than 100 Myr. Fast gas expulsion is required to explain the radius distribution of the observed OCs, but to account for the mass distribution, slow gas expulsion is necessary; the real gas expulsion speed should be somewhere in between. In any case, to explain the mass and radius distributions of the observed OCs, initial embedded clusters with masses higher than 3000 $M_\odot$ are necessary. However, in Fig. 1, the upper limit of the embedded cluster sample is less than 1000 $M_\odot$, and only a few clumps have a mass higher than 3000 $M_\odot$. The progenitor of an OC should have a significantly higher mass. Thus, the currently observed clumps cannot be the "direct" precursors of the currently observed OCs.

If the star formation history of the Milky Way is burst-like and time-dependent, the currently observed OCs may be remnants of massive clumps from the past. At present, only relatively low-mass clumps remain. Another explanation is that those massive and relatively old OCs come from post-gas expulsion coalescence of multiple embedded clusters. We compared the separation of OCs and the typical size of molecular clouds in the Milky Way and find that most molecular clouds can only form one OC. Each molecular cloud typically includes many embedded clusters; the typical separation between them is $\approx$1 pc (Zhou et al. 2024d). Thus, after expansion, they should be able to undergo mergers. Further study is necessary to distinguish between the different scenarios.

Acknowledgements. Thanks the referee for providing detailed and constructive review comments, which have helped to improve and clarify this work.

References

Aarseth, S. J., Henon, M., & Wielen, R. 1974, A&A, 37, 183
Alfaro, E. J., & Román-Zúñiga, C. G. 2018, MNRAS, 478, L110
André, P. J., Palmeirim, P., & Arzoumanian, D. 2022, A&A, 667, L1
Banerjee, S., & Kroupa, P. 2013, ApJ, 764, 29
Banerjee, S., & Kroupa, P. 2014, ApJ, 787, 158
Banerjee, S., & Kroupa, P. 2015, MNRAS, 447, 728
Banerjee, S., & Kroupa, P. 2017, A&A, 597, A28
Banerjee, S., Kroupa, P., & Oh, S. 2012, ApJ, 746, 15
Banerjee, S., Belczynski, K., Fryer, C. L., et al. 2020, A&A, 639, A41
Bate, M. R., Tricco, T. S., & Price, D. J. 2014, MNRAS, 437, 77
Baumgardt, H., & Kroupa, P. 2007, MNRAS, 380, 1589
Belloni, D., Askar, A., Giersz, M., Kroupa, P., & Rocha-Pinto, H. J. 2017, MNRAS, 471, 2812
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
Brinkmann, N., Banerjee, S., Motwani, B., & Kroupa, P. 2017, A&A, 600, A49
Cantat-Gaudin, T., Jordi, C., Wright, N. J., et al. 2019, A&A, 626, A17
Carpenter, J. M. 2000, AJ, 120, 3139
Carpenter, J. M., Snell, R. L., Schloerb, F. P., & Skrutskie, M. F. 1993, ApJ, 407, 657
Chen, L., de Grijs, R., & Zhao, J. L. 2007, AJ, 134, 1368
Cournoyer-Cloutier, C., Sills, A., Harris, W. E., et al. 2023, MNRAS, 521, 1338
Csengeri, T., Weiss, A., Wyrowski, F., et al. 2016, A&A, 585, A104
Das, S. R., Gupta, S., Prakash, P., Samal, M., & Jose, J. 2023, ApJ, 948, 7
Della Croce, A., Dalessandro, E., Livernois, A., et al. 2023, A&A, 674, A93
Della Croce, A., Dalessandro, E., Livernois, A. R., & Vesperini, E., 2024, A&A, 683, A10
Dib, S. 2023, ApJ, 959, 88
Dib, S., Kim, J., & Shadmehri, M. 2007, MNRAS, 381, L40
Dib, S., Shadmehri, M., Padoan, P., et al. 2010, MNRAS, 405, 401
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61
Dobbs, C. L., Bending, T. J. R., Pettitt, A. R., Buckner, A. S. M., & Bate, M. R. 2022, MNRAS, 517, 675
Fahrion, K., & De Marchi, G. 2024, A&A, 681, A20
Farias, J. P., Offner, S. S. R., Grudić, M. Y., Guszejnov, D., & Rosen, A. L. 2024, MNRAS, 527, 6732
Faustini, F., Molinari, S., Testi, L., & Brand, J. 2009, A&A, 503, 801
Feigelson, E. D., Townsley, L. K., Broos, P. S., et al. 2013, ApJS, 209, 26
Flaccomio, E., Micela, G., Peres, G., et al. 2023, A&A, 670, A37
Fujii, M. S., Saitoh, T. R., & Portegies Zwart, S. F. 2012, ApJ, 753, 85
Fujii, M. S., Saitoh, T. R., Hirai, Y., & Wang, L. 2021, PASJ, 73, 1074
Fujii, M. S., Wang, L., Hirai, Y., et al. 2022, MNRAS, 514, 2513
Getman, K. V., Feigelson, E. D., Kuhn, M. A., et al. 2014, ApJ, 787, 108
Getman, K. V., Kuhn, M. A., Feigelson, E. D., et al. 2018, MNRAS, 477, 298
Geyer, M. P., & Burkert, A. 2001, MNRAS, 323, 988
Giacobbo, N., Mapelli, M., & Spera, M. 2018, MNRAS, 474, 2959
Haghi, H., Safaei, G., Zonoozi, A. H., & Kroupa, P. 2020, ApJ, 904, 43
Heggie, D., & Hut, P. 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics
Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142
Howard, C. S., Pudritz, R. E., & Harris, W. E. 2018, Nat. Astron., 2, 725
Hunt, E. L., & Reffert, S. 2023, A&A, 673, A114
Hunt, E. L., & Reffert, S. 2024, A&A, 686, A42
Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, MNRAS, 315, 543
Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, MNRAS, 329, 897
Kirk, H., & Myers, P. C. 2011, ApJ, 727, 64
Kroupa, P. 1995a, MNRAS, 277, 1491
Kroupa, P. 1995b, MNRAS, 277, 1507
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2005, in ESA Special Publication, 576, The Three-Dimensional Universe with Gaia, eds. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 629
Kroupa, P. 2008, in The Cambridge *N*-Body Lectures, 760, eds. S. J. Aarseth, C. A. Tout, & R. A. Mardling, 181
Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2014, ApJ, 787, 107
Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2015a, ApJ, 812, 131
Kuhn, M. A., Getman, K. V., & Feigelson, E. D. 2015b, ApJ, 802, 60
Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32
Kuhn, M. A., Hillenbrand, L. A., Carpenter, J. M., & Avelar Menendez, A. R. 2020, ApJ, 899, 128
Kumar, M. S. N., Keto, E., & Clerkin, E. 2006, A&A, 449, 1033
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87
Küpper, A. H. W., Maschberger, T., Kroupa, P., & Baumgardt, H. 2011, MNRAS, 417, 2300
Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57
Lada, E. A., Depoy, D. L., Evans, Neal J., I., & Gatley, I. 1991, ApJ, 371, 171
Lane, J., Kirk, H., Johnstone, D., et al. 2016, ApJ, 833, 44
Lim, B., Hong, J., Yun, H.-S., et al. 2020, ApJ, 899, 121
Lim, B., Nazé, Y., Hong, J., et al. 2022, AJ, 163, 266
Littlefair, S. P., Naylor, T., Jeffries, R. D., Devey, C. R., & Vine, S. 2003, MNRAS, 345, 1205
Machida, M. N., & Matsumoto, T. 2012, MNRAS, 421, 588
Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106
Malinen, J., Juvela, M., Rawlings, M. G., et al. 2012, A&A, 544, A50
Marks, M., & Kroupa, P. 2012, A&A, 543, A8
Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, AJ, 151, 5
Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57
Mor, R., Robin, A. C., Figueras, F., Roca-Fàbrega, S., & Luri, X. 2019, A&A, 624, L1
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Mužić, K., Almendros-Abad, V., Bouy, H., et al. 2022, A&A, 668, A19
Nony, T., Robitaille, J. F., Motte, F., et al. 2021, A&A, 645, A94
Oh, S., & Kroupa, P. 2016, A&A, 590, A107
Oh, S., & Kroupa, P. 2018, MNRAS, 481, 153
Oh, S., Kroupa, P., & Pflamm-Altenburg, J. 2015, ApJ, 805, 92
Pang, X., Li, Y., Yu, Z., et al. 2021, ApJ, 912, 162
Pang, X., Tang, S.-Y., Li, Y., et al. 2022, ApJ, 931, 156
Pavlik, V., Kroupa, P., & Šubr, L. 2019, A&A, 626, A79
Plunkett, A. L., Fernández-López, M., Arce, H. G., et al. 2018, A&A, 615, A9
Polak, B., Mac Low, M.-M., Klessen, R. S., et al. 2024, A&A, 690, A207
Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431
Rahner, D., Pellegrini, E. W., Glover, S. C. O., & Klessen, R. S. 2017, MNRAS, 470, 4453
Reina-Campos, M., Gnedin, O. Y., Sills, A., & Li, H. 2024, arXiv e-prints [arXiv:2408.04694]

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **537,** 845–857 (2025) https://doi.org/10.1093/mnras/staf076
Advance Access publication 2025 January 15

The post-gas expulsion coalescence of embedded clusters as an origin of open clusters

Jian-wen Zhou [1]★ Sami Dib [2]★ and Pavel Kroupa [3,4]★

[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[2]*Max Planck Institute für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[3]*Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, D-53115 Bonn, Germany*
[4]*Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic*

Accepted 2025 January 10. Received 2024 November 18; in original form 2024 September 26

ABSTRACT

The mismatch between the mass function of the Milky Way's embedded clusters (ECs) and that of open clusters (OCs) raises the question of whether each OC originates from a single EC. In this work, we explore a scenario in which OCs form as a result of post-gas expulsion coalescence of ECs within the same parental molecular cloud. We model this process using *N*-body simulations of ECs undergoing expansion due to gas expulsion. Our initial conditions are based on the observed spatial, kinematic, and mass distributions of ECs in three representative massive star-forming regions. Initially, ECs are isolated. After further expansion, interactions between ECs begin, mutually influencing their evolution. We examine this process as a function of gas expulsion time-scales, spatial separations between ECs, and their relative velocities. Our results demonstrate that, within a reasonable range of these parameters, the coalescence of ECs is robust and largely insensitive to initial conditions. The mass of ECs plays a critical role in the coalescence process. More massive ECs form stable gravitational cores, which greatly facilitate coalescence and help the resulting cluster resist expansion and Galactic tidal forces. Additionally, the number of ECs also enhances coalescence. The current mass distribution of clumps in the Milky Way suggests that directly forming massive ECs is challenging. However, the coalescence of multiple low-mass ECs can account for the observed parameter space of OCs in the Milky Way.

Key words: methods: numerical – ISM: clouds – open clusters and associations: general – Galaxy: stellar content – galaxies: star clusters: general – galaxies: star formation.

1 INTRODUCTION

Massive star-forming regions (MSFRs) are a major contributor to star formation in the Galaxy. The young stellar clusters that form within these regions may eventually develop into open clusters (OCs; Lada & Lada 2003; Motte, Bontemps & Louvet 2018). It has been suggested that most, if not all, observed stars originate from embedded clusters (Kroupa 1995a, b; Lada & Lada 2003; Kroupa 2005; Megeath et al. 2016; Dinnbier, Kroupa & Anderson 2022). Despite their importance, the properties of embedded clusters and their relationship to OCs remain unclear. The mechanisms driving the fragmentation of gas that forms these clusters, as well as the subsequent dynamics of their stars, are not yet well understood (Megeath et al. 2016; Dib 2023; Li 2024; Zhou et al. 2024f). Although embedded clusters are the likely precursors to OCs, Lada & Lada (2003) discovered that only 7 per cent of embedded clusters survive the gas dispersal phase. In *N*-body simulations of Kroupa (2001) and Brinkmann et al. (2017), massive clusters containing many O stars most likely survive as OCs. However, low-mass embedded clusters or groups dissolve quickly, partly owing to the loss of their residual gas.

In Zhou, Dib & Kroupa (2024a), we collected samples of Galactic clumps, embedded clusters and OCs to compare their physical properties. We showed that the mass distribution of OCs covers a significantly larger mass range than that of embedded clusters, by about one order of magnitude. Given the current mass distribution of clumps in the Milky Way, the evolutionary sequence from a single clump evolving into an embedded cluster and subsequently into an OC cannot account for the observed OCs with old ages and high masses. This fact is also supported by *N*-body simulations of individual embedded clusters. In order to explain the mass and radius distributions of the observed OCs, embedded clusters with initial masses of more than 3000 $M_\odot$ are necessary. However, for the embedded cluster sample collected in Zhou et al. (2024a), the upper limit of the mass is less than 1000 $M_\odot$. Additionally, only a few clumps have masses larger than 3000 $M_\odot$. For an OC, its progenitor should have a significantly larger mass. Thus, the currently observed clumps cannot be the 'direct' precursors of the currently observed OCs with relatively older ages and larger masses.

A possibility is that those massive and relatively old OCs form from the post-gas expulsion coalescence of multiple embedded

★ E-mail: jwzhou@mpifr-bonn.mpg.de (JWZ); dib@mpia.de (SD); pkroupa@uni-bonn.de (PK)

© 2025 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

clusters. In Zhou et al. (2024a), we compared the separation of OCs and the typical size of molecular clouds in the Milky Way, and found that most molecular clouds may only form one OC. Since a molecular cloud typically includes many embedded clusters, this result supports the coalescence scenario. Moreover, the typical separation between embedded clusters in MSFRs can be ≈1 pc (Zhou, Kroupa & Dib 2024g). Therefore, after these clusters expand, they should be able to undergo coalescence. There is now extensive literature arguing that star clusters form through the merging of subclusters, both from simulations (Vázquez-Semadeni et al. 2009; Fujii, Saitoh & Portegies Zwart 2012; Vázquez-Semadeni, González-Samaniego & Colín 2017; Howard, Pudritz & Harris 2018; Sills et al. 2018; Dobbs et al. 2022; Fujii et al. 2022; Guszejnov et al. 2022; Cournoyer-Cloutier et al. 2023; Reina-Campos et al. 2025; Polak et al. 2024b) and observations (Sabbi et al. 2012; Dalessandro et al. 2021; Pang et al. 2022; Della Croce et al. 2023; Fahrion & De Marchi 2024). This study will examine the physical processes involved in the coalescence of embedded clusters after gas expulsion and expansion, and utilize simulations to fit the observed physical parameters of OCs. This paper is organized as follows: in Section 2, we summarize the sample of OCs that is used in this work. In Section 3, we present our results. In the first subsection (i.e. Section 3.1), we summarize the different phases at which the the mergers of OCs progenitors could possibly occur and in the subsection that follows, we present and discuss the results from numerical simulations that we employ in order to simulate this process. In Section 4, we discuss our findings and in Section 5, we present our conclusions.

2 SAMPLE

Hunt & Reffert (2023) performed a blind, all-sky search for OCs, utilizing 729 million Gaia DR3 sources down to a magnitude of $G \approx 20$. This effort resulted in a homogeneous catalogue comprising 7167 clusters. By assessing the masses and Jacobi radii of these clusters, the catalogue's members were classified into two categories: bound OCs and unbound moving groups (MGs; Hunt & Reffert 2024). There are 5647 OCs and 1309 MGs, with 3530 OCs and 539 MGs classified as high quality, based on median colour–magnitude diagram classifications of greater than 0.5 and signal-to-noise ratio (S/N) of greater than 5σ (Hunt & Reffert 2023). In this work, we only focus on the high-quality sample of OCs. To calculate the photometric masses of the clusters, Hunt & Reffert (2024) first derived the photometric masses of the member stars in each cluster using the PARSEC isochrone (Bressan et al. 2012). They then corrected for selection effects and applied a correction for unresolved binary stars. Subsequently, mass functions were fitted and integrated to determine the total cluster mass.

3 RESULTS

3.1 Mergers at different stages

During different stages of cluster formation and evolution, various types of mergers can occur.

3.1.1 Formation of clumps

This stage should be dominated by the hydrodynamics of the gas. Clumps are locally dense gas structures in molecular cloud, which serve as local star-forming sites within molecular cloud.

3.1.2 Formation of embedded clusters

Once clumps form, their evolution and internal star formation proceed together until the formation of embedded clusters. There may be an isolated evolution from clumps to embedded clusters. Clumps are dynamically decoupled from their parent molecular clouds (Peretto et al. 2023; Zhou et al. 2023b). As discussed in Urquhart et al. (2022), the differences observed for the physical parameters (such as radius, mass, density) of clumps at different evolutionary stages are in fact the result of observational biases, rather than the evolution of the embedded objects. This implies that the clumps remain stable throughout their evolutionary process and do not merge. It also suggests that the material lost through outflow and accretion on to stars is roughly balanced by the inflow of new material on to the clumps during star formation. Zhou et al. (2024g) used the minimum spanning tree method to measure the separations between embedded clusters or very young star clusters and clumps in MSFRs investigated in the Massive Young Star-Forming Complex Study in Infrared and X-ray (MYStIX) project (Feigelson et al. 2013; Kuhn, Getman & Feigelson 2015). The separations between clusters, between clumps, and between clusters and clumps are comparable, which indicates that the evolution from clump to embedded cluster proceeds in isolation and locally, and does not affect the surrounding objects significantly. These findings align with the conclusions of Zhou et al. (2024d, e), which indicate that feedback from embedded clusters does not notably alter the physical characteristics of the surrounding dense gas structures. Using Gaia DR2 data, Kuhn et al. (2019) investigated the kinematics of subclusters identified in Kuhn et al. (2014) for the same MSFRs, and found no evidence that these groups are merging.

3.1.3 Evolution of embedded clusters

As the effects of feedback from the embedded clusters becomes increasingly stronger, they are responsible for gradually dispersing the gas of the clumps (Dale & Bonnell 2008; Dib et al. 2011, 2013; Dale, Ercolano & Bonnell 2012; Zhou et al. 2022; Lewis et al. 2023). Due to the gas expulsion, the embedded clusters will expand and subsequently interact and coalescence. Recent observational progress has highlighted the crucial role of early expansion in the evolution of young star clusters. By leveraging Gaia DR2 and DR3 data in combination with multi-object spectroscopy, studies have revealed expansion processes occurring in very young clusters within star-forming regions through both case studies (Cantat-Gaudin et al. 2019; Wright et al. 2019; Kuhn et al. 2020; Lim et al. 2020, 2022; Swiggum et al. 2021; Mužić et al. 2022; Das et al. 2023; Flaccomio et al. 2023) and statistical works (Kuhn et al. 2019; Della Croce et al. 2024; Jadhav et al. 2024; Wright et al. 2024).

Seven of the 17 MSFRs investigated in the MYStIX project are covered in the ATLASGAL survey (Schuller et al. 2009), which are the Lagoon nebula, NGC 6334, NGC 6357, the Eagle nebula, M 17, the Carina nebula, and the Trifid nebula. As shown in appendix A of Zhou et al. (2024a), the ATLASGAL + Planck 870 μm data (Csengeri et al. 2016) was used to trace the gas distribution around embedded or very young star clusters. On the maps, clusters and clumps are well separated. In stage3, feedback from the newly formed embedded clusters will disperse the surrounding gas, then these clusters will be located within the cavities of H II regions. Finally, the clusters or their coalescence remnants will be gas-free. In this work, we only focus on the coalescence of the formed embedded clusters following their expansion, which occurs in stage3.

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

The dynamics of both stars and gas have an important influence on the structure and evolution of embedded star clusters. A comprehensive simulation should combine both components. However, considering the limitations of computational resources, we have to simplify the simulation by focusing only one the first aspect, namely that of stellar dynamics. The complexity of the simulation arises from the interaction between the cluster and the surrounding gas during the formation and evolution of the cluster. If the evolution of the cluster is relatively independent of the gas, as in stage3, we can primarily focus on *N*-body simulations. Sills et al. (2018) presents simulations of embedded clusters that model the simultaneous evolution of both the stellar and gas components, beginning with initial conditions that are motivated and directly drawn from the MYStIX observations in Kuhn et al. (2014). By using live gas particles, they account for the gravitational interactions between the gas and stars within these clusters. Their results suggest that the subsequent evolution of these regions is largely insensitive to variations in the parameters within reasonable ranges. The initially highly sub-structured systems rapidly evolves into spherical, monolithic, and smooth configurations, both in terms of spatial and velocity distributions. The evolution of young embedded clusters appears to be primarily driven by gravitational interactions among the stars, with the surrounding gas having only a minor influence.

3.2 Coalescence of embedded clusters in simulations

3.2.1 N-body simulation

The parameters for the simulation are summarized in previous works, i.e. Kroupa, Aarseth & Hurley (2001), Baumgardt & Kroupa (2007), Banerjee, Kroupa & Oh (2012), Banerjee & Kroupa (2013, 2014, 2015), Oh, Kroupa & Pflamm-Altenburg (2015), Oh & Kroupa (2016, 2018), Banerjee & Kroupa (2017); Brinkmann et al. (2017), Pavlik, Kroupa & Šubr (2019), Wang, Kroupa & Jerabkova (2019), Dinnbier et al. (2022), Zhou et al. (2024a), Zhou, Kroupa & Wu (2024h). Previous simulations have demonstrated the effectiveness and appropriateness of the parameter settings (see below). The impact of varying parameter settings on simulation outcomes, as well as the exploration of multidimensional parameter space, are also discussed in the referenced literature. In this study, we primarily apply the established simulation framework to interpret observational data. For a detailed description and discussion of the initial conditions and parameters of the simulation can refer to Zhou et al. (2024a). Here, we provide only a brief overview.

The initial density profile of all clusters is the Plummer profile (Aarseth, Henon & Wielen 1974; Heggie & Hut 2003; Kroupa 2008). The half-mass radius $r_{\rm h}$ of the cluster is given by the $r_{\rm h}$–$M_{\rm ecl}$ relation (Marks & Kroupa 2012). All clusters are fully mass segregated ($S = 1$), no fractalization, and in the state of virial equilibrium ($Q = 0.5$). S and Q are the degree of mass segregation and the virial ratio of the cluster, respectively. More details can be found in Küpper et al. (2011) and the user manual for the MCLUSTER code. Discussions regarding these parameters can be found in Banerjee & Kroupa (2013), Zhou et al. (2024h). The initial mass functions of the clusters are chosen to be canonical (Kroupa 2001) with the most massive star following the $m_{\rm max}$–$M_{\rm ecl}$ relation of Weidner, Kroupa & Pflamm-Altenburg (2013). We assume the clusters to be at solar metallicity, i.e. $Z = 0.02$ (von Steiger & Zurbuchen 2016). The clusters traverse circular orbits within the Galaxy, positioned at a Galactocentric distance of 8.5 kpc, moving at a speed of 220 km s^{-1}. The initial binary setup follows the method described in Wang et al. (2019). All stars are initially in binaries, i.e. $f_{\rm b} = 1$, where $f_{\rm b}$ is the primordial binary fraction.

Modelling gas removal from embedded clusters poses significant challenges due to the complexities of radiation hydrodynamical processes, which involve intricate and uncertain physical mechanisms. To simplify the approach, we simulate the primary dynamical effects of gas expulsion by applying a diluting, spherically symmetric external gravitational potential to a model cluster, following the methods of Kroupa et al. (2001) and Banerjee & Kroupa (2013). This analytical method receives partial validation from Geyer & Burkert (2001), who conducted comparison simulations using smoothed particle hydrodynamics to represent the gas. Additionally, hydrodynamics + *N*-body simulations by Farias et al. (2024) affirm that the exponential decay model described in equation (1) effectively captures the gas removal process driven by radiation and wind feedback. Specifically, we adopt the potential of a spherically symmetric, time-dependent mass distribution

$$M_{\rm g}(t) = M_{\rm g}(0)\, t \leq \tau_{\rm d},$$
$$M_{\rm g}(t) = M_{\rm g}(0) \exp\left(-\frac{(t-\tau_{\rm d})}{\tau_{\rm g}}\right) t > \tau_{\rm d}. \quad (1)$$

Here, $M_{\rm g}(t)$ is the total mass in gas which is spatially distributed with the Plummer density distribution (Kroupa 2008) and starts depleting with time-scale $\tau_{\rm g}$ after a delay of $\tau_{\rm d}$. The Plummer radius of the gas distribution is kept time-invariant (Kroupa et al. 2001; Zhou et al. 2024a). We use an average velocity of the outflowing gas of $v_{\rm g} \approx 10\,{\rm km\,s^{-1}}$, which is the typical sound-speed in an H II region. This gives $\tau_{\rm g} = r_{\rm h}(0)/v_{\rm g}$, where $r_{\rm h}(0)$ is the initial half-mass radius of the cluster. As for the delay-time, we take the representative value of $\tau_{\rm d} \approx 0.6$ Myr (Kroupa et al. 2001), this being about the lifetime of the ultracompact H II region. In this work, we assume a star formation efficiency SFE ≈ 0.33 as a benchmark (Lada & Lada 2003; Megeath et al. 2016; Zhou et al. 2024b, 2024c), i.e. $M_{\rm g}(0) = 2M_{\rm ecl}(0)$.

The MCLUSTER program (Küpper et al. 2011) was used to set the initial configurations. The dynamical evolution of the model clusters was computed using the state-of-the-art PETAR code (Wang et al. 2020). PETAR employs well-tested analytical single and binary stellar evolution recipes (SSE/BSE; Hurley, Pols & Tout 2000; Hurley, Tout & Pols 2002; Giacobbo, Mapelli & Spera 2018; Banerjee et al. 2020; Tanikawa et al. 2020).

3.2.2 Initial conditions

For the 17 MSFRs investigated in the MYStIX project, nine of them are subcluster complexes. As shown in Kuhn et al. (2014), some of these nine MSFRs exhibit similar morphologies. Finally, we selected NGC 1893, NGC 6334 and the Carina nebula as representatives. In these three MSFRs, the spatial distribution of their subclusters varies, particularly with significant differences in the total mass of the subclusters, which helps to fit the observed physical parameters of OCs. We have identified the subclusters in these three MSFRs using the dendrogram algorithm according to the surface density distributions of stars in Zhou et al. (2024g), and calculated the physical parameters of the subclusters. Considering that the mass-radius relation of embedded clusters can be well fitted by the ≈1 Myr expanding line in Zhou et al. (2024h), we simulated the evolution of the subclusters in each MSFR for the first 1 Myr using the recipe described above. The initial masses of the subclusters strictly follow the observed values, and each subcluster is simulated individually. Then, all subclusters were collected together to create an initial configuration similar to the observation, as shown in Fig. 1. The

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

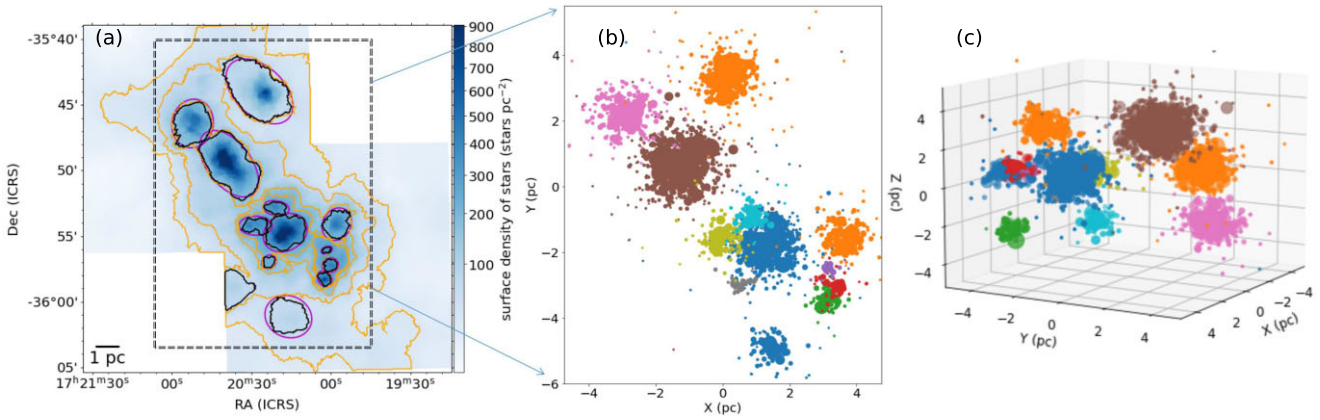


Figure 1. The initial conditions of the simulations drawn from the observations of NGC 6334. (a) Background is the surface density of stars. Black contours show the masks of embedded clusters identified by the dendrogram algorithm in Zhou et al. (2024g). Magenta ellipses are the approximate ellipses that encompass the embedded clusters; (b) The collection of all simulated embedded clusters which are used to create an initial configuration of the embedded cluster complex similar to the observation; (c) Considering spatial separations between embedded clusters based on panel (b). In panels (b) and (c), the size of the dot scales with the stellar mass using an arbitrary scaling.

separations between subclusters in Fig. 1(b) strictly adhere to the observed values.

3.2.3 *Gas expulsion*

The physical processes involved in stage1, stage2, and part of stage3 lead to the current spatial and mass distributions of clumps and star clusters in MSFRs that we observe today. In this work, our simulations start directly from the current state of MSFRs. Star clusters have separated from the gas (see appendix A of Zhou et al. 2024a) and their subsequent evolution is primarily driven by the dynamical interactions within and between star clusters, such that we can focus solely on the N-body dynamics. However, the evolution of MSFRs is a continuous process, and events that occurred in the past still have an impact on the present. The main physical process related to the gas in stage3 is the initial gas expulsion due to feedback, which also provides the initial conditions for the subsequent evolution of the cluster in N-body simulations. The initially embedded clusters are gravitationally bound by the surrounding gas. After the gas expulsion, the clusters will expand and subsequently interact and coalescence. Modelling the gas expulsion process is crucial for understanding the current physical state of the star clusters.

As discussed in Zhou et al. (2024c) and Zhou et al. (2024a), more massive clumps are capable of forming more massive clusters, which results in stronger feedback and higher gas expulsion velocity. Consequently, we expect correlations between the strength of feedback, the mass of the clumps (or clusters), and the gas expulsion velocity. Low-mass clusters should experience slower gas expulsion compared to high-mass clusters. The SFE determines the total amount of the remaining gas, which also influences the time-scale of gas expulsion. Additionally, the complex, non-spherical spatial distribution of gas can further alter this time-scale. In summary, the uncertainties in these parameters (gas expulsion velocity, SFE and spatial distribution of gas) can ultimately be factored into the gas expulsion time-scale. Moreover, the amount of residual gas not only influences the expulsion process but also determines the gravitational potential of the gas. As discussed in Zhou et al. (2024a), a lower SFE leads to more residual gas, which drives more intense expansion and greater mass loss in the cluster after gas expulsion. Essentially, a lower SFE is equivalent to a shorter gas expulsion time-scale. Thus, apart from the fast gas expulsion with the gas depletion time of $\tau_{\rm g}$, we also simulated a moderate gas expulsion with the gas depletion time of $5\tau_{\rm g}$ for each embedded cluster.

3.2.4 *Spatial distribution and velocity difference*

Currently, we assume all embedded clusters are initially at rest and located on the same plane. A more realistic scenario is that clusters have relative velocities and line-of-sight spatial separations. To take these two factors into account, we approximate the molecular cloud containing the clusters as an ellipsoid. To derive the lengths of the three axes of the ellipsoid (semimajor axis, a, intermediate axis, b, semiminor axis, c), we fit the coordinates of the clusters located at the periphery to determine a and b. Given that the real cloud shape should be more sheet-like instead of spherical (Shetty & Ostriker 2006; Inutsuka et al. 2015; Arzoumanian et al. 2018, 2022; Kohno et al. 2021; Rezaei Kh. & Kainulainen 2022; Clarke et al. 2023; Ganguly et al. 2023; Zhou et al. 2023a, b), we assume $c = b$. Then we randomly distribute clusters along the Z-axis (line-of-sight). As shown in Fig. 1, Figs 2 and 3, the spatial distribution of clusters is nearly spherical. Therefore, we have overestimated the spacing between clusters.

For molecular clouds in the catalogue of Miville-Deschênes, Murray & Lee (2017) across the Galactic plane, the most probable cloud size is ≈30 pc, and the most probable value of the velocity dispersion is ≈1.95 km s^{-1} on the scale of clouds. In our case, the sizes of the MSFRs are ≈5 pc for NGC 6334 and NGC 1893, and ≈20 pc for the Carina MSFR. As a conservative estimate, we take the velocity dispersion of the original molecular clouds in the three MSFRs as 2 km s^{-1} and assume that the clusters inherit the velocity dispersion of the clouds. The system's centre has a velocity of 0 km s^{-1}, and the velocity of the outermost cluster is 2 km s^{-1}. The velocities of other clusters are distributed according to the Larson relation, i.e. $v \propto L^{0.5}$, where L is the distance to the system's centre.

3.2.5 *Simulation results analysis*

The main objective of this study is to assess whether the post-gas expulsion coalescence of embedded clusters can explain the observed physical parameters of OCs. Fig. A1, Fig. A2, Fig. A3,

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

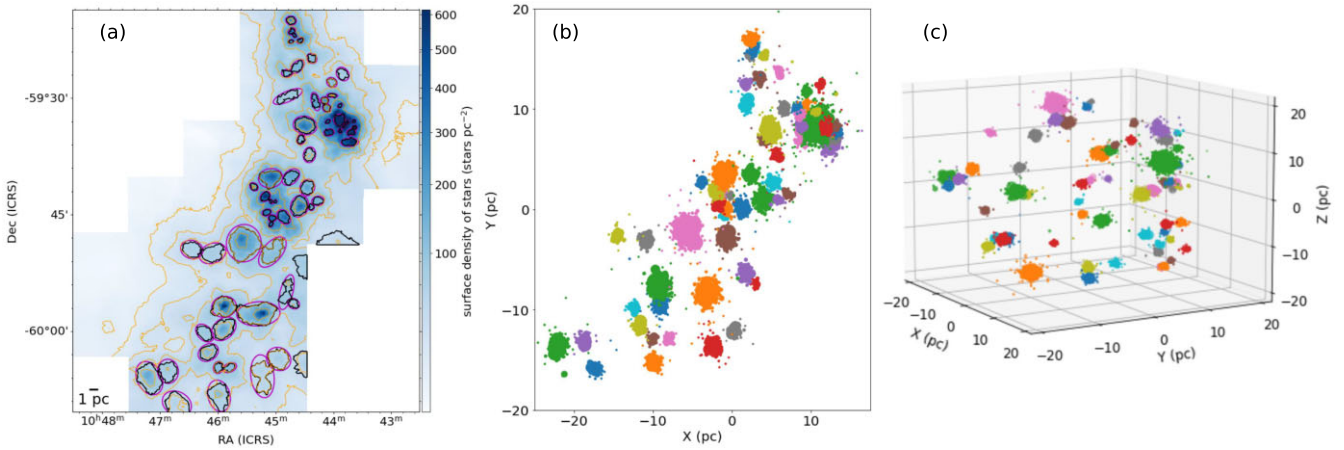


Figure 2. Same as Fig. 1, but for Carina.

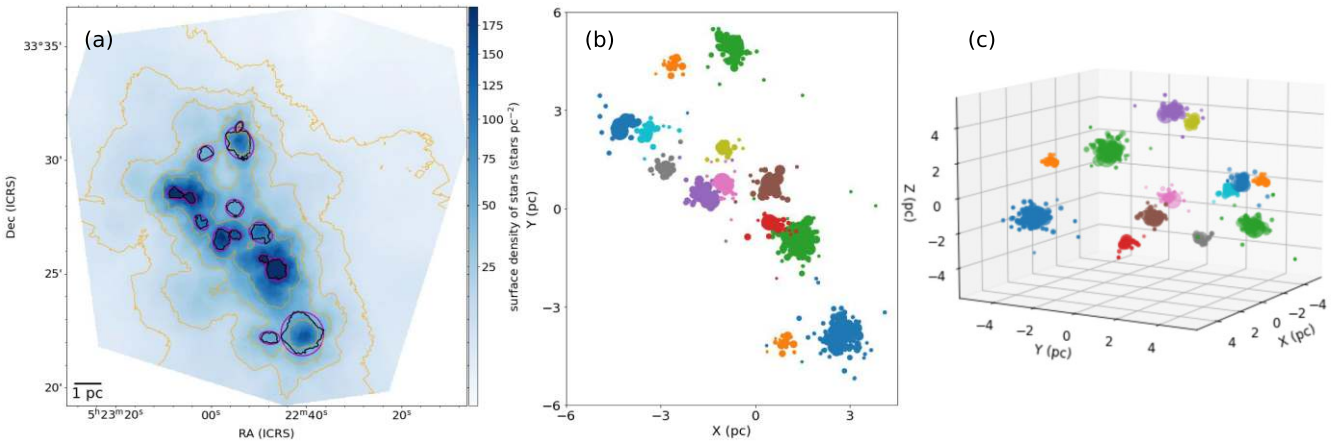


Figure 3. Same as Fig. 1, but for NGC 1893.

Fig. A4, Figs A5 and A7 present the evolution of the embedded cluster complexes over time. In the simulations, before 1 Myr, the evolution of each cluster is approximately independent. After further expansion, they start to interact upon contact, leading to subsequent mutual influence in their evolution. The spatial distribution, relative velocities, mass distribution, and gas expulsion modes of embedded clusters will affect the dynamics of their coalescence as well as the stability of the coalesced products. We simulated all possible cases listed in Table 1 for NGC 1893. The embedded clusters in NGC 1893 have low mass (see Fig. 4) and are few in number (13). They will quickly disperse after expansion regardless of the initial conditions, as shown in Fig. 5(a) and Fig. A1. There may not be a real coalescence between them, but rather the formation of a loose association.

