Star formation in galaxies and star clusters

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Thomas Maschberger

aus
Bad Neustadt

Bonn 2010

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

1. Gutachter: Prof. Dr. Pavel Kroupa
2. Gutachter: Prof. Dr. Cathie Clarke

Tag der Promotion: 17.02.2011

Erscheinungsjahr: 2011

Zusammenfassung

Diese Dissertation widmet sich der Entstehung von Sternen, von Galaxien-weiten Skalen (ca. 10 000 pc) bis zu Sub-Sternhaufen Skalen (ca. 0.1 pc). Ich entwickle und teste ein Verfahren zur Bestimmung der Sternentstehungsgeschichten von Galaxien aus deren Sternhaufenpopulation, insbesondere deren schwerste Sternhaufen. Es folgt die Anwendung dieser Methode auf die Große Magellansche Wolke, wo die Ergebnisse der neuen Methode bestätigt werden durch die Ergebnisse die man erhält vom Zählen einzelner Sterne. Sterne entstehen in Gruppen, in der Basisieinheit eines Sternhaufens, so daß sowohl die Massenfunktion von Sternen, die sich in einem Sternhaufen gebildet haben, als auch die Massenfunktion von Sternhaufen in einer Galaxie bestimmt werden müssen um galaxien-weite Sternentstehung quantitativ zu Beschreiben. Diese Massenfunktion folgen Potenzgesetzen, für welche ich statistische Methoden zur Parameterschätzung und Hypothesentests entwickle unter besonderer Berücksichtigung des oberen Massenbereichs. Die statistischen Methoden werden angewandt auf die Sternhaufen der Galaxie M51 um die Form der anfäglichen Haufenmassenfunktion und deren frühe Entwicklung zu untersuchen. Ich untersuche ebenfalls den oberen Massenbereich der stellaren Anfangsmassenfunktion, sowohl in Beobachtungen und in Theorie. Die Analyse einer numerischen Simulation der Entstehung eines Sternhaufens (Zeitentwicklung von Substruktur, Massensegregation und stellare Massenfunktion) beschließt diese Dissertation.

Summary

This thesis is devoted to star formation from galaxy-scales ($\approx$ 10 000 pc) to sub-star cluster-scales ($\approx$ 0.1 pc). I develop and test a new method to derive star formation histories of Galaxies from their star cluster content, in particular the most massive star clusters. This is followed by an application of this method to the Large Magellanic Cloud, where the results of this new method are confirmed with results obtained from counting individual stars. Stars form in a grouped way, in the basic unit of a star cluster, so that both the mass function of stars formed within a star cluster and the mass function of star clusters in a galaxy need to be known for a quantitative description of galaxy-wide star formation. These mass functions have a power-law functional form, for which I develop statistical tools to estimate the parameters and to perform goodness-of-fit tests, with a particular emphasis on the upper mass end. The statistical methods are applied to the star clusters in M51 to investigate the shape of the initial star cluster mass function and the early evolution of it. I also investigate the upper mass end of the stellar initial mass function, both in observations and theory. The analysis of a numerical simulation of star cluster formation (time-evolution of substructure, mass segregation, stellar mass function) concludes this thesis.

List of Publications

Maschberger, Th. & Kroupa, P. (2007).
A new method to derive star formation histories of galaxies from their star cluster distributions.
Monthly Notices of the Royal Astronomical Society, 379:34–42.2
(Chapter 3).

Maschberger, Th. & Clarke, C. J. (2008).
Maximum stellar mass versus cluster membership number revisited.
Monthly Notices of the Royal Astronomical Society, 391:711–717.
(Chapter 7).

Maschberger, Th. & Kroupa, P. (2009).
Estimators for the exponent and upper limit, and goodness-of-fit tests for (truncated) power-law distributions.
Monthly Notices of the Royal Astronomical Society, 395:931–942.
(Chapter 5).

Maschberger, Th., Clarke, C. J., Bonnell, I. A. & Kroupa, P. (2010).
Properties of hierarchically forming star clusters.
Monthly Notices of the Royal Astronomical Society, 404:1061–1080.
(Chapter 8).

Maschberger, Th. & Kroupa, P. (2011)
The star formation history of the Large Magellanic Cloud as seen by star clusters and stars.
Monthly Notices of the Royal Astronomical Society, 411:1495–1502
(Chapter 4).

Contents

Chapter 1

Prologue

Quippe mihi non multo minus admirandae videntur occasiones, quibus homines in cognationem rerum coelestium deveniunt; quam ipsa Natura rerum coelestium.
Johannes Kepler[⋆]

"For me the circumstances under which man comes to understand the celestial things appear not less admirable as the nature of these celestial things themselves." — Following this sentence of Kepler I would like to share in this Prologue my fascination for the historical development of Astronomy —inspired by the library in Bonn which holds among other originals Kepler's Harmonices Mundi, Huyghen's Systema Saturnium, the Astronomicum Caesareum by Apian or Hevelius' Machinae Coelestis— by giving an outline how our current picture of star formation emerged in time.

The sky at night is a spectacular display of sparkling stars, which were already in antiquity grouped into constellations representing mythological characters. Orion, the hunter, is a prominent feature of the northern sky, and Ptolemy catalogued in the Almagest 38 stars, one of which is classified as "nebulous" (Ptolemaius, 1816). Galilei (1610) chose Orion as one of the first targets for his telescope, and he found that the "nebulous" star in Orion's head was actually a group of 21 stars (The open cluster Collinder 69 around λ Orionis, see e.g. Murdin & Penston, 1977). Surprisingly, he did not find the famous "Orion Nebula" in the sword as it should be detectable with his telescope, his drawing leaves this region empty, which Humboldt commented with "Es war derselbe ohnedies wenig zur Annahme von Nebeln geneigt" — "He was anyhow hardly inclined to accept nebulae" (Humboldt, 1850, p. 336). Indeed, Galilei writes "Stellae ab Astronomis singulis in hanc usque diem NEBULOSAE appellatae, Stellularum mirum immodum consitarum greges sunt" (All stars that until this very day have been called "nebulous" by every single astronomers are swarms of small stars placed exceedingly closely together. Galilei, 1610, opposite p. 17). The nebula in the sword was presumably

[⋆]Kepler (1609, to Chapter XLV in Argumenta singulorum capitum)

described for the first time by Peiresc in 1610, but only noted in his unpublished observational notes which where brought to the public by Bigourdan (1916, recently the discovery of Peiresc has been questioned by Siebert, 2009).