For NGC 6334, we simulated four cases, i.e. 'fast', 'fast-vd', 'fast-vd-v0', and 'moderate-vd'. Overall, they produce similar results, but there are also differences. The line-of-sight spatial separations and relative velocities between embedded clusters ('fast-vd') lead to a larger radius and faster mass loss of the coalesced cluster, but this effect can be counteracted by a slower gas expulsion of the embedded clusters ('moderate-vd'). Therefore, there is a degeneracy between the line-of-sight spatial separations and the relative velocities between embedded clusters and the gas expulsion mode of embedded clusters. Thus, 'moderate-vd' and 'fast' give comparable simulation results, as shown in Figs 5(c) and (d). Comparing 'fast-vd' and 'fast-vd-v0', we can see that the line-of-sight spatial separations are more important than the relative velocities in determining the mass loss and radius of the coalesced cluster.

NGC 6334 and NGC 1893 both have a small number of embedded clusters, but NGC 6334 contains more massive embedded clusters (see Fig. 4). By comparing Figs A1 and A2, it is evident that the mass of embedded clusters plays a critical role in the coalescence process. Massive embedded clusters provide a more stable gravitational core, significantly aiding the coalescence and enabling the resulting cluster to remain stable within the Galactic tidal field. Strong gravitational forces between the embedded cluster complexes are essential to counteract the system's expansion and the Galactic tidal influence, ensuring the coalescence process continues. This strong local gravitational field is predominantly generated by the massive embedded clusters. In Fig. A2, we can see the coalescence between embedded clusters in NGC 6334.

Most of the embedded clusters in the Carina complex are low-mass ($<100\,M_\odot$, see Fig. 4), but they are numerous and densely distributed. Additionally, the Carina also contains a few massive embedded clusters ($>300\,M_\odot$). Therefore, even though the embedded clusters in Carina were initially given large line-of-sight spatial separations (the maximum $\sim$40 pc), they are still able to coalesce after expansion, which is significantly different from the case of NGC 1893. However, this coalescence complex is not stable enough and gradually dissipates in the Galactic tidal field, as shown in Fig. A7. The large line-of-sight spatial separations lead to insufficient

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

Table 1. 'Terminated time' is the time when the simulation is terminated. The cluster may or may not have dissolved by this time. We consider two types of gas expulsion modes, i.e. 'fast' (τ_g) and 'moderate' ($5\tau_g$). τ_g is the gas depletion time in equation 1. 'Yes' or 'no' indicate whether spatial separations or relative velocities between embedded clusters are accounted for in the simulations. Both 'NGC1893-fast-vd-d0.5' and 'NGC1893-fast-vd' incorporate spatial separations between embedded clusters, but the separation of the former is half of the latter.

Cases	Gas expulsion	Spatial separation	Relative velocity	Terminated time (Myr)
NGC1893-fast	Fast	No	No	100
NGC1893-fast-vd	Fast	Yes	Yes	100
NGC1893-fast-vd-v0	Fast	Yes	No	100
NGC1893-fast-vd-d0.5	Fast	Yes, $*0.5$	Yes	100
NGC1893-moderate-vd	Moderate	Yes	Yes	100
NGC6334-fast	Fast	No	No	614.25
NGC6334-fast-vd	Fast	Yes	Yes	463.75
NGC6334-fast-vd-v0	Fast	Yes	No	100
NGC6334-moderate-vd	Moderate	Yes	Yes	259.25
Carina-fast	Fast	No	No	556.5
Carina-fast-vd	Fast	Yes	Yes	324.5

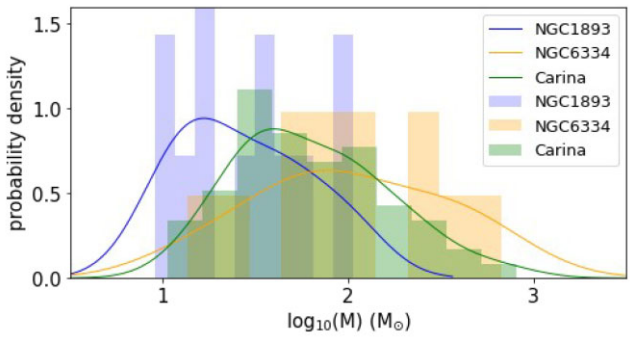


Figure 4. Mass distribution of embedded clusters in three MSFRs.

coalescence between embedded clusters, preventing the formation of a stable core. In contrast, since there is no line-of-sight spatial separation in the case of 'Carina-fast', the embedded clusters are able to sufficiently coalesce, forming a stable cluster core that can persist for over 300 Myr, as shown in Fig. A5.

3.3 Comparison with observations

The primary goal of this work is to evaluate whether the post-gas expulsion coalescence of embedded clusters can account for the observed physical parameters of OCs. Therefore, we only consider two extreme cases, i.e. 'fast' and 'fast-vd'. The simulation results from other parameter choices should lie in between these two extremes, as shown in Fig. 5.

As presented in Zhou et al. (2024a), to explain the mass and radius distributions of the observed OCs, initial embedded clusters with masses larger than 3000 $M_\odot$ are necessary. However, the upper limit of the embedded cluster sample collected in Zhou et al. (2024a) is less than 1000 $M_\odot$, and only few ATLASGAL clumps have a mass larger than 3000 $M_\odot$. The current mass distribution of clumps in the Milky Way suggests that the evolutionary process, where a single clump develops into an embedded cluster and eventually into an OC, cannot account for the observed OCs with relatively old ages and large masses. However, each molecular cloud typically contains multiple clumps, which are the local sites of star formation within the molecular cloud and the precursors of embedded clusters. As presented above, the formed embedded clusters can coalesce after expanding. If we require that OCs evolve directly from single embedded clusters, then the embedded cluster that serve as the precursor must have a sufficiently high mass to account for the observed physical parameters of OCs. However, the currently observed clump masses make it difficult to directly form massive embedded clusters. Nevertheless, a composite of multiple low-mass embedded clusters could, through coalescence, cover the parameter space observed in OCs, as shown in Fig. 6, where the mass and radius of the clusters are displayed as a function of their age. In the post-gas expulsion coalescence scenario, massive embedded clusters are not necessary. This is consistent with the current mass distribution of clumps in the Milky Way. We note that, as presented in Zhou et al. (2024a), the evolution of individual embedded clusters can also account for part of the observed parameter space of OCs.

4 DISCUSSION

The structures observed in the surface density maps of stars from the MYStIX project represent the projected distributions of stars within star-forming regions, leading to possible overlaps of physically separate star groups on the map. Additionally, some of the small dense embedded clusters might not be visible in the surface density maps or not be identified by the dendrogram algorithm. Moreover, we may underestimate the mass of the embedded clusters, due to assuming an average mass of 0.5 $M_\odot$ for all stars in each embedded cluster. However, in this work, we simply aim to obtain a reasonable spatial distribution of embedded clusters in MSFRs based on observations. This work does not seek to accurately represent the evolution of embedded clusters within NGC 6334, NGC 1893 and the Carina.

As described in Section 3.3, based on the current mass distribution of clumps and embedded clusters in the Milky Way, combining the evolution of individual embedded clusters and the coalescence of multiple embedded clusters can encompass the parameter space of the observed OCs. This work focuses on the coalescence of the formed embedded clusters after expansion. Actually, there is now extensive literature arguing that star clusters form through the mergers of subclusters, both from simulations and observations. The *N*-body simulations of Fujii et al. (2012) revealed that clusters formed through the hierarchical merging of multiple subclusters can account for the characteristics (such as mass distribution, mass segregation, runaway stars, massive stars, and massive binaries) observed in young dense clusters like NGC 3603 and Westerlund 1 and 2 in the Milky Way, as well as R136 in the Large Magellanic Cloud (LMC). The idea that runaway stars originate from subcluster mergers is also explored in Polak et al. (2024a). Sills et al. (2018) conducted simulations of young, embedded star clusters, modelling

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

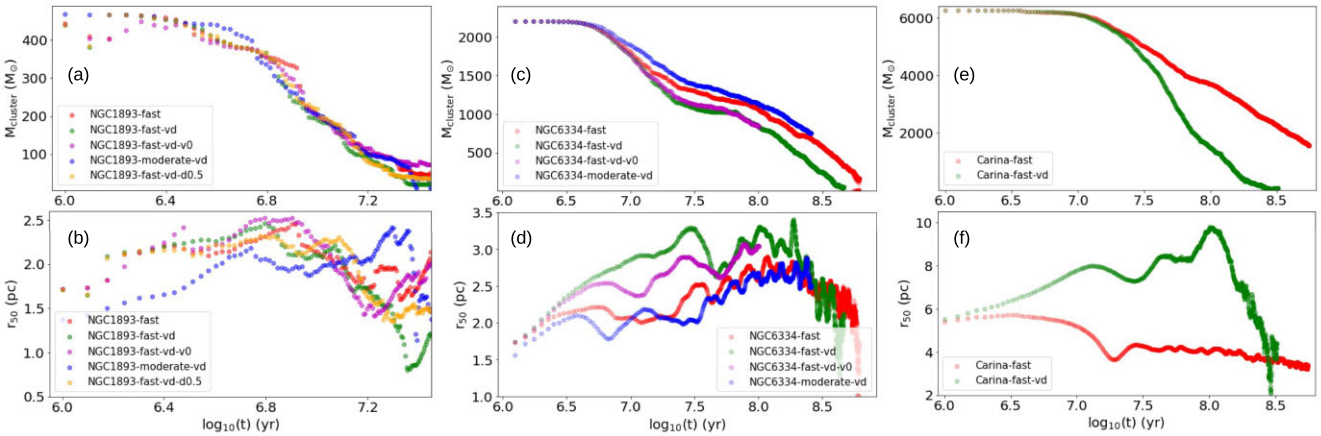


Figure 5. Mass and radius evolution of the embedded cluster complexes over time in different cases listed in Table 1. r_{50} is the radius containing 50 per cent of the members within the tidal radius of the cluster.

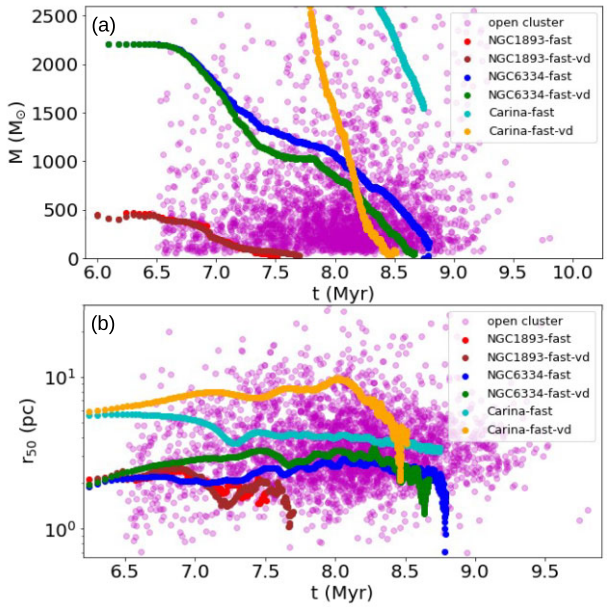


Figure 6. Fitting the physical parameters of OCs in Hunt & Reffert (2024) using the coalescence simulations. r_{50} is the radius containing 50 per cent of the members within the tidal radius of the OC. M is the mass of the OC.

the dynamical evolution of stars and gas simultaneously, with initial conditions derived directly from observations in the MYStIX project. They found that initially highly sub-structured systems rapidly evolved into spherical, monolithic, and smooth distributions in both spatial and velocity distributions, similar to those shown in Fig. A2, Fig. A3, Fig. A4, Figs A5, and A7. Although the simulations of Sills et al. (2018) and those presented in this work differ in some aspects, the results are similar, indicating that the coalescence of subclusters is not sensitive to the initial conditions.

On the observational side, Gennaro et al. (2011) proposed that the elongation and mass segregation observed in Westerlund 1 could be the result of a merger of multiple subclusters that formed nearly coevally within the parent giant molecular cloud. Sabbi et al. (2012) reported the identification of two distinct stellar populations in the core of the giant H II region 30 Doradus in the LMC. The differences in morphology and age between the two subclusters suggest that an ongoing merger may be taking place within the core of 30 Doradus. Della Croce et al. (2023) identified a massive, substructured stellar system (called 'LISCA II') in the Galactic Perseus complex that is likely undergoing hierarchical cluster assembly. This system consists of nine star clusters. They observed a dominant contraction pattern toward the centre of the system, primarily driven by the external regions of the clusters' complex, along with a milder central expansion. These observations align with expectations for the early evolutionary phases of stellar systems forming into a coherent massive structure, as predicted by their N-body models. In Fig. A5, we can see some similar physical processes described above. Another sub-structured stellar system (called 'LISCA I') was identified in Dalessandro et al. (2021), which is populated by seven co-moving clusters with signs of potential interactions. After comparing with direct N-body simulation, the authors suggest that LISCA I could be in an intermediate stage of cluster assembly, which may eventually evolve into a more massive stellar system. Although the subcluster merger simulations in Dalessandro et al. (2021) used very different initial conditions, the hierarchical cluster formation path presented in their work (fig. 7 of Dalessandro et al. 2021) is comparable with Figs A5 and A7 in this work. Thus, within a reasonable range of parameters, the coalescence of subclusters is a robust process, not sensitive to the initial conditions.

5 CONCLUSION

Based on the spatial and mass distributions of embedded clusters in three MSFRs from the observations (NGC 1893, NGC 6334, and the Carina Nebula), we simulated the evolution of the embedded clusters in each MSFR to ≈1 Myr, and then collected them together to create an initial configuration of the embedded cluster complex similar to that in the observation of each MSFR. In the simulations, before ≈1 Myr, the evolution of each embedded cluster is approximately independent. After further expansion, they start to interact upon contact, leading to subsequent mutual influence in their evolution. The spatial distribution, relative velocities, mass distribution and gas expulsion modes of embedded clusters will affect the dynamics of their coalescence process as well as the stability of the coalesced products. The large line-of-sight spatial separations lead to insufficient coalescence between embedded clusters, preventing the formation of a stable core. Greater spatial separations and higher relative velocities between embedded clusters lead to larger radius and faster mass loss of the coalesced star cluster, but this effect

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

can be counteracted by a slower gas expulsion of the embedded clusters. By comparing with other independent works, we conclude that, within a reasonable range of parameters, the coalescence of embedded clusters is a robust process, not sensitive to the initial conditions.

The embedded clusters in NGC 1893 have low masses and are few in numbers. They will quickly disperse after expansion regardless of the initial conditions of the simulations. There may not be a real coalescence between them, but rather the formation of a loose association. NGC 6334 and NGC 1893 both have a small number of embedded clusters, but NGC 6334 contains more massive embedded clusters, leading to the coalescence of embedded clusters in NGC 6334. Thus, the embedded cluster mass plays a critical role in the coalescence process. Massive embedded clusters provide a more stable gravitational core, significantly aiding the coalescence and enabling the resulting cluster to counteract the system's expansion and the Galactic tidal influence. By comparing the Carina and NGC 1893 MSFRs, it seems that a larger number of embedded clusters can also facilitate their coalescence.

The currently observed clumps in the Milky Way are not massive enough to directly form massive embedded clusters. However, a composite of multiple low-mass embedded clusters could, through coalescence, cover the parameter space of the observed OCs in the Milky Way. In the post-gas expulsion coalescence scenario, OCs do not necessarily require massive embedded clusters as precursors.

ACKNOWLEDGEMENTS

We would like to thank the referee for the detailed comments and suggestions, which have helped to improve and clarify this work.

DATA AVAILABILITY

All the data used in this work are available from the first author.

REFERENCES

Aarseth S. J., Henon M., Wielen R., 1974, A&A, 37, 183
Arzoumanian D., Shimajiri Y., Inutsuka S.-I., Inoue T., Tachihara K., 2018, PASJ, 70, 96
Arzoumanian D. et al., 2022, A&A, 660, A56
Banerjee S., Kroupa P., 2013, ApJ, 764, 29
Banerjee S., Kroupa P., 2014, ApJ, 787, 158
Banerjee S., Kroupa P., 2015, MNRAS, 447, 728
Banerjee S., Kroupa P., 2017, A&A, 597, A28
Banerjee S., Kroupa P., Oh S., 2012, ApJ, 746, 15
Banerjee S., Belczynski K., Fryer C. L., Berczik P., Hurley J. R., Spurzem R., Wang L., 2020, A&A, 639, A41
Baumgardt H., Kroupa P., 2007, MNRAS, 380, 1589
Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, MNRAS, 427, 127
Brinkmann N., Banerjee S., Motwani B., Kroupa P., 2017, A&A, 600, A49
Cantat-Gaudin T. et al., 2019, A&A, 626, A17
Clarke S. D., Sánchez-Monge Á., Williams G. M., Howard A. D. P., Walch S., Schneider N., 2023, MNRAS, 519, 3098
Cournoyer-Cloutier C. et al., 2023, MNRAS, 521, 1338
Csengeri T. et al., 2016, A&A, 585, A104
Dale J. E., Bonnell I. A., 2008, MNRAS, 391, 2
Dale J. E., Ercolano B., Bonnell I. A., 2012, MNRAS, 424, 377
Dalessandro E. et al., 2021, ApJ, 909, 90
Das S. R., Gupta S., Prakash P., Samal M., Jose J., 2023, ApJ, 948, 7
Della Croce A. et al., 2023, A&A, 674, A93
Della Croce A., Dalessandro E., Livernois A., Vesperini E., 2024, A&A, 683, A10
Dib S., 2023, MNRAS, 524, 1625
Dib S., Piau L., Mohanty S., Braine J., 2011, MNRAS, 415, 3439
Dib S., Gutkin J., Brandner W., Basu S., 2013, MNRAS, 436, 3727
Dinnbier F., Kroupa P., Anderson R. I., 2022, A&A, 660, A61
Dobbs C. L., Bending T. J. R., Pettitt A. R., Buckner A. S. M., Bate M. R., 2022, MNRAS, 517, 675
Fahrion K., De Marchi G., 2024, A&A, 681, A20
Farias J. P., Offner S. S. R., Grudić M. Y., Guszejnov D., Rosen A. L., 2024, MNRAS, 527, 6732
Feigelson E. D. et al., 2013, ApJS, 209, 26
Flaccomio E. et al., 2023, A&A, 670, A37
Fujii M. S., Saitoh T. R., Portegies Zwart S. F., 2012, ApJ, 753, 85
Fujii M. S., Wang L., Hirai Y., Shimajiri Y., Kumamoto J., Saitoh T., 2022, MNRAS, 514, 2513
Ganguly S., Walch S., Seifried D., Clarke S. D., Weis M., 2023, MNRAS, 525, 721
Gennaro M., Brandner W., Stolte A., Henning T., 2011, MNRAS, 412, 2469
Geyer M. P., Burkert A., 2001, MNRAS, 323, 988
Giacobbo N., Mapelli M., Spera M., 2018, MNRAS, 474, 2959
Guszejnov D., Markey C., Offner S. S. R., Grudić M. Y., Faucher-Giguère C.-A., Rosen A. L., Hopkins P. F., 2022, MNRAS, 515, 167
Heggie D., Hut P., 2003, The Gravitational Million-Body Problem: a Multidisciplinary Approach to Star Cluster Dynamics. Cambridge Univ. Press, Cambridge, p. 372
Howard C. S., Pudritz R. E., Harris W. E., 2018, Nat. Astron., 2, 725
Hunt E. L., Reffert S., 2023, A&A, 673, A114
Hunt E. L., Reffert S., 2024, A&A, 686, A42
Hurley J. R., Pols O. R., Tout C. A., 2000, MNRAS, 315, 543
Hurley J. R., Tout C. A., Pols O. R., 2002, MNRAS, 329, 897
Inutsuka S.-i., Inoue T., Iwasaki K., Hosokawa T., 2015, A&A, 580, A49
Jadhav V. V., Kroupa P., Wu W., Pflamm-Altenburg J., Thies I., 2024, A&A, 687, A89
Kohno M. et al., 2021, PASJ, 73, S129
Kroupa P., 1995a, MNRAS, 277, 1491
Kroupa P., 1995b, MNRAS, 277, 1507
Kroupa P., 2001, MNRAS, 322, 231
Kroupa P., 2005, in Turon C., O'Flaherty K. S., Perryman M. A. C., eds, Proceedings of the Gaia Symposium (ESA SP-576), The Three-Dimensional Universe with Gaia. Observatoire de Paris-Meudon, Meudon, France, p. 629
Kroupa P., 2008, in Aarseth S. J., Tout C. A., Mardling R. A., eds, Lecture Notes in Physics, Vol. 760, The Cambridge N-Body Lectures. Springer-Verlag, Berlin, p. 181
Kroupa P., Aarseth S., Hurley J., 2001, MNRAS, 321, 699
Kuhn M. A. et al., 2014, ApJ, 787, 107
Kuhn M. A., Getman K. V., Feigelson E. D., 2015, ApJ, 802, 60
Kuhn M. A., Hillenbrand L. A., Sills A., Feigelson E. D., Getman K. V., 2019, ApJ, 870, 32
Kuhn M. A., Hillenbrand L. A., Carpenter J. M., Avelar Menendez A. R., 2020, ApJ, 899, 128
Küpper A. H. W., Maschberger T., Kroupa P., Baumgardt H., 2011, MNRAS, 417, 2300
Lada C. J., Lada E. A., 2003, ARA&A, 41, 57
Lewis S. C. et al., 2023, ApJ, 944, 211
Li G.-X., 2024, MNRAS, 532, 1126
Lim B., Hong J., Yun H.-S., Hwang N., Kim J. S., Lee J.-E., Park B.-G., Park S., 2020, ApJ, 899, 121
Lim B., Nazé Y., Hong J., Yoon S.-y., Lee J., Hwang N., Park B.-G., Lee J.-E., 2022, AJ, 163, 266
Marks M., Kroupa P., 2012, A&A, 543, A8
Megeath S. T. et al., 2016, AJ, 151, 5
Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, ApJ, 834, 57
Motte F., Bontemps S., Louvet F., 2018, ARA&A, 56, 41
Mužić K., Almendros-Abad V., Bouy H., Kubiak K., Peña Ramírez K., Krone-Martins A., Moitinho A., Conceição M., 2022, A&A, 668, A19
Oh S., Kroupa P., 2016, A&A, 590, A107
Oh S., Kroupa P., 2018, MNRAS, 481, 153
Oh S., Kroupa P., Pflamm-Altenburg J., 2015, ApJ, 805, 92

Downloaded from https://academic.oup.com/mnras/article/537/2/845/7954750 by guest on 15 April 2025

MNRAS **000**, 000–000 (2024) Preprint 7 February 2025 Compiled using MNRAS LATEX style file v3.2

Origin of open clusters revealed by the evolution of the $m_{\rm max} - M_{\rm ecl}$ relation

J. W. Zhou ⋆[1] Pavel Kroupa, [2,3]
[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany*
[2]*Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany*
[3]*Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Analysis of the Gaia DR3 open cluster catalog reveals that as clusters age, the $m_{\rm max} - M_{\rm cluster}$ relation of open clusters increasingly deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation observed in embedded clusters. Over time, this relation develops a clear stratification based on cluster age. Older open clusters tend to exhibit a higher maximum mass limit for $M_{\rm cluster}$ compared to younger ones. If open clusters originate from individual embedded clusters, relatively old and massive open clusters should trace back to earlier epochs, as younger counterparts with similar properties are absent. Alternatively, the formation of such clusters may result from the coalescence of multiple embedded clusters after gas expulsion, which negates the requirement for massive embedded clusters as direct progenitors. Given the current mass distribution of clumps in the Milky Way, it is also impossible for those massive open clusters' precursors to form today. We compared the observed $m_{\rm max} - M_{\rm cluster}$ relations in open clusters with results from N-body simulations modeling both the evolution of individual clusters and the coalescence of multiple clusters. The simulations reveal that the mode of gas expulsion has a significant impact on the $m_{\rm max} - M_{\rm cluster}$ relation, as cluster mass loss is highly sensitive to the details of gas removal. Observationally, in open clusters younger than 10 Myr, $m_{\rm max}$ values are systematically lower than those predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation. Furthermore, open clusters around 10 Myr of age exhibit a clear detachment from the initial $m_{\rm max} - M_{\rm ecl}$ trend. These two observational facts align with the predictions from cluster coalescence simulations but contradict those from simulations of individual cluster evolution. This suggests that the post-gas expulsion coalescence of embedded clusters may be a dominant mechanism in the formation of open clusters, consistent with the findings of our previous studies.

Key words: – ISM: clouds – galaxies: structure – galaxies: star clusters – techniques: simulation

1 INTRODUCTION

The question of whether newborn stars in star clusters are drawn randomly from the Initial Mass Function (IMF) is critical for various fields of astrophysics. The stochastic (or random) sampling approach posits that each star's formation in an embedded cluster is an independent event, with its mass being randomly selected from a probability density function (PDF), effectively treating the IMF as a PDF. Conversely, the optimal sampling theory (Kroupa et al. 2013; Schulz et al. 2015) suggests a deterministic relationship between the mass of an embedded cluster and the mass of each individual star within it, due to a self-regulation star formation process in the cluster (Kroupa 2015; Yan et al. 2023). A key factor in determining whether a star cluster should be modeled using stochastic sampling or optimal sampling from the IMF is the presence of a significant relationship between the mass of the most massive star ($m_{\rm max}$) and the total mass of the embedded cluster in stars ($M_{\rm ecl}$). If a statistically significant correlation exists between $m_{\rm max}$ and $M_{\rm ecl}$, it would suggest that physical processes, such as self-regulation during star formation within the cluster, are at play. Actually, the $m_{\rm max} - M_{\rm ecl}$ relation has been well established analytically (Elmegreen 2000; Weidner & Kroupa 2004; Oey & Clarke 2005), observationally (Larson 2003; Weidner & Kroupa 2006; Weidner et al. 2010; Zhou et al. 2024a) and numerically (Bonnell et al. 2003, 2004; Peters et al. 2011). Yan et al. (2023) provided a deterministic tool for sampling stellar mass that reproduces the observed $m_{\rm max} - M_{\rm ecl}$ relation and aligns well with the luminosities of star-forming molecular clumps (also refer to Zhou et al. (2024c)). The embedded star cluster samples in Yan et al. (2023) are younger than 5 Myr. In this study, we explore how the initial $m_{\rm max} - M_{\rm ecl}$ relation for embedded clusters evolves over time, with a particular focus on the similarities and differences between this relation and the $m_{\rm max} - M_{\rm cluster}$ relation for open clusters. Furthermore, we discuss the indications of the temporal changes in the initial $m_{\rm max} - M_{\rm ecl}$ relation to the dynamics of cluster evolution and the formation processes of open clusters.

2 SAMPLE AND SIMULATION DATA

2.1 Sample

Hunt & Reffert (2023) conducted a blind, all-sky search for open clusters using 729 million sources from Gaia DR3 down to magnitude G≈20, creating a homogeneous catalog of 7167 clusters. By measuring cluster masses and Jacobi radii, the members in the star

⋆ E-mail: jwzhou@mpifr-bonn.mpg.de

© 2024 The Authors

Table 1. N-body simulations modeling both the evolution of individual clusters and the coalescence of multiple clusters under different gas expulsion modes. More details can refer to Zhou et al. (2024b) and Zhou et al. (2025). "Time" is the time when the simulation is terminated. The cluster may or may not have dissolved by this time.

Cases	Mass ($M_\odot$)	Time (Myr)	Dissolved
10000-fast	10000	263	yes
10000-moderate1	10000	100	no
10000-moderate2	10000	270	no
10000-slow	10000	359.25	no
3000-fast	3000	100	yes
3000-moderate1	3000	441	no
3000-moderate2	3000	256.25	yes
3000-slow	3000	950.25	yes
1000-fast	1000	67.5	yes
1000-moderate1	1000	145.5	yes
1000-moderate2	1000	78	yes
1000-slow	1000	533.5	yes
300-fast	300	69	yes
300-moderate1	300	100	yes
300-moderate2	300	45.25	yes
300-slow	300	346.5	yes
NGC1893-fast	468	305.75	yes
NGC1893-fast-vd	468	283.25	yes
NGC6334-fast	2207	614.25	yes
NGC6334-fast-vd	2207	463.75	yes
Carina-fast	6271	556.5	no
Carina-fast-vd	6271	324.5	yes

cluster catalog are classified into two categories, i,e, bound open clusters (OCs) and unbound moving groups (MGs) (Hunt & Reffert 2024). In Hunt & Reffert (2024), the catalog includes 5647 OCs and 1309 MGs, where 3530 OCs and 539 MGs are of high-quality with median color-magnitude diagram (CMD) classifications greater than 0.5 and signal-to-noise ratio (S/N) greater than 5σ (Hunt & Reffert 2023). In this work, we only focus on high-quality OCs and MGs.

To calculate the photometric masses of the clusters, Hunt & Reffert (2024) first derived the photometric masses of the member stars in each cluster using the PARSEC isochrones (Bressan et al. 2012). They then corrected for selection effects and applied a correction for unresolved binary stars. Subsequently, mass functions were fitted and integrated to determine the total cluster mass.

2.2 N-body simulations

In Zhou et al. (2024b), we computed the dynamical evolution of three clusters with the masses in stars of 300 $M_\odot$, 1000 $M_\odot$ and 3000 $M_\odot$ for the first 100 Myr under four different gas expulsion modes. In this work, we continue the simulations from the previous work until the clusters dissolve, as shown in Table.1. We also simulate a 10000 $M_\odot$ cluster using the same recipe as in Zhou et al. (2024b). In Zhou et al. (2025), we simulated the post-gas expulsion coalescence of embedded clusters formed within the same parental molecular cloud that undergo a phase of expansion caused by the effects of gas expulsion. In this work, we only consider two extreme cases in the coalescence simulation, i.e. "fast" and "fast-vd". The simulation results from other parameter choices should lie in between these two extremes, as presented in Zhou et al. (2025).

In simulations of both individual cluster and cluster coalescence, the initial mass functions of embedded clusters were chosen to be canonical (Kroupa 2001), with the most massive star following the

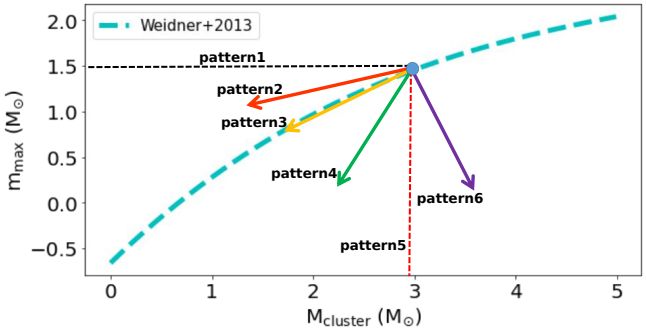


Figure 1. Evolutionary patterns of the $m_{\rm max} - M_{\rm ecl}$ relation over time.

observational $m_{\rm max} - M_{\rm ecl}$ relation of Weidner et al. (2013):

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \quad (1)$$

where $y = \log_{10}(m_{\rm max}/M_\odot)$, $x = \log_{10}(M_{\rm ecl}/M_\odot)$, a_0 = -0.66, a_1 = 1.08, a_2 = -0.150, and a_3 = 0.0084. In Zhou et al. (2024a), we also derived a theoretical $m_{\rm max} - M_{\rm ecl}$ relation,

$$M_{\rm ecl} \approx \frac{5.37}{\frac{0.77}{m_{\rm max}^{1.3}} - 0.001} - \frac{3.33}{m_{\rm max}^{0.3}\left(\frac{0.77}{m_{\rm max}^{1.3}} - 0.001\right)}. \quad (2)$$

This theoretical relation is used for comparison in the following discussion but is not applied in the simulations.

3 RESULTS AND DISCUSSION

3.1 Evolutionary patterns of the $m_{\rm max} - M_{\rm ecl}$ relation

In the simulations, an embedded cluster evolves from a point on the initial $m_{\rm max} - M_{\rm ecl}$ relation. During the evolution, both the mass of the most massive star in the cluster and the total mass of the cluster decrease over time. As illustrated in Fig.1, the possible evolutionary patterns are as follows: If the cluster's total mass decreases more rapidly, it will deviate above the initial $m_{\rm max} - M_{\rm ecl}$ relation (red arrow, pattern2); If they decrease at the same rate, the evolution will exactly follow the initial $m_{\rm max} - M_{\rm ecl}$ relation (orange arrow, pattern3); If the mass of the most massive star decreases more rapidly, it will deviate below the initial $m_{\rm max} - M_{\rm ecl}$ relation (green arrow, pattern4); The black (pattern1) and red (pattern5) dashed lines represent two extremes, where either the mass of the most massive star or the cluster's total mass does not change over time; The purple arrow (pattern6) represents an impossible evolutionary pattern for individual embedded clusters.

3.2 The $m_{\rm max} - M_{\rm cluster}$ relation in observation

We picked out the most massive star in each observed open cluster. Fig.2 displays the $m_{\rm max} - M_{\rm cluster}$ relations of open clusters in different age ranges. Overall, as the cluster ages increase, the $m_{\rm max} - M_{\rm cluster}$ relation gradually deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation, and ultimately manifests as age stratification. The clusters are gradually loosing the most massive stars through evolution.

The dataset in Weidner et al. (2013), used to fit the initial $m_{\rm max} - M_{\rm ecl}$ relation, requires clusters with age estimates of less than 5 Myr. However, some of these estimates come with significant uncertainties, suggesting that certain clusters could be as old as 9 Myr within the margin of error (Yan et al. 2023). In Fig.2(a), we focus on the

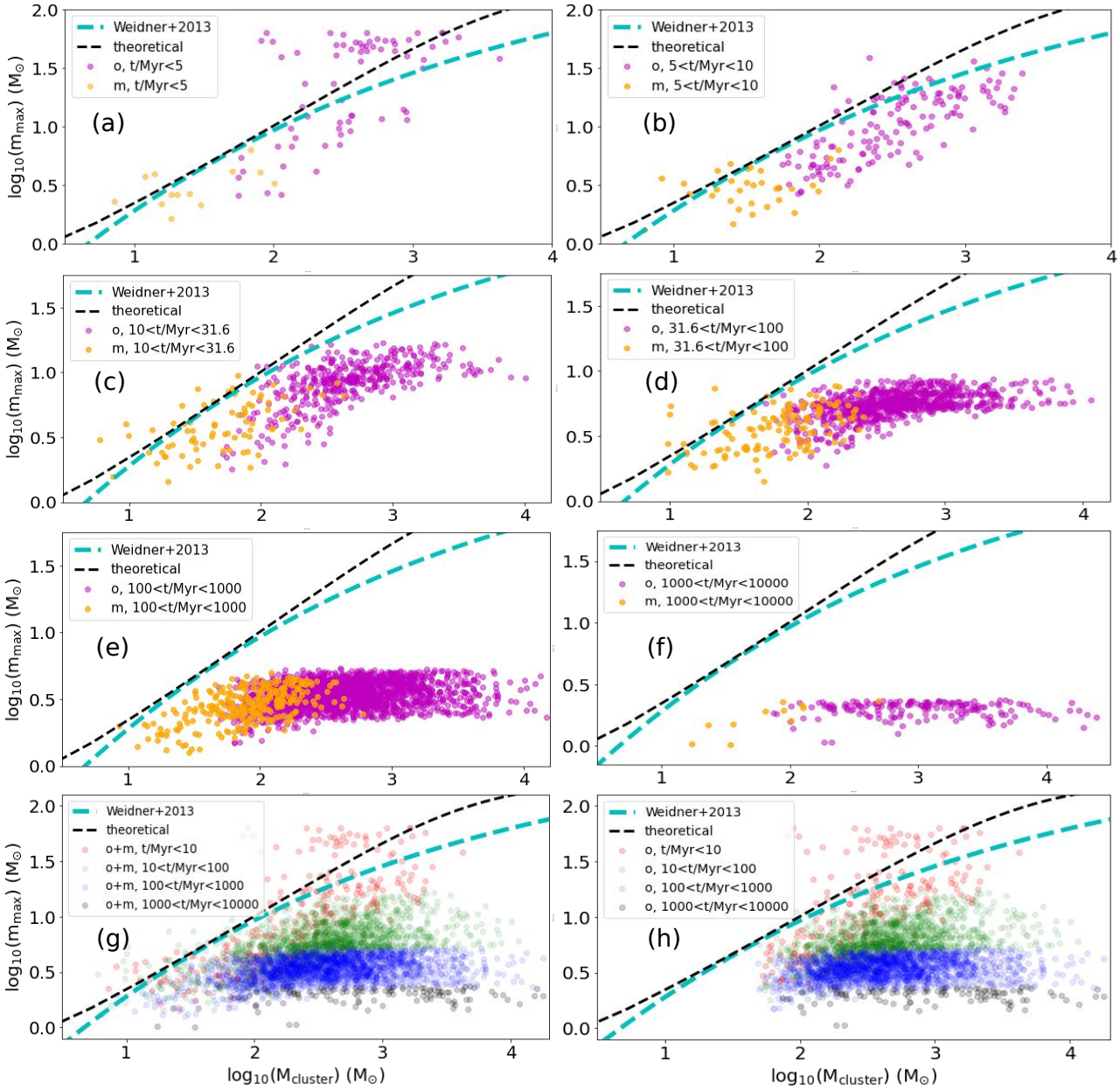


Figure 2. The $m_{\rm max} - M_{\rm cluster}$ relation of the observed open clusters evolves over time. The observational and theoretical $m_{\rm max} - M_{\rm ecl}$ relations of equation.1 and equation.2 correspond to the cyan and black dashed curves, respectively. "o" and "m" are the bound open clusters and unbound moving groups in the catalog of Hunt & Reffert (2024).

open clusters younger than 5 Myr. The $m_{\rm max} - M_{\rm cluster}$ relation of these clusters roughly follows the initial $m_{\rm max} - M_{\rm ecl}$ relation.

As shown in Fig.2(h), in different age ranges, the older the open clusters are, the higher the upper mass limit of $M_{\rm cluster}$, similar to the impossible "pattern6" evolutionary pattern for individual embedded clusters in Fig.1. One might argue that this is a selection effect, where only massive clusters can survive for a long time. If that's the case, these old, massive open clusters must originate from the past, as there are no corresponding younger progenitors in Fig.2(h). Given the current mass distribution of clumps in the Milky Way, it is also impossible for such massive open clusters' precursors to form today (Zhou et al. 2024b). However, as verified in Zhou et al. (2025), the post-gas expulsion coalescence of a group of embedded clusters can form massive open clusters. In this scenario, massive embedded clusters are not necessarily required as precursors for the formation of massive open clusters.

3.3 Comparison of observation and simulation

In Fig.2, the $m_{\rm max} - M_{\rm cluster}$ relation of the observed high-quality OCs and MGs is mainly located between "pattern3" and "pattern6" (see Fig.1). For the simulations shown in Fig.3, the gas expulsion modes can determine the $m_{\rm max} - M_{\rm cluster}$ relation, because the mass loss of the clusters strongly depends on the gas expulsion modes. From fast to slow gas expulsions, the $m_{\rm max} - M_{\rm cluster}$ relation shifts from "pattern2" or "pattern1" to "pattern5". To some extent, the mass of a cluster and the gas expulsion mode are equivalent. Em-

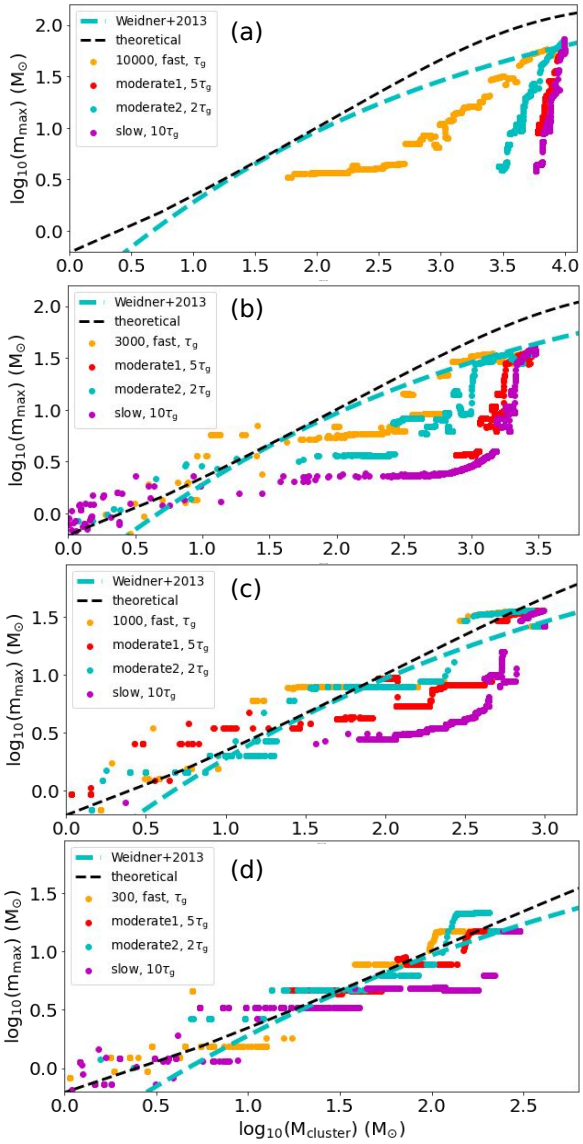


Figure 3. The initial $m_{max} - M_{ecl}$ relation evolves over time in individual cluster simulations. "fast", "moderate1", "moderate2" and "slow" are four different gas expulsion modes. More details can refer to Zhou et al. (2024b).

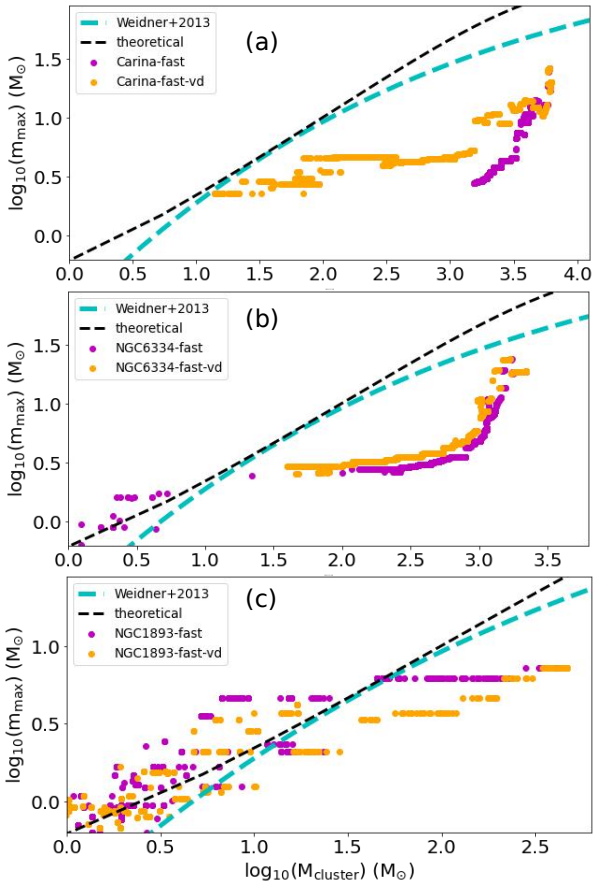


Figure 4. The initial $m_{max} - M_{ecl}$ relation evolves over time in cluster coalescence simulations. More details can refer to Zhou et al. (2025).

bedded clusters with different masses may achieve similar evolutionary trajectories on the $m_{max} - M_{cluster}$ diagram under different gas expulsion modes. For example, the evolutionary track on the $m_{max} - M_{cluster}$ diagram of a 1000 $M_\odot$ embedded cluster undergoing slow gas expulsion is similar to that of a 3000 $M_\odot$ embedded cluster with moderate gas expulsion. Comparing Fig.3 and Fig.5 with Fig.4 and Fig.6, although each subcluster in the coalescence simulation is in fast gas expulsion, the evolution of their coalesced product is similar to a cluster with slow gas expulsion in the individual cluster simulation.

For the fast gas expulsion, the clusters ($\lesssim$3000 $M_\odot$) cannot survive more than 100 Myr, as shown in Fig.5. And the $m_{max} - M_{ecl}$ relation evolves following "pattern3" or "pattern2" in Fig.3, which can only cover the upper boundary of the observed $m_{max} - M_{cluster}$ relation of open clusters in Fig.2. As shown in Fig.7 and Fig.8, a slow gas expulsion mode in individual cluster simulation is necessary to fit the observed $m_{max} - M_{cluster}$ relation of open clusters.

In Fig.7 and Fig.8, both observation and simulation show the mass of the most massive stars can be larger than the prediction of the initial $m_{max} - M_{ecl}$ relation. However, our simulations started from the initial $m_{max} - M_{ecl}$ relation of Weidner et al. (2013). It means the merger of stars during the cluster evolution. Similar results were also presented in Oh & Kroupa (2018), where they found that more than half of the simulated clusters host a merger product as its most massive member within the first 5 Myr of cluster evolution.

3.4 Individual cluster simulation vs. cluster coalescence simulation

In Fig.2, open clusters with the age around 10 Myr have clearly detached from the initial $m_{max} - M_{ecl}$ relation in observation. This is consistent with the coalescence simulation shown in Fig.9. However, in the individual cluster simulation, the clusters with the age around 10 Myr still roughly follow the initial $m_{max} - M_{ecl}$ relation, as presented in Fig.7 and Fig.8, which is inconsistent with the observation.

The youngest cluster in the catalog of Hunt & Reffert (2024) has the age of $\approx$2.9 Myr. In Fig.2, overall, when the age <10 Myr, m_{max} in open clusters are systematically lower than those predicted by the initial $m_{max} - M_{ecl}$ relation. This may not be due to the evolution of embedded clusters. In Fig.7 and Fig.8, even for clusters with ages around 10 Myr in individual cluster simulations, m_{max} still remains close to the initial $m_{max} - M_{ecl}$ relation, if they initially followed it as we set in the simulations. A possible explanation is that m_{max} in these clusters are smaller than those predicted by the initial

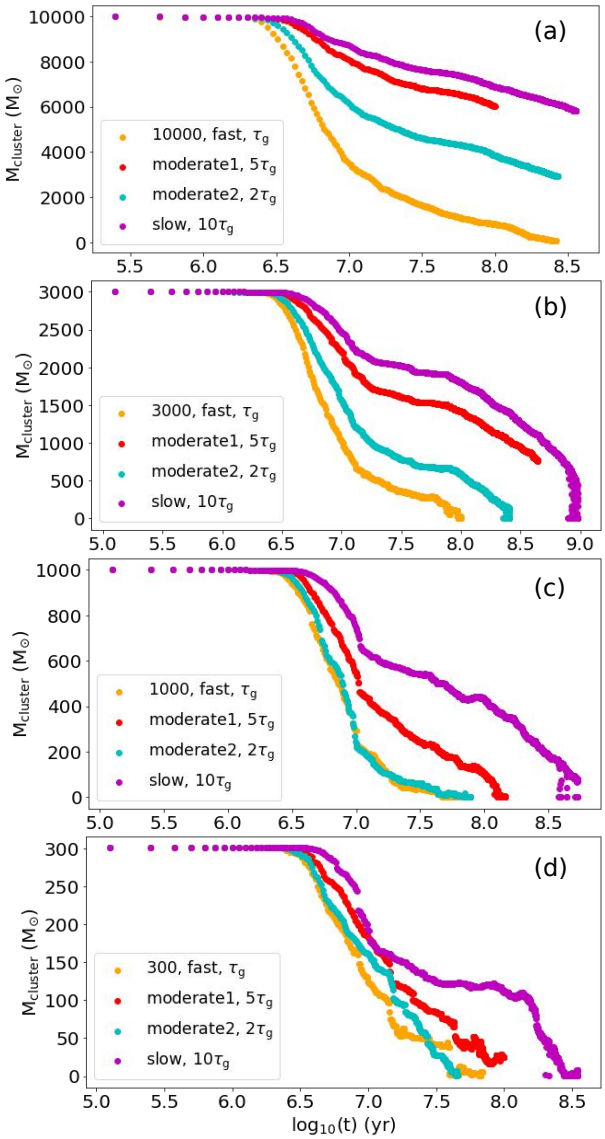


Figure 5. The mass of individual embedded cluster evolves over time in individual cluster simulations.

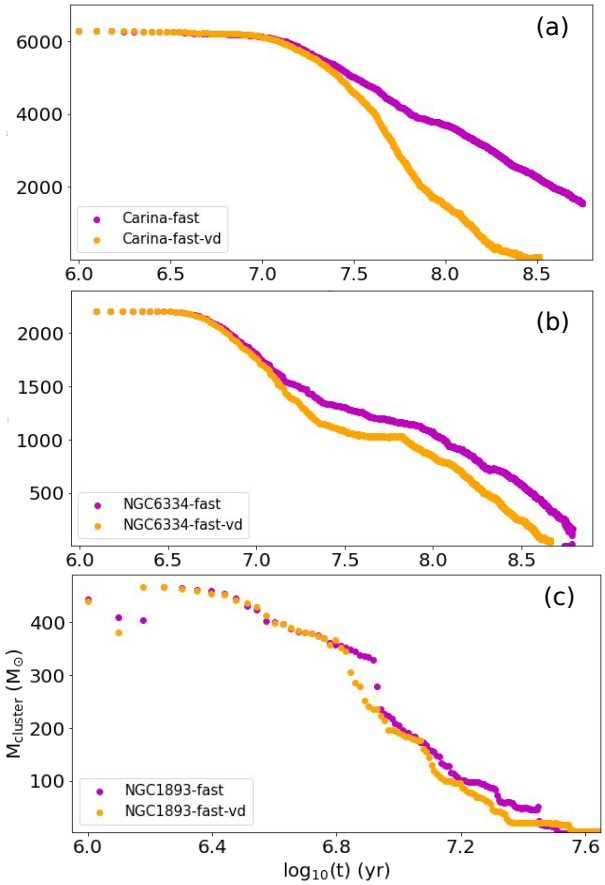


Figure 6. The mass of embedded cluster complex evolves over time in cluster coalescence simulations.

$m_{\rm max} - M_{\rm ecl}$ relation from the beginning. If open clusters are formed through the coalescence of embedded clusters, and each embedded cluster satisfies the initial $m_{\rm max} - M_{\rm ecl}$ relation, then $m_{\rm max}$ in the coalesced cluster will naturally fall below the value predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation, as shown in Fig.9.

In short, after comparing Fig.7, Fig.8 and Fig.9, we conclude that the coalescence simulation matches the observation better than the individual cluster simulation.

4 CONCLUSION

A crucial factor in deciding whether a star cluster should be modeled using stochastic or optimal sampling from the initial mass function (IMF) is the presence of a significant correlation between the mass of the most massive star ($m_{\rm max}$) and the total stellar mass of the embedded cluster ($M_{\rm ecl}$). Examination of the Gaia DR3 open cluster catalog indicates that as clusters evolve, the $m_{\rm max} - M_{\rm cluster}$ relation in open clusters progressively deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation seen in embedded clusters, ultimately exhibiting an age-dependent stratification. Across different age groups, older open clusters tend to have a higher upper mass limit of $M_{\rm cluster}$. If open clusters originate from individual embedded clusters, those relatively old and massive open clusters should form in earlier epochs, as no younger counterparts with similar properties are observed. Alternatively, massive open clusters may emerge from the post-gas expulsion coalescence of multiple embedded clusters, removing the need for massive embedded clusters as their direct progenitors. Furthermore, given the current mass distribution of clumps in the Milky Way, the formation of such massive open cluster precursors would not be feasible today (Zhou et al. 2024b).

We compared the observed $m_{\rm max} - M_{\rm cluster}$ relation of open clusters with results from N-body simulations of individual cluster evolution and cluster coalescence. In the individual cluster simulations, gas expulsion modes play a crucial role in shaping the $m_{\rm max} - M_{\rm cluster}$ relation, as cluster mass loss is highly dependent on the gas expulsion mechanism. In the case of fast gas expulsion, clusters with masses $\lesssim$3000 $M_\odot$ cannot survive beyond 100 Myr, and the resulting $m_{\rm max} - M_{\rm cluster}$ relation only matches the upper boundary of the observed relation for open clusters. Therefore, a slow gas expulsion mode is required to reproduce the observed $m_{\rm max} - M_{\rm cluster}$ relation of open clusters. When the age <10 Myr, $m_{\rm max}$ in open clusters is systematically lower than the values predicted by the initial $m_{\rm max} - M_{\rm ecl}$ relation. This discrepancy is unlikely to result from the natural evolution of embedded clusters. In the individual cluster simulation, even for clusters around 10 Myr old, $m_{\rm max}$ still remains

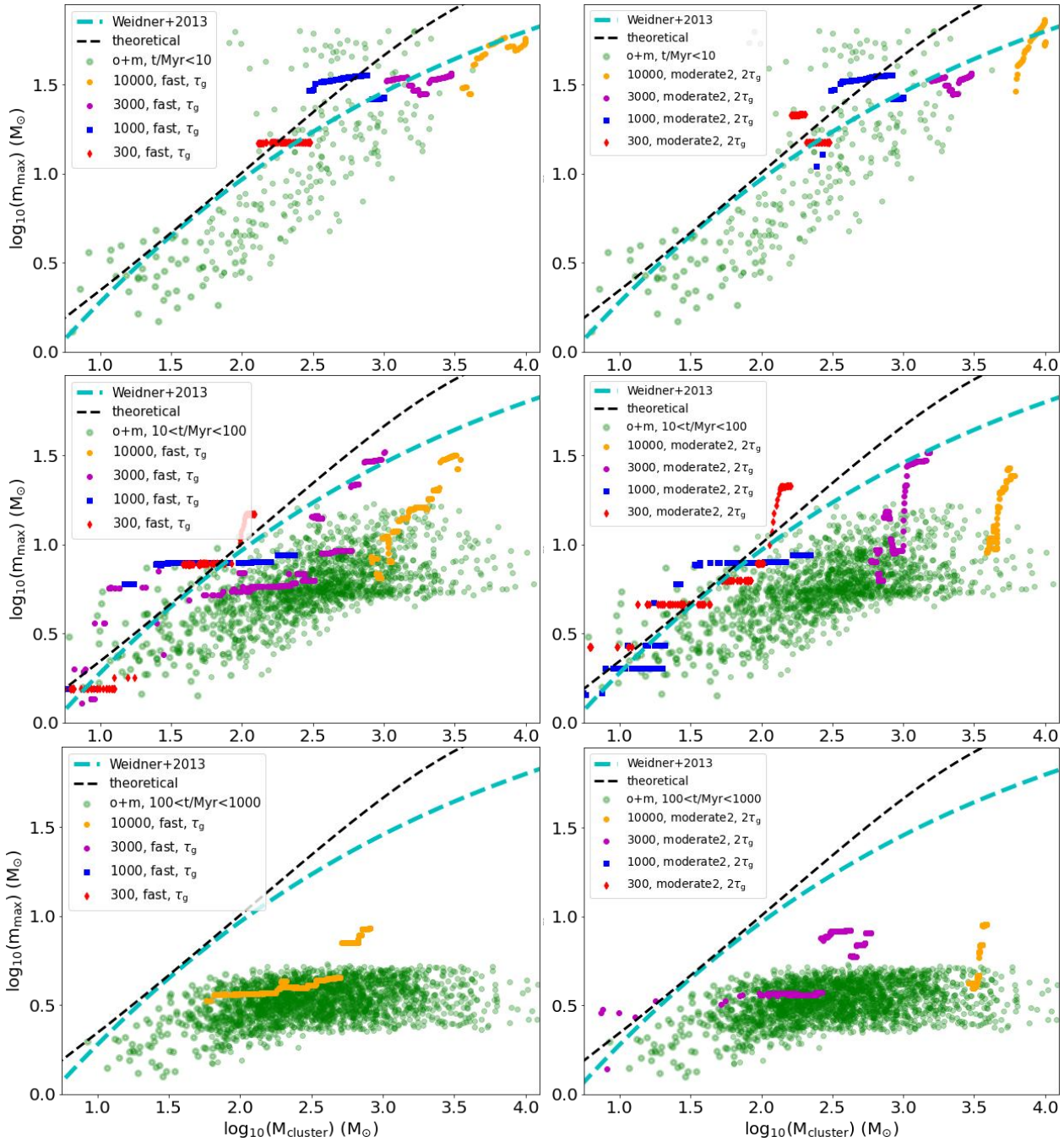


Figure 7. Fitting the observed $m_{max} - M_{cluster}$ relations of open clusters in different age ranges using the individual cluster simulations in the corresponding age ranges. The left and right panels represent "fast" and "moderate2" gas expulsion modes, respectively.

close to the initial $m_{max} - M_{ecl}$ relation. One possible explanation is that m_{max} in these clusters was already lower than that predicted by the initial $m_{max} - M_{ecl}$ relation from the beginning. This could indicate coalescence between embedded clusters, assuming that each individual embedded cluster strictly adheres to the initial $m_{max} - M_{ecl}$ relation. Moreover, observations show that open clusters around 10 Myr old have clearly deviated from the initial $m_{max} - M_{ecl}$ relation. However, in the individual cluster simulation, clusters of similar age still roughly follow this initial relation, which contradicts observations. These two inconsistencies in the individual cluster simulation are effectively resolved in the cluster coalescence simulation. This suggests that post-gas expulsion coalescence of embedded clusters should be a dominant mechanism in the formation of open clusters, aligning with the conclusions of Zhou et al. (2024b) and Zhou et al. (2025).

5 DATA AVAILABILITY

All the data used in this work are available from the first author.

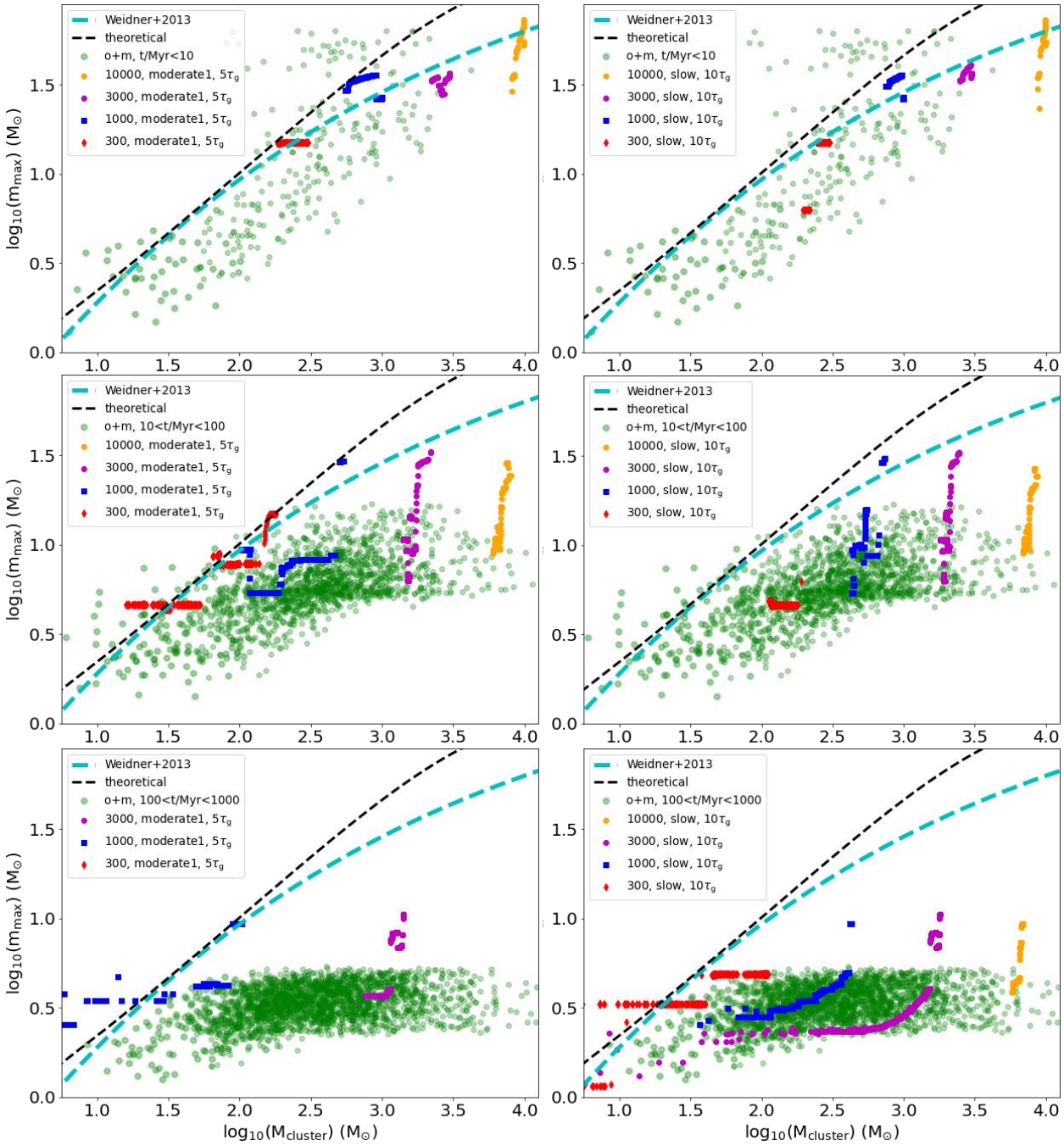


Figure 8. Same as Fig.7. The left and right panels represent "moderate1" and "slow" gas expulsion modes, respectively.

REFERENCES

Bonnell, I. A., Bate, M. R., & Vine, S. G. 2003, MNRAS, 343, 413, doi: 10.1046/j.1365-8711.2003.06687.x

Bonnell, I. A., Vine, S. G., & Bate, M. R. 2004, MNRAS, 349, 735, doi: 10.1111/j.1365-2966.2004.07543.x

Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127, doi: 10.1111/j.1365-2966.2012.21948.x

Elmegreen, B. G. 2000, ApJ, 539, 342, doi: 10.1086/309204

Hunt, E. L., & Reffert, S. 2023, A&A, 673, A114, doi: 10.1051/0004-6361/202346285

—. 2024, A&A, 686, A42, doi: 10.1051/0004-6361/202348662

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

—. 2015, Canadian Journal of Physics, 93, 169, doi: 10.1139/cjp-2014-0179

Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, Vol. 5, 115, doi: 10.1007/978-94-007-5612-0_4

Larson, R. B. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 287, Galactic Star Formation Across the Stellar Mass Spectrum, ed. J. M. De Buizer & N. S. van der Bliek, 65–80, doi: 10.48550/arXiv.astro-ph/0205466

Oey, M. S., & Clarke, C. J. 2005, ApJ, 620, L43, doi: 10.1086/428396

Oh, S., & Kroupa, P. 2018, MNRAS, 481, 153, doi: 10.1093/mnras/

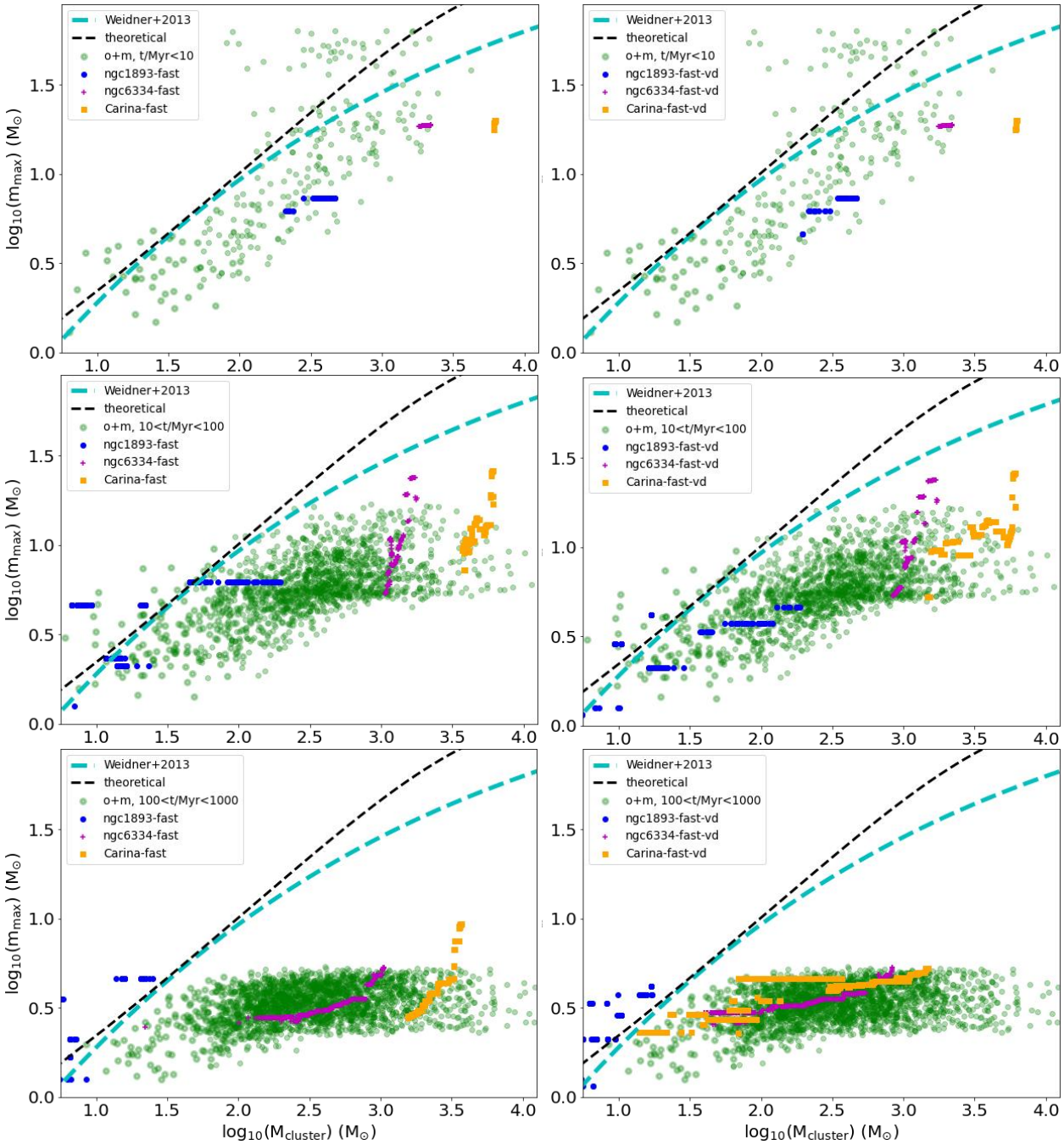


Figure 9. Fitting the observed $m_{\rm max} - M_{\rm cluster}$ relations of open clusters in different age ranges using the cluster coalescence simulations in the corresponding age ranges.

sty2245

Peters, T., Banerjee, R., Klessen, R. S., & Mac Low, M.-M. 2011, ApJ, 729, 72, doi: 10.1088/0004-637X/729/1/72

Schulz, C., Pflamm-Altenburg, J., & Kroupa, P. 2015, A&A, 582, A93, doi: 10.1051/0004-6361/201425296

Weidner, C., & Kroupa, P. 2004, MNRAS, 348, 187, doi: 10.1111/j.1365-2966.2004.07340.x

—. 2006, MNRAS, 365, 1333, doi: 10.1111/j.1365-2966.2005.09824.x

Weidner, C., Kroupa, P., & Bonnell, I. A. D. 2010, MNRAS, 401, 275, doi: 10.1111/j.1365-2966.2009.15633.x

Weidner, C., Kroupa, P., & Pflamm-Altenburg, J. 2013, MNRAS, 434, 84, doi: 10.1093/mnras/stt1002

Yan, Z., Jerabkova, T., & Kroupa, P. 2023, A&A, 670, A151, doi: 10.1051/0004-6361/202244919

Zhou, J. W., Dib, S., & Kroupa, P. 2024a, PASP, 136, 094302, doi: 10.1088/1538-3873/ad77f4

Zhou, J.-w., Dib, S., & Kroupa, P. 2024b, A&A, 691, A293, doi: 10.1051/0004-6361/202451728

—. 2025, MNRAS, 537, 845, doi: 10.1093/mnras/staf076

Zhou, J. W., Kroupa, P., & Dib, S. 2024c, PASP, 136, 094301, doi: 10.1088/1538-3873/ad6f44

Astronomy & Astrophysics manuscript no. output
February 26, 2025
©ESO 2025

The dependence of black hole formation in open clusters on the cluster formation process

Jian-Wen Zhou[1]

Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: jwzhou@mpifr-bonn.mpg.de

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

This study explores how the star cluster formation process influences the evolutionary outcomes of star clusters. Since black holes originate from massive stars, the method used to model a star cluster—whether through stochastic sampling or optimal sampling from the initial mass function (IMF)—directly determines the initial population of massive stars. The gas expulsion mode (timescale) of initially embedded star clusters play a crucial role in their dynamical state and subsequent evolution. These factors significantly influence the likelihood of massive stars and compact objects being ejected from the cluster, thereby affecting its black hole formation rate. Furthermore, different gas expulsion modes impact the survival time of star clusters, determining how long the evolutionary remnants of massive stars remain within the cluster environment. This, in turn, may have a profound effect on the formation of black hole multiple systems. Regarding the cluster formation mechanism, star clusters formed through the coalescence of multiple subclusters exhibit markedly different dynamical behaviors and evolutionary pathways compared to those formed as single entities. Given that the post-gas expulsion coalescence of embedded clusters within the same parental molecular cloud is likely a dominant mechanism in open cluster formation, it is promising to investigate black hole formation in different coalescence scenarios. A cluster's evolution is intrinsically linked to its formation process. Variations in cluster mass, gas expulsion modes, and formation mechanisms ultimately shape the cluster's expansion speed, which in turn influences its mass loss rate and black hole formation rate. Therefore, these physical factors should be carefully incorporated into the initial conditions of N-body simulations. Achieving this requires reconstructing a comprehensive physical picture of star cluster formation and evolution.

Key words. stars: black holes – stars: evolution – ISM: clouds – (Galaxy:) open clusters and associations: general

1. Introduction

Recent proper motion data from *Gaia*'s Data Release 3 (DR3; Gaia Collaboration et al. (2023)) has led to the discovery of two dormant black holes (BHs, i.e. *Gaia* BH1 and *Gaia* BH2) in BH–star binary systems within the Galactic field (Chakrabarti et al. 2023; El-Badry et al. 2023b,a; Tanikawa et al. 2023). As part of the validation process for the upcoming fourth Gaia data release (DR4), Gaia Collaboration et al. (2024) analyzed preliminary astrometric binary solutions and identified a black hole with a mass of ∼33 $M_\odot$ in a binary system, now referred to as *Gaia* BH3. Following the detection of *Gaia* BH1 and *Gaia* BH2, discussions and hypotheses emerged regarding the origins of these systems. The formation histories of *Gaia* BH1 and *Gaia* BH2 are challenging to reconcile with isolated binary (IB) evolution, as their orbital properties do not align with the expected outcomes of common envelope evolution (El-Badry et al. 2023a,b). Their near-solar metallicities and inferred orbits within the Galactic plane suggest that they were likely formed in open clusters (Rastello et al. 2023; Di Carlo et al. 2024; Marín Pina et al. 2024; Tanikawa et al. 2024).

Young and open star clusters host most massive stars, which are the precursors of compact objects. Consequently, the majority of black holes in the Milky Way likely spent their early years in these clusters, engaging in dynamic interactions. N-body simulations are powerful tools for studying star cluster dynamics and black hole formation within clusters. Physically, the initial conditions of N-body simulations are determined by the star cluster formation process. Given the significant progress in observational studies of massive star and star cluster formation, these observational results should be translated into the initial conditions for star cluster dynamical simulations. This work also examines how initial conditions and the cluster formation process influence the evolutionary outcomes of star clusters.

2. N-body simulations

A portion of the N-body simulations for star clusters presented in Zhou et al. (2024a, 2025a); Zhou & Kroupa (2025) were summarized in Table.1. The details of the simulation recipes were also thoroughly documented in these studies. These simulations were previously used to explain the observed physical parameters of open clusters. Here, we focus on the final products of these simulations, i.e. black holes. Overall, these simulations can be divided into two types: the cluster coalescence simulation (Zhou et al. 2025a) and the individual cluster simulation (Zhou et al. 2024a). In line with the theme of this work, we now briefly describe the initial conditions used in the simulations, based on observations of the star cluster formation process.

The question of whether newborn stars in star clusters are randomly drawn from the initial mass function (IMF) is crucial for various fields of astrophysics. The stochastic (or random) sampling approach assumes that each star forms independently within an embedded cluster, with its mass randomly selected from a probability density function (PDF), effectively treating the IMF as a PDF. In contrast, the optimal sampling

Table 1. N-body simulations modeling both the evolution of individual clusters and the coalescence of multiple clusters under different gas expulsion modes. More details can refer to Zhou et al. (2024a) and Zhou et al. (2025a). "Time" is the time when the simulation is terminated. The cluster may or may not have dissolved by this time.

Cases	Type	Mass ($M_\odot$)	Time (Myr)	Dissolved
10000-fast	individual	10000	263	yes
3000-fast	individual	3000	100	yes
3000-moderate1	individual	3000	441	no
3000-moderate2	individual	3000	256.25	yes
3000-slow	individual	3000	950.25	yes
1000-fast	individual	1000	67.5	yes
1000-moderate1	individual	1000	145.5	yes
1000-moderate2	individual	1000	78	yes
1000-slow	individual	1000	533.5	yes
300-fast	individual	300	69	yes
300-moderate1	individual	300	100	yes
300-moderate2	individual	300	45.25	yes
300-slow	individual	300	346.5	yes
NGC6334-fast	coalescence	2207	614.25	yes

theory (Kroupa et al. 2013; Schulz et al. 2015) proposes a deterministic link between the total mass of an embedded cluster ($M_{\rm ecl}$) and the mass of the most massive star ($m_{\rm max}$) within it, driven by a self-regulated star formation process in the cluster (Kroupa 2015; Yan et al. 2023). In fact, the $m_{\rm max} - M_{\rm ecl}$ relation has been well established analytically (Elmegreen 2000; Weidner & Kroupa 2004; Oey & Clarke 2005), observationally (Larson 2003; Weidner & Kroupa 2006; Weidner et al. 2010; Zhou et al. 2024b) and numerically (Bonnell et al. 2003, 2004; Peters et al. 2011). In our N-body simulations, the initial mass functions of the clusters were chosen to be canonical (Kroupa 2001), with the most massive star following the $m_{\rm max} - M_{\rm ecl}$ relation of Weidner et al. (2013):

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \quad (1)$$

where $y = \log_{10}(m_{\rm max}/M_\odot)$, $x = \log_{10}(M_{\rm ecl}/M_\odot)$, a_0 = -0.66, a_1 = 1.08, a_2 = -0.150, and a_3 = 0.0084.

We modeled the key dynamical effects of gas expulsion by introducing a spherically symmetric, diluting external gravitational potential to a simulated cluster, following the approach outlined in Kroupa et al. (2001); Banerjee & Kroupa (2013); Zhou et al. (2024a). Specifically, we adopted the gravitational potential of a spherically symmetric, time-dependent mass distribution:

$$\begin{aligned} M_g(t) &= M_g(0) \qquad t \le \tau_d, \\ M_g(t) &= M_g(0)\exp\left(-\tfrac{(t-\tau_d)}{\tau_g}\right) \quad t > \tau_d. \end{aligned} \quad (2)$$

Here, $M_g(t)$ is the total mass of the gas. It starts depleting after a delay of τ_d, and is totally depleted within a timescale of τ_g. We used an average velocity of the outflowing gas of $v_g \approx 10$ km s^{-1}, which is the typical sound speed in an HII region. This gives $\tau_g = r_h(0)/v_g$, where $r_h(0)$ is the initial half-mass radius of the cluster. As discussed in Zhou et al. (2024a), uncertainties in parameters such as gas expulsion velocity, star formation efficiency (SFE), and the spatial distribution of gas contribute to variations in the gas expulsion timescale. Consequently, this timescale does not solely reflect the rate at which gas is removed. In addition to the rapid gas expulsion characterized by the gas decay time, τ_g, described above, we also simulated scenarios with slow and moderate gas expulsions. For the slow gas expulsion, the gas decay time was set to $10\tau_g$. Moderate gas expulsion falls between the fast and slow cases, and given the broad parameter space between τ_g and $10\tau_g$, we considered two intermediate scenarios: $2\tau_g$ ("moderate2") and $5\tau_g$ ("moderate1").