Cysat (1619) mentions the nebula in a book on a comet, and the first drawing of the nebula appears in the unfortunately never widely recognised work by Hodierna (1654, brought to wider attention by Foderà Serio et al., 1985). Huyghens independently discovered the nebula again and described and drew in the Systema Saturnium (Huyghens, 1659), and this work is still often quoted as the first observation of the nebula. He describes that three stars near to each other, and four others shine as if through a cloud, which appears much brighter than all "the other very black rest of the sky". He saw this, "quidquid est" — "whatever it may be", several times at the same location. The Orion nebula appears in the perhaps first catalogue of nebulae by Halley (1716, containing six nebulae), where Huyghens is given credit for discovering the first nebula at all. A Nebula is understood by Halley as "extraordinary great Space in the Ether", in which a medium "shines with its own proper Lustre" (Halley, 1716).

The number of observed nebulae increased with time, Messier's first catalogue contains 45 nebulae (Messier, 1771, the Orion Nebula, M42, is individually described after the catalogue), and his last catalogue comprises 103 nebulae (Messier, 1784). Messier was more interested in Comets, and he only searched for nebulae "so that astronomers would not confuse the same nebulae with faint comets" (Messier, 1801, translation form Gingerich, 1987). The first astronomer to observe nebulae for their own sake is William Herschel, turning to Astronomy after a prolific career in music (he wrote 24 symphonies, for example), and becoming a maker of excellent telescopes. A few years after Messier, Herschel published his catalogues of in total 2500 nebulae (Herschel, 1786, 1789, 1802). The list of Herschel was extended by his son John to a catalogue of 5079 objects (the "General Catalogue", Herschel, 1864), and later to the"New General Catalogue" (Dreyer, 1888), whose NGC-numbers are still widely in use today.

William Herschel is the first to classify nebulae according to their appearance. His classes are: I Bright nebulae, II Faint nebulae, III Very faint nebulae, IV Planetary Nebulae (a term he coined), V Very large nebulae, VI Very compressed and rich clusters of stars, VII Pretty much compressed clusters of large or small stars, and VIII Coarsely scattered clusters of stars (Herschel, 1786). A few years later he discusses nebulosity and states about Orion that it is an object "where we are still inclined to remain in the once adopted idea, of stars exceedingly remote, and inconceivably crowded, as being the occasion of that remarkable appearance" (Herschel, 1791, p. 73). However, in the cited work he also argues that they are most likely consisting of a "shining fluid".

The gaseous nature of nebulae could only be observationally established after the technique of spectroscopy has been introduced to Astronomy. A pioneer on this field is W. Huggins, who obtained the first spectra of nebulae in 1864 (Huggins, 1864, amongst the first spectroscopically observed nebulae are a number of planetary nebulae, among them M57 and M27, two star clusters and the galaxy Andromeda with its companion). He was very puzzled when he saw the first nebular spectrum: "At first I suspected some derangement of the instrument had taken place; for no spectrum was seen, but only a short line of light perpendicular to the direction of dispersion. I then found that the light of this nebula [GC. 4373, 37 H IV.] , unlike any other ex-terrestrial light which had yet been subjected by me to prismatic analysis, was not composed of light of different refrangibilities, an therefore could not form a spectrum" (Huggins, 1864, p. 438). The Orion Nebula was the test-case for Huggins (1865) to investigate whether a nebula containing several stars would entirely consist of closely aggregated stars, as this was strongly supported by observations with Lord Rosse's "Leviathan"

telescope (Hoskin, 1990). From his observations he finds that gas is present in the nebula, which establishes the Orion Nebula as the first (gas-) embedded star cluster. Interestingly, Huggins also connects his finding to W. Herschel's ideas of star formation.

The (theoretical) connection of nebulae with star formation, in particular the formation of the solar system, the *nebular hypothesis*, goes back to Kant (1755) and, independently, Laplace (1796). William Herschel comes to a surprisingly modern sounding conclusion in his thoughts about the nature of nebulae: "If, therefore, this matter [the nebulous fluid] is self-luminous, it seems more fit to produce a star by its condensation than to depend on the star for its existence" (Herschel, 1791, p.85). The intent of these early theories is to explain the solar system, even though Jeans' work on the stability of gas clouds is motivated by the idea of forming planets from a meteorite swarm, which could be approximated as a gas (Jeans, 1902). It is perhaps typical for the working of the human mind that the understanding of the star formation process emerged in reverse to the time-arrow of the physical events. After realising that nuclear processes provide the energy of a star to shine and understanding the main sequence of stars (around 1930-1940), the pre-main-sequence of stellar evolution was first modelled by Henyey et al. (1955) and Hayashi (1961). The first object discovered to show a pre-main-sequence of stars is the Orion Nebula (Parenago, 1954 in Russian, Parenago & Sharov, 1961 in English), although the results in NGC 2264 became better known as they were not published in Russian (Walker, 1956). An even earlier stage of stellar evolution, protostars, were also first discovered in Orion (Becklin & Neugebauer, 1967; Kleinmann & Low, 1967).

Further ingredients in the star formation recipe were mentioned by McNally (1971): 1) collapse under gravity (Jeans, 1902), 2) random accretion (McCrea, 1960; von Weizsäcker, 1951) 3) gravitational clustering (Layzer, 1964), 4) the ideas of Ambartsumian (1960) combining star formation, formation of clouds and galaxy evolution. But McNally (1971) still states "Nevertheless the mechanism whereby stars form is still unknown" . Only in the recent years has it become possible to follow the hydrodynamical collapse and fragmentation of a gas cloud forming more than one star in detail (and three dimensions). Examples are the simulations of Bonnell et al. (2003), forming one star cluster like the Orion Nebula Cluster from an initial cloud of 1000 $M_\odot$, and Bonnell et al. (2008) which, starting with 10000 $M_\odot$ of gas, illustrates evolution of a larger star forming region that contains several clusters. A star forming region is perhaps the basic unit of star formation from a galactic point of view. Star clusters are regular characters appearing in star forming regions. For a better overview it is best to study them not in the Milky Way, but in external galaxies.

The first discovered extragalactic nebula was 30 Doradus in the Large Magellanic Cloud (La Caille, 1755), although at that time he assumed that the Magellanic Clouds are just separated parts of the Milky Way. By coincidence this is also a star forming region, about ten times larger than the Orion Nebula Cluster. Dunlop (1828) and John Herschel (1847) discovered many more nebulae and star clusters in the Large Magellanic Cloud, and by the early 20th century the Magellanic Clouds were recognised as galaxies, small Milky Ways of their own. Hodge (1988) in his catalogue of 255 new star clusters in the Large Magellanic Cloud gives a list of star cluster searches and comes to the number of 2053 catalogued star clusters in the Large Magellanic Cloud. He also estimates that the Large Magellanic Cloud may host about 4200 star clusters in total.