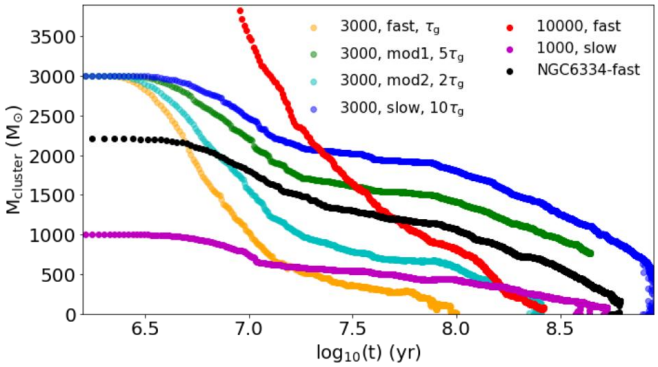


Fig. 1. The evolution of cluster mass over time.

3. Results

3.1. Cluster evolution

The evolution of clusters under different simulation scenarios has been extensively analyzed in Zhou et al. (2024a, 2025a); Zhou & Kroupa (2025). Here, we summarize some key conclusions. For individual cluster simulations, the mass loss of a cluster is strongly influenced by the mode of gas expulsion: the faster the gas expulsion, the greater the mass loss. However, there exists a degeneracy between cluster mass and gas expulsion timescale. For instance, a massive cluster undergoing rapid gas expulsion will expand quickly and lose mass at a high rate. In contrast, a lower-mass cluster experiencing slow gas expulsion will evolve more gradually, leading to a slower mass loss. Despite these differences, clusters of varying initial masses can exhibit similar evolutionary outcomes. As shown in Fig. 1, a 1000 $M_\odot$ cluster undergoing slow gas expulsion evolves in a manner similar to a 3000 $M_\odot$ cluster experiencing fast gas expulsion. Conversely, a 10000 $M_\odot$ cluster with fast gas expulsion exhibits similar evolutionary behavior to a 3000 $M_\odot$ cluster with slow gas expulsion.

In cluster coalescence simulations, factors such as the spatial distribution, relative velocities, mass distribution, and gas expulsion modes of subclusters influence both the dynamics of their coalescence process and the stability of the final coalesced product. As shown in Fig.1, despite each subcluster undergoing fast gas expulsion, the evolution of the coalesced cluster closely resembles that of a single cluster with moderate gas expulsion in individual cluster simulations. This suggests a degeneracy between cluster mass, the initial gas expulsion mode, and the cluster formation mechanism (whether monolithic or through coalescence).

3.2. Black hole formation

In Fig.2, Fig.3, Fig.4 and Fig.5, the 300 $M_\odot$ cluster is unable to form a black hole at its center under any gas expulsion mode. A black hole can form at the center of the 1000 $M_\odot$ cluster only if the gas expulsion is slow. For the 3000 $M_\odot$ cluster, black holes can form at the center under moderate or slow gas expulsion modes. The 10000 $M_\odot$ cluster can form central black holes even with fast gas expulsion, whereas the 3000 $M_\odot$ cluster cannot under the same conditions. Moreover, despite each initial subcluster undergoing fast gas expulsion, the coalesced cluster with the mass of 2207 $M_\odot$ can still form a black hole at its center. These results reflect the aforementioned degeneracy again.

Variations in cluster mass, gas expulsion modes, and formation mechanisms ultimately influence the cluster's expansion speed, which in turn affects its mass loss rate and black hole formation rate. Fig.6 shows the velocity distribution of stars within the 3000 $M_\odot$ cluster under different gas expulsion modes. A star cluster's expansion speed reflects the overall distribution of stellar velocities and plays a crucial role in determining whether massive stars and their remnants remain bound to the cluster.

In Fig.4, we can see closely spaced black holes. Their ability to form multiple systems (binary or triple) depends on the dynamical state of the star cluster. The initial gas expansion mode is crucial to the cluster's dynamical evolution and has a significant impact on its lifetime. Black holes in slowly expanding clusters experience a more stable dynamical environment and have substantially more time to evolve within the cluster. The number of black holes at the cluster centers of Fig.4(a) and (b) is 2 and 3, respectively. Fig.7 shows the evolution of the distance between each pair of black holes over time. Three black holes in Fig.4(b) eventually form a triple system, whereas two black holes in Fig.4(a) fail to form a binary system even until the cluster dissolves.

4. Discussion

In Sec.3, we demonstrate the crucial influence of the gas expulsion mode and cluster formation mechanism on the dynamical evolution of star clusters and black hole formation. Therefore, these factors should be carefully incorporated into the initial conditions of N-body simulations.

4.1. Gas expulsion mode

Giant molecular clouds (GMCs) are widely recognized as the primary gas reservoirs that fuel star formation and serve as the birthplaces of nearly all stars. Within these clouds, clumps are dense, localized structures that act as the main sites of star formation (Urquhart et al. 2018, 2022). The process of star cluster formation in molecular clouds involves several stages, including the formation of clumps in diffuse gas, the evolution of these clumps, the formation of embedded clusters within the clumps, and the subsequent evolution of the embedded clusters along with clump destruction due to feedback (Krumholz et al. 2019; Krause et al. 2020; Zhou et al. 2022, 2024c). In Urquhart et al. (2018, 2022), the ATLASGAL clumps at various stages encompass all of these physical processes. A clump eventually forms an embedded cluster with an IMF (Yan et al. 2017; Zhou et al. 2024d,b). An ATLASGAL clump with an HII region (HII-clump) consists of an embedded cluster surrounded by a gas envelope, with additional gas present within the cluster. Zhou et al. (2024c) presents high-resolution ALMA (Atacama Large Millimeter/submillimeter Array) observations of HII-clumps, revealing clear gas shells surrounding the protoclusters. HII-clumps represent the final stage of embedded cluster formation in the clumps, which are critical to build the bridge between the stellar kinematics and abundances community and the star formation community, because protoclusters or embedded clusters in HII-clumps are the first stage of gas-free star clusters.

Following gas expulsion driven by feedback from the embedded cluster, the cluster will undergo expansion (Kuhn et al. 2019; Wright et al. 2024; Della Croce et al. 2024; Zhou et al. 2024f). Understanding the detailed feedback processes within clumps is essential for tracing the evolutionary pathway from clumps to embedded clusters and ultimately to open clusters. This knowledge is crucial for setting the initial conditions of N-body simulations. The strength of feedback depends on the mass of the embedded cluster, leading to variations in the gas expulsion timescale. More massive clumps can give rise to more massive clusters, which, in turn, generate stronger feedback and higher gas expulsion velocities (Dib et al. 2013). There should be a correlation between feedback strength, clump (or cluster) mass, and gas expulsion velocity. Low-mass clusters are expected to experience a slower gas expulsion process compared to high-mass clusters (Pang et al. 2021).

For embedded clusters, integrating both hydrodynamics and N-body simulations is crucial for advancing our understanding of early cluster feedback and gas expulsion processes (Sills et al. 2018; Suin et al. 2022; Rieder et al. 2022; Fujii et al. 2022; Lewis et al. 2023; Cournoyer-Cloutier et al. 2023; Rodriguez et al. 2023; Jo et al. 2024; Polak et al. 2024). The spherically symmetric external gravitational potential from the gas, adopted in Sec.2, is partially supported by Geyer & Burkert (2001), who performed comparison simulations using the smoothed particle hydrodynamics method to model the gas. Additionally, hydrodynamics+N-body simulations in Farias et al. (2024) confirm that the exponential decay model presented in Eq.2 generally provides a reliable description of gas removal driven by radiation and wind feedback.

In Zhou et al. (2025b), we tried to constrain the gas expulsion mode of initially embedded star clusters by inversely analyzing the expansion speed of very young open clusters (< 5 Myr) using the *Gaia* DR3 data. After comparing with N-body simulations, we found that all four gas expulsion modes in Sec.2 are possible. As discussed in Sec.3.2, slowly expanding clusters are more likely to retain black holes and black hole-formed multiple systems, making them promising candidates for black hole searches with *Gaia* observations.

4.2. Cluster formation mechanism

In Zhou et al. (2024a), we collected samples of Galactic clumps, embedded clusters and open clusters to compare their physical properties. We showed that the mass distribution of open clus-

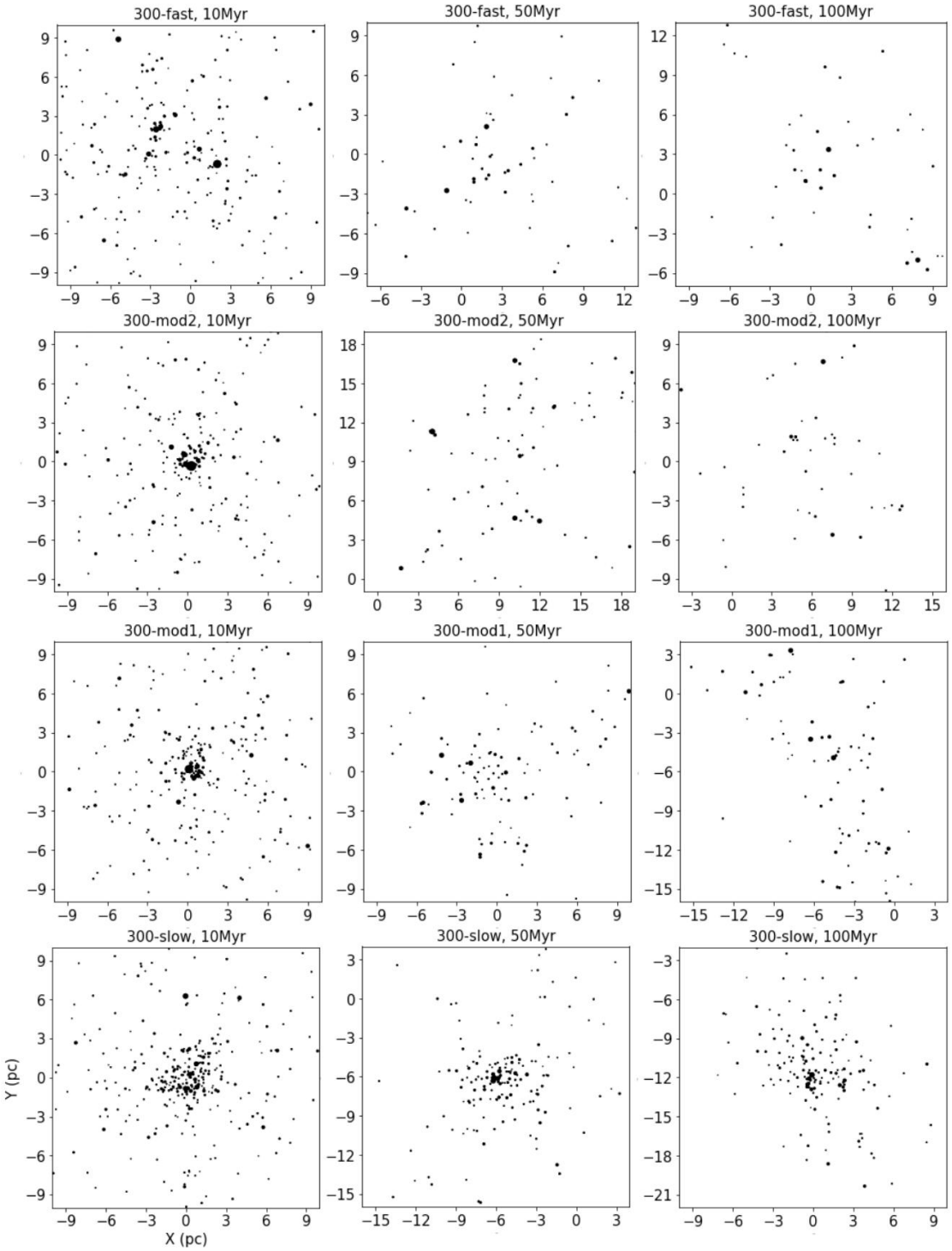


Fig. 2. The evolution of the 300 $M_\odot$ cluster under different gas expulsion modes. Snapshots projected on the XY plane (20 pc × 20 pc around the cluster center) at ages 10 Myr, 50 Myr, and 100 Myr are shown here. The size of the dots reflects the stellar mass.

ters covers a significantly larger mass range than that of embedded clusters, by about one order of magnitude. Given the current mass distribution of clumps in the Milky Way, the evolutionary sequence from a single clump evolving into an embedded cluster and subsequently into an open cluster cannot account for the observed open clusters with old ages and high masses. This fact is also supported by N-body simulations of individual embedded clusters. Moreover, we compared the separation of open clusters and the typical size of molecular clouds in the Milky Way, and found that most molecular clouds may only form one open cluster. Since a molecular cloud typically forms many embedded clusters, this result supports the coalescence scenario. The typical separation between embedded clusters in massive star-forming regions can be ≈1 pc (Zhou et al. 2024e). Therefore, after these clusters expand, they should be able to undergo coalescence. There is now extensive literature arguing that star clusters form through the merging of subclusters, both from simulations (Vázquez-Semadeni et al. 2009; Fujii et al. 2012; Vázquez-Semadeni et al. 2017; Howard et al. 2018; Sills et al. 2018; Fujii et al. 2022; Dobbs et al. 2022; Guszejnov et al. 2022; Cournoyer-

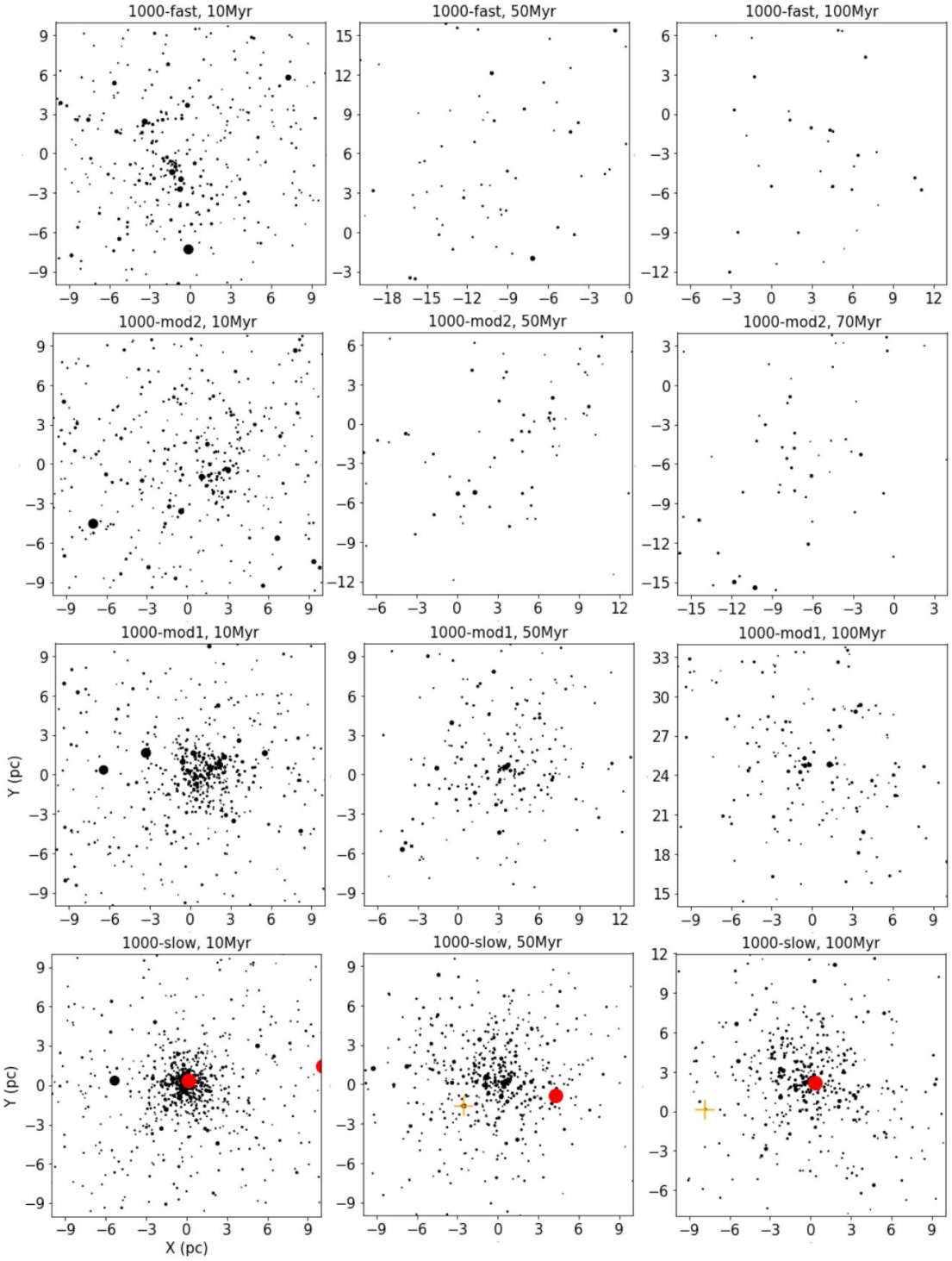


Fig. 3. Same as Fig.2, but for the 1000 $M_{\odot}$ cluster. Orange pluses and red dots represent neutron stars and black holes, respectively.

Cloutier et al. 2023; Polak et al. 2024) and observations (Sabbi et al. 2012; Dalessandro et al. 2021; Pang et al. 2022; Della Croce et al. 2023; Fahrion & De Marchi 2024).

Continuing the work of Zhou et al. (2024a), in Zhou et al. (2025a), we explored a scenario in which open clusters form as a result of post-gas expulsion coalescence of embedded clusters within the same parental molecular cloud, and found that the coalescence of multiple low-mass embedded clusters can account for the observed parameter space of open clusters in the Milky Way. In Zhou & Kroupa (2025), analysis of the *Gaia* DR3 open cluster catalog reveals that as clusters age, the $m_{\rm max} - M_{\rm cluster}$ relation of open clusters increasingly deviates from the initial $m_{\rm max} - M_{\rm ecl}$ relation observed in embedded clusters. Over time, this relation develops a clear stratification based on cluster age. We compared the observed $m_{\rm max} - M_{\rm cluster}$ relations in open clusters with results from N-body simulations modeling both the evolution of individual clusters and the coalescence of multiple clusters, and found that the coalescence simulation can better explain the observational results. The simulations also reveal that the mode of gas expulsion has a significant impact on the $m_{\rm max} - M_{\rm cluster}$ relation, as cluster mass loss is highly sensitive to the details of gas removal, as shown in Fig.1. As discussed in

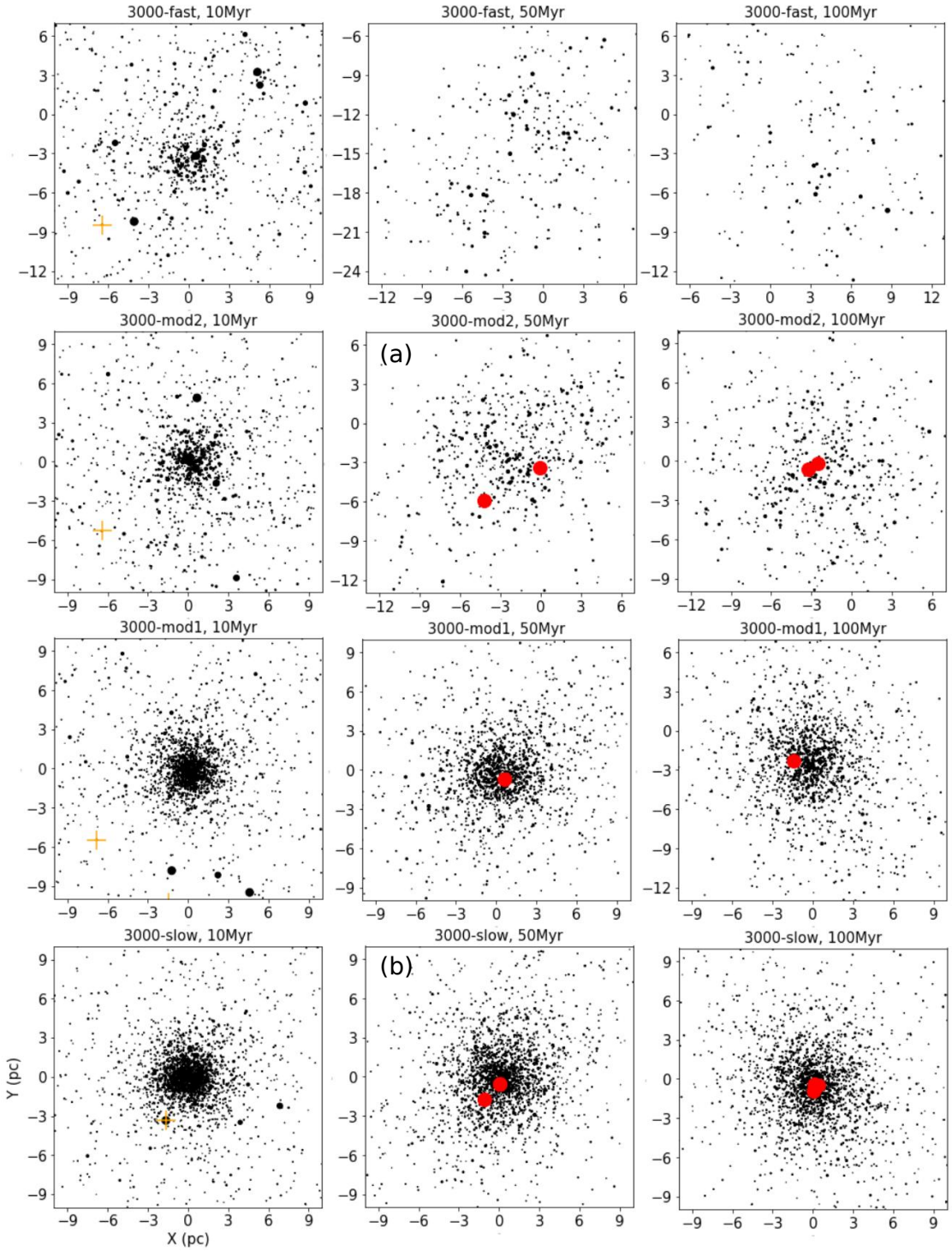


Fig. 4. Same as Fig.3, but for the 3000 $M_{\odot}$ cluster. The central regions of the star clusters in Fig.4(a) and (b) contain 2 and 3 black holes, respectively. In panel (b), two of the three black holes are too close to be distinguished on the map.

Zhou et al. (2025b), the coalescence of expanding star clusters can lead to a reduction in overall expansion speed, which may explain the slow expansion of very young star clusters (< 5 Myr) in the *Gaia* DR3 open cluster catalog of Hunt & Reffert (2024). In summary, these studies suggest that the post-gas expulsion coalescence of embedded clusters may be a key mechanism in the formation of open clusters.

In the cluster coalescence and individual cluster simulations by Zhou et al. (2024a, 2025a); Zhou & Kroupa (2025), the dynamical states of coalesced and single star clusters differ significantly throughout their evolution, leading to distinct outcomes. Moreover, under different coalescence scenarios, the dynamical states and evolutionary products of the coalesced clusters vary considerably. As discussed above, since embedded cluster coalescence is a key pathway for open cluster formation, studying black hole formation across various coalescence scenarios would be promising. In coalescence scenarios, cluster dynamics and black hole formation are influenced by various factors, including the mass ratio, relative velocity, collision angle, and initial expansion speed of the subclusters. Possible coalescence

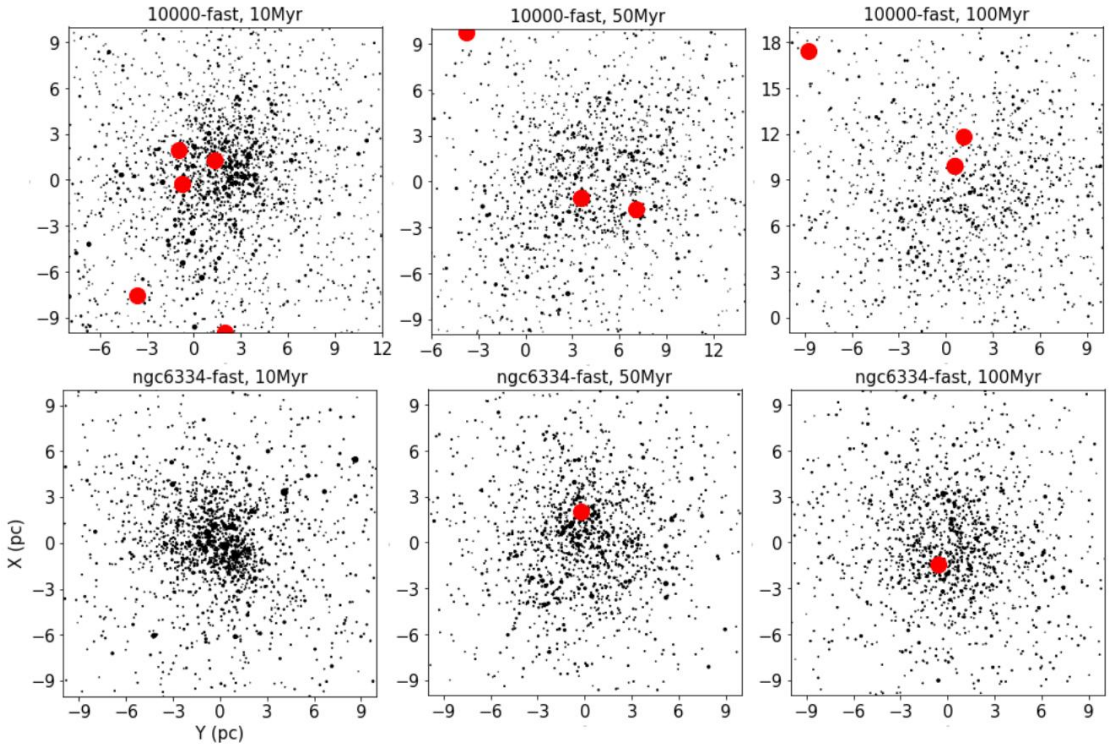


Fig. 5. Same as Fig.3, but for the 10000 $M_\odot$ cluster and the coalesced cluster with the mass of 2207 $M_\odot$.

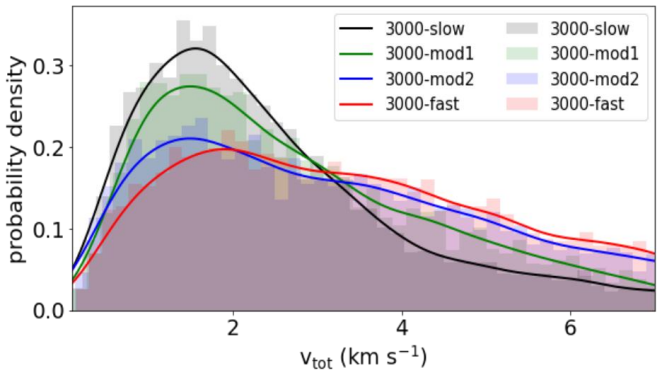


Fig. 6. Velocity distribution of stars within the 3000 $M_\odot$ cluster under different gas expulsion modes.

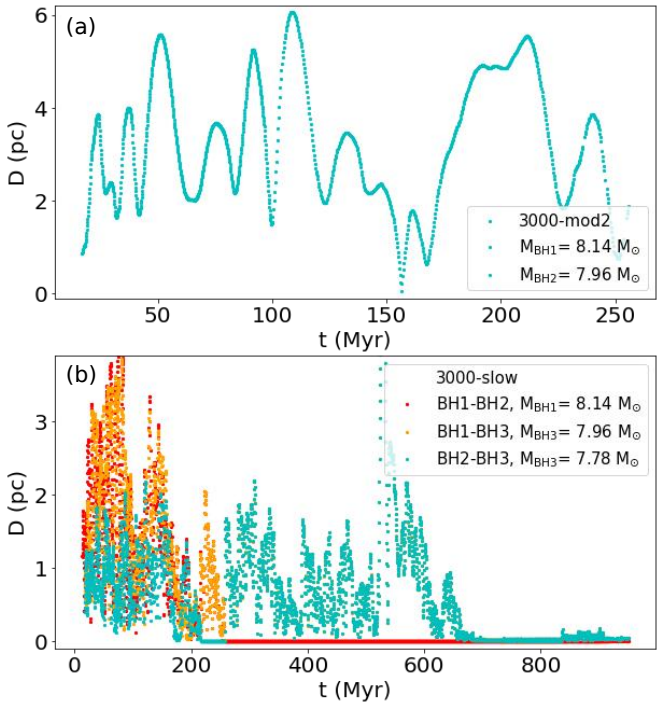


Fig. 7. The temporal evolution of the distance between each pair of black holes.

scenarios can be inferred from the physical parameters and spatial distribution of clumps in molecular clouds, as these clumps serve as precursors to embedded clusters.

5. Conclusion

This study examines how the star cluster formation process influences the evolutionary outcomes of star clusters. The progenitors of black holes are massive stars. Modeling a star cluster using stochastic sampling or optimal sampling from the IMF will directly determine the content of massive stars in the initial star cluster. As discussed in **?**, a significant relationship between the mass of the most massive star ($m_{\rm max}$) and the total mass of the embedded cluster in stars ($M_{\rm ecl}$) supports the optimal sampling. Recent observational progress has highlighted the crucial role of early expansion in the evolution of young star clusters. In N-body simulations, the gas expulsion mode (timescale) of initially embedded star clusters has an almost decisive impact on the dynamical state and subsequent evolution of the clusters. This significantly impacts the likelihood of massive stars and compact objects being ejected from the cluster, thereby affecting the cluster's black hole formation rate. Additionally, the survival time of star clusters varies under different gas expulsion modes, which determines how long the evolutionary products of massive stars remain in the cluster environment. This may have

a significant impact on the formation of multiple systems involving black holes. As to the cluster formation mechanism, star clusters formed through the coalescence of multiple subclusters exhibit significantly different cluster dynamics and evolutionary outcomes compared to those formed as single entities. Since the post-gas expulsion coalescence of embedded clusters formed in the same parental molecular cloud may be a dominant mechanism for the formation of open clusters, it is promising to investigate black hole formation in different coalescence scenarios.

A cluster's evolution is closely tied to its formation process rather than being independent of it. Variations in cluster mass, gas expulsion modes, and formation mechanisms ultimately influence the cluster's expansion speed, which in turn affects its mass loss rate and black hole formation rate. The expansion speed reflects the overall distribution of stellar velocities, playing a crucial role in determining whether massive stars and their remnants remain bound to the cluster. Therefore, these physical factors should be carefully incorporated into the initial conditions of N-body simulations. This requires us to consider a complete physical picture of star cluster formation and evolution. The evolution of star clusters and the formation of compact objects, as well as the formation of massive stars and star clusters, are typically studied by two separate research communities. However, to understand black hole formation in open cluster, it is necessary to integrate the knowledge from both communities to reconstruct a complete physical picture.

References

Banerjee, S. & Kroupa, P. 2013, ApJ, 764, 29
Bonnell, I. A., Bate, M. R., & Vine, S. G. 2003, MNRAS, 343, 413
Bonnell, I. A., Vine, S. G., & Bate, M. R. 2004, MNRAS, 349, 735
Chakrabarti, S., Simon, J. D., Craig, P. A., et al. 2023, AJ, 166, 6
Cournoyer-Cloutier, C., Sills, A., Harris, W. E., et al. 2023, MNRAS, 521, 1338
Dalessandro, E., Varri, A. L., Tiongco, M., et al. 2021, ApJ, 909, 90
Della Croce, A., Dalessandro, E., Livernois, A., & Vesperini, E. 2024, A&A, 683, A10
Della Croce, A., Dalessandro, E., Livernois, A., et al. 2023, A&A, 674, A93
Di Carlo, U. N., Agrawal, P., Rodriguez, C. L., & Breivik, K. 2024, ApJ, 965, 22
Dib, S., Gutkin, J., Brandner, W., & Basu, S. 2013, MNRAS, 436, 3727
Dobbs, C. L., Bending, T. J. R., Pettitt, A. R., Buckner, A. S. M., & Bate, M. R. 2022, MNRAS, 517, 675
El-Badry, K., Rix, H.-W., Cendes, Y., et al. 2023a, MNRAS, 521, 4323
El-Badry, K., Rix, H.-W., Quataert, E., et al. 2023b, MNRAS, 518, 1057
Elmegreen, B. G. 2000, ApJ, 539, 342
Fahrion, K. & De Marchi, G. 2024, A&A, 681, A20
Farias, J. P., Offner, S. S. R., Grudić, M. Y., Guszejnov, D., & Rosen, A. L. 2024, MNRAS, 527, 6732
Fujii, M. S., Saitoh, T. R., & Portegies Zwart, S. F. 2012, ApJ, 753, 85
Fujii, M. S., Wang, L., Hirai, Y., et al. 2022, MNRAS, 514, 2513
Gaia Collaboration, Panuzzo, P., Mazeh, T., et al. 2024, A&A, 686, L2
Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, A&A, 674, A1
Geyer, M. P. & Burkert, A. 2001, MNRAS, 323, 988
Guszejnov, D., Markey, C., Offner, S. S. R., et al. 2022, MNRAS, 515, 167
Howard, C. S., Pudritz, R. E., & Harris, W. E. 2018, Nature Astronomy, 2, 725
Hunt, E. L. & Reffert, S. 2024, A&A, 686, A42
Jo, Y., Kim, S., Kim, J.-h., & Bryan, G. L. 2024, ApJ, 974, 193
Krause, M. G. H., Offner, S. S. R., Charbonnel, C., et al. 2020, Space Sci. Rev., 216, 64
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2015, Canadian Journal of Physics, 93, 169
Kroupa, P., Aarseth, S., & Hurley, J. 2001, MNRAS, 321, 699
Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, Vol. 5, 115
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Kuhn, M. A., Hillenbrand, L. A., Sills, A., Feigelson, E. D., & Getman, K. V. 2019, ApJ, 870, 32
Larson, R. B. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 287, Galactic Star Formation Across the Stellar Mass Spectrum, ed. J. M. De Buizer & N. S. van der Bliek, 65–80
Lewis, S. C., McMillan, S. L. W., Low, M.-M. M., et al. 2023, ApJ, 944, 211
Marín Pina, D., Rastello, S., Gieles, M., et al. 2024, A&A, 688, L2
Oey, M. S. & Clarke, C. J. 2005, ApJ, 620, L43
Pang, X., Li, Y., Yu, Z., et al. 2021, ApJ, 912, 162
Pang, X., Tang, S.-Y., Li, Y., et al. 2022, ApJ, 931, 156
Peters, T., Banerjee, R., Klessen, R. S., & Mac Low, M.-M. 2011, ApJ, 729, 72
Polak, B., Mac Low, M.-M., Klessen, R. S., et al. 2024, A&A, 690, A207
Rastello, S., Iorio, G., Mapelli, M., et al. 2023, MNRAS, 526, 740
Rieder, S., Dobbs, C., Bending, T., Liow, K. Y., & Wurster, J. 2022, MNRAS, 509, 6155
Rodriguez, C. L., Hafen, Z., Grudić, M. Y., et al. 2023, MNRAS, 521, 124
Sabbi, E., Lennon, D. J., Gieles, M., et al. 2012, ApJ, 754, L37
Schulz, C., Pflamm-Altenburg, J., & Kroupa, P. 2015, A&A, 582, A93
Sills, A., Rieder, S., Scora, J., McCloskey, J., & Jaffa, S. 2018, MNRAS, 477, 1903
Suin, P., Shore, S. N., & Pavlík, V. 2022, A&A, 667, A69
Tanikawa, A., Cary, S., Shikauchi, M., Wang, L., & Fujii, M. S. 2024, MNRAS, 527, 4031
Tanikawa, A., Hattori, K., Kawanaka, N., et al. 2023, ApJ, 946, 79
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
Vázquez-Semadeni, E., Gómez, G. C., Jappsen, A. K., Ballesteros-Paredes, J., & Klessen, R. S. 2009, ApJ, 707, 1023
Vázquez-Semadeni, E., González-Samaniego, A., & Colín, P. 2017, MNRAS, 467, 1313
Weidner, C. & Kroupa, P. 2004, MNRAS, 348, 187
Weidner, C. & Kroupa, P. 2006, MNRAS, 365, 1333
Weidner, C., Kroupa, P., & Bonnell, I. A. D. 2010, MNRAS, 401, 275
Weidner, C., Kroupa, P., & Pflamm-Altenburg, J. 2013, MNRAS, 434, 84
Wright, N. J., Jeffries, R. D., Jackson, R. J., et al. 2024, MNRAS, 533, 705
Yan, Z., Jerabkova, T., & Kroupa, P. 2017, A&A, 607, A126
Yan, Z., Jerabkova, T., & Kroupa, P. 2023, A&A, 670, A151
Zhou, J.-w., Dib, S., & Kroupa, P. 2024a, A&A, 691, A293
Zhou, J. W., Dib, S., & Kroupa, P. 2024b, PASP, 136, 094302
Zhou, J.-w., Dib, S., & Kroupa, P. 2025a, MNRAS, 537, 845
Zhou, J. W., Dib, S., Wyrowski, F., et al. 2024c, A&A, 682, A173
Zhou, J. W. & Kroupa, P. 2025, Origin of open clusters indicated by the evolution of the $m_{\rm max} - M_{\rm ecl}$ relation, submitted
Zhou, J. W., Kroupa, P., & Dib, S. 2024d, PASP, 136, 094301
Zhou, J.-w., Kroupa, P., & Dib, S. 2024e, A&A, 688, L19
Zhou, J. W., Kroupa, P., & Dib, S. 2025b, Constraining the gas expulsion modes of embedded clusters by analyzing the expansion speed of open clusters, submitted
Zhou, J.-W., Kroupa, P., & Wu, W. 2024f, A&A, 691, A204
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038

Original papers of Chapter.6

The original paper of Sec.6.2,

> Zhou, J. W., Kroupa, P., & Dib, S. 2025, A&A, 695, L17
> *Inversely synthesizing the core mass function of high-mass star-forming regions from the canonical initial mass function*

The original paper of Sec.6.3,

> Zhou, J.-W., Kroupa, P., under revision
> *Clump populations in molecular clouds*

The original paper of Sec.6.4,

> Zhou, J.-W., Kroupa, P., & Dib, S. 2025, MNRAS, 541, 1276
> *The most massive star clusters in molecular clouds: Insights from the integrated cloud-wide initial mass function (ICIMF) theory*

A&A, 695, L17 (2025)
https://doi.org/10.1051/0004-6361/202452924
© The Authors 2025

Astronomy & Astrophysics

LETTER TO THE EDITOR

Inversely synthesizing the core mass function of high-mass star-forming regions from the canonical initial mass function

J. W. Zhou[1,⋆], Pavel Kroupa[2,3,⋆], and Sami Dib[4,⋆]

[1] Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
[2] Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany
[3] Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic
[4] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Received 8 November 2024 / Accepted 1 March 2025

ABSTRACT

Many studies have revealed that the core mass function (CMF) in high-mass star-forming regions is top-heavy. In this work, we start from the canonical initial mass function (IMF) to inversely synthesize the observed CMFs of high-mass star formation regions, taking into account variations in multiplicity and mass conversion efficiency from core to star ($\epsilon_{\rm core}$). To match the observed CMFs, cores of different masses should have varying $\epsilon_{\rm core}$, with $\epsilon_{\rm core}$ increasing as the core mass decreases. However, the multiplicity fraction does not affect the synthesized CMFs. To accurately fit the high-mass end of the CMF, it is essential to determine whether the CMF shows a slope transition from the low-mass end to the high-mass one. If the CMF truly undergoes a slope transition but observational biases obscure it, leading to a combined fit with a shallower slope, this could artificially create a top-heavy CMF.

Key words. binaries: general – stars: formation – stars: protostars – ISM: clouds

1. Introduction

Giant molecular clouds (GMCs) are widely recognized as the primary gas reservoirs that fuel star formation and serve as the birthplaces of nearly all stars. A dense core is a localized fragment or overdensity within a molecular cloud, representing a local minimum in gravitational potential and serving as the precursor to an individual star or stellar system (André et al. 2014). A prestellar core is a dense core that remains starless while being gravitationally bound. In star-forming regions within the solar neighborhood that produce solar-type stars, the prestellar core mass function (CMF) closely resembles the initial mass function (IMF) (Alves et al. 2007; Enoch et al. 2008; André et al. 2014; Könyves et al. 2015; Di Francesco et al. 2020). This similarity has led to the hypothesis that the IMF's shape may be directly inherited from the CMF. However, the Milky Way includes both low-mass and high-mass star-forming regions. A thorough analysis of large statistical samples of massive protoclusters is essential to observationally determine whether the origin of the IMF is truly independent of cloud properties. Thanks to observations with the Atacama Large Millimeter/submillimeter Array (ALMA), numerous studies have now examined CMFs in high-mass star-forming regions (Motte et al. 2018; Liu et al. 2018; Cheng et al. 2018; Sanhueza et al. 2019; Lu et al. 2020; Sadaghiani et al. 2020; O'Neill et al. 2021; Pouteau et al. 2022; Nony et al. 2023; Louvet et al. 2024), revealing that the CMFs in these regions are flatter than the ones in nearby, low-mass star-forming regions. The excess of high-mass cores in these regions indicates a top-heavy CMF, which may challenge the universality of the IMF. In the literature, the CMF fitting for high-mass star-forming regions typically does not differentiate between core types (such as prestellar or protostellar). However, the differentiation was done in Sanhueza et al. (2019), Nony et al. (2023). Sanhueza et al. (2019) derived a power law index of -2.17 ± 0.10 for the prestellar CMF, which is slightly shallower than the Salpeter IMF. In Nony et al. (2023), on the other hand, the prestellar CMF has a power law index of -2.46 ± 0.15. The protostellar CMF in Nony et al. (2023) is significantly shallower than the Salpeter value, with an index of -1.96 ± 0.09. As is shown in Sect. 3.1, the observed CMF of high-mass star-forming regions may not be robust.

It is now widely accepted that the majority of stars both form and exist on the main sequence with at least one stellar companion; "single stars are the exception, not the rule" (see Offner et al. 2023 and references therein). Goodwin & Kroupa (2005) argued that the short ($<10^5$ yr) dynamical decay times of nonhierarchical multiple systems imply that most stars must form in binary systems or dynamically stable hierarchical higher-order systems to explain the observed very high fraction of binaries in about 1 Myr old star-forming regions. Multiple star systems are formed during the earliest stages of star formation. Observations have offered evidence that the multiplicity fraction of protostars is higher than that of pre-main-sequence and main-sequence stars (Looney et al. 2000; Chen et al. 2013), indicating that most systems initially form as multiples. Observations of multiplicity in the youngest stellar systems are key to understanding their origins (Izquierdo et al. 2018; Olguin et al. 2021, 2022; Tobin et al. 2022; Li et al. 2024). As is shown in Fig. 1 of Offner et al. (2023), the likelihood of massive protostars forming as multiple systems may approach nearly 100%.

⋆ Corresponding authors; jwzhou@mpifr-bonn.mpg.de, pkroupa@uni-bonn.de, dib@mpia.de

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Open access funding provided by Max Planck Society.

The multiplicity of protostars is tightly correlated with the origin of the IMF or the transformation from the CMF to the IMF. Despite the high multiplicity fraction among protostars, it is often assumed that a single core will produce only one star. A core at low resolution may actually consist of multiple smaller-scale cores, which should form a multiple system rather than a single star. Assuming a one-to-one mapping between cores and stars artificially shifts the cores to the higher-mass range, potentially making the CMF appear top-heavy. Apart from the multiplicity, the mass conversion efficiency from core to star ($\epsilon_{\rm core}$) will also have a significant impact on the transformation between the CMF and the IMF. For the cores in star-forming regions within the solar neighborhood, the values of $\epsilon_{\rm core}$ are ≈20−40% (Alves et al. 2007; Könyves et al. 2010, 2015; Megeath et al. 2016). In this work, we start from the canonical IMF to inversely synthesize the observed CMF of high-mass star formation regions, taking into account variations in multiplicity and $\epsilon_{\rm core}$.

2. Sample

For three high-mass star-forming regions, W43-MM1&MM2&MM3, we used the core catalogs in Motte et al. (2018), Pouteau et al. (2022), Nony et al. (2023), Louvet et al. (2024). For the same observation in the ALMA-IMF survey (Motte et al. 2022), the core catalogs of W43-MM1&MM2&MM3 are presented in Pouteau et al. (2022), Nony et al. (2023), Louvet et al. (2024). For W43-MM1, there is an independent catalog derived from the observation in Motte et al. (2018). For the same region, different catalogs have different selection criteria and identifications, varying in the number of cores and their mass distribution, which is reflected in the differences in the CMF, as is shown in Fig. 1.

3. Results and discussion

3.1. Observed CMFs

For W43-MM2&MM3, the core catalogs in Nony et al. (2023) were directly selected from those in Pouteau et al. (2022), which only retains a portion of the original lists. The CMFs derived from the core catalogs in Pouteau et al. (2022) and Nony et al. (2023) have similar slopes, but the catalogs in Nony et al. (2023) with fewer cores shift systematically upward. For W43-MM1, three independent identifications in different observations give comparable CMFs. For W43-MM3, two independent identifications in the same observation give quite different CMFs. Therefore, numerous observational constraints affect the CMF, including the calculation of the core's physical parameters, observational resolution, identification methods, and the completeness of the core population. However, the shallow nature of the CMF slope in W43 seems fairly robust at least.

3.2. Clumps and embedded clusters

Clumps are localized high-density structures within a molecular cloud and serve as the primary sites of star formation within the cloud (Urquhart et al. 2018, 2022). Cores are assumed to be overdense regions with respect to their parental clump. Generally, a cloud includes many clumps, and a clump includes many cores. They appear as hierarchical, multiscale hub-filament structures (Kumar et al. 2020; Zhou et al. 2022, 2023, 2024a; Zhou & Davis 2024). To trace the evolution from the CMF to the IMF, the observation should cover the entire clump, as was done in the ALMA-IMF survey. A clump eventually forms an embedded cluster with an IMF (Yan et al. 2017; Zhou et al. 2024b,c). This IMF corresponds to the initial CMF in the clump. Therefore, clumps should be regarded as the fundamental units for studying the origin of the IMF (Kroupa et al. 2024). W43 is a molecular cloud complex. The main star-forming regions in W43 are the clumps. The three subregions of W43 (W43-MM1&MM2&MM3) correspond to three independent clumps displayed in Table 1. They have independent star formation (Zhou et al. 2024d), and thus different CMFs, as is shown in Fig. 1.

The ATLASGAL clumps in Urquhart et al. (2022) corresponding to each region are listed in Table 1. Zhou et al. (2024b,c) systematically studied the relationship between the physical parameters of the embedded clusters within ATLASGAL clumps with HII regions (HII-clumps) and the physical parameters of the clumps themselves. We derived a $M_{\rm cl}$−$M_{\rm ecl}$ relation in Zhou et al. (2024c):

$$\log_{10}(M_{\rm cl}/M_\odot) = (1.02 \pm 0.02) \times \log_{10}(M_{\rm ecl}/M_\odot) + (0.52 \pm 0.05). \quad (1)$$

Therefore, according to the clump mass ($M_{\rm cl}$), we can directly estimate the embedded cluster mass in stars ($M_{\rm ecl}$) in the clump. Although this relation is derived from HII-clumps, considering that the clumps in different evolutionary stages have similar mass distributions (Urquhart et al. 2022), it can also be used for younger clumps.

3.3. Initial stellar populations

Based on the recipe described in Appendix B, the initial stellar population in each high-mass star-forming region was generated using the updated `McLuster` program (Küpper et al. 2011; Wang et al. 2019), according to the embedded cluster mass ($M_{\rm ecl}$) estimated in Sect. 3.2. Assuming that the average mass of stars in a cluster is 0.5 $M_\odot$, the total number of stars in a cluster ($N_{\rm star}$) is $M_{\rm ecl}/(0.5\,M_\odot)$. The proportion of binaries in the initial stellar population ($f_{\rm b}$) is an adjustable parameter in the `McLuster` code. Since the primary and secondary mass of the binary was selected randomly from the canonical IMF, for each embedded cluster, we ran `McLuster` 100 times, and overlapped these 100 stellar populations to derive a synthesized core population. For binaries, we summed the masses of each pair corresponding to one core. For single stars, we retained the generated mass. Then the stellar mass was divided by $\epsilon_{\rm core}$ to construct the mass distribution of the cores.

To derive the CMF, one common approach is to fit a linear relationship to data that has been grouped into evenly sized bins in logarithmic space. However, binning can lead to substantial information loss, as it aggregates data points into groups rather than using each one individually, which can obscure finer details in the distribution. This issue is especially pronounced in the higher mass range, where bins may contain very few data points, resulting in slope estimates that are often skewed and unreliable (Maíz Apellániz & Úbeda 2005; Maschberger & Kroupa 2009). To address these limitations, Koen (2006) employed a complementary cumulative distribution function (CCDF) plot, which avoids binning and retains more detail than traditional binning methods. Taking $m_{\rm max} = 100\,M_\odot$, equation (B.1) was used to generate mock data of the IMF. According to the theoretical $m_{\rm max}$−$M_{\rm ecl}$ relation derived in Zhou et al. (2024c), $m_{\rm max} = 100\,M_\odot$ corresponds to $M_{\rm ecl} \approx 5000\,M_\odot$; thus, the number of

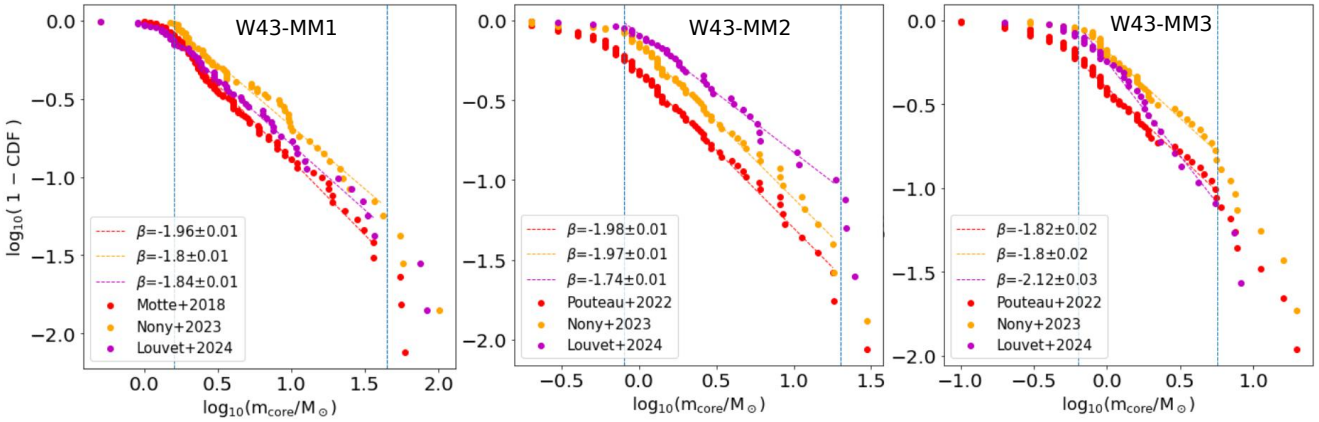


Fig. 1. Complementary cumulative distribution function (CCDF) plot and the linear fitting of the observed CMF in the mass range marked by two vertical dashed lines. β is the slope of the corresponding dN/dM mass function derived from linear fitting of the CCDF. The lower mass of the adopted fitting range corresponds to the observational mass completeness level estimated by Louvet et al. (2024).

Table 1. Physical parameters of the corresponding ATLASGAL clumps.

Region	Clump	$\log_{10}(M_{\rm fwhm})$ ($M_\odot$)	$M_{\rm ecl}$ ($M_\odot$)	$m_{\rm max}$ ($M_\odot$)	$N_{\rm star}$
W43-MM3	AGAL030.718−00.082	3.173	399	26	798
W43-MM2	AGAL030.703−00.067	3.211	435	27	870
W43-MM1	AGAL030.818−00.056	3.248	473	29	946

Notes. $M_{\rm fwhm}$ is the clump mass defined in Urquhart et al. (2022). $M_{\rm ecl}$ is the mass of the embedded cluster finally formed in a clump. $m_{\rm max}$ is the mass of the most massive star in an embedded cluster. $N_{\rm star}$ is the total number of stars in a cluster.

stars ($N_{\rm star}$) in the realization is 10 000. Fig. B.1 shows the CCDF plot of the IMF.

3.4. Comparison of synthesized and observed CMFs

For all three regions in Fig. 1, the low-mass end of the CMFs does not exhibit the clear transition seen in the IMF, which may be caused by the uncertainty in core mass estimation. As is discussed in Louvet et al. (2024), the primary sources of error in the mass estimates stem from uncertainties in the opacity index and core temperature. To account for these uncertainties, they conducted a Monte Carlo simulation on 330 cores exceeding the 1.64 $M_\odot$ completeness threshold, allowing simultaneous variations in the opacity index, temperature, and source flux. Fig. 6 of Louvet et al. (2024) presents the cumulative CMFs obtained from 10^3 simulation trials, where we can see a clear slope transition.

According to the observed CMF, we have fixed the mass range of the linear fitting for each clump in Fig. 1. For the synthesized core population of each clump, we strictly follow the fixed mass range to do linear fitting (the dashed black segments in Figs. 2, 3 and 4). A direct comparison between the synthesized and observed CMFs is only meaningful above the estimated mass completeness level of the observed sample. Therefore, for the purpose of direct comparison, all CCDFs (both observed and synthesized) were normalized to 1 at the mass completeness level marked by the left vertical dashed line in Fig. 1.

In Fig. 2f, there is a noticeable transition region in the synthesized CMFs, where the slope changes from flat to steep; that is, segment "2" in Fig. 5. As is illustrated by the dashed red line in Fig. 5, if the fitting range lies precisely within this transition region, it encompasses both the flat and steep slopes, resulting in an overall flat slope. If an observed CMF truly exhibits such a slope transition but observational biases obscure it, leading to a combined fit with a shallower slope, then the apparent top-heavy CMF would be merely an artifact. To accurately fit the high-mass end of the CMF, it is essential to identify whether the critical masses for the transition – "m1" and "m2" in Fig. 5 – exist. If so, this critical mass, m2, should be used as the starting point for fitting the slope of the high-mass end, which can then be compared to the Salpeter value. These issues appear in Fig. 6 of Louvet et al. (2024), in which a combined fit results in a shallower CMF slope. When focusing only on the high-mass end of this figure, the CMF slope is actually comparable to the Salpeter value.

By setting $f_{\rm b} = 1$ and $\epsilon_{\rm core} = 0.2$, in Figs. 2a–c, we explored how the selected $m_{\rm max}$ and $N_{\rm star}$ (or $M_{\rm ecl}$) influence the synthesized CMFs, and found that the synthesized CMFs are not sensitive to these parameters. A larger $m_{\rm max}$ will increase the mass at the high-mass end, which is beyond the observed core mass range. For the same reason, we do not consider systems of a higher order than binaries (such as triples and quadruples) in this work. In Fig. 3, the synthesized CMFs with $f_{\rm b} = 1$ do not provide better fits than those with $f_{\rm b} = 0.5$ or $f_{\rm b} = 0$. A high multiplicity fraction in the synthesis is not necessary.

However, in Figs. 2a–c, the synthesized CMFs do not align with the observed ones. The reason may be that, in each case, we apply a constant $\epsilon_{\rm core}$ for all the synthesized cores. A likely possibility is that cores of different masses have different $\epsilon_{\rm core}$. In Figs. 2d–e, we compare two scenarios in which $\epsilon_{\rm core}$ either increases or decreases with decreasing core mass, respectively. In each case, $\epsilon_{\rm core}$ was randomly selected from a range, then, all selected $\epsilon_{\rm core}$ were sorted in descending or ascending order. We found that in order to match the observed CMFs, $\epsilon_{\rm core}$ must increase with decreasing core mass. A possible explanation is that higher-mass cores contain more gas, leading to the

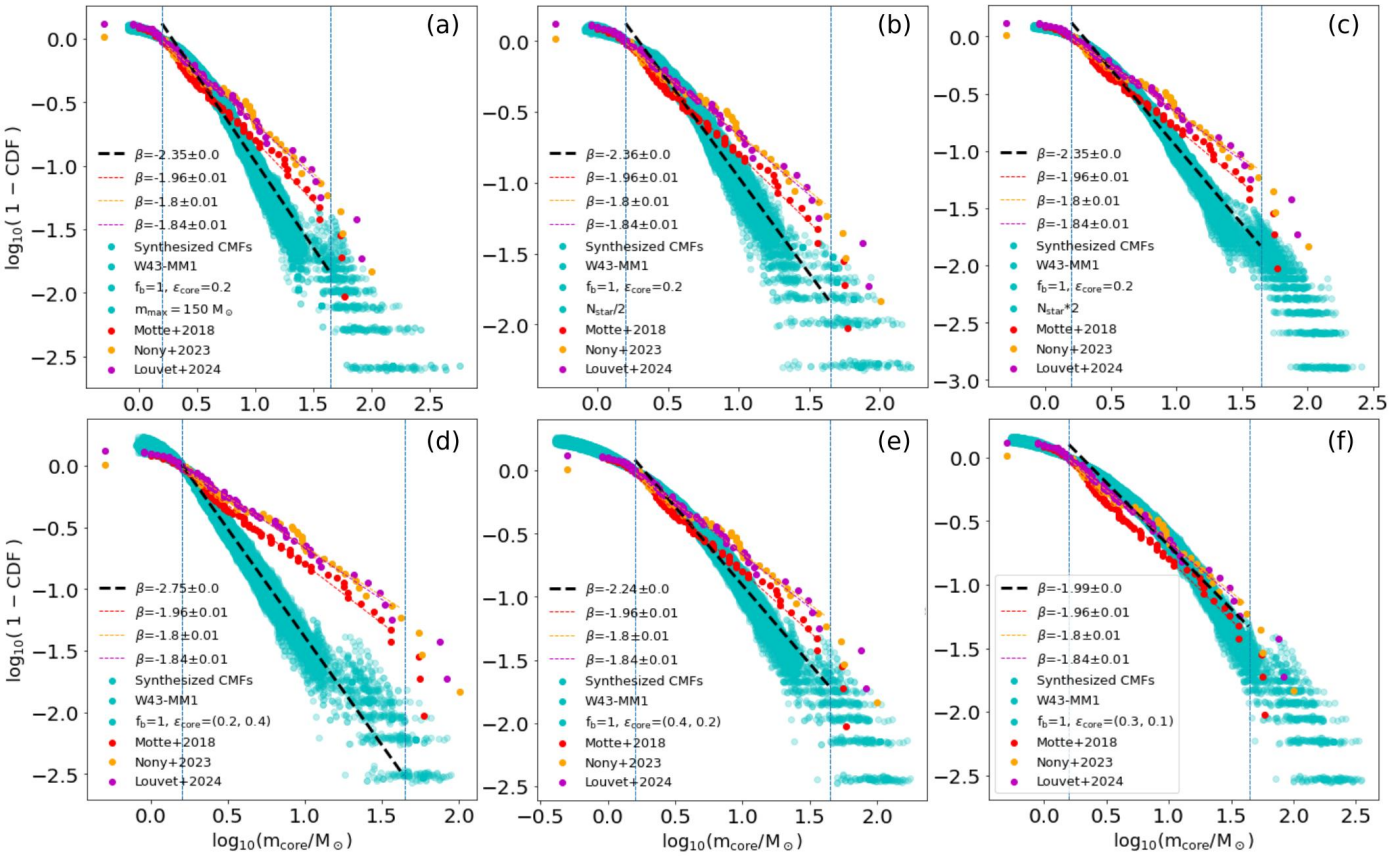


Fig. 2. Complementary cumulative distribution function (CCDF) plot and the linear fitting of the synthesized CMFs (in cyan) for W43-MM1 by setting $f_b = 1$ and $\epsilon_{core} = 0.2$. (a)–(c) The cases of $m_{max} = 150\,M_\odot$, half the original number of stars ($N_{star}/2$) and twice the original number of stars ($N_{star} * 2$), respectively; (d) The case in which ϵ_{core} decreases with decreasing core mass. ϵ_{core} was randomly selected from the range (0.2, 0.4); (e) The case in which ϵ_{core} increases with decreasing core mass. ϵ_{core} was randomly selected from the range (0.2, 0.4); (f) Same as panel e, but ϵ_{core} is in the range (0.1, 0.3). All selected ϵ_{core} in panels d, e, and f are sorted in descending or ascending order. All CCDFs (both observed and synthesized) were normalized to 1 at the mass completeness level marked by the left vertical dashed line.

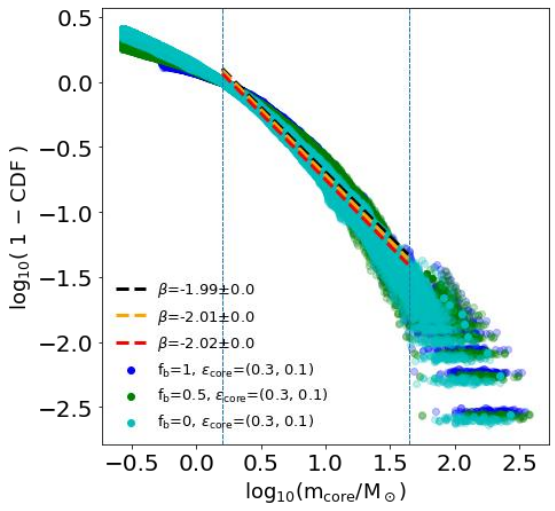


Fig. 3. Comparison of the synthesized CMFs for W43-MM1 for $f_b = 1$, 0.5 and 0. All synthesized CCDFs were normalized to 1 at the mass completeness level marked by the left vertical dashed line.

formation of more massive stars, and stronger feedback ultimately disperses more gas. This is similar to the observational results at the clump scale (Zhou et al. 2024c), where the star formation efficiency (SFE) is strongly correlated with the clump mass: higher-mass clumps have a lower SFE. Due to observational resolution limitations, many of the so-called high-mass cores may actually be sub-clumps, which eventually form multiple systems, rather than just a single star. Furthermore, we should keep in mind that the cores in high-mass star formation regions can continue to accrete and accumulate mass during the process of star formation. These need to be considered when estimating ϵ_{core}.

To align the synthesized CMF with the observed CMF, we need to assume lower-mass cores that have higher ϵ_{core}, and then adjust the range of ϵ_{core} accordingly. As is shown in Figs. 2f and 4, this approach successfully aligns the synthesized CMFs of W43-MM1/MM2/MM3 with the observed ones. Since the currently observed CMF may be not robust, we do not discuss the selection of optimal parameters further, as this depends on the specific CMF. Ideally, the mass dependence of ϵ_{core} would be directly determined from observations in future work.

In this work, we have assumed a canonical IMF and inversely reconstructed the observed CMFs to investigate potential factors contributing to a top-heavy CMF. The factors discussed in this work may or may not be valid, implying that the top-heavy CMF could be genuine. The variation in the IMF with environmental conditions has been extensively studied (Lee et al. 2009; Gunawardhana et al. 2011; Mor et al. 2019; Dib 2023; Kroupa et al. 2024; del Alcázar-Julià et al. 2025). Rather than asserting the universality of the IMF, this Letter emphasizes

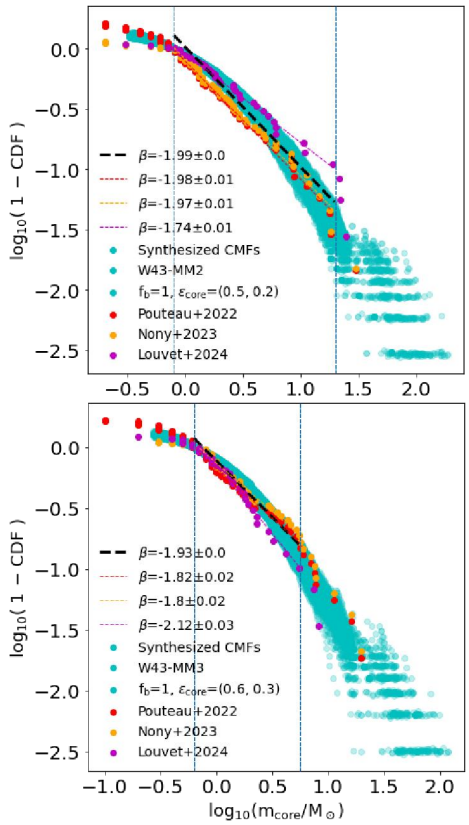


Fig. 4. Same as Fig. 2f for W43-MM1. Here is the comparison of the synthesized CMFs and the observed CMFs for W43-MM2/MM3, assuming that $\epsilon_{\rm core}$ increases with decreasing core mass. All CCDFs (both observed and synthesized) were normalized to 1 at the mass completeness level marked by the left vertical dashed line.

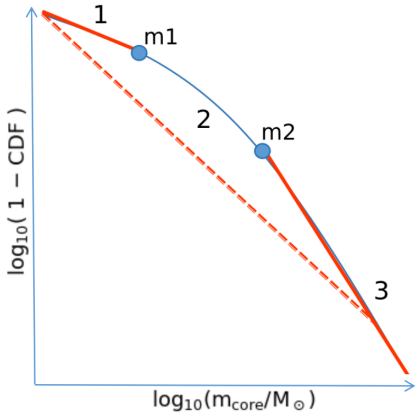


Fig. 5. The synthesized CMF can be roughly divided into three segments, i.e., "1", "2", and "3". Segment "2" is the transition part. "m1" and "m2" mark the mass range of the transition. Segment "3" is the real CMF of high-mass cores. The dashed red line is the observed CMF, for which observational biases prevent one from distinguishing the transition, resulting in a combined fit of three different segments with a shallower slope. Thus, the observed top-heavy CMF is an artifact.

potential issues in the CMF fitting that warrant careful consideration.

Acknowledgements. We thank the referee for providing constructive review comments, which have helped to improve and clarify this work. PK thanks the DAAD Eastern European Exchange Programme at Bonn and Charles University for support.

References

Alves, J., Lombardi, M., & Lada, C. J. 2007, A&A, 462, L17
André, P., Di Francesco, J., Ward-Thompson, D., et al. 2014, in Protostars and Planets VI, eds. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 27
Belloni, D., Askar, A., Giersz, M., Kroupa, P., & Rocha-Pinto, H. J. 2017, MNRAS, 471, 2812
Chen, X., Arce, H. G., Zhang, Q., et al. 2013, ApJ, 768, 110
Cheng, Y., Tan, J. C., Liu, M., et al. 2018, ApJ, 853, 160
Dabringhausen, J., Kroupa, P., & Baumgardt, H. 2009, MNRAS, 394, 1529
del Alcázar-Julià, M., Figueras, F., Robin, A. C., Bienaymé, O., & Anders, F. 2025, A&A, submitted
Dib, S. 2023, ApJ, 959, 88
Dib, S., Kim, J., & Shadmehri, M. 2007, MNRAS, 381, L40
Di Francesco, J., Keown, J., Fallscheer, C., et al. 2020, ApJ, 904, 172
Elmegreen, B. G., & Shadmehri, M. 2003, MNRAS, 338, 817
Enoch, M. L., Evans, N. J., Sargent, A. I., et al. 2008, ApJ, 684, 1240
Goodwin, S. P., & Kroupa, P. 2005, A&A, 439, 565
Gunawardhana, M. L. P., Hopkins, A. M., Sharp, R. G., et al. 2011, MNRAS, 415, 1647
Izquierdo, A. F., Galván-Madrid, R., Maud, L. T., et al. 2018, MNRAS, 478, 2505
Koen, C. 2006, MNRAS, 365, 590
Könyves, V., André, P., Men'shchikov, A., et al. 2010, A&A, 518, L106
Könyves, V., André, P., Men'shchikov, A., et al. 2015, A&A, 584, A91
Kroupa, P. 1995, MNRAS, 277, 1507
Kroupa, P. 2001, MNRAS, 322, 231
Kroupa, P. 2011, in Computational Star Formation, eds. J. Alves, B. G. Elmegreen, J. M. Girart, & V. Trimble, IAU Symp., 270, 141
Kroupa, P., Gjergo, E., Jerabkova, T., & Yan, Z. 2024, ArXiv e-prints [arXiv:2410.07311]
Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87
Küpper, A. H. W., Maschberger, T., Kroupa, P., & Baumgardt, H. 2011, MNRAS, 417, 2300
Lee, J. C., Gil de Paz, A., Tremonti, C., et al. 2009, ApJ, 706, 599
Li, S., Sanhueza, P., Beuther, H., et al. 2024, Nat. Astron., 8, 472
Liu, M., Tan, J. C., Cheng, Y., & Kong, S. 2018, ApJ, 862, 105
Looney, L. W., Mundy, L. G., & Welch, W. J. 2000, ApJ, 529, 477
Louvet, F., Sanhueza, P., Stutz, A., et al. 2024, A&A, 690, A33
Lu, X., Cheng, Y., Ginsburg, A., et al. 2020, ApJ, 894, L14
Maíz Apellániz, J., & Úbeda, L. 2005, ApJ, 629, 873
Marks, M., & Kroupa, P. 2012, A&A, 543, A8
Marks, M., Kroupa, P., Dabringhausen, J., & Pawlowski, M. S. 2012, MNRAS, 422, 2246
Maschberger, T., & Kroupa, P. 2009, MNRAS, 395, 931
Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, AJ, 151, 5
Mor, R., Robin, A. C., Figueras, F., Roca-Fàbrega, S., & Luri, X. 2019, A&A, 624, L1
Motte, F., Nony, T., Louvet, F., et al. 2018, Nat. Astron., 2, 478
Motte, F., Bontemps, S., Csengeri, T., et al. 2022, A&A, 662, A8
Nony, T., Galván-Madrid, R., Motte, F., et al. 2023, A&A, 674, A75
Offner, S. S. R., Moe, M., Kratter, K. M., et al. 2023, ASP Conf. Ser., 534, 275
Oh, S., Kroupa, P., & Pflamm-Altenburg, J. 2015, ApJ, 805, 92
Olguin, F. A., Sanhueza, P., Guzmán, A. E., et al. 2021, ApJ, 909, 199
Olguin, F. A., Sanhueza, P., Ginsburg, A., et al. 2022, ApJ, 929, 68
O'Neill, T. J., Cosentino, G., Tan, J. C., Cheng, Y., & Liu, M. 2021, ApJ, 916, 45
Pouteau, Y., Motte, F., Nony, T., et al. 2022, A&A, 664, A26
Sadaghiani, M., Sánchez-Monge, Á., Schilke, P., et al. 2020, A&A, 635, A2
Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444
Sanhueza, P., Contreras, Y., Wu, B., et al. 2019, ApJ, 886, 102
Tobin, J. J., Offner, S. S. R., Kratter, K. M., et al. 2022, ApJ, 925, 39
Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059
Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389
Wang, L., Kroupa, P., & Jerabkova, T. 2019, MNRAS, 484, 1843
Yan, Z., Jerabkova, T., & Kroupa, P. 2017, A&A, 607, A126
Yan, Z., Jerabkova, T., & Kroupa, P. 2023, A&A, 670, A151
Zhou, J.-W., & Davis, T. 2024, PASA, 41, e076
Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038
Zhou, J. W., Wyrowski, F., Neupane, S., et al. 2023, A&A, 676, A69
Zhou, J. W., Dib, S., Juvela, M., et al. 2024a, A&A, 686, A146
Zhou, J. W., Kroupa, P., & Dib, S. 2024b, PASP, 136, 094301
Zhou, J. W., Dib, S., & Kroupa, P. 2024c, PASP, 136, 094302
Zhou, J.-W., Kroupa, P., & Dib, S. 2024d, A&A, 688, L19

Appendix A: Supplementary figures

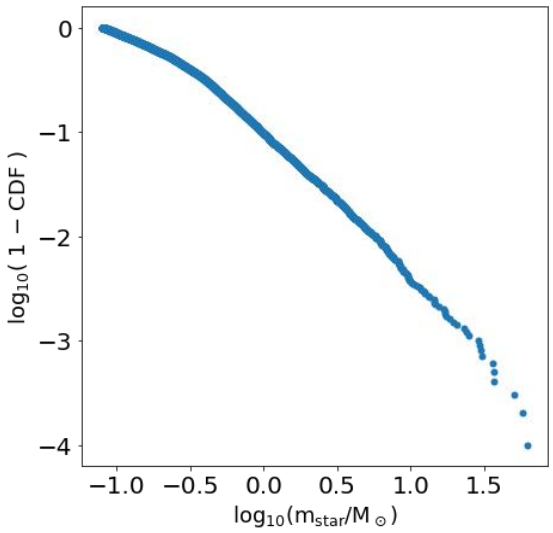


Fig. A.1. Complementary cumulative distribution function (CCDF) plot of the IMF in equation B.1 by taking $m_{\rm max} = 100\,M_\odot$ and $\alpha_3 = 2.3$.

Appendix B: Initial stellar populations

Kroupa (1995) proposed a model for the redistribution of energy and angular momentum within proto-binary systems, which directly results in binary characteristics such as mass ratio, eccentricity and period by the time the dynamical evolution of the star cluster begins to take effect. This process effectively generates the initial binary population (IBP), which serves as the starting point for star cluster simulations. The model has been validated against observational data and has successfully explained the binary properties observed in young clusters, associations, and Galactic field late-type binaries (Kroupa 2011; Marks & Kroupa 2012). The process of transforming birth binaries into initial binaries is commonly referred to as pre-main-sequence eigenevolution (Kroupa 1995), as it takes place during the pre-main-sequence phase of binary stars. The term "birth" refers to all protostars still embedded in circumprotostellar material. In contrast, "initial" refers to pre-main-sequence stars that are no longer embedded in such material, as it has been either accreted or expelled during the redistribution of energy and angular momentum. Over a period of approximately 10^5 years, this pre-main-sequence eigenevolution converts the birth population into the initial population. Belloni et al. (2017) presents an updated version of pre-main-sequence eigenevolution, with the primary difference being the mass ratio distribution in the initial binary populations.

To undergo pre-main-sequence eigenevolution, binaries at birth are formed with specific distributions. The primary mass is selected randomly from the canonical IMF (Kroupa 2001):

$$\xi_\star(m) = \begin{cases} 0, & m < 0.08\,\mathrm{M}_\odot, \\ 2k_\star m^{-1.3}, & 0.08\,\mathrm{M}_\odot \le m < 0.5\,\mathrm{M}_\odot, \\ k_\star m^{-2.3}, & 0.5\,\mathrm{M}_\odot \le m < 1\mathrm{M}_\odot, \\ k_\star m^{-\alpha_3}, & 1\,\mathrm{M}_\odot \le m < m_{\rm max}, \\ 0, & m_{\rm max} \le m, \end{cases} \tag{B.1}$$

where $\alpha_3 = 2.3$ is the constant Salpeter-Massey index for the invariant canonical IMF but will change for larger $\rho_{\rm cl}$ (the density of star-forming clump) to account for IMF variation under star-burst conditions (Elmegreen & Shadmehri 2003; Dib et al. 2007; Dabringhausen et al. 2009; Marks et al. 2012; Dib 2023; Kroupa et al. 2024). As verified in Zhou et al. (2024c), we take $\alpha_3 = 2.3$ in this work. The value of $m_{\rm max}$ is decided by the theoretical $m_{\rm max} - M_{\rm ecl}$ relation (Yan et al. 2023; Zhou et al. 2024c). The secondary mass is also randomly drawn from the same IMF, meaning the binary components are paired randomly. It's important to note that the primary and secondary stars are distinguished only after both stars have been independently chosen from the canonical IMF.