The connection between galaxy-wide star formation and the star cluster population is the starting point of this thesis, from which we go on to increasingly smaller scales to end with an investigation of the formation of an individual star cluster like Orion.

Chapter 2

Introduction

This thesis is devoted to star formation physics from galaxy-wide scales ($\approx$ 10 000 pc) down to sub-star-cluster scales ($\approx$ 0.1 pc). The galaxy-wide star formation and its connection with star clusters is the starting point of the research presented in this work (Ch. 3 & 4). From there we continue to discuss the evolution of star cluster populations (Ch. 6) and the stellar "population" in a star cluster (Ch. 7 & 8). The means for our investigation are mass distribution functions, for whose analysis we develop statistical techniques (Ch. 5). Finally, we finish with the detailed properties of an individual star forming region (Ch. 8). The following Sections elaborate the connections between the covered subjects.

2.1 Galaxy-wide star formation

The first part of this thesis (Chapters 3 and 4) lays out and applies a method to derive star formation histories of galaxies from their star cluster content. Detailed star formation histories trace the stellar mass build-up of a galaxy, and are an important link of observational and speculative cosmology. The choice of star clusters to derive a star formation history has an important practical implication, as they can be observed up to distances where individual stars are not resolvable any more, such that potentially a larger and more varied sample of galaxies can be studied than has been possible until now.

Massive star clusters have life-times comparable to that of a galaxy, so that in the star cluster population the star formation history of a galaxy is quasi frozen-in. In Chapter 3 we develop and test a method based on the most massive star cluster in the time-interval chosen to discretise the star formation history. The basis for the conversion of star cluster masses into star formation rates is the observed relation between the brightest cluster in a galaxy (which is very likely the most massive) and its present star formation rate as determined from far-infrared fluxes, which serves as a proxy for the total star formation rate (Billett et al., 2002; Larsen, 2002). With the probabilistic nature of the star cluster mass function the most massive star cluster also follows a distribution function, which has to be taken into account. Thus, rather than having to just invert a function, a probability distribution has to be inverted in order to derive star formation rates from the cluster masses. Chapter 3 is devoted to the procedure of this method, accounting for the statistical aspects as well as for the dynamical evolution of star clusters and observational restrictions. Using Monte-Carlo simulations of star cluster populations, the performance and limitations of the method are shown, finding that the star cluster ansatz has the potential to derive star formation histories.

A "real world" example is given in Chapter 4, where we compare the star cluster results with the results of the standard technique based on individual stars and their colour-magnitude diagram. The Large Magellanic Cloud serves us as a "guinea pig", it has a large star cluster population and is sufficiently nearby to allow one to resolve stars. Both the star formation history derived from either the most-massive star clusters or the colour-magnitude diagram show a similar structure and absolute star formation rate; however, the agreement starts to disappear for ages larger than one Gyr. The gap in the star cluster population of the Large Magellanic Cloud between roughly 1 and 10 Gyr is well known, but it is puzzling that there was a time of enhanced star formation without producing the appropriate number of star clusters. Perhaps the incomplete spatial coverage may account for this and the missing clusters are located in the regions that are not observed. Also, the star cluster sample has not very well defined selection criteria, there are for example numerous young star clusters missing, among them the prominent cluster R136 in 30 Doradus.

Due to the proximity of the Large Magellanic Cloud star clusters can be detected down to a few hundred $M_{\odot}$at young ages. Therefore it is possible to straightforwardly calculate the amount of mass in star clusters and compare this to the mass of all stars in the selected age range. The star formation rates derived in this way are lower by a factor ≈ 10. This discrepancy can be explained by two arguments; firstly, not all stars need to form in compact star clusters; and secondly, star clusters undergo a phase of violent evolution shortly after their formation that leads to their dissolution. The second proposed solution should be evident from the star formation histories, the discrepancy should increase within the first, say, 30 Myr, as star clusters dissolve caused by gas expulsion. Unfortunately, we cannot do this as the star cluster sample in the Large Magellanic Cloud is just missing the youngest clusters. A further discussion of the early star cluster evolution is given in Chapter 6.

The observational basis for deriving the star formation histories is, as already said above, the relation between the brightest cluster and the galaxy-wide star formation rate. To understand this relation several ideas have been proposed which could be distinguished observationally from the properties of the star cluster mass function. The extent to which the star cluster mass function (a power law) is populated depends on the star formation rate in the galaxy. Under the hypotheses that the brightest cluster is the most massive one and that all stars form in star clusters, and with choosing a time scale, this correlation is equivalent to a correlation of the most massive cluster with the total mass of the cluster population. The total mass follows simply as a product of the star formation rate and the time scale, and, with a universal cluster mass function and thus a universal mean mass, the total mass is proportional to the total number of star clusters. A correlation of the most massive star cluster with the total number of star clusters is expected; this is usually coined "size-of-sample effect". The upper limit of the star cluster mass function is universal in this scenario.

The purely statistical explanation has been questioned by Weidner et al. (2004), suggesting that actually the upper limit of the star cluster mass function depends on the star formation rate of the galaxy. In this scenario the formation of a massive star cluster is prohibited at low star formation rates, even when one waits for a very long time. Weidner et al. (2004) need a rather steep exponent of the cluster mass function ($\beta = 2.4$) to reproduce the observed relation. As the star cluster mass function follows from the mass function of giant molecular clouds (which has a smaller exponent), it has been suggested that the star cluster mass function has analogously an exponential turn-down at the high mass end (a Schechter function, Gieles et al., 2006a), such that $\beta = 2$ for most of the cluster mass function. Bastian (2008) investigated the size-of-sample effect in the $M_{\rm max}$–$N_{\rm tot}$ relation for a Schechter function, and

find a high (universal) truncation mass, but also stated that the truncation mass might vary from galaxy to galaxy. Larsen (2009) pursues the question of a varying truncation mass by analysing several spiral galaxies, and also finds that a universal truncation is unlikely. The $M_{\rm max}$–$N_{\rm tot}$ relation is perhaps not the best way to answer the question, as a varying and universal upper limit lead to very similar appearances of this relation, so that a large number of galaxies is needed to arrive at a statistically significant conclusion. Therefore, we investigate more appropriate statistical methods in Chapter 5, and apply them to the star cluster mass function of M51 in Chapter 6, with an emphasis on the upper mass end.

2.2 Statistical methods

The star cluster mass function is, as seen above, one ingredient in an explanation of the brightest cluster–star formation rate relation. For the maximum exploitation of the available data it is necessary to have good statistical techniques. As the masses of stars and star clusters obey a power-law distribution function, we therefore present and test in Chapter 5 statistical techniques for power laws. Power-law distributions seem to be an ubiquitous feature in nature. The statistical analysis has two aspects, estimating the parameters and goodness-of-fit tests. For a power law the parameters needing estimation are the exponent and also the upper limit. By simply performing a linear regression to a histogram of constant logarithmic bin size a substantial bias in the estimated exponent can be introduced. This can be remedied by using variable-size bins with a constant number of data points in them (Maíz Apellániz & Úbeda, 2005), but binning does not directly lead to an estimate for the upper limit. An estimate for the upper limit can be obtained with the Maximum Likelihood method, although it is very biased (the estimate is only the largest data point). But we found a way to correct for this bias.

As a second step in the data analysis the assumption of a power law as the parent distribution function has to be verified using a goodness-of-fit test. A histogram is a first and simple way to do this, but by grouping the data much information is lost. Especially the upper mass end is often hidden in only one or two bins, and a histogram is no statistical test. An alternative graphical display of all available data is the stabilised probability–probability plot, discussed in detail in Chapter 5. The formal "measurement" for agreement of model and data is done with a goodness-of-fit test, the Kolmogorov-Smirnov test is one of the more widely known. The behaviour of the star cluster mass function at the upper mass end is an important question, on which a goodness-of-fit test should shed light. Unfortunately the hypotheses between which should be distinguished are very similar, a (universal) truncation at a high mass, a truncation at a lower mass (depending on the star formation rate), or an exponential turn-down. Not all goodness-of-fit tests possess the necessary power to distinguish between them, the Kolmogorov-Smirnov test, for example, is not a good choice. Therefore, Chapter 5 contains a comparative study of a number of goodness-of-fit tests for finding the best choice, and a method to improve standard tests, such as the Kolmogorov-Smirnov test.

Applications follow in Chapter 6 with the star cluster mass function in the galaxy M51 and in Chapter 8 where the predictions for the stellar mass function from hydrodynamical simulations are discussed.

2.3 Star cluster mass function

The star cluster mass function holds information about both the global, galaxy-wide and local, cluster-wide star formation. The variations of the shape of the (initial or embedded) star cluster mass function of different galaxies give evidence of different star forming environments. Equally, a universal star cluster mass function (or more specifically a universal exponent) shows that star cluster formation is a local process. The upper end of the star cluster mass function is perhaps where an environment dependence can be detected. As described above, from the $M_{\rm max}$–$n_{\rm tot}$ relation, the upper limit presumably varies with star formation rate. Unfortunately the cluster sample of the Large Magellanic Cloud cannot be used for a statistical analysis as it has selection issues. However, in Chapter 6 we analyse the star cluster population of M51. We investigate whether an infinite, a (sharply) truncated or an exponentially decaying power law fits the data. Also, we investigate how much the choice of isochrones and metallicities for the age and mass determination influences the results for the star cluster mass function.

Cluster-internal aspects of star formation can be obtained from the time-evolution of the star cluster mass function, which is expected due to the dynamical evolution of star clusters, also investigated in Chapter 6. Massive clusters evolve slower than less massive clusters, so that the mass function becomes flatter in time. The speed of flattening is set by the environment in the galaxy, the tidal field, encounters with clouds and spiral arms etc. In the first, say, thirty million years the evolution of star clusters is mainly determined by the transition from the embedded to the gas-free state. If this transition depends on the star cluster mass, then the star cluster mass function should also change. The changes could be gradual, so that the overall appearance of a power law is unchanged, only the exponent changes. Alternatively there is the scenario of Kroupa & Boily (2002): Low-n clusters have no massive stars and therefore gas expulsion occurs gradually and the star cluster survives. Similarly, in very massive clusters several supernovae are necessary to expel the gas, which thus also happens gradually. However, in the intermediate mass range the gas is removed violently with a single supernova, which likely destroys the cluster. The cluster mass function in this scenario would exhibit a feature in the mass range around a few $10^3 M_\odot$ and a broad maximum at 10^5 $M_\odot$.

2.4 Stellar initial mass function

The stellar initial mass function (IMF) is an important phenomenological distribution function with a wide range of applications. It has been found to be very similar in diverse environments, so that its universality is assumed (Kroupa, 2001, 2002). We focus on the "high"-mass regime ($m > 0.5$ $M_\odot$) where the distribution follows a power law with the Salpeter exponent of 2.35 (Salpeter, 1955) up to masses of ≈ 150 $M_\odot$ (the physical upper limit for stars Weidner & Kroupa, 2004; Figer, 2005; Oey & Clarke, 2005). Whereas the universality of the IMF exponent is generally agreed upon the universality of the upper limit has been questioned (Kroupa & Weidner, 2003; Weidner & Kroupa, 2006). A possible explanation could go as follows. The build-up of massive stars occurs via accretion starting from lower-mass seeds leading to a power law mass distribution, with an ever increasing upper mass limit. At some point the accretion is stopped by feedback, such as radiation, stellar winds or even a supernova, which freezes-in the mass function. The effectiveness of feedback depends on the amount of gas

and the number of massive stars. These two quantities, however, scale with each other, so that an upper limit of the IMF depending on the number of stars formed may be expected. The details of this are further discussed in Chapter 8.

Observationally a "richness"-dependent upper mass limit can be investigated in two ways. In Chapter 7 we follow the method of Weidner & Kroupa (2006), analysing the relation between maximum stellar mass and cluster membership number. We extend their sample of data with data from the literature and present a new way to statistically analyse the data. The advantage of this approach is its simplicity —only counting stars and one mass determination per cluster— but a major drawback is the difficulty in obtaining a homogeneous, complete sample of star clusters. The predictions for a universal upper limit and a cluster-dependent upper limit differ only slightly, so that a very large sample of clusters is necessary for a distinction (Selman & Melnick, 2008, state that 4000 clusters of 75 stars each would be needed). However, the additional data presented by Weidner et al. (2010) are not consistent with a universal upper limit (for clusters in the mass range 1000–10000 $M_\odot$). Even with invoking some selection bias it is hard to explain that 97% of the data lie below the median. A solution to this puzzle can probably only be found by an individual inspection of each of the clusters.

The most direct way to study the upper mass limit is to statistically analyse the upper stellar mass function in a star cluster. We follow this twice, in Chapter 5 we use the stars of R136 in 30 Doradus to test for a truncation at the physical upper mass limit. For a mass-dependency a number of star clusters need to be analysed, a project for which pilot studies have been commenced (see Chapter 9). In Chapter 8 we apply the methods of Chapter 5 on the outcome of hydrodynamical simulations of star cluster formation.