In Kroupa (1995), the pre-main-sequence eigenevolution theory was developed mainly to explain observed properties of Galactic field low-mass binaries (A, F, G, K and M-dwarfs). We note that low-mass binaries in this work correspond to all binaries whose primary mass is smaller than $5\,\mathrm{M}_\odot$. Binaries with primary masses greater than this are called high-mass binaries. For high-mass binaries with O and B-dwarfs as primaries (greater than $5\,\mathrm{M}_\odot$), we utilize observational distributions derived from O and B-dwarfs in Sana et al. (2012) to generate the birth population. Here we assume that the process of transforming the birth population into the initial population for high-mass binaries has already occurred, and directly utilize observational distributions, since this transformation occurs rapidly for massive stars. The mass ratio distribution of high-mass binaries is a uniform distribution (Oh et al. 2015; Belloni et al. 2017). The eccentricity and period distributions of both low- and high-mass binaries are described in Belloni et al. (2017).

MNRAS **000**, 000–000 (2024) Preprint 13 April 2025 Compiled using MNRAS LATEX style file v3.2

Clump populations in molecular clouds

J. W. Zhou ⋆[1] Pavel Kroupa, [2,3]
[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany*
[2]*Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany*
[3]*Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We investigated the physical properties of the 31 largest clump complexes and their associated molecular clouds in the Milky Way. The physical parameters of the clump mass function for a given clump complex exhibit strong correlations with the mass of its associated molecular cloud. We constructed an analytical model for the clump mass function of a cloud, where the model parameters depend solely on the cloud mass. For molecular clouds in nearby galaxies, this model can be used to generate their internal clump populations based on mass. Almost all clump complexes have the Q parameters less than 0.8, with an average value of ∼0.64, implying the widespread presence of substructures within clump complexes. Clump complexes exhibiting clear mass segregation are linked to more massive molecular clouds, which are typically more evolved. This mass segregation may result from gravitational mass settling or the enhanced gas accretion efficiency of centrally located clumps, allowing them to accumulate mass more effectively.

Key words: Submillimeter: ISM – ISM: structure – ISM: evolution –stars: formation – stars: luminosity function, mass function – method: statistical

1 INTRODUCTION

Galaxies act as stellar factories in the universe, producing stars through the gravitational collapse of the densest regions within molecular clouds in the interstellar medium (ISM). Molecular clouds are widely distributed within galaxies and serve as localized sites for star formation. Each molecular cloud typically contains multiple clumps, which are the local sites of star formation within the cloud and the precursors of embedded star clusters (Miville-Deschênes et al. 2017; Rosolowsky et al. 2021; Urquhart et al. 2022; Yan et al. 2017; Zhou et al. 2024d,c). As presented in Motte et al. (2018); Vázquez-Semadeni et al. (2019); Kumar et al. (2020); Henshaw et al. (2020); Zhou et al. (2022, 2023, 2024b); Zhou & Davis (2024); Zhou et al. (2024a); Zhou & Li (2025a), the interstellar medium (ISM) from galaxy to dense core scales presents multi-scale/hierarchical hub-filament structures. Like dense core as the hub in clump, clump as the hub in molecular cloud, molecular cloud can be the hub in galaxy. Hubs at different scales as local gravitational centers can continuously accrete surrounding diffuse gas, thereby accumulating mass. Gaining insight into the characteristics of GMCs is crucial for deciphering the relationship between gas and star formation in galaxies (Kennicutt & Evans 2012; Schinnerer & Leroy 2024). With high-resolution observations across various wavelengths, we are now able to break down star formation into individual molecular clouds in nearby galaxies (Leroy et al. 2021a,b; Lee et al. 2022; Emsellem et al. 2022; Lee et al. 2023). High-resolution CO imaging data from ALMA (Atacama Large Millimeter/submillimeter Array) allow for systematic studies of molecular clouds in nearby galaxies.

Since only molecular clouds within the Milky Way can be resolved at the level of their internal structures, they serve as a crucial reference for understanding the internal composition and star formation processes of molecular clouds beyond our galaxy. In this study, we explore the connection between molecular clouds and their internal clumps in the Milky Way, which may serve as a template for dissecting molecular clouds in nearby galaxies. Especially, we focus on the clump mass function of individual molecular clouds, which allows us to generate the population of clumps within a molecular cloud based on its mass. This helps us study the formation and evolution of stellar populations within molecular clouds.

2 DATA AND CATALOGS

Urquhart et al. (2022) provides an up-to-date version of the ATLASGAL clump catalogue. The individual ATLASGAL clumps were organized into complexes using friends-of-friends clustering analysis. This method associates clumps that are closely situated in terms of longitude, latitude, and velocity. The study of Miville-Deschênes et al. (2017) presents a catalog of 8107 molecular clouds identified by CO (1-0) emission that covers the entire Galactic plane. We matched the ATLASGAL clump complexes in Urquhart et al. (2022) with the molecular clouds in Miville-Deschênes et al. (2017) by minimizing the differences in their central coordinates and systemic velocities. We also used the CO (1-0) cube of Dame et al. (2001)[1].

⋆ E-mail: jwzhou@mpifr-bonn.mpg.de

[1] https://lweb.cfa.harvard.edu/rtdc/CO/

© 2024 The Authors

Table 1. Physical parameters of the clump mass function for the 31 largest clump complexes in the Milky Way.

Name	N	l (deg)	b (deg)	D (kpc)	Q	C_1	C_2	β_1	β_2	M_1 ($M_\odot$)	M_2 ($M_\odot$)	M_3 ($M_\odot$)	$M_{\rm clump,tot}$ ($M_\odot$)	$M_{\rm clump,min}$ ($M_\odot$)	$M_{\rm clump,max}$ ($M_\odot$)	$M_{\rm cloud,as}$ ($M_\odot$)	$M_{\rm cloud,re}$ ($M_\odot$)
G030.653-00.017	250	30.653	-0.017	4.90	0.64	4.71	7.88	1.68	2.79	2.15	2.70	3.20	5.10	1.55	3.47	6.06	6.73
G333.121-00.353	245	333.121	-0.353	3.57	0.60	4.23	7.22	1.89	3.09	2.05	2.45	3.05	4.82	1.49	3.22	6.24	6.43
G024.322+00.187	61	24.322	0.187	7.60	0.66	4.09	6.83	1.29	1.89	2.25	2.75	3.30	4.89	2.09	3.82	6.01	6.01
G336.996-00.002	107	336.996	-0.002	4.70	0.66	4.18	6.90	1.60	2.23	2.25	2.95	3.60	4.99	1.56	4.09	5.76	5.98
G012.960-00.168	85	12.96	-0.168	2.60	0.63	3.41	6.09	1.53	3.08	1.70	2.25	2.80	4.24	1.67	2.85	5.82	5.93
G331.389-00.106	87	331.389	-0.106	3.90	0.64	3.66	6.34	1.91	3.38	2.20	2.65	3.25	4.50	1.59	3.23	5.46	5.92
G018.936-00.350	58	18.936	-0.35	4.00	0.68	3.46	6.13	1.41	3.14	1.80	2.40	2.95	4.30	1.71	3.22	5.75	5.91
G305.453+00.065	102	305.453	0.065	4.00	0.63	3.91	6.58	1.91	3.58	2.35	2.80	3.15	4.76	1.88	3.45	5.83	5.90
G023.072-00.349	41	23.072	-0.349	5.00	0.52	3.57	6.24	1.70	2.61	2.15	2.75	3.30	4.42	1.89	3.40	5.91	5.89
G049.261-00.318	93	49.261	-0.318	5.20	0.59	4.14	6.77	1.34	2.51	2.10	2.85	3.70	5.01	2.08	3.90	5.09	5.83
G006.183-00.358	78	6.183	-0.358	3.00	0.61	3.46	6.03	1.26	2.40	1.60	2.20	2.75	4.37	1.51	3.20	5.34	5.72
G332.124-00.070	67	332.124	-0.07	3.10	0.47	3.32	5.84	1.48	2.94	1.80	2.30	2.95	4.24	1.44	2.99	5.28	5.64
G340.248-00.273	61	340.248	-0.273	3.56	0.72	3.30	5.75	1.23	2.71	1.85	2.20	2.80	4.26	1.85	2.98	5.43	5.52
G010.147-00.237	65	10.147	-0.237	2.90	0.71	3.16	5.61	1.41	2.11	1.65	2.10	2.55	4.13	1.47	2.82	4.98	5.51
G345.479+00.954	51	345.479	0.954	2.38	0.63	3.18	5.62	1.61	2.64	1.95	2.40	2.70	4.15	1.48	3.12	5.35	5.49
G331.094-00.414	53	331.094	-0.414	3.98	0.52	3.42	5.85	1.34	2.22	1.85	2.40	3.10	4.39	1.86	3.24	4.98	5.49
G337.690-00.053	38	337.688	-0.053	3.57	0.76	3.03	5.46	1.38	3.75	1.75	2.35	2.70	4.00	1.45	3.07	5.22	5.48
G014.477-00.149	40	14.477	-0.149	3.08	0.54	3.13	5.54	1.37	2.54	1.65	2.40	2.80	4.11	1.67	3.14	5.34	5.44
G327.006-00.161	48	327.006	-0.161	3.46	0.65	3.13	5.49	1.37	2.04	1.90	2.25	2.75	4.12	1.70	2.87	5.47	5.37
G343.620-00.165	45	343.62	-0.165	2.60	0.57	2.83	5.18	1.58	3.07	1.60	2.10	2.50	3.83	1.39	2.71	5.52	5.33
G008.581-00.329	38	8.581	-0.329	4.45	0.52	3.29	5.56	1.25	2.60	1.80	2.45	3.10	4.32	1.51	3.60	5.23	5.20
G018.179-00.303	41	18.179	-0.303	3.41	0.86	2.93	5.18	1.48	2.68	1.70	2.25	2.80	3.96	1.43	2.97	4.82	5.16
G328.166-00.540	41	328.166	-0.54	2.73	0.69	3.02	5.26	1.45	2.20	1.85	2.25	2.68	4.06	1.89	2.87	5.20	5.14
G335.294-00.261	36	335.294	-0.261	2.92	0.76	3.00	5.24	1.43	2.99	1.80	2.25	2.65	4.04	1.84	3.19	5.10	5.13
G351.206+00.703	159	351.206	0.703	1.30	0.52	3.10	5.32	2.03	2.91	1.40	2.30	2.70	4.14	0.72	3.02	4.84	5.11
G016.996+00.719	62	16.996	0.719	1.86	0.58	2.77	4.93	1.77	2.87	1.45	2.10	2.45	3.83	1.28	2.81	4.96	4.99
G015.079-00.625	55	15.079	-0.625	2.00	0.70	2.96	5.09	1.76	2.44	1.75	2.10	2.80	4.02	1.33	3.33	5.02	4.94
G353.118+00.702	136	353.118	0.702	1.35	0.72	2.81	4.88	1.49	2.91	1.20	1.70	2.15	3.87	0.84	2.24	4.80	4.84
G326.616+00.648	67	326.616	0.648	1.82	0.64	2.76	4.81	1.35	2.50	1.20	1.80	2.53	3.83	1.13	2.61	4.72	4.80
G344.102-00.632	38	344.102	-0.632	2.21	0.72	2.94	4.98	1.34	2.23	1.75	2.25	2.70	4.01	1.76	2.76	4.78	4.78
G014.227-00.571	66	14.227	-0.571	1.50	0.61	2.77	4.79	1.15	2.18	1.20	1.75	2.16	3.84	1.20	2.52	4.19	4.75

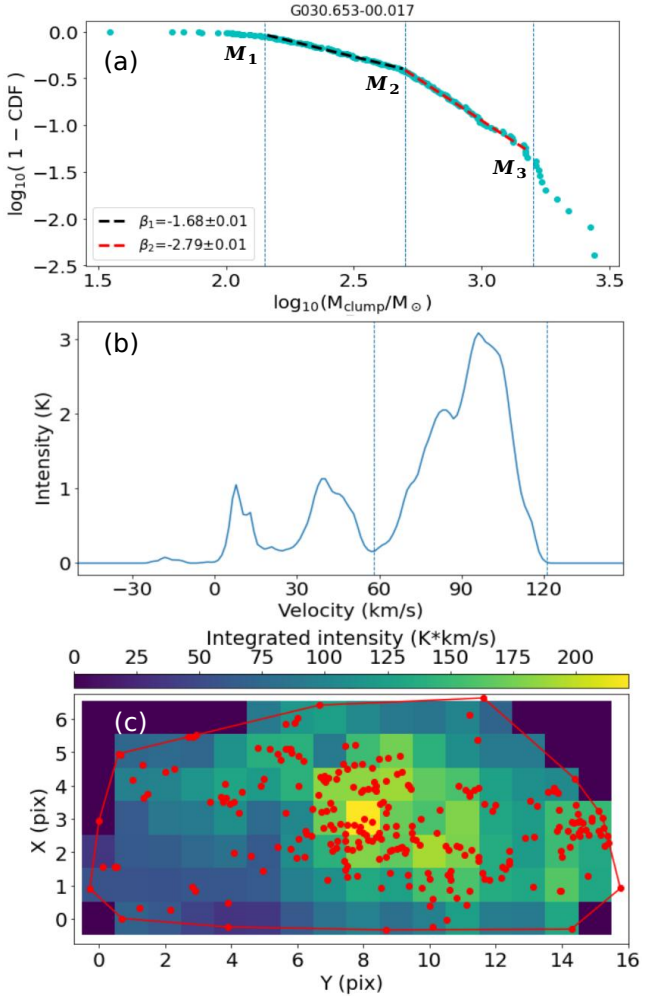


Figure 1. Clump complex G030.653-00.017 as an example. (a) Clump mass function. The complementary cumulative distribution function (CCDF) plot and the linear fitting of the observed CMF in the mass range marked by the vertical dashed lines. The two segments of the clump mass function correspond to three mass endpoints, i.e. M_1, M_2 and M_3; (b) Average spectrum of CO (1-0) emission for all molecular gas in the region marked in panel (c). Two vertical lines mark the velocity range of the molecular gas covering the clump complex; (c) The convex hull of G030.653-00.017, which is the smallest convex shape that encloses all given clumps. Red dots are the clumps.

3 RESULTS AND DISCUSSION

3.1 Clump mass function

3.1.1 The broken masses and the slopes of the clump mass function

A common method for deriving the clump mass function (CMF) is to fit a linear relationship to data that has been grouped into uniformly sized bins in logarithmic space. However, this binning process can lead to significant information loss, as it clusters individual data points into groups, potentially masking finer details in the distribution. This effect is particularly problematic at the higher mass end, where bins may contain only a few data points, making slope estimates prone to bias and uncertainty (Maíz Apellániz & Úbeda 2005; Maschberger & Kroupa 2009). To overcome these drawbacks, Koen (2006) utilizes a complementary cumulative distribution function (CCDF) plot, which preserves more detail by avoiding binning altogether.

In order to fit the CMF of each clump complex, it is necessary to ensure a sufficient number of clumps. For all clump complexes, we first sorted them by the number of clumps and then performed fitting sequentially. We found that when the number of clumps in a clump complex is less than 35, the CMF can not exhibit a clear shape. Therefore, we only considered clump complexes with more than 35 clumps, as listed in Table.1.

Typically, the power-law part of the CMF for each clump complex is divided into two segments (flat+steep), as shown in Fig.1(a), Fig.A1 and Fig.A2. We fitted the two segments with different mass ranges separately, and obtained the slopes β_1 and β_2, respectively. As marked in Fig.1(a), the two mass ranges correspond to three mass endpoints, denoted as M_1, M_2 and M_3.

3.1.2 Correlation with molecular cloud mass

The mass of the associated molecular clouds ($M_{\rm cloud,as}$) in the catalog of Miville-Deschênes et al. (2017) has a clear correlation with M_2 in Fig.2(a). However, the molecular cloud associated with a clump complex may not fully encompass the entire complex. In other words, it may not accurately reflect the gaseous environment in which the clump complex resides. Therefore, based on the spatial distribution and coordinates of the clumps, we computed the convex hull of each clump complex, which is the smallest convex shape that encloses all given clumps, as shown in Fig.1(c). Then we used the same convex shape of each clump complex to extract the average spectrum according to the CO cube and fixed the velocity range of the CO emission covering the clump complex to calculate the total molecular gas mass ($M_{\rm cloud,re}$), following the method presented in Miville-Deschênes et al. (2017). Comparing Fig.2 (a) and (b), $M_{\rm cloud,re}$ exhibits a slightly better correlation with M_2 than $M_{\rm cloud,as}$. Moreover, M_1, M_3, $M_{\rm clump,min}$ and $M_{\rm clump,max}$ also have clear correlations with $M_{\rm cloud,re}$. $M_{\rm clump,min}$ and $M_{\rm clump,max}$ are the smallest and largest clump masses in the clump complex. Both β_1 and β_2 have weak correlations with $M_{\rm cloud,re}$. For all clouds, the median values of β_1 and β_2 are 1.45 and 2.64.

3.1.3 Analytical model of the clump mass function

As discussed in Urquhart et al. (2022), the mass distribution of the ATLASGAL clumps at different evolutionary stages are quite comparable. Thus, we assume that the clump mass remains constant during the evolution process. Since all clumps within a single molecular cloud follow the two-segment distribution presented above, an analytical expression of the clump mass function for a single molecular cloud is,

$$\xi_{\rm clump}(M_{\rm clump}, M_{\rm cloud}) = \begin{cases} C_1 * M_{\rm clump}^{-\beta_1}, \\ M_{\rm clump,min} \le M_{\rm clump} \le M_2; \\ C_2 * M_{\rm clump}^{-\beta_2}, \\ M_2 \le M_{\rm clump} \le M_{\rm clump,max}, \end{cases} \quad (1)$$

where C_1 and C_2 are the normalization constants. According to the fitted correlations in Fig.2 and Fig.3,

$$M_{\rm clump,min} = 10^{-0.02\pm0.59} * M_{\rm cloud}^{0.29\pm0.11}, \quad (2)$$

$$M_{\rm clump,max} = 10^{0.35\pm0.66} * M_{\rm cloud}^{0.50\pm0.12}, \quad (3)$$

$$M_2 = 10^{-0.21\pm0.43} * M_{\rm cloud}^{0.46\pm0.08}, \quad (4)$$

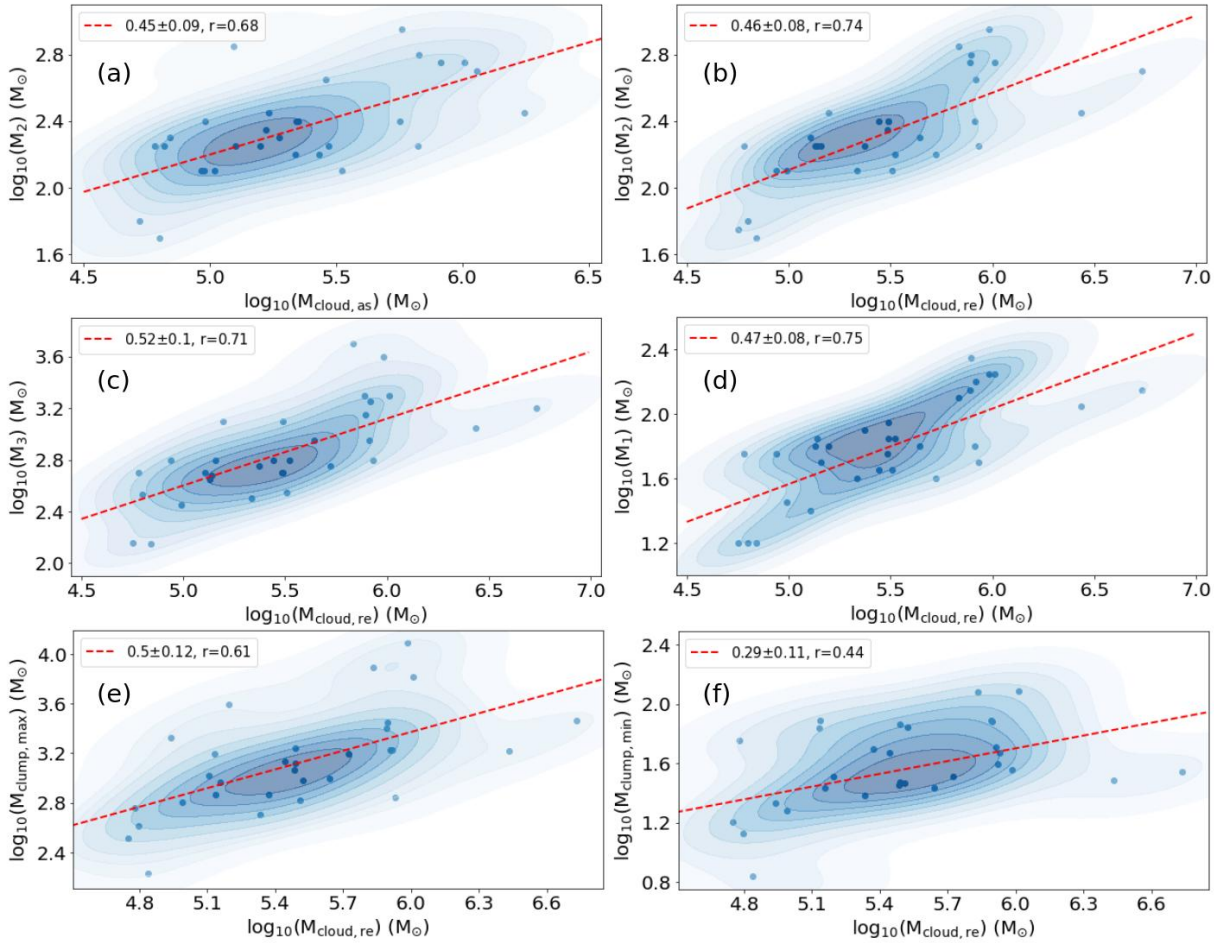


Figure 2. Correlations between the broken masses of the clump mass function (M_1, M_2 and M_3), the smallest and largest clump masses in the clump complex ($M_{\rm clump,min}$ and $M_{\rm clump,max}$) and the cloud mass ($M_{\rm cloud,as}$ or $M_{\rm cloud,re}$). r is the Pearson correlation coefficient.

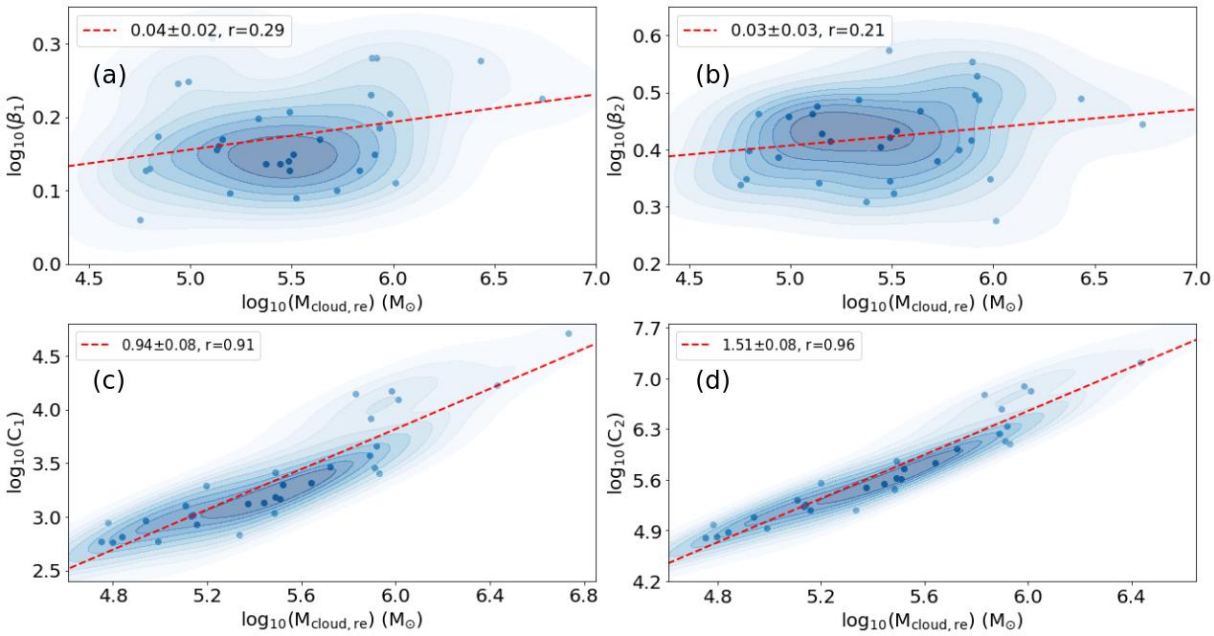


Figure 3. Correlations between the slopes (β_1 and β_2) and the normalization constants (C_1 and C_2) of the clump mass function and the cloud mass. r is the Pearson correlation coefficient.

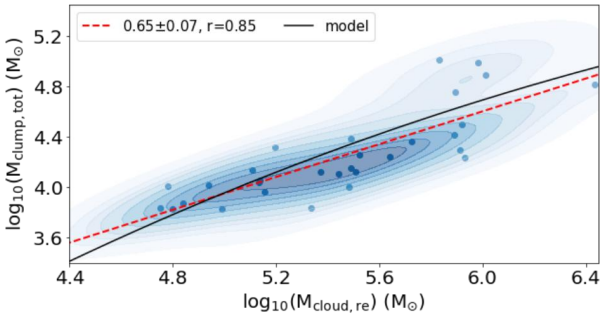


Figure 4. Correlation between the total mass of the clumps within a cloud ($M_{\rm clump,tot}$) and the cloud mass. r is the Pearson correlation coefficient. Red dashed line and black line represent the linear fitting and the model, respectively.

$$\beta_1 = 10^{-0.03\pm0.13} * M_{\rm cloud}^{0.04\pm0.02}, \tag{5}$$

and

$$\beta_2 = 10^{0.25\pm0.15} * M_{\rm cloud}^{0.03\pm0.03}. \tag{6}$$

The total mass of the clumps within a cloud is

$$M_{\rm clump,tot} = \int_{M_{\rm clump,min}}^{M_{\rm clump,max}} M_{\rm clump}\, \xi_{\rm clump}(M_{\rm clump}, M_{\rm cloud})\, {\rm d}M_{\rm clump}. \tag{7}$$

The continuity condition of the clump mass function is

$$C_1 * M_2^{-\beta_1} = C_2 * M_2^{-\beta_2}. \tag{8}$$

Solving equations (1)-(6), we derived a correlation between $M_{\rm clump,tot}$ and $M_{\rm cloud}$. Combined with equation (7), after substituting the values of $M_{\rm clump,tot}$ and $M_{\rm cloud,re}$ in Table.1, we obtained the values of C_1 and C_2 for each cloud, as listed in Table.1. In Fig.3, C_1 and C_2 have strong correlations with $M_{\rm cloud,re}$,

$$C_1 = 10^{-1.81\pm0.44} * M_{\rm cloud}^{0.94\pm0.08}, \tag{9}$$

$$C_2 = 10^{-2.51\pm0.44} * M_{\rm cloud}^{1.51\pm0.08}. \tag{10}$$

3.1.4 Model validation

Fig.4 shows a strong correlation between $M_{\rm clump,tot}$ and $M_{\rm cloud,re}$. The linear fitting gives

$$M_{\rm clump,tot} = 10^{0.7\pm0.41} * M_{\rm cloud}^{0.65\pm0.07}. \tag{11}$$

To test the validity of the model in Sec.3.1.3, we compared the theoretical and observed $M_{\rm clump,tot} - M_{\rm cloud,re}$ relations, and found that the model can roughly follow the observation, despite the significant uncertainty in the model parameters. Alternatively, in addition to the analytical formalism in Sec.3.1.3, we provided the specific parameters of the clump mass function for each molecular cloud in Table.1, allowing for the selection of an appropriate set of parameters based on the cloud mass.

3.2 Spatial distribution of clumps

Next, we further investigated the spatial distribution of clump complexes within molecular clouds. We first utilized the Q parameter to examine the spatial structure of the cluster complex. The Q parameter defined in Cartwright & Whitworth (2004) is used to quantify the spatial distribution of stars in a cluster, which is given by

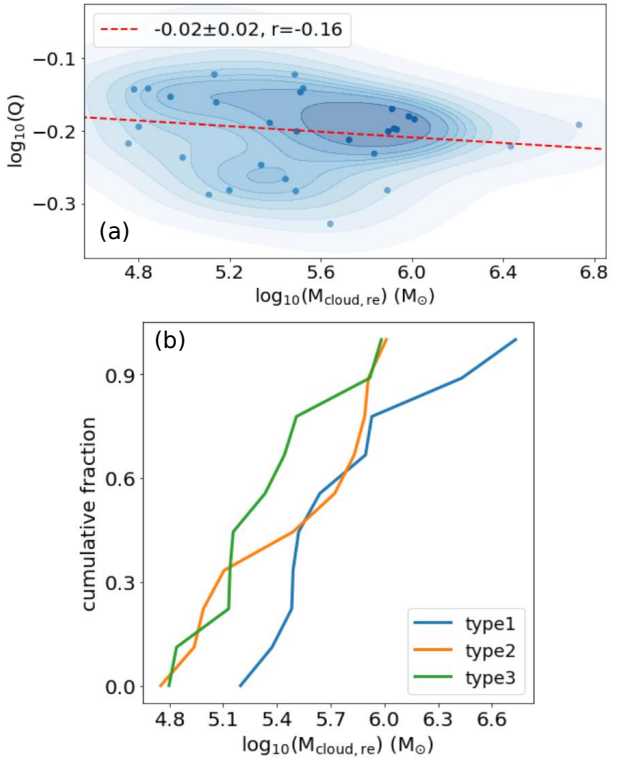


Figure 5. The spatial structure of clump complexes. (a) Correlation between the Q parameter and the cloud mass. r is the Pearson correlation coefficient; (b) Mass distribution of molecular clouds associated with clump complexes exhibiting different levels of mass segregation.

$$Q = \frac{\bar{m}}{\bar{s}}. \tag{12}$$

$\bar{m}$ is the normalized mean edge length of the minimum spanning tree (MST). As discussed in Cartwright & Whitworth (2004), the mean edge length of the MST depends on the number of stars in the cluster (N). As N increases, more short edges are formed on the MST, causing $\bar{m}$ to decrease. Therefore, a factor $(NA)^{1/2}/(N-1)$ should be applied to normalize the mean edge length for clusters with different numbers of stars, then we have

$$\bar{m} = \frac{\sum_{i=1}^{N-1} L_i}{(NA)^{1/2}}, \tag{13}$$

where L_i is the length of each edge, and A is the area of cluster as $A = \pi r_{\rm cluster}^2$. $\bar{s}$ is the normalized correlation length,

$$\bar{s} = \frac{L_{\rm a}}{r_{\rm cluster}}, \tag{14}$$

where $L_{\rm a}$ is the the mean separation length between all stars. Finally, we have

$$Q = \frac{\sum_{i=1}^{N-1} L_i}{L_{\rm a}(N\pi)^{1/2}}. \tag{15}$$

The Q value quantifies the degree of subclustering and the overall radial density gradient within a specific region. Q $\geq$ 0.8 indicates a centrally concentrated spatial distribution, while Q $<$ 0.8 suggests the presence of subclustering, with smaller values representing a higher degree of subclustering.

By analogy between "clumps and clump complexes" with "stars and star clusters", the above analysis can also be applied to study the spatial structure of clump complexes. As shown in Table.1, almost all clump complexes have the Q parameters less than 0.8, with an average value ~0.64, which implies the widespread presence of substructures within clump complexes. In Fig.5(a), there is no clear correlation between the Q parameter and the cloud mass.

3.3 Mass segregation

We further investigated the mass segregation of clumps within the clump complex. Allison et al. (2009) proposed the mass segregation ratio (Λ) as a metric for detecting and quantifying mass segregation in star clusters, which is defined as

$$\Lambda = \frac{\langle l_{\rm r} \rangle \pm \sigma_{\rm r}}{l_{\rm m}}, \tag{16}$$

where $l_{\rm m}$ represents the average MST edge length among the N most massive stars in the cluster, and $l_{\rm r}$ denotes the average MST edge length among N randomly selected stars in each set. $\sigma_{\rm r}$ is the standard deviation of the 100 different sets of $l_{\rm r}$, and $\langle l_{\rm r} \rangle$ is the final average length of 100 random sets. When the N massive stars are segregated, $l_{\rm m}$ should be smaller than $l_{\rm r}$, i.e. Λ>1.

Fig.A3 and Fig.A4 show the mass segregation ratio for the 35 most massive members of each clump complex. Typically, they can be divided into three categories: (1) Clear mass segregation (type1). For almost all 35 most massive members, $\Lambda > 1$; (2) Vague (type2). For most of the 35 most massive members, Λ is around 1; (3) No clear mass segregation (type3). For a significant proportion of the 35 most massive members, $\Lambda < 1$.

As shown in Fig.5(b), clump complexes with clear mass segregation are associated with more massive molecular clouds. As verified in Zhou et al. (2024a), more massive molecular clouds are more evolved. Therefore, as a molecular cloud evolves, clumps within it gradually become segregated. A molecular cloud is a gravitational network formed by multi-scale hub-filament structures, with clumps acting as the nodes of this network (Zhou et al. 2025; Zhou & Li 2025b). The emergence of segregation in this gravitational network over time may be a natural process under gravity. As presented above, more massive molecular clouds can form more massive clumps. The mass growth of molecular clouds and their internal clumps occurs simultaneously, which aligns with the picture of hierarchical, multi-scale hub-filament structures and gas inflows from large to small scales (Motte et al. 2018; Vázquez-Semadeni et al. 2019; Kumar et al. 2020; Henshaw et al. 2020; Zhou et al. 2022, 2023, 2024b; Zhou & Davis 2024). In the case of mass segregation, more massive clumps are located closer to the center of the clump complex, which may be due to gravitational mass settling or the fact that centrally located clumps can more efficiently accrete gas and thus accumulate mass more effectively. The latter is similar to a clump-scale "competitive accretion" scenario (Bonnell et al. 1997, 2001), which originally proposed to explain the formation of massive stars in the clump.

4 DATA AVAILABILITY

All the data used in this work are available from the first author.

REFERENCES

Allison, R. J., Goodwin, S. P., Parker, R. J., et al. 2009, Monthly Notices of the Royal Astronomical Society, 395, 1449, doi: 10.1111/j.1365-2966.2009.14508.x

Bonnell, I. A., Bate, M. R., Clarke, C. J., & Pringle, J. E. 1997, MNRAS, 285, 201, doi: 10.1093/mnras/285.1.201

—. 2001, MNRAS, 323, 785, doi: 10.1046/j.1365-8711.2001.04270.x

Cartwright, A., & Whitworth, A. P. 2004, MNRAS, 348, 589, doi: 10.1111/j.1365-2966.2004.07360.x

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792, doi: 10.1086/318388

Emsellem, E., Schinnerer, E., Santoro, F., et al. 2022, A&A, 659, A191, doi: 10.1051/0004-6361/202141727

Henshaw, J. D., Kruijssen, J. M. D., Longmore, S. N., et al. 2020, Nature Astronomy, 4, 1064, doi: 10.1038/s41550-020-1126-z

Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531, doi: 10.1146/annurev-astro-081811-125610

Koen, C. 2006, MNRAS, 365, 590, doi: 10.1111/j.1365-2966.2005.09739.x

Kumar, M. S. N., Palmeirim, P., Arzoumanian, D., & Inutsuka, S. I. 2020, A&A, 642, A87, doi: 10.1051/0004-6361/202038232

Lee, J. C., Whitmore, B. C., Thilker, D. A., et al. 2022, ApJS, 258, 10, doi: 10.3847/1538-4365/ac1fe5

Lee, J. C., Sandstrom, K. M., Leroy, A. K., et al. 2023, ApJ, 944, L17, doi: 10.3847/2041-8213/acaaae

Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021a, ApJS, 257, 43, doi: 10.3847/1538-4365/ac17f3

Leroy, A. K., Hughes, A., Liu, D., et al. 2021b, ApJS, 255, 19, doi: 10.3847/1538-4365/abec80

Maíz Apellániz, J., & Úbeda, L. 2005, ApJ, 629, 873, doi: 10.1086/431458

Maschberger, T., & Kroupa, P. 2009, MNRAS, 395, 931, doi: 10.1111/j.1365-2966.2009.14577.x

Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57, doi: 10.3847/1538-4357/834/1/57

Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41, doi: 10.1146/annurev-astro-091916-055235

Rosolowsky, E., Hughes, A., Leroy, A. K., et al. 2021, MNRAS, 502, 1218, doi: 10.1093/mnras/stab085

Schinnerer, E., & Leroy, A. K. 2024, ARA&A, 62, 369, doi: 10.1146/annurev-astro-071221-052651

Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389, doi: 10.1093/mnras/stab3511

Vázquez-Semadeni, E., Palau, A., Ballesteros-Paredes, J., Gómez, G. C., & Zamora-Avilés, M. 2019, MNRAS, 490, 3061, doi: 10.1093/mnras/stz2736

Yan, Z., Jerabkova, T., & Kroupa, P. 2017, A&A, 607, A126, doi: 10.1051/0004-6361/201730987

Zhou, J.-W., & Davis, T. 2024, Publ. Astron. Soc. Australia, 41, e076, doi: 10.1017/pasa.2024.47

Zhou, J. W., Dib, S., & Davis, T. A. 2024a, MNRAS, 534, 683, doi: 10.1093/mnras/stae2101

Zhou, J. W., Dib, S., Juvela, M., et al. 2024b, A&A, 686, A146, doi: 10.1051/0004-6361/202449514

Zhou, J. W., Dib, S., & Kroupa, P. 2024c, PASP, 136, 094302, doi: 10.1088/1538-3873/ad77f4

Zhou, J. W., Kroupa, P., & Dib, S. 2024d, PASP, 136, 094301, doi: 10.1088/1538-3873/ad6f44

Zhou, J. W., & Li, G.-X. 2025a, MNRAS, 537, 2630, doi: 10.1093/mnras/staf199

—. 2025b, MNRAS, doi: 10.1093/mnras/staf199

Zhou, J. W., Wyrowski, F., Urquhart, J. S., et al. 2025, High-resolution APEX/LAsMA ^{12}CO and ^{13}CO (3-2) observation of the G333 giant molecular cloud complex : III. Decomposition of molecular clouds into multi-scale hub-filament structures

Zhou, J.-W., Liu, T., Evans, N. J., et al. 2022, MNRAS, 514, 6038, doi: 10.