A further reason to investigate the upper limit of the star cluster mass function in Chapter 6 and the upper limit of the stellar mass function in Chapters 7 and 8 are the consequences of the *integrated galactic initial mass function* (IGIMF) concept (Kroupa & Weidner, 2003; Weidner & Kroupa, 2005), which have important implications for the evolution of galaxies with small star formation rates. A varying upper stellar limit implies that the mass distribution of all newly formed stars together (the IGIMF) differs from the initial mass function within a star forming region. This follows because there are many more small star clusters without massive stars than massive clusters with massive stars, so that in the sum the massive stars are underrepresented in the total sample, i.e. the integrated galactic initial mass function has a steeper massive end. The extent of this effect depends on the distribution of star cluster masses. With the upper limit of the star cluster mass function depending on the star formation rate, this effect is strongest for small star formation rates. Consequences of this effect are a non-linear dependence of the Hα luminosity (Pflamm-Altenburg et al., 2007a) and metal yields (Weidner & Kroupa, 2005, supernova rate) on the star formation rate in a galaxy.

2.5 Formation of a star cluster

The last research Chapter of this work (Chapter 8) is concerned with the small-scale properties of star formation within an individual star forming region. We have mentioned the analysis of the upper mass function already above. In this Chapter we analyse hydrodynamical simulations of a single star cluster and a larger star forming region. The general picture is that the initial cloud fragments into filaments which contain subclusters of newly-formed stars that later merge into larger entities.

We consider here several aspects of the initial stellar mass function (or, more precisely, the sink particle mass function as proxy for the stellar mass function). We can follow the time-evolution of the mass function, and we find that the exponent decreases in time. This flattening is caused by the continuous formation of new stars and accretion of matter onto them. Also, the exponents are smaller than the empirical Salpeter exponent, caused by mass segregation and, presumably, by the missing feedback, which has not been included for computational feasibility. Not only the exponent, but also the upper limit of the mass function increases with time and increasing number (and mass) of stars, and a universal upper limit of 150 $M_{\odot}$ is not compatible with the simulation results.

The larger simulation leads not only to the formation of a single cluster, but is designed to mimic the Orion region with several star clusters and an unbound region of gas with more dispersed star formation. The stellar mass function of this whole simulation is not compatible with a power law truncated at a mass similar to what would be expected from an individual cluster of that size. It rather shows a deficiency of massive stars compared to a truncated power law. The turn-down in the mass function is a signature of the IGIMF effect, and it is very interesting to see that it manifests itself in a model of a larger star forming region. The results of the simulations are a strong incentive for an observational project to look for similar behaviour.

A further investigated feature of the simulations is the development of mass segregation. Massive stars are found to be the stars that first form in a subcluster, further stars form around them. This and the fast dynamical evolution times lead to subclusters that are already mass segregated. During mergers of subclusters the mass segregation is disturbed for a short while, but quickly re-establishes itself. The positions of the massive stars are important for the evolution of a star cluster. For example, a system of only a few massive stars in the centre of a star cluster is dynamically very unstable, and three-body interactions can lead to ejections of massive stars from the star cluster. Similarly, if massive stars as main sources of feedback are in the central parts of a cluster, the time-scales of gas expulsion could be shorter than in the case of an extended population of massive stars. Future research will perhaps reveal that only certain configurations of star clusters survive the gas expulsion phase. The ratio of surviving vs. destroyed star clusters after gas expulsion is the ratio of clustered vs. field star formation, discussed Chapter 4.

Chapter 3

A new method to derive star formation histories of galaxies from their star cluster distributions

Th. Maschberger & P. Kroupa
MNRAS 379:34–42 (2007)

Star formation happens in a clustered way which is why the star cluster population of a particular galaxy is closely related to the star formation history of this galaxy. From the probabilistic nature of a mass function follows that the mass of the most-massive cluster of a complete population, $M_{\max}$, has a distribution with the total mass of the population as a parameter. The total mass of the population is connected to the star formation rate (SFR) by the length of a formation epoch.
Since due to evolutionary effects only massive star clusters are observable up to high ages it is convenient to use this $M_{\max}(\mathrm{SFR})$ relation for the reconstruction of a star formation history. The age-distribution of the most-massive clusters can therefore be used to constrain the star formation history of a galaxy. The method, including an assessment of the inherent uncertainties, is introduced with this contribution, while following papers will apply this method to a number of galaxies.

3.1 Introduction

During the last few years it has been recognised that most and probably all stars form in embedded clusters (Lada & Lada, 2003). The observational work, notably by Larsen (2002), has established that star clusters ranging from the oldest globular clusters to the youngest low-mass objects have to be regarded as a continuous distribution by mass — globular clusters are not fundamentally different from open clusters but merely the upper mass end of the distribution. This has been shown explicitly by Kroupa & Boily (2002) on the example of the Milky Way population II spheroid. Today star clusters must be viewed as the *"fundamen-*

tal building blocks" of galaxies because they also determine the morphological appearance of whole galaxies through the physics of their formation (Kroupa, 2005).

Star formation is therefore closely connected to the star cluster distribution in a galaxy and thus it is expected that the star formation history of a galaxy leaves its imprint on the star cluster distribution. Since star clusters can live for a long time the star cluster distribution of a galaxy can be compared to a diary: Events like interactions of galaxies lead to an enhanced star formation activity. From this results a larger number of clusters being formed during the interaction time. This qualitative statement is well known and in this work we present a quantitative method to derive the star formation history (SFH) of a galaxy directly from its star cluster content.

Until now detailed SFHs can be determined only for galaxies which are at a distance allowing individual stars to be resolved, i.e. within the local group up to ≈ 1 Mpc away. The colour-magnitude diagram (CMD) which is thusly obtained then allows the construction of a SFH using theoretical isochrones.

In distant galaxies, however, star clusters appear as compact sources whereas the individual stars give a homogeneous distribution over the area. Modern instruments like the Hubble space telescope make it possible to obtain a cluster age- and mass-distribution for galaxies beyond the distance where individual stars can be resolved, e.g. for M51 (Bastian et al., 2005; cluster formation rate: Gieles et al., 2005) and M101 (Bianchi et al., 2005). These galaxies lie at a distance of about 7 Mpc (M51: Takáts & Vinkó, 2006; M101: Kelson et al., 1996) which demonstrates the potential of our new method.

A first approach to derive the star formation history from a cluster age- and mass-distribution could simply be to use all observed clusters in given time intervals. But this would lead to wrong results, clusters evolve and thus there are fewer clusters at higher ages. However, massive clusters evolve slowly, clusters with masses $\gtrapprox 10^5$ $M_\odot$ have lifetimes comparable to a Hubble time (Baumgardt & Makino, 2003). Weidner et al. (2004) found a relation which establishes the connection between massive clusters and the galaxy-wide star formation rate: During a *formation epoch* a *complete population* of star clusters is formed, for which the mass of the most massive cluster, $M_{\rm max}$, depends on the star formation rate (the $M_{\rm max}(\mathrm{SFR})$ relation). Consequently, the star formation history of a galaxy can be regarded as a sequence of such formation epochs, in which the most massive cluster carries the information about the star formation rate.

Since Weidner et al. (2004) used a deterministic law for the $M_{\rm max}(\mathrm{SFR})$ relation, we briefly re-analyse their argumentation, allowing a distribution of the most massive cluster rather than a fixed value for a given SFR. Then we present the new method and conclude with testing it for some typical cases.

3.2 Complete Populations and the $M_{\rm max}(\mathrm{SFR})$-Relation

3.2.1 Complete Populations

The distribution of star clusters in a star cluster population is characterised by the shape of the distribution function (shown to be a power law, e.g. Weidner & Kroupa (2006) and references therein), the mass of the population, $M_{\rm tot}$, and the mass limits, $M_{\rm L}$ and $M_{\rm U}$. The cluster mass function can be written as

$$\xi_{\rm cl}(M) = \frac{dN}{dM} \quad = \quad kM^{-\beta}, \tag{3.1}$$

where the normalisation constant k is determined by $M_{\rm tot}$, $M_{\rm L}$ and $M_{\rm U}$.

In this context *complete population* denotes a statistically meaningful representation of the embedded cluster mass function. The underlying distribution function is defined on a mass interval appropriate to the mass of the population. This is a very general concept which is applicable to other cases where objects obey a distribution function, e.g. stars in a star cluster.

The lower mass limit, $M_{\rm L}$, is given by the physical minimum mass of a star cluster and is independent of the mass of the population. Until now, the physical minimal mass of a star cluster is not well known, here it is assumed to be 5 $M_\odot$ corresponding to groups of about a dozen stars such as those forming in Taurus-Auriga (Briceño et al., 2002).

For the upper mass limit, $M_{\rm U}$, two cases have to be distinguished: If a complete population has a mass which is larger than the limiting physical maximum cluster mass, then the distribution function is defined up to the physical maximum mass, a hitherto not well understood quantity. Star clusters with masses larger than about 10^6 $M_\odot$ show complex stellar populations and are probably distinct from the "normal" star cluster content of a galaxy (Weidner et al., 2004). Such arguments based on the structural properties of clusters would imply the same physical maximum mass in all galaxies. Gieles et al. (2006a) derived maximal cluster masses from the cluster luminosity function, obtaining masses between 0.5–$2.5 \times 10^6 M_\odot$ for NGC 6946, M51 and the "Antennae" (NGC 4038/39). Thus, while arguments exist for a limiting maximum star-cluster mass below which stellar populations are simple (mono-metallic and -age), Mieske et al. (2002) and Martini & Ho (2004) show that ultra-compact dwarf galaxies ($M \gtrapprox 10^6$ $M_\odot$) may be an extension of the "cluster" formation process to large masses. Consequently, we do not limit the "cluster" masses but allow these to formally reach $M_{\rm U} = 10^9$ $M_\odot$ for sufficiently high SFRs.

The case that $M_{\rm tot}$ is smaller than the physical maximal mass implies that no star clusters more massive than $M_{\rm tot}$ can exist in this population. Therefore $M_{\rm tot}$ is the upper mass limit of the cluster mass function. This includes the case that a population can consist only of one cluster with the mass $M_{\rm tot}$, but this is very improbable.

Thus by using the total mass as the normalisation criterion,

$$M_{\rm tot} = \int_{M_{\rm L}}^{M_{\rm U}} M \xi_{\rm cl}(M) dM, \tag{3.2}$$

the normalisation constant becomes

$$k = \frac{M_{\rm tot}(2-\beta)}{M_{\rm U}^{2-\beta} - M_{\rm L}^{2-\beta}}. \tag{3.3}$$

The total number of clusters in a population, $N_{\rm tot}$, follows from $N_{\rm tot}(M_{\rm tot}) = \int_{M_{\rm L}}^{M_{\rm U}} \xi_{\rm cl}(M) dM$.

Weidner et al. (2004) argued that a complete population of star clusters is not made up by all clusters ever formed in a galaxy, but by a subset of clusters formed during a *formation epoch*. Assuming that all stars form in star clusters the total mass of a complete population is then given by the product of the SFR and the length of the formation epoch,

$$M_{\rm tot} = {\rm SFR} \times \delta t. \tag{3.4}$$

Thus, given a certain (short) δt, the total mass and thereby implicitly the upper mass limit of the distribution function depend on the current SFR. With this description massive clusters can only form if there is much star forming activity, while quiescent phases only produce low-mass clusters.

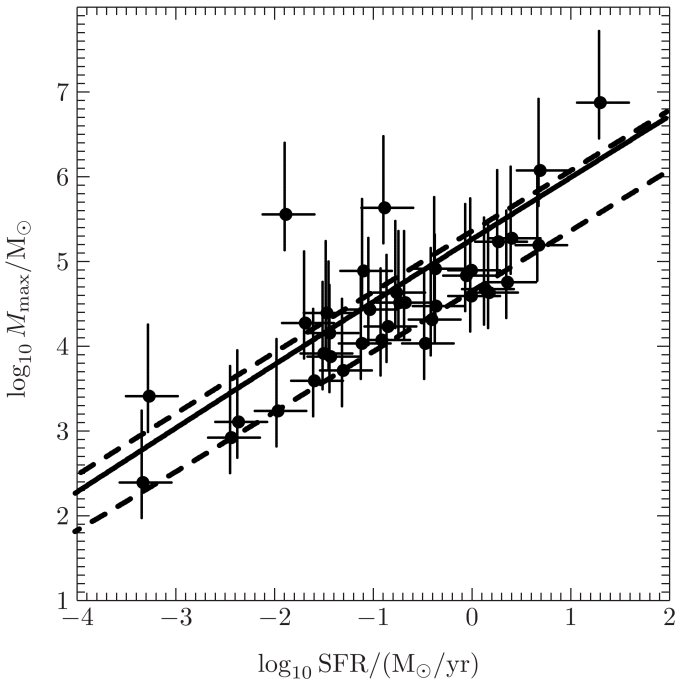


Figure 3.1: Masses of the brightest clusters vs. presen-day SFR in galaxies, data from Larsen (2002) (absolute magnitudes converted to masses by Weidner et al., 2004). The lines shown with the data are based on our statistical point of view, presented in Sec. 3.2.2. The parameters δt and β are chosen to fit to the data with $\delta t = 10$ Myr, $\beta = 2.4$, and $M_{\rm L} = 5\ {\rm M}_\odot$. The solid line is the expectation value for the distribution of the most massive cluster (eq. 3.6) in dependence of the total mass of the population. The dashed lines are the borders of the region in which 2/3rd of all most massive clusters are expected, calculated using eq. 3.5.