MNRAS **000**, 000–000 (2024) Preprint 29 January 2025 Compiled using MNRAS LATEX style file v3.2

The most massive star clusters in molecular clouds: Insights from the integrated cloud-wide initial mass function (ICIMF) theory

J. W. Zhou ⋆[1] Pavel Kroupa †[2,3] Sami Dib ‡[4]
[1]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany*
[2] *Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), Universität Bonn, Nussallee 14–16, 53115 Bonn, Germany*
[3] *Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic*
[4]*Max Planck Institute für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The combination of the high-resolution ALMA, JWST and HST observations provides unprecedented insights into the connection between individual molecular clouds and their internal stellar populations in nearby galaxies. The molecular clouds in five nearby galaxies were identified based on the integrated intensity maps of CO (2−1) emission from ALMA observations. We used the JWST 21 μm data to estimate the star formation rate (SFR) surface density of the clouds and calculate the masses of the embedded stellar populations in the clouds. After matching the star cluster and stellar association catalogs derived from the HST observations with the identified molecular clouds, we found clear correlations between the physical parameters of molecular clouds and their internal stellar populations. Based on the masses of the total stellar populations and their corresponding clouds, we obtained a typical value of the cloud-scale star formation efficiency (SFE), ≈1.4%. The mass of the most massive cluster ($M_{\rm cluster,max}$) in a cloud is positively proportional to the mass ($M_{\rm cloud}$), the column density, the SFR sand the SFR surface density of the cloud. The observed $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation can be interpreted theoretically on the basis of the integrated cloud-wide IMF (ICIMF) theory, which provides a quantitative framework for understanding the correlations between molecular clouds and their internal stellar populations. Star formation in the Milky Way and nearby galaxies follows similar patterns on the cloud scale. The most massive young star clusters in the clouds of nearby galaxies may be cluster complexes rather than single clusters.

Key words: – ISM: clouds – galaxies: clusters – galaxies: star formation

1 INTRODUCTION

Galaxies act as stellar factories in the universe, producing stars through the gravitational collapse of the densest regions within molecular clouds in the interstellar medium (ISM). Molecular clouds are widely distributed within galaxies and serve as localized sites for star formation. Each molecular cloud typically contains multiple clumps, which are the local sites of star formation within the cloud and the precursors of embedded star clusters (Dib et al. 2012; Miville-Deschênes et al. 2017; Rosolowsky et al. 2021; Urquhart et al. 2022; Zhou et al. 2023). The hierarchical structure of molecular gas in a galaxy determines the spatial distribution of stellar populations (Gouliermis et al. 2015; Elmegreen et al. 2014; Grasha et al. 2017; Chevance et al. 2023). Young clusters with different ages constitute a cluster population within the parent molecular cloud. The evolution of the molecular cloud is closely related to the formation and evolution of the internal star clusters. The star clusters become fully exposed after the molecular cloud is dispersed by feedback from the star clusters (Kim et al. 2016; Kruijssen et al. 2019; Chevance et al. 2022; Watkins et al. 2023; Zhou et al. 2024g).

It has been suggested that most if not all observed stars originate from embedded clusters (Kroupa 1995a,b; Lada & Lada 2003; Kroupa 2005; Megeath et al. 2016; Dinnbier et al. 2022). Based on this viewpoint, the integrated galaxy-wide initial mass function (IGIMF) theory has been developed to calculate galaxy-wide IMFs, starting with the IMF that is observationally constrained at the star cluster scale (Kroupa & Weidner 2003; Kroupa et al. 2013; Yan et al. 2017; Jerabkova et al. 2018; Dib 2022; Haslbauer et al. 2024). The IMFs assembled in star-forming regions combine to create the IMF of the entire galaxy. This process involves first generating a series of embedded star clusters from the embedded cluster mass function (ECMF), then producing a series of stars for each embedded cluster from the IMF, and finally summing the stars from all clusters to construct the galaxy-wide IMF. Star clusters first form in molecular clouds. Molecular clouds, serving as intermediaries between star clusters and the galaxy, should also be incorporated into the IGIMF theory. Molecular clouds and their internal star clusters could be quantitatively characterized using a localized IGIMF theory, which can be called the integrated cloud-wide initial mass function (ICIMF) theory. Fig.1 illustrates the correspondence between hierarchical gas structures (star-forming regions) and hierarchically integrated IMFs, providing a depiction of their relationship.

With high-resolution observations across various wavelengths, we are now able to break down star formation into individual molecular clouds in nearby galaxies (Leroy et al. 2021a,b; Lee et al. 2022;

⋆ E-mail: jwzhou@mpifr-bonn.mpg.de
† E-mail: pkroupa@uni-bonn.de
‡ E-mail: dib@mpia.de

© 2024 The Authors

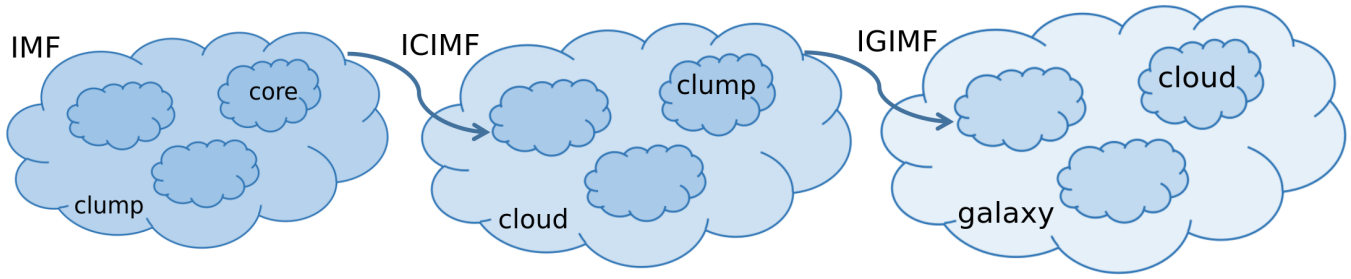


Figure 1. Correspondence between hierarchical gas structures (star-forming regions) and hierarchically integrated IMFs. "ICIMF" and "IGIMF" are the integrated cloud-wide IMF and the integrated galaxy-wide IMF, respectively.

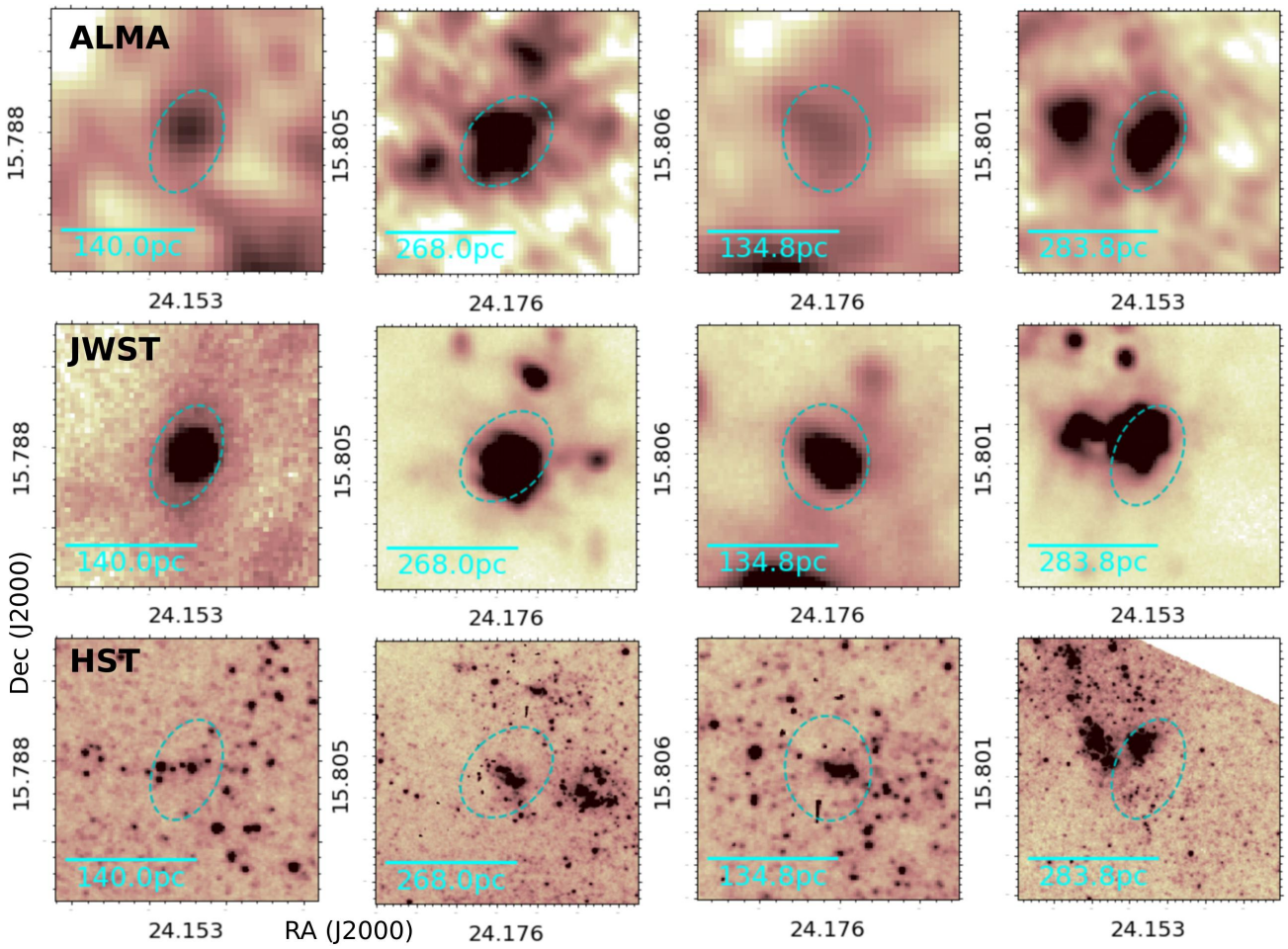


Figure 2. Some examples from Zhou et al. (2024a) are used to demonstrate the spatial distribution of molecular clouds in ALMA observation (the first row) and stellar populations in JWST (the second row) and HST (the third row) observations. Dashed cyan ellipses represent the effective ellipses of the molecular clouds.

Emsellem et al. 2022; Lee et al. 2023). High-resolution CO imaging data from ALMA (Atacama Large Millimeter/submillimeter Array) observations enable systematic studies of molecular clouds across nearby galaxies. The JWST (James Webb Space Telescope) and HST (Hubble Space Telescope) observations can reveal embedded clusters, young star clusters, and stellar associations within molecular clouds (Gouliermis et al. 2017; Whitmore et al. 2021; Thilker et al. 2022; Larson et al. 2023; Levy et al. 2024). The combination of these high-resolution observations provides an unprecedented look at the connection between individual clouds and the stellar populations (Turner et al. 2022). Some examples of multi-band observations are shown in Fig.2.

In this work, we first look for correlations between the physical parameters of molecular clouds and their internal stellar populations, and then use the ICIMF theory to interpret the observed correlations. Finally, we try to unify star formation in the Milky Way and nearby galaxies on the cloud scale based on the ICIMF theory.

2 DATA AND CATALOGS

2.1 Star cluster and stellar association catalogs

The catalogs for compact clusters and multi-scale associations across five galaxies (NGC 1433, NGC 1566, NGC 1559, NGC 3351 and NGC 3627) are being made available as the DR3 release of the PHANGS-HST survey. A detailed description of these catalogs can be found in Lee et al. (2022); Whitmore et al. (2021); Thilker et al. (2022); Larson et al. (2023). Based on the classification technique, there are two kinds of compact cluster catalogs, i.e. human classified and machine learning (ML) classified. The ML classified technique greatly increases the number of identified star clusters within each galaxy, allowing for a much better statistical analysis. Thus, we use the ML classified catalog in this work. According to NUV–U–B–V–I photometry of the star clusters, the publicly available spectral energy distribution (SED) fitting code CIGALE was used to characterize their physical properties, such as their age, mass, and reddening (Turner et al. 2021). For the stellar association catalogs, as described in Larson et al. (2023), they firstly selected point sources from the DOLPHOT source catalogs (Thilker et al. 2022) to serve as 'tracer' stars, using both the HST NUV-band (F275W) and V-band (F555W). The positions of 'tracer' stars are used to create star density maps on which they used the watershed algorithm to identify stellar associations. All light from within the bounds of an association is used in the photometry for that structure. The measured five-band NUV, U, V, B, I photometry gives the total flux of the structure. Using the SED fitting (Turner et al. 2021), they estimated the ages, masses, and reddening of the stellar associations according to the total flux. The associations are identified at three physical scale levels (16 pc, 32 pc and 64 pc) for all five galaxies. The NUV-band selection primarily identifies the youngest stellar populations, while the V-band traces both young and old stars, enabling the detection of stellar associations across a wider range of ages. As described in Larson et al. (2023), the V-band selection identifies more associations at all scale levels. Most of the associations are identified through V-band selection, with 75-85 percent of the young associations found in the NUV-band also overlapping with the larger set of associations detected in the V-band.

In this work, we exclude clusters and associations whose estimated masses are less than their mass estimation errors, as their masses are not well determined.

2.2 CO (2-1) cube and 21 μm map

For the five galaxies, we used the combined 12m+7m+TP PHANGS-ALMA CO (2−1) data cubes to investigate gas kinematics and dynamics, which have a spectral resolution of 2.5 km s^{-1}. We also use the JWST 21 μm (F2100W) maps from the PHANGS-JWST survey. The full width at half maximum (FWHM) of the original F2100W point spread function (PSF) is ≈0.65" (Williams et al. 2024). The angular resolutions or linear resolutions of these datasets are listed in Table.1. Overviews of the PHANGS-ALMA and PHANGS-JWST surveys' science goals, sample selection, observation strategy, and data products are described in Leroy et al. (2021a,b); Lee et al. (2023). All the data are available on the PHANGS team website [1].

[1] https://sites.google.com/view/phangs/home

3 RESULTS

3.1 Identification of molecular clouds

For the five galaxies considered in this work, we conducted a direct identification of hierarchical (sub-)structures based on the 2D intensity maps. As described in Rosolowsky et al. (2008), the dendrogram algorithm decomposes density or intensity data into hierarchical structures called leaves, branches, and trunks. Using the *astrodendro* package [2], there are three major input parameters for the dendrogram algorithm: *min_value* for the minimum value to be considered in the dataset, *min_delta* for a leaf that can be considered as an independent entity, and *min_npix* for the minimum area of a structure. For the CO (2−1) data cube, there are two types of Moment 0 maps (strictly masked and broadly masked) in the data product of the PHANGS-ALMA survey [3]. The strictly masked maps only include emissions that are identified as signals with high confidence in the data cube, which might filter out the relatively faint structures. The broadly masked maps offer superior completeness and cover larger areas compared to the strictly masked maps. However, due to the inclusion of more regions with faint emissions or areas close to bright emissions, they tend to be noisier and may contain false positives. In order to ensure the reliability of the identified structures and because we are only interested in local dense structures, we select the strictly masked Moment 0 maps to identify structures. Since all the retained structures on the strictly masked Moment 0 maps are reliable, we only require the smallest area of the identified structure be larger than 1 beam area. We do not set additional parameters in the algorithm to minimize the dependence of the identification on parameter settings. Finally, the numbers of leaf structures in the five galaxies are listed in Table.2. In Fig.3, the CO structures identified by the dendrogram algorithm exhibit a strong correspondence with the background intensity maps. For a structure with an area A and a total integrated intensity $I_{\rm CO}$, the mass of the structure can be calculated by

$$M_{\rm cloud} = \alpha_{\rm CO}^{2-1} \times I_{\rm CO} \times A, \quad (1)$$

where $\alpha_{\rm CO}^{2-1} \approx 6.7\ {\rm M_\odot pc^{-2}(K\ km\ s^{-1})^{-1}}$ (Leroy et al. 2022).

For the 21 μm emission, apart from the *min_npix*= 1 beam area, we also take the values of *min_value*= 3*$\sigma_{\rm rms}$, *min_delta* = 3*$\sigma_{\rm rms}$, where $\sigma_{\rm rms}$ is the background intensity. We first retained as many structures as possible using these lower standards, then eliminated the diffuse structures, as done in Zhou et al. (2024a).

The algorithm characterizes the morphology of each structure by approximating it as an ellipse. Within the dendrogram, the root mean square (rms) sizes (second moments) of the intensity distribution along the two spatial dimensions define the long and short axes of the ellipse, denoted as a and b. As described in Zhou et al. (2024e), the initial ellipse with a and b is smaller, so a multiplication factor of two is applied to appropriately enlarge the ellipse. Then the effective physical radius of an ellipse is $R_{\rm eff} = \sqrt{2a \times 2b} * D$, where D is the distance of the galaxy.

3.1.1 Velocity component

For the identified molecular structures, we extracted the average spectrum of each structure to investigate its velocity components and

[2] https://dendrograms.readthedocs.io/en/stable/index.html

[3] Details of the masking strategy and completeness statistics are presented in the PHANGS pipeline paper (Leroy et al. 2021b).

Table 1. Resolutions of the datasets. NGC 1559 is not observed by JWST currently. "D" and "i" are the distances and the inclination angles of the galaxies, respectively.

	D (Mpc)	i (deg)	angular (") ALMA	linear (pc) ALMA	linear (pc) JWST
NGC 1433	18.6	28.6	1.10	99.1	58.6
NGC 1559	19.4	65.4	1.25	117.5	–
NGC 1566	17.7	29.5	1.25	107.6	55.7
NGC 3351	10.0	45.1	1.46	70.7	31.5
NGC 3627	11.3	57.3	1.63	89.2	35.6

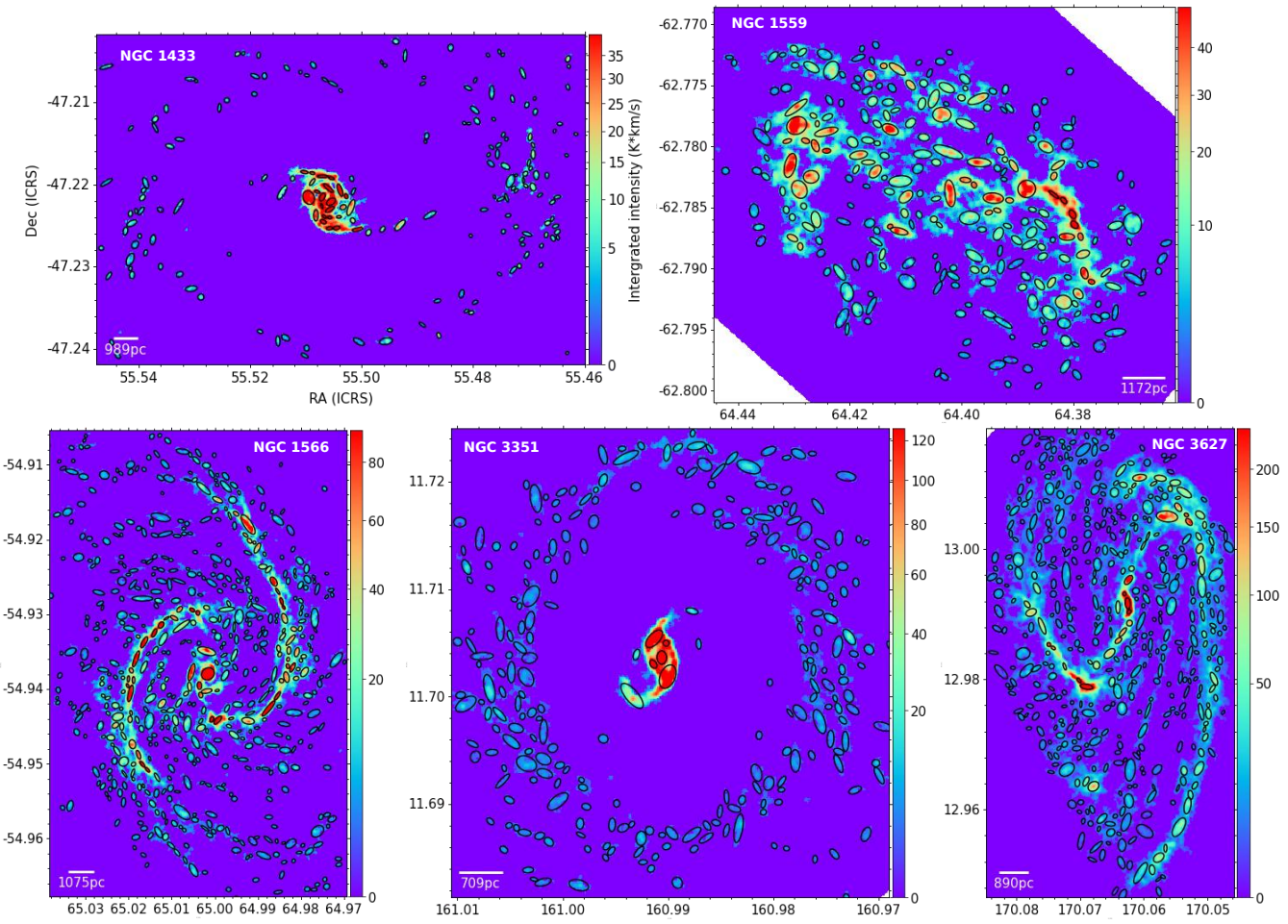


Figure 3. Background is the intergrated intensity map of CO (2−1) emission. Black ellipses represent leaf structures identified by the dendrogram algorithm. The scalebar shows the scale of 10 beam sizes.

gas kinematics. Large-scale velocity gradients from the galaxy's rotation contribute to the non-thermal velocity dispersion. To address this, before extracting the average spectra, we removed the bulk motion due to galaxy rotation by creating gas dynamical models using the Kinematic Molecular Simulation (KinMS) package (Davis et al. 2013). According to the averaged spectra, 67 structures with poor line profiles were eliminated. Then, following the procedure described in Zhou et al. (2024f), we fitted the averaged spectra of 1712 leaf structures individually using the fully automated Gaussian decomposer `GAUSSPY+` algorithm (Lindner et al. 2015; Riener et al. 2019). As shown in Table.2, ≈96% structures have a single velocity component, i.e. type1 (following the nomenclature introduced in Zhou et al. (2024f)). For some structures with more than one velocity component (i.e. type3), as discussed in Zhou et al. (2024g), we treat type3 structures as independent structures like type1 structures, but with more complex gas motions.

3.2 Correlations between molecular clouds and their internal stellar populations

3.2.1 Stellar associations

Since the beam sizes of all five galaxies are more than 64 pc, we only focus on the associations identified on the scale level of 64 pc.

Table 2. Leaf structures in each galaxy, see Sec.3.1.1 for more details.

	leaves	fitted	type1	type3
NGC 1433	187	178	175	3
NGC 1559	262	253	246	7
NGC 1566	588	559	540	19
NGC 3351	271	264	263	1
NGC 3627	471	458	414	44
total	1779	1712	1638	74

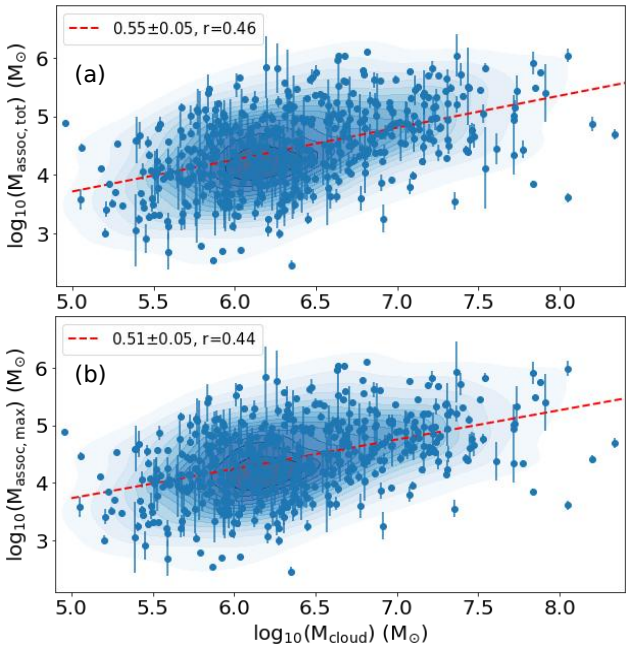


Figure 4. (a) The correlation between the mass of the cloud ($M_{\rm cloud}$) and the total mass of the associations in the cloud ($M_{\rm assoc,tot}$), i.e. the $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relation; (b) The correlation between $M_{\rm cloud}$ and the mass of the most massive association ($M_{\rm assoc,max}$) in the cloud, i.e. the $M_{\rm assoc,max}$-$M_{\rm cloud}$ relation. r is the Pearson correlation coefficient.

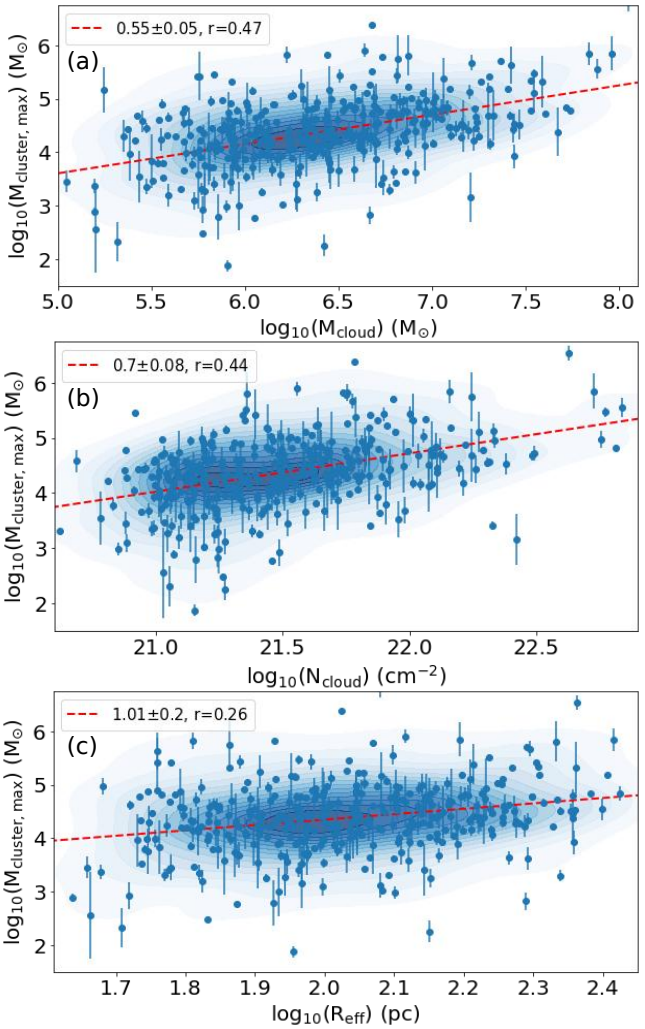


Figure 5. (a) The correlation between the mass of the cloud and the mass of the most massive cluster ($M_{\rm cluster,max}$) in the cloud, i.e. the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation; (b) The correlation between the column density of the cloud and $M_{\rm cluster,max}$, i.e. the $M_{\rm cluster,max}$-$N_{\rm cloud}$ relation; (c) The correlation between the effective radius of the cloud and $M_{\rm cluster,max}$. r is the Pearson correlation coefficient.

If the distances between the center coordinates of the associations and a molecular cloud are less than the effective radius of the cloud, we distribute these associations to the cloud. If the associations from the V-band selection and the NUV-band selection overlap, we select the one with a larger area which includes more stars. Actually, the association from the V-band selection always has a larger area, as described in Sec.2.1. Finally, we added the masses of all independent associations in each cloud. Combining the data of all five galaxies together, Fig.4(a) shows a clear correlation between the mass of the cloud and the total mass of the associations in the cloud, i.e. the $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relation. For each cloud, we also singled out the most massive association with the mass $M_{\rm assoc,max}$. As displayed in Fig.4(b), the $M_{\rm assoc,max}$-$M_{\rm cloud}$ and $M_{\rm assoc,tot}$-$M_{\rm cloud}$ relations are very similar, and this is because one cloud typically corresponds to one association identified at the 64 pc level.

3.2.2 Star clusters

Using the same criterion, if the distances between the center coordinates of the clusters and a molecular cloud are less than the effective radius of the cloud, we distribute these clusters to the cloud. Generally, one cloud will contain multiple star clusters. We first focus on the most massive one. Fig.5(a) displays a clear correlation between the mass of the cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation. Moreover, there is also a clear correlation between the column density of the cloud ($N_{\rm cloud}$) and $M_{\rm cluster,max}$, i.e. the $M_{\rm cluster,max}$-$N_{\rm cloud}$ relation. However, there is only a weak correlation between the size of the cloud ($R_{\rm eff}$) and $M_{\rm cluster,max}$.

3.2.3 Total stellar populations

A molecular cloud may contain only star clusters (case1), only stellar associations (case2), or both star clusters and stellar associations (case3). For case3, the clusters may be inside the associations or independent of them. To estimate the total mass of the stellar populations within the cloud, we need to sum the masses of both the independent clusters and the associations. To avoid redundant calculations of mass, the clusters that fall within the associations should be excluded. Similarly, if the distances between the center coordi-

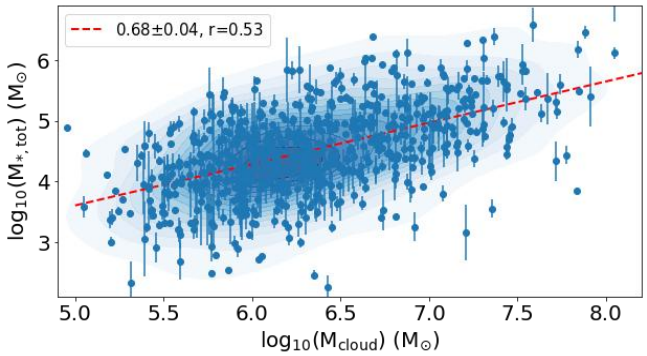


Figure 6. The correlation between the mass of the cloud and the mass of the total stellar populations ($M_{*,\rm tot}$) in the cloud, i.e. the $M_{*,\rm tot}$-$M_{\rm cloud}$ relation. r is the Pearson correlation coefficient.

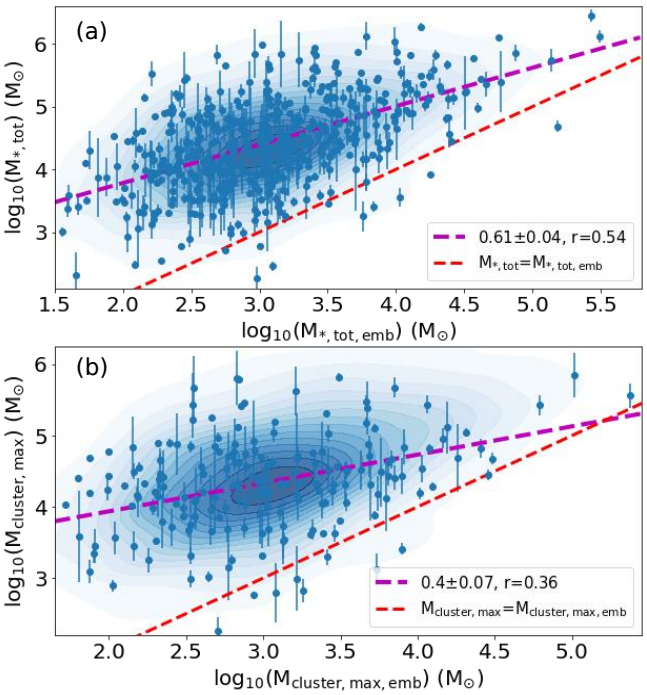


Figure 7. (a) Comparison between $M_{*,\rm tot,emb}$ and $M_{*,\rm tot}$; (b) Comparison between $M_{\rm cluster,max,emb}$ and $M_{\rm cluster,max}$. r is the Pearson correlation coefficient.

nates of the clusters and an association are less than the effective radius of the association, we distribute these clusters to the association and do not count them. Fig.6 displays a clear correlation between the mass of the cloud and the mass of the total stellar populations in the cloud, i.e. the $M_{*,\rm tot}$-$M_{\rm cloud}$ relation.

Since star clusters and associations being associated with molecular clouds inherently indicates their youth, we did not impose age restrictions on the samples in the above analysis, as age estimates are subject to significant uncertainties.

3.3 Embedded stellar populations

The embedded stellar populations in molecular clouds may not be observed in the PHANGS-HST survey, which is more sensitive to the exposed stellar populations. As done in Zhou et al. (2024a), the JWST 21 μm emission was used to estimate the total mass of the embedded stellar populations in a cloud. This method gives quite comparable masses with those obtained from the SED fitting, as compared in Zhou et al. (2024a). Using the prescriptions described in Leroy et al. (2021a), the WISE 22 μm data can be used to calculate the local star formation rate (SFR) surface density via

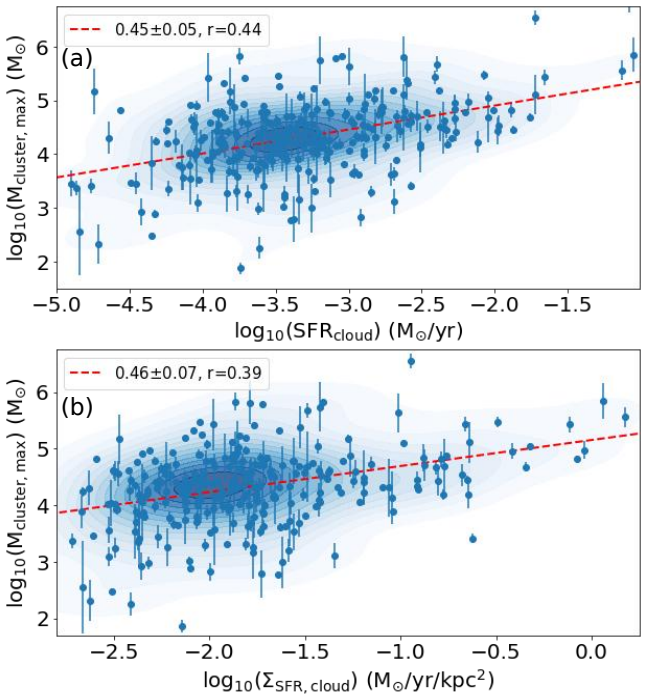


Figure 8. (a) The correlation between the star formation rate (SFR) of the cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\rm cluster,max}$-$\rm SFR_{cloud}$ relation; (b) The correlation between the SFR surface density of the cloud and $M_{\rm cluster,max}$, i.e. the $M_{\rm cluster,max}$-$\Sigma_{\rm SFR,cloud}$ relation. r is the Pearson correlation coefficient.

$$\frac{\Sigma_{\rm SFR}}{\rm M_\odot\ yr^{-1}\ kpc^{-2}} = 3.8\times10^{-3}\left(\frac{I_{22\mu\rm m}}{\rm MJy\ sr^{-1}}\right)\cos i, \quad (2)$$

where i is the inclination angle of the galaxy. Here, we used the JWST 21 μm data to estimate the SFR surface density. For NGC 628, Kim et al. (2023) defined the duration of the embedded phase of star formation as the time during which CO and 21 μm emissions are found to be overlapping, i.e. $t_{\rm fb,21\mu m}$, and they obtained $t_{\rm fb,21\mu m} \approx 5.1$ Myr. We assume this timescale is also applicable to the five galaxies in this work. Following the same method as in Zhou et al. (2024a), the spatial separation between 21 μm and CO emission peaks is used to estimate the typical age ($t_{\rm age}$) of the embedded stellar populations in the cloud. Then, the total mass of the embedded stellar populations is

$$M_{*,\rm tot,emb} = \Sigma_{\rm SFR}\times A\times t_{\rm age}, \quad (3)$$

where A is the cloud's area. Here, we consider all 21 μm emission enclosed in the cloud. The same method is also applied to calculate the mass of bright 21 μm sources ($M_{\rm cluster,emb}$) identified in Sec.3.1, which represent the embedded clusters. As shown in Fig.7, both $M_{*,\rm tot,emb}$ and $M_{\rm cluster,max,emb}$ are systematically smaller than the corresponding $M_{*,\rm tot}$ and $M_{\rm cluster,max}$ observed in the PHANGS-HST survey. This suggests a short duration of the embedded phase. Kim et al. (2021) found that the embedded phase of star formation in six nearby galaxies lasts for 2−7 Myr and constitutes 17−47 percent of the cloud lifetime.

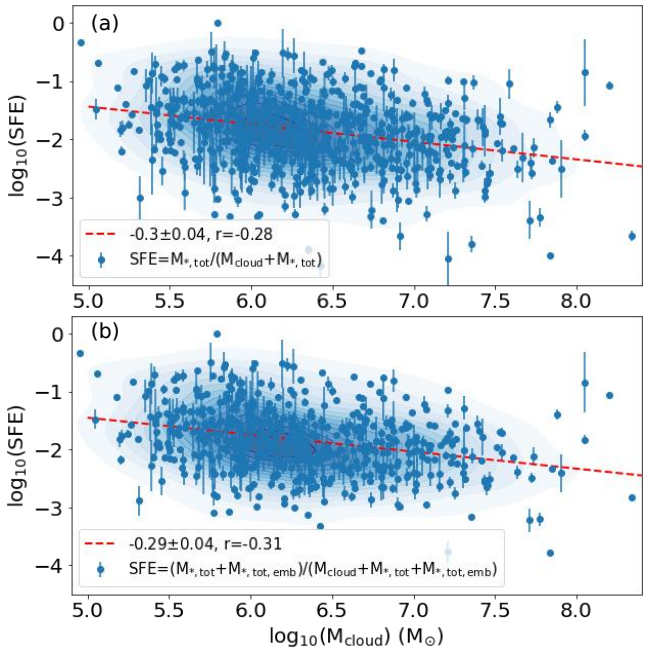


Figure 9. The correlation between the star formation efficiency (SFE) of the cloud and the mass of the cloud. (a) Embedded stellar populations were not considered in the calculation of SFE; (b) Embedded stellar populations were considered. r is the Pearson correlation coefficient.

Fig.8(a) and (b) show clear correlations between the SFR and the $\Sigma_{\rm SFR}$ of the cloud and the mass of the most massive cluster in the cloud, i.e. the $M_{\rm cluster,max}$-SFR$_{\rm cloud}$ relation and the $M_{\rm cluster,max}$-$\Sigma_{\rm SFR,cloud}$ relation.

3.4 Star formation efficiency

The star formation efficiency (SFE) of each cloud is defined as

$$\mathrm{SFE} \approx \frac{M_{*,\rm tot}}{M_{\rm cloud}+M_{*,\rm tot}+M_{\rm I}}, \tag{4}$$

where $M_{\rm I}$ is the mass of the ionized gas. First, we need to assess the significance of $M_{\rm I}$ in the total mass budget. Based on the HII region catalog in Santoro et al. (2022); Groves et al. (2023), if the distances between the center coordinates of the HII regions and a molecular cloud is less than the effective radius of the cloud, we distributed these HII regions to the cloud. Generally, HII regions associated with molecular clouds are compact. Thus, we used Equation 5 of Egorov et al. (2023) to calculate the mass of the ionized gas and this is given by

$$M_{\rm I} \approx 29.3 \frac{F(\mathrm{H}\alpha)_{\rm tot}}{10^{-16}\ \mathrm{erg\ s^{-1}\ cm^{-2}}} \left(\frac{D}{\mathrm{Mpc}}\right)^2 \left(\frac{n_{\rm e}}{\mathrm{cm^{-3}}}\right)^{-1} M_\odot, \tag{5}$$

where $F(\mathrm{H}\alpha)_{\rm tot}$ is the total H_α flux in a HII region. We also use Equation 6 of Egorov et al. (2023) to calculate the volume-averaged electron density,

$$n_{\rm e} \approx 1.36\times10^{-16} \left(\frac{L(\mathrm{H}\alpha)}{\mathrm{erg\ s^{-1}}}\right)^{0.5} \left(\frac{R_{\rm eff}}{\mathrm{pc}}\right)^{-1.5} \mathrm{cm^{-3}}, \tag{6}$$

where $L(\mathrm{H}\alpha) = 4\pi D^2 F(\mathrm{H}\alpha)_{\rm tot}$.

Due to the resolution limitations, the radius of the HII region may be overestimated. Additionally, electrons might not be fully distributed throughout the HII region. These factors can lead to an underestimation of the electron density (Egorov et al. 2023). For our samples, the typical value of the calculated electron density is only $\approx$1 cm^{-3}. Then we have the typical mass of ionized gas in HII regions being $\approx 10^{4.5} M_\odot$, similar to the value presented in Figure 5 of Egorov et al. (2023) calculated by the minimum electron density. Using the maximum electron density, the typical mass of the ionized gas in Figure 5 of Egorov et al. (2023) is $\approx 10^{2.5} M_\odot$, which means that the electron density has increased by 2 orders of magnitude overall, i.e. to $\approx$100 cm^{-3}. Therefore, when a reasonable estimate of the electron density is available, $M_{\rm I}$ can be neglected relative to $M_{\rm cloud}$. A typical value of $M_{\rm cloud}$ is $\approx 10^{6.3} M_\odot$. Thus, we have

$$\mathrm{SFE} \approx \frac{M_{*,\rm tot}}{M_{\rm cloud}+M_{*,\rm tot}}. \tag{7}$$

Using Eq.7, we estimate the SFE for the clouds in five galaxies. The median value of the SFE in Fig.9 is $\approx$1.4%, which is comparable with the values listed in Kim et al. (2022); Chevance et al. (2023).

4 DISCUSSION

4.1 The ICIMF theory

Table.3 summarizes all the fitted correlations between the physical parameters of molecular clouds and their internal stellar populations. In this section, we will verify that the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation can be understood theoretically on the basis of the integrated cloud-wide IMF (ICIMF) theory. We start by assuming all the stars in a cloud form in a population of embedded clusters with the mass function

$$\xi_{\rm ecl}(M_{\rm ecl}) = k_{\rm ecl} M_{\rm ecl}^{-\beta} (M_{\rm ecl,min} \leqslant M_{\rm ecl} \leqslant M_{\rm ecl,max}), \tag{8}$$

where $k_{\rm ecl}$ is a normalization constant and $M_{\rm ecl}$ is the stellar mass of the embedded cluster. The mass of the smallest stellar group observed in Kuhn et al. (2015) is $\approx 5M_\odot$. Thus, we take $M_{\rm ecl,min} = 5M_\odot$, which is also consistent with the value adopted in Yan et al. (2017). The embedded cluster population mass conservation gives

$$M_{*,\rm tot} = \int_{M_{\rm ecl,min}}^{M_{\rm ecl,max}} M_{\rm ecl}\, \xi_{\rm ecl}(M_{\rm ecl})\, \mathrm{d}M_{\rm ecl}, \tag{9}$$

where $M_{*,\rm tot}$ is the total stellar mass formed in a cloud. $M_{*,\rm tot}$ and the cloud mass are linked by the star formation efficiency, i.e.

$$M_{*,\rm tot} \approx \mathrm{SFE} * M_{\rm cloud}. \tag{10}$$

To determine the stellar masses within a stellar system based on a given IMF, one needs to apply a specific sampling method. Here, we adopt the optimal sampling. The details of this method and its validity are available in Kroupa et al. (2013); Schulz et al. (2015); Kroupa (2015); Yan et al. (2017, 2023). The optimal sampling normalization condition is

$$1 = \int_{M_{\rm ecl,max}}^{M_{\rm ecl,up}} \xi_{\rm ecl}(M_{\rm ecl})\, \mathrm{d}M_{\rm ecl}. \tag{11}$$

We take the upper integration limit (the hypothetical physical upper bound of embedded cluster mass) as $M_{\rm ecl,up} \approx 10^9$ M$_\odot$, following Yan et al. (2017). This normalization form is based on the assumption that there exists exactly one object within the mass range from $M_{\rm ecl,max}$ to $M_{\rm ecl,up}$ (Schulz et al. 2015). By solving Eqs.(8)-(11), we can build the relations between $M_{\rm ecl,max}$ ($\approx M_{\rm cluster,max}$), $M_{*,\rm tot}$ and $M_{\rm cloud}$. Setting $x \equiv M_{\rm ecl,max}$, $y \equiv M_{*,\rm tot}$, $z \equiv M_{\rm ecl,min}$ and

Table 3. All the correlations are fitted with simple linear functions in the form of $y = k * x + b$. r is the Pearson correlation coefficient.

y	x	k	Δk	b	Δb	r
$\log_{10}(M_{\rm assoc,tot}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.546	0.047	0.983	0.304	0.464
$\log_{10}(M_{\rm assoc,max}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.510	0.047	1.177	0.302	0.442
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.548	0.054	0.864	0.350	0.469
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(N_{\rm cloud}/{\rm cm}^{-2})$	0.695	0.075	-10.573	1.632	0.435
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(R_{\rm eff}/{\rm pc})$	1.009	0.201	2.329	0.410	0.255
$\log_{10}(M_{*,\rm tot}/M_\odot)$	$\log_{10}(M_{\rm cloud}/M_\odot)$	0.681	0.042	0.194	0.269	0.534
$\log_{10}(M_{*,\rm tot}/M_\odot)$	$\log_{10}(M_{*,\rm tot,emb}/M_\odot)$	0.612	0.041	2.558	0.131	0.544
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}(M_{\rm cluster,max,emb}/M_\odot)$	0.397	0.074	3.141	0.230	0.362
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}[{\rm SFR}_{\rm cloud}/(M_\odot{\rm yr}^{-1})]$	0.446	0.054	5.796	0.183	0.442
$\log_{10}(M_{\rm cluster,max}/M_\odot)$	$\log_{10}[\Sigma_{\rm SFR,cloud}/(M_\odot{\rm yr}^{-1}{\rm kpc}^{-2})]$	0.459	0.065	5.153	0.124	0.388
$\log_{10}({\rm SFE})$	$\log_{10}(M_{\rm cloud}/M_\odot)$	-0.293	0.039	0.013	0.250	-0.306

$C \equiv M_{\rm ecl,up}$, the solution of the equations has two cases: When $\beta \neq 2$,

$$y = \frac{z^{2-\beta} - x^{2-\beta}}{\frac{\beta-2}{1-\beta}\left(C^{1-\beta} - x^{1-\beta}\right)} \approx \frac{(\beta-1)(x^{2-\beta} - z^{2-\beta})}{(2-\beta)x^{1-\beta}}. \quad (12)$$

When $\beta = 2$,

$$y = \frac{\ln(x) - \ln(z)}{\frac{1}{1-\beta}\left(C^{1-\beta} - x^{1-\beta}\right)} \approx x \ln\left(\frac{x}{z}\right). \quad (13)$$

In Eq.12 and Eq.13, when $\beta > 1$, as long as $M_{\rm ecl,up}$ is sufficiently larger than $M_{\rm ecl,max}$, its specific value is not important.

For the slope β of the cluster mass function in Eq.8, generally, the value is ≈2 (Lada & Lada 2003; Krumholz et al. 2019; Mok et al. 2020; Wainer et al. 2022). For massive and compact clusters, such as Super Star Clusters (SSCs), the slope can be ≈1.8 (Cuevas-Otahola et al. 2023). Using the most massive clusters associated with molecular clouds presented in Fig.5(a), we fitted the cluster mass function in Fig.10(a). It indeed provides a slightly flatter slope, with $\beta \approx 1.9$. For all clusters associated with molecular clouds, the slope is ≈2. However, as shown in Fig.10, the most massive clusters and all clusters have comparable mass distributions, rather than the most massive clusters having a systematically larger mass. Since we will use the model to fit the $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation, taking the slope fitted in Fig.10(a) is more appropriate. As shown in Fig.11, when fixing SFE=1.4%, a smaller β can provide a better fit. To some extent, a smaller β and a larger SFE are equivalent. For low-mass molecular clouds, larger SFE will yield a better fit. The fact that low-mass molecular clouds have higher SFEs can be observed in Fig.9. When the age of the cluster is less than 10 Myr, the model also provides a better fit due to the stronger connection between young and embedded star clusters.

On the galaxy scale, Weidner et al. (2013) suggests the embedded cluster mass function to be flatter (top heavy) for galaxies with a high SFR, with a dependence that is given by

$$\beta = -0.106 \log_{10} {\rm SFR} + 2, \quad (14)$$

where the SFR is in unit of $[{\rm M}_\odot/{\rm yr}]$. In Weidner et al. (2004); Weidner & Kroupa (2005), the observed most massive embedded star cluster in a galaxy is determined through the total galaxy-wide SFR

$$\frac{M_{\rm ecl,max}}{M_\odot} \propto \left(\frac{\rm SFR}{{\rm M}_\odot/{\rm yr}}\right)^{0.75}. \quad (15)$$

Fig.8(a) shows a clear correlation between $M_{\rm cluster,max}$ and ${\rm SFR}_{\rm cloud}$, although the slope is shallower, ≈0.5. Given that the estimation of the SFR in Sec.3.3 is different from that in Weidner et al. (2013), the comparison is not straightforward. But similar to Eq.14, a β-SFR relation may also exist on the cloud scale. Surveying adequately complete cluster populations in molecular clouds of nearby galaxies to fit the cloud-scale β remains challenging, but it is feasible for molecular clouds in the Milky Way. In order to better fit the observational data in Fig.11, we need to understand how β varies with the properties of the molecular clouds.

Both SFE and β vary with the physical properties of the molecular clouds. Therefore, using the same parameter settings for all molecular clouds can only result in a rough fit. Moreover, in this work, we use an unified CO−to−H_2 conversion factor (i.e. $\alpha_{\rm CO}^{2-1} \approx 6.7\ {\rm M}_\odot{\rm pc}^{-2}({\rm K\ km\ s}^{-1})^{-1}$) for all five galaxies. But $\alpha_{\rm CO}^{2-1}$ may vary with the galaxy, and even different locations within the same galaxy may have different $\alpha_{\rm CO}^{2-1}$. Therefore, it is important to correct $\alpha_{\rm CO}^{2-1}$ for each molecular cloud to accurately estimate the mass of the cloud. This will be left for future work.

4.2 Cluster complex rather than single cluster

In Sec.3.2.2, the typical number of clusters associated with the clouds is only ≈2. For open clusters in the Milky Way, after comparing the separation of open clusters and the typical size of molecular clouds in Zhou et al. (2024b), we found that most molecular clouds may only form one open cluster. The pc-scale clumps within molecular clouds in the Milky Way are the basic units of star cluster formation. One molecular cloud can produce multiple embedded clusters in the Milky Way, which can coalesce into more massive clusters. The N-body simulations and the Gaia DR3 observations suggest that the post-gas expulsion coalescence of embedded clusters may be a dominant mechanism in the formation of open clusters in the Milky Way (Zhou & Kroupa 2024; Zhou et al. 2025). As shown in Zhou et al. (2024d), the typical separation of embedded clusters in a cloud can be ≈1 pc. Due to the limited resolution of the observation, the observed clusters in nearby galaxies may be cluster complexes rather than single clusters. To further explore this possibility, we need to refer to molecular clouds within the Milky Way, as only molecular clouds within the Milky Way can have their internal structures resolved. We matched the ATLASGAL clump complex catalog of Urquhart et al. (2022) and the molecular cloud catalog of Miville-Deschênes et al. (2017) by minimizing the differences in their central coordinates and systemic velocities. In Fig.12(a), there is a strong correlation between the mass of the most massive clump ($M_{\rm clump,max}$) in a cloud and the mass of the cloud. In Zhou et al. (2024c), we derived a relation between the clump mass and the mass

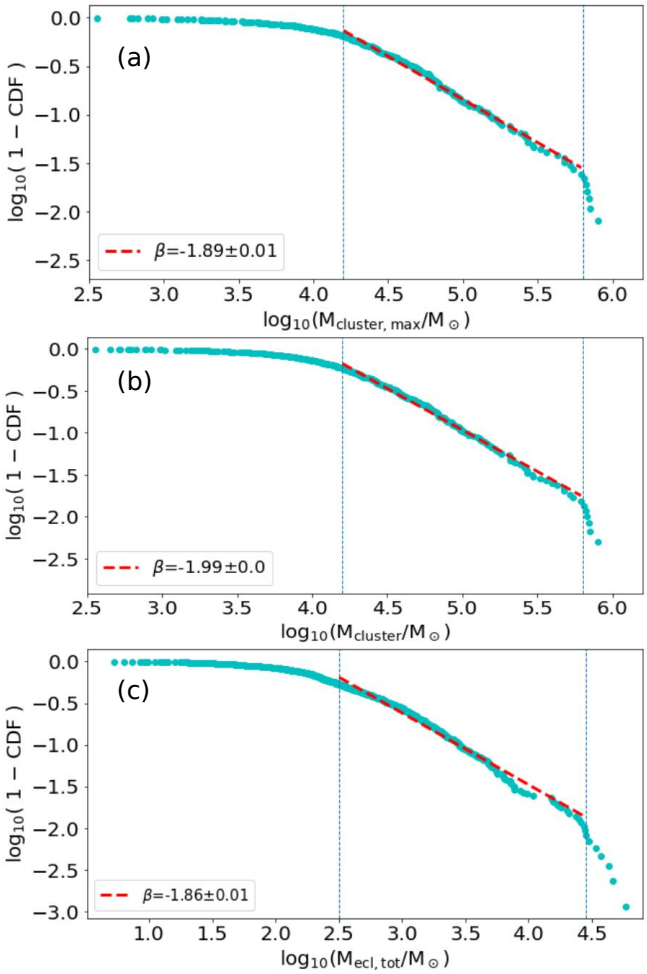


Figure 10. The complementary cumulative distribution function (CCDF) plot and the linear fitting of the cluster mass function. The vertical dashed lines mark the mass range of the linear fitting. (a) The mass function of the most massive clusters associated with molecular clouds in Fig.5(a); (b) The mass function of all clusters associated with molecular clouds; (c) The mass function of $M_{\rm ecl,tot}$, which represents the total mass of all embedded clusters in a cloud in the Milky Way.

of the embedded cluster formed in the clump, i.e.

$$\log_{10}(M_{\rm clump}/M_\odot) = (1.02\pm0.02)\times\log_{10}(M_{\rm ecl}/M_\odot)+(0.52\pm0.05). \quad (16)$$

Naturally, one would think that the most massive embedded cluster ($M_{\rm ecl,max}$) formed in the most massive clump will be the most massive exposed cluster in the cloud presented in Sec.3.2.2. However, when we directly compare the $M_{\rm ecl,max}-M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max}-M_{\rm cloud}$ relation in nearby galaxies in Fig.12(b), the former is systematically lower than the latter. Given that the coalescence of embedded clusters could be a key mechanism in the formation of open clusters in the Milky Way, we need to add the mass of neighboring embedded clusters in a cloud to approximate the mass of the final coalesced product ($M_{\rm ecl,tot}$). As a simplification, we sum the masses of all embedded clusters within a cloud. In Fig.12(c), the $M_{\rm ecl,tot}-M_{\rm cloud}$ and $M_{\rm cluster,max}-M_{\rm cloud}$ relations can link together. The mass function of $M_{\rm ecl,tot}$ is also similar to that of $M_{\rm cluster,max}$, as shown in Fig.10(c). In actual observations, as long as the objects are enclosed within the resolution limit, they are thought to be a single cluster. For $M_{\rm cluster,max}$ in Sec.3.2.2, they could be the products of coalescence or currently in the process of coalescing. The substructures within them need to be inferred based on observations in the Milky Way. We should be cautious with the very massive and young star clusters in nearby galaxies, as they may not actually be real clusters. This fact should be taken into account in the study of stellar population synthesis as well as star cluster formation and evolution in nearby galaxies.

Although the overall star formation rates between galaxies will be different, according to the ICIMF theory or the clear correlations between $M_{\rm cloud}$ and $M_{\rm cluster,max}$ (or $M_{\rm ecl,max}$ and $M_{\rm clump,max}$), star formation in the galaxy is localized to the molecular cloud. The connection in Fig.12(d) means that star formation in the Milky Way and nearby galaxies follow similar patterns on the cloud scale. Currently, only molecular clouds within the Milky Way can be studied in detail. Combined with the ICIMF theory, we can infer the internal details of molecular clouds in nearby galaxies using the template from the Milky Way.

5 CONCLUSION

In this study, we first investigated the connections between the physical characteristics of molecular clouds and their stellar populations. We then utilized the ICIMF theory to interpret the observed patterns, with the goal of developing a unified framework for star formation at the cloud scale in the Milky Way and nearby galaxies. The main conclusions are as follows:

1. The molecular clouds in five galaxies (NGC 1433, NGC 1566, NGC 1559, NGC 3351 and NGC 3627) were identified based on the integrated intensity (Moment 0) maps of CO (2−1) emission from ALMA observations. The catalogs for compact clusters and multi-scale associations across these galaxies were matched with the identified clouds. We used the JWST 21 μm data to estimate the star formation rate (SFR) surface density of the clouds and calculate the masses of the embedded stellar populations in the clouds, and found that the embedded stellar population mass is systematically smaller than that of the exposed stellar population. This suggests a short duration of the embedded phase. We found that the mass of the most massive cluster ($M_{\rm cluster,max}$) in a cloud is positively proportional to the mass ($M_{\rm cloud}$), the column density, the SFR and the SFR surface density of the cloud.

2. The mass of the ionized gas ($M_{\rm I}$) in each cloud can be neglected relative to $M_{\rm cloud}$. Based on the masses of the total stellar populations and their corresponding clouds, we obtained a typical value of the cloud-scale SFE, ≈1.4%. There is a trend that the SFE decreases as the cloud mass increases. More massive clouds lead to a larger stellar population and more massive star clusters, resulting in stronger feedback. This may suppress star formation in massive clouds, leading to a relatively lower SFE. This self-regulated star formation also supports the choice of the optimal sampling.

3. We constructed an integrated cloud-wide IMF (ICIMF) model to fit the observed $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation. When fixing SFE=1.4%, there is only one free parameter β (the slope of the cluster mass function) in the solutions. For low-mass molecular clouds, a higher SFE can yield a better fit. When the cluster's age is under 10 Myr, the model also provides a better fit due to the stronger connection between young and embedded star clusters. The ICIMF theory provides a quantitative framework for understanding the correlations between molecular clouds and their internal stellar populations.

4. For molecular clouds in the Milky Way, there is a strong correlation between the mass of the most massive clump ($M_{\rm clump,max}$)

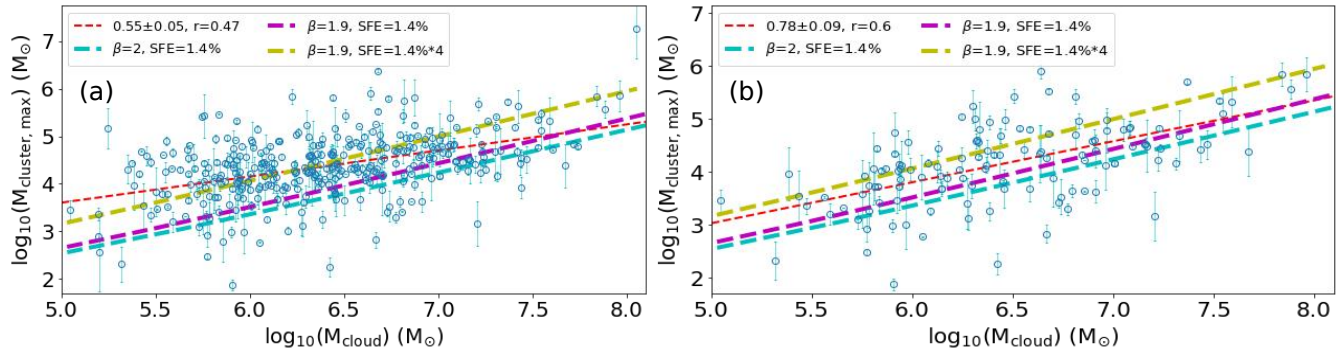


Figure 11. The $M_{\rm cluster,max}$-$M_{\rm cloud}$ relation fitted by the model based on the ICIMF theory presented in Sec.4.1. r is the Pearson correlation coefficient. β is the slope of the cluster mass function. SFE is the star formation efficiency. (a) All clusters selected in Sec.3.2.2; (b) Only clusters younger than 10 Myr.

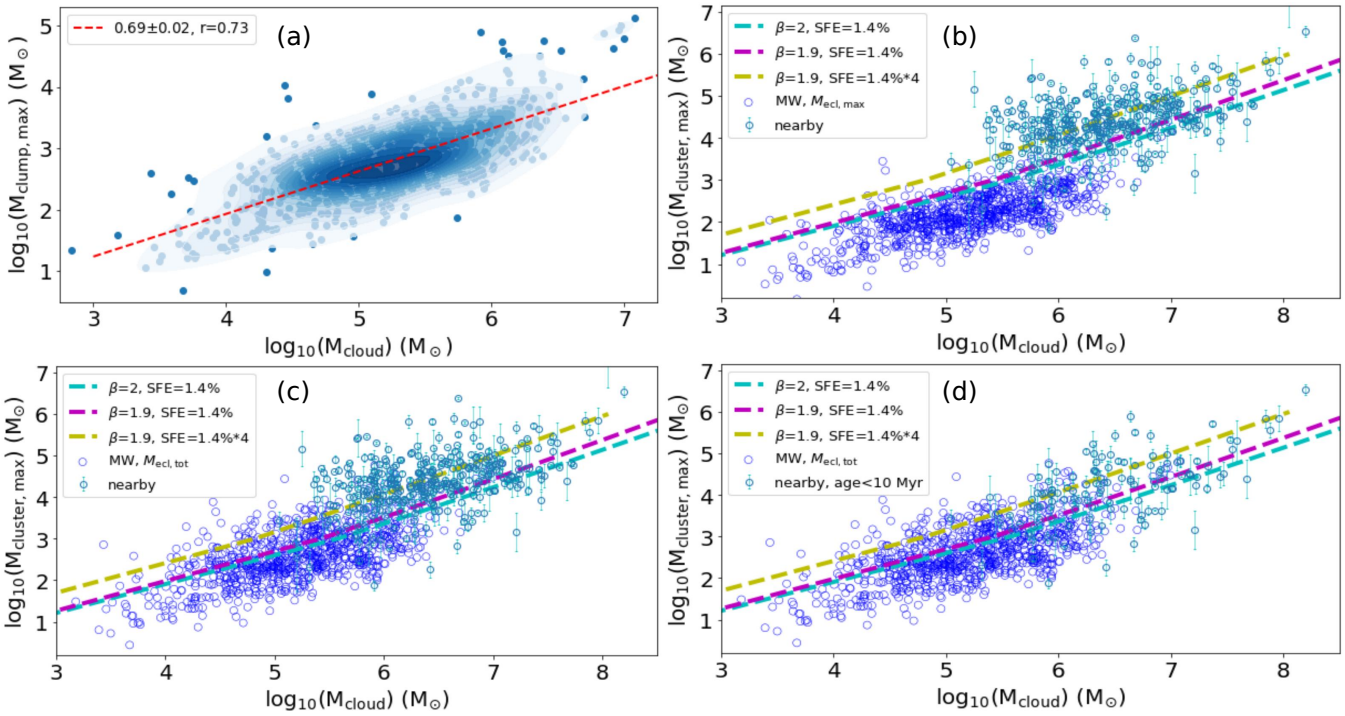


Figure 12. (a) Correlation between the mass of the most massive clump ($M_{\rm clump,max}$) in a cloud and the mass of the cloud in the Milky Way; (b) Comparison of the $M_{\rm ecl,max} - M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max} - M_{\rm cloud}$ relation in nearby galaxies. $M_{\rm ecl,max}$ is the mass of the most massive embedded cluster formed in the most massive clump. (c) Comparison of the $M_{\rm ecl,tot} - M_{\rm cloud}$ relation in the Milky Way and the $M_{\rm cluster,max} - M_{\rm cloud}$ relation in nearby galaxies. $M_{\rm ecl,tot}$ is the total mass of all embedded clusters in a cloud; (d) Same as panel (c), but $M_{\rm cluster,max}$ represents the clusters younger than 10 Myr. r is the Pearson correlation coefficient. β is the slope of the cluster mass function. SFE is the star formation efficiency.

in a cloud and the mass of the cloud. For the most massive embedded cluster ($M_{\rm ecl,max}$) formed in the most massive clump, when we directly compare $M_{\rm ecl,max} - M_{\rm cloud}$ and $M_{\rm cluster,max} - M_{\rm cloud}$ relations, the former is systematically lower than the latter. Given that the coalescence of embedded clusters could be a key mechanism in the formation of open clusters in the Milky Way, we need to add the mass of neighboring embedded clusters in a cloud to approximate the mass of the final coalesced product ($M_{\rm ecl,tot}$). Then the $M_{\rm ecl,tot} - M_{\rm cloud}$ and $M_{\rm cluster,max} - M_{\rm cloud}$ relations can link together. The mass function of $M_{\rm ecl,tot}$ is also similar to that of $M_{\rm cluster,max}$. This connection means that star formation in the Milky Way and nearby galaxies follow similar patterns on the cloud scale. Research on molecular clouds within the Milky Way can serve as a template for understanding molecular clouds and their internal stellar populations in nearby galaxies. The most massive young star clusters in the clouds of nearby galaxies should be cluster complexes rather than single clusters. The substructures within them need to be inferred based on observations in the Milky Way.

6 DATA AVAILABILITY

All the data used in this work are available on the PHANGS team website. [4].

[4] https://sites.google.com/view/phangs/home

ACKNOWLEDGEMENTS

We thank Z. Q. Yan for helpful comments. It is a pleasure to thank the PHANGS team, the data cubes and other data products shared by the team make this work can be carried out smoothly.

REFERENCES

Chevance, M., Krumholz, M. R., McLeod, A. F., et al. 2023, in Astronomical Society of the Pacific Conference Series, Vol. 534, Protostars and Planets VII, ed. S. Inutsuka, Y. Aikawa, T. Muto, K. Tomida, & M. Tamura, 1, doi: 10.48550/arXiv.2203.09570

Chevance, M., Kruijssen, J. M. D., Krumholz, M. R., et al. 2022, MNRAS, 509, 272, doi: 10.1093/mnras/stab2938

Cuevas-Otahola, B., Mayya, Y. D., Arriaga-Hernández, J., Puerari, I., & Bruzual, G. 2023, MNRAS, 525, 5262, doi: 10.1093/mnras/stad2643

Davis, T. A., Alatalo, K., Bureau, M., et al. 2013, MNRAS, 429, 534, doi: 10.1093/mnras/sts353

Dib, S. 2022, A&A, 666, A113, doi: 10.1051/0004-6361/202243793

Dib, S., Helou, G., Moore, T. J. T., Urquhart, J. S., & Dariush, A. 2012, ApJ, 758, 125, doi: 10.1088/0004-637X/758/2/125

Dinnbier, F., Kroupa, P., & Anderson, R. I. 2022, A&A, 660, A61, doi: 10.1051/0004-6361/202142082

Egorov, O. V., Kreckel, K., Glover, S. C. O., et al. 2023, A&A, 678, A153, doi: 10.1051/0004-6361/202346919

Elmegreen, D. M., Elmegreen, B. G., Adamo, A., et al. 2014, ApJ, 787, L15, doi: 10.1088/2041-8205/787/1/L15

Emsellem, E., Schinnerer, E., Santoro, F., et al. 2022, A&A, 659, A191, doi: 10.1051/0004-6361/202141727

Gouliermis, D. A., Thilker, D., Elmegreen, B. G., et al. 2015, MNRAS, 452, 3508, doi: 10.1093/mnras/stv1325

Gouliermis, D. A., Elmegreen, B. G., Elmegreen, D. M., et al. 2017, MNRAS, 468, 509, doi: 10.1093/mnras/stx445

Grasha, K., Calzetti, D., Adamo, A., et al. 2017, ApJ, 840, 113, doi: 10.3847/1538-4357/aa6f15

Groves, B., Kreckel, K., Santoro, F., et al. 2023, MNRAS, 520, 4902, doi: 10.1093/mnras/stad114

Haslbauer, M., Yan, Z., Jerabkova, T., et al. 2024, arXiv e-prints, arXiv:2405.05313, doi: 10.48550/arXiv.2405.05313

Jerabkova, T., Hasani Zonoozi, A., Kroupa, P., et al. 2018, A&A, 620, A39, doi: 10.1051/0004-6361/201833055

Kim, J., Chevance, M., Kruijssen, J. M. D., et al. 2021, MNRAS, 504, 487, doi: 10.1093/mnras/stab878

—. 2022, MNRAS, 516, 3006, doi: 10.1093/mnras/stac2339

—. 2023, ApJ, 944, L20, doi: 10.3847/2041-8213/aca90a

Kim, J.-G., Kim, W.-T., & Ostriker, E. C. 2016, ApJ, 819, 137, doi: 10.3847/0004-637X/819/2/137

Kroupa, P. 1995a, MNRAS, 277, 1491, doi: 10.1093/mnras/277.4.1491

—. 1995b, MNRAS, 277, 1507, doi: 10.1093/mnras/277.4.1507

Kroupa, P. 2005, in ESA Special Publication, Vol. 576, The Three-Dimensional Universe with Gaia, ed. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 629, doi: 10.48550/arXiv.astro-ph/0412069

—. 2015, Canadian Journal of Physics, 93, 169, doi: 10.1139/cjp-2014-0179

Kroupa, P., & Weidner, C. 2003, ApJ, 598, 1076, doi: 10.1086/379105

Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, Vol. 5, 115, doi: 10.1007/978-94-007-5612-0_4

Kruijssen, J. M. D., Schruba, A., Chevance, M., et al. 2019, Nature, 569, 519, doi: 10.1038/s41586-019-1194-3

Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227, doi: 10.1146/annurev-astro-091918-104430

Kuhn, M. A., Feigelson, E. D., Getman, K. V., et al. 2015, ApJ, 812, 131, doi: 10.1088/0004-637X/812/2/131

Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57, doi: 10.1146/annurev.astro.41.011802.094844

Larson, K. L., Lee, J. C., Thilker, D. A., et al. 2023, MNRAS, 523, 6061, doi: 10.1093/mnras/stad1600

Lee, J. C., Whitmore, B. C., Thilker, D. A., et al. 2022, ApJS, 258, 10, doi: 10.3847/1538-4365/ac1fe5

Lee, J. C., Sandstrom, K. M., Leroy, A. K., et al. 2023, ApJ, 944, L17, doi: 10.3847/2041-8213/acaaae

Leroy, A. K., Schinnerer, E., Hughes, A., et al. 2021a, ApJS, 257, 43, doi: 10.3847/1538-4365/ac17f3

Leroy, A. K., Hughes, A., Liu, D., et al. 2021b, ApJS, 255, 19, doi: 10.3847/1538-4365/abec80

Leroy, A. K., Rosolowsky, E., Usero, A., et al. 2022, ApJ, 927, 149, doi: 10.3847/1538-4357/ac3490

Levy, R. C., Bolatto, A. D., Mayya, D., et al. 2024, arXiv e-prints, arXiv:2408.04135, doi: 10.48550/arXiv.2408.04135

Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138, doi: 10.1088/0004-6256/149/4/138

Megeath, S. T., Gutermuth, R., Muzerolle, J., et al. 2016, AJ, 151, 5, doi: 10.3847/0004-6256/151/1/5

Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, ApJ, 834, 57, doi: 10.3847/1538-4357/834/1/57

Mok, A., Chandar, R., & Fall, S. M. 2020, ApJ, 893, 135, doi: 10.3847/1538-4357/ab7a14

Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78, doi: 10.1051/0004-6361/201935519

Rosolowsky, E., Hughes, A., Leroy, A. K., et al. 2021, MNRAS, 502, 1218, doi: 10.1093/mnras/stab085

Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, ApJ, 679, 1338, doi: 10.1086/587685

Santoro, F., Kreckel, K., Belfiore, F., et al. 2022, A&A, 658, A188, doi: 10.1051/0004-6361/202141907

Schulz, C., Pflamm-Altenburg, J., & Kroupa, P. 2015, A&A, 582, A93, doi: 10.1051/0004-6361/201425296

Thilker, D. A., Whitmore, B. C., Lee, J. C., et al. 2022, MNRAS, 509, 4094, doi: 10.1093/mnras/stab3183

Turner, J. A., Dale, D. A., Lee, J. C., et al. 2021, MNRAS, 502, 1366, doi: 10.1093/mnras/stab055

Turner, J. A., Dale, D. A., Lilly, J., et al. 2022, MNRAS, 516, 4612, doi: 10.1093/mnras/stac2559

Urquhart, J. S., Wells, M. R. A., Pillai, T., et al. 2022, MNRAS, 510, 3389, doi: 10.1093/mnras/stab3511

Wainer, T. M., Johnson, L. C., Seth, A. C., et al. 2022, ApJ, 928, 15, doi: 10.3847/1538-4357/ac51cf

Watkins, E. J., Kreckel, K., Groves, B., et al. 2023, A&A, 676, A67, doi: 10.1051/0004-6361/202346075

Weidner, C., & Kroupa, P. 2005, ApJ, 625, 754, doi: 10.1086/429867

Weidner, C., Kroupa, P., & Larsen, S. S. 2004, MNRAS, 350, 1503, doi: 10.1111/j.1365-2966.2004.07758.x

Weidner, C., Kroupa, P., Pflamm-Altenburg, J., & Vazdekis, A. 2013, MNRAS, 436, 3309, doi: 10.1093/mnras/stt1806

Whitmore, B. C., Lee, J. C., Chandar, R., et al. 2021, MNRAS, 506, 5294, doi: 10.1093/mnras/stab2087

Williams, T. G., Lee, J. C., Larson, K. L., et al. 2024, ApJS, 273, 13, doi: 10.3847/1538-4365/ad4be5

Yan, Z., Jerabkova, T., & Kroupa, P. 2017, A&A, 607, A126, doi: 10.1051/0004-6361/201730987

—. 2023, A&A, 670, A151, doi: 10.1051/0004-6361/202244919

Zhou, J. W., Dib, S., & Davis, T. A. 2024a, MNRAS, 534, 683, doi: 10.1093/mnras/stae2101

Zhou, J.-w., Dib, S., & Kroupa, P. 2024b, A&A, 691, A293, doi: 10.1051/0004-6361/202451728

Zhou, J. W., Dib, S., & Kroupa, P. 2024c, PASP, 136, 094302, doi: 10.1088/1538-3873/ad77f4

—. 2025, arXiv e-prints, arXiv:2501.08831, doi: 10.48550/arXiv.2501.08831