3.2.2 The $M_{\rm max}({\rm SFR})$ relation

Observations give evidence that the brightness of the brightest cluster in a galaxy depends on the present-day SFR in the galaxy (Larsen, 2002). Since the brightest cluster in a galaxy is usually young it can be interpreted as the most massive cluster of the current formation epoch. Weidner et al. (2004) converted the luminosities to masses, as shown in Fig. 1. The data show a large scatter which in our interpretation results from the distribution of the most massive cluster. Since the distribution of $M_{\rm max}$ for a given $M_{\rm tot}$ is known the data can be used to determine the length of the formation epoch.

In our description of star cluster populations the upper mass limit, $M_{\rm U}$ and the most massive cluster, $M_{\rm max}$, are not identical. For an ensemble of populations with the same total mass, $M_{\rm max}$ has a distribution parametrised by $M_{\rm tot}$ and β (cf. Oey & Clarke, 2005). This distribution can be written as a probability density,

$$\phi(M_{\rm max}) \quad = \quad \left(\frac{1}{N_{\rm tot}}\int_{M_{\rm L}}^{M_{\rm max}} \xi_{\rm cl}(M)dM\right)^{N_{\rm tot}-1} \xi_{\rm cl}(M_{\rm max}), \tag{3.5}$$

where $N_{\rm tot}$ and the normalisation of $\xi_{\rm cl}$ depend on $M_{\rm tot}$. This is different to the ansatz of Weidner et al. (2004) where $M_{\rm max}$ was assumed to be identical for all populations with the same $M_{\rm tot}$, i.e. not distributed. The distribution of $M_{\rm max}$ is asymmetric because of the asymmetric cluster mass function and is characterised by the average mass of the most massive cluster, $\overline{M_{\rm max}}$, given by

$$\overline{M_{\rm max}} \quad = \quad \int_{M_{\rm L}}^{M_{\rm U}} M'_{\rm max}\phi(M'_{\rm max})dM'_{\rm max}. \tag{3.6}$$

Figure 3.2 shows the distribution of $M_{\rm max}$ and the location of $\overline{M_{\rm max}}$. Due to the asymmetry the median, $M_{1/2}$, does not have the same location as the average, $\overline{M_{\rm max}}$, but lies below it. Therefore it is expected that in an ensemble more $M_{\rm max}$ lie below $\overline{M_{\rm max}}$ than above.

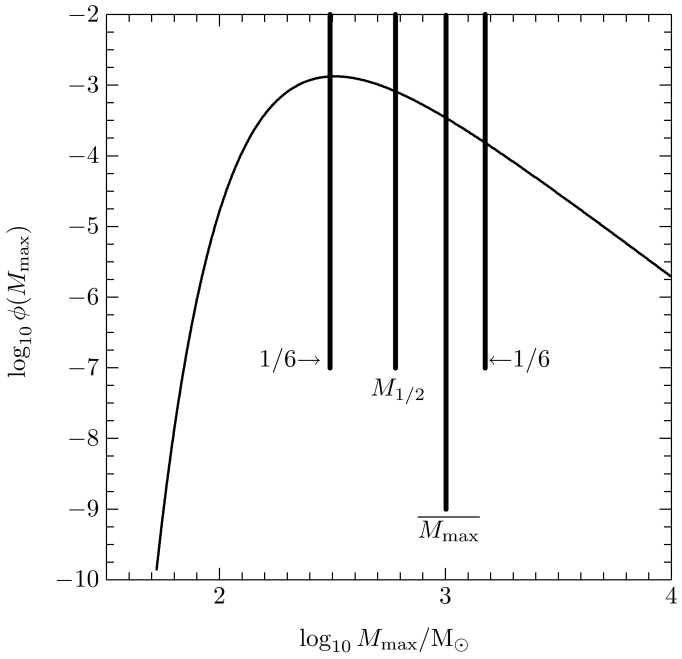


Figure 3.2: The distribution of $M_{\rm max}$ for a population with $M_{\rm tot} = 10^4$ M$_\odot$, $\beta = 2.4$, and $M_{\rm L} = 5$ M$_\odot$. The number of $M_{\rm max}$ expected to lie below and above the average $\overline{M_{\rm max}}$ is different due to the asymmetry. The median, $M_{1/2}$, has a different value than the average. 2/3rd of all $M_{\rm max}$ are expected in the region delimited by lines marked with $1/6 \rightarrow$ and $\leftarrow 1/6$.

The observations shown in Fig. 3.1 are an ensemble of most massive clusters of populations with different total masses. Since $\phi(M_{\rm max})$ is parametrised by $M_{\rm tot}$ it is possible to derive the average $\overline{M_{\rm max}}$ in dependence of $M_{\rm tot}$ and via eq. 3.4 also in dependence of the SFR. This analytical curve could be compared to a line derived from observations e.g. by least squares fitting. The least squares fitting procedure would give a line that leads through the region where most of the $M_{\rm max}$ are, but this does not match $\overline{M_{\rm max}}$ due to the asymmetry of $\phi(M_{\rm max})$. Most of the $M_{\rm max}$ lie below $\overline{M_{\rm max}}$.

To constrain δt from the observations it is more convenient to use the region where a certain fraction of the data is expected. The location of this region —at high or low values of $M_{\rm max}$— depends on the duration of the formation epoch: A long formation epoch pushes the region towards high masses, a short formation epoch to lower. A larger β steepens the relation. The best fitting region where 2/3rd of the most massive clusters are expected is shown in Fig. 3.1. As found by Weidner et al. (2004) the values of δt = 10 Myr and $\beta = 2.4$ provide a good fit to the data. Complete populations of star clusters form in formation epochs lasting for 10 Myr which is comparable to the time-scale of the emergence of cluster populations from spiral arms (Egusa et al., 2004; Bonnell et al., 2006b).

3.3 A method to derive star formation histories using star clusters

3.3.1 Concept and restrictions

The analysis of the $M_{\rm max}$(SFR) relation suggests that the SFH of a galaxy can be interpreted as a sequence of formation epochs lasting for $\approx$ 10 Myr. A star cluster population with a total mass determined by the current SFR emerges during each formation epoch. Therefore it should be possible to infer the SFR from the properties of this population, but for the largest part of the lifetime of a galaxy only a small fraction of coeval clusters is observable making the determination of the population mass difficult.

On this account we propose a different approach using only the most massive clusters of each formation epoch. The mass of the most massive cluster also depends on the current SFR, but the distribution of them does not allow the derivation of the SFR from only *one* most massive cluster. If it is assumed that the galaxy-wide SFR changes significantly only on a time-scale that includes a number of formation epochs, then the clusters of this set of formation epochs can be seen as an ensemble of identical cluster populations. For the ensemble average of the most massive clusters the mass of the population can be calculated using the $M_{\rm max}(\mathrm{SFR})$ relation: The probability density $\phi(M_{\rm max})$ is parametrised by $M_{\rm tot} = \mathrm{SFR} \times \delta t$ and β and consequently also $\overline{M_{\rm max}}$, written symbolically as

$$\overline{M_{\rm max}} \quad = \quad f(\mathrm{SFR}), \tag{3.7}$$

where $f(\mathrm{SFR})$ is given by the integral of eq. 3.6. Thus the SFR corresponding to this ensemble follows by inversion of the previous equation,

$$\mathrm{SFR} \quad = \quad f^{-1}(\overline{M_{\rm max}}). \tag{3.8}$$

The inversion of the integral (f) is done numerically.

This gives the general idea to reconstruct SFHs: The lifetime of a galaxy is divided into time windows containing a number of formation epochs. For each of the windows the average mass of the most massive clusters is calculated and from this mass the underlying SFR is derived.

The age determinations of available cluster data for galaxies, as e.g. the Large Magellanic Cloud (de Grijs & Anders, 2006), usually have uncertainties of ≈ 0.4 dex. If the observational situation is optimal, i.e. observations in the most suitable filters could be made, then the age uncertainty can be much smaller. de Grijs et al. (2005) obtained $\Delta\log(\mathrm{age/yr}) \lessapprox 0.15$ "in the majority of cases" for conditions as for NGC 3310. Since it does not make any sense to try to detect variations of the SFR on time-scales shorter than the age uncertainties of the data, the averaging window has to be chosen to have the same length (or longer) than the available age uncertainty. Because the error is constant in logarithm the length of the averaging window depends on the age. We chose a length of 0.5 dex for the averaging window. A sequence of neighbouring, independent averaging windows leads to a SFH determined at discrete points in time over the lifetime of a galaxy. For a continuous SFH the averaging window is moved in 10 Myr-steps. At old ages the length of the averaging window is reduced to ensure that the oldest formation epoch used in the averaging window contains a cluster. An averaging window filled only halfway would lead to a systematically underestimated SFR.

As the averaging window increases with age, the minimum duration of resolvable events in the SFH also increases. This is an inherent restriction of our proposed method. Short bursts of star formation which happened at large ages cannot be resolved.

3.3.2 Corrections for dynamical cluster evolution

The cluster ages and masses for large samples of clusters are usually determined by fitting models to observed spectral energy distributions. The age and the "initial" mass of a cluster are parameters for the fitting routine. In all cases the cluster mass is determined by scaling a model "initial" mass. Since the applied models usually only consider mass loss due to stellar evolution and not due to dynamical evolution, the fitted mass of a cluster does not correspond to the initial cluster mass for which the $M_{\rm max}(\mathrm{SFR})$ relation is valid.

Star clusters are no isolated static objects: the gravitational force keeps the stars in constant motion with respect to the centre of mass, which itself moves on an orbit through the host galaxy. Stars can evaporate from the region dominated by the cluster potential and leave their star cluster. This mass loss of the cluster due to dynamical evolution depends on the eccentricity of the orbit and the distance to the galactic centre. Furthermore, it depends on the mass of the host galaxy which determines the strength of the tidal field.

Since these parameters are mostly not available, the analytic model from Lamers et al. (2005a) for the average mass loss suffered by a cluster in a galaxy is used. Lamers et al. (2005a) find a good agreement of their model with the results of Baumgardt & Makino (2003), who give a formula derived from *N*-body experiments.

For the statistical model adopted here it is assumed that cluster disruption depends only on the initial mass (Boutloukos & Lamers, 2003). In this case, the initial mass can be calculated from the observed mass and age using

$$M_i(t) = \left(\left(\frac{M}{\mathrm{M}_\odot} \right)^\gamma + \frac{\gamma t}{t_0} \right)^{\frac{1}{\gamma}}, \tag{3.9}$$

where $\gamma = 0.62$ is identical for all galaxies. t_0 describes the tidal field and can be determined from the dissolution time of a 10^4 $\mathrm{M}_\odot$ cluster,

$$t_0 = \left(\frac{t_4}{660} \right)^{\frac{1}{0.967}}. \tag{3.10}$$

The parameter t_4 has been determined for a number of galaxies (cf. Boutloukos & Lamers, 2003; Lamers et al., 2005b).

3.3.3 Upper and lower limit for the SFH

The observed cluster content of a galaxy usually does not provide a cluster for every formation epoch. The number of observed clusters older than a few Gyr is much smaller than the number of formation epochs. de Grijs & Anders (2006) found only ≈ 10 clusters older than 4 Gyr in the Large Magellanic Cloud, as similarly Bastian et al. (2005) for M51. This originates from a SFR which was so low that no clusters were formed being massive enough to be visible today. The brightness limit of the observations is therefore an upper limit for the brightness of the actually formed most massive cluster. With a cluster evolution model the brightness limit at a given time can be converted to an initial mass which then is the upper mass limit for the most massive cluster. The GALEV models (Schulz et al., 2002) give the luminosity evolution $\mathscr{M}(t)$ of a "simple stellar population" (i.e. single burst, single metallicity) with a mass of 1.6×10^9 $\mathrm{M}_\odot$ for different photometric bands. By scaling $\mathscr{M}(t)$ to the limiting magnitude of the observation, $\mathscr{M}_{\mathrm{lim}}$, with the band chosen according to the bands used in the cluster-mass fitting, the limiting mass can be derived (cf. Hunter et al., 2003):

$$M_{\mathrm{lim}}(t) = 1.6 \times 10^{9+0.4(\mathscr{M}(t)-\mathscr{M}_{\mathrm{lim}})}\ \mathrm{M}_\odot. \tag{3.11}$$

The GALEV models do not take dynamical evolution into account, therefore M_{lim} has also to be corrected for dynamical evolution as described in the previous section ($M = M_{\mathrm{lim}}$ in eq. 3.9).

Now the SFH can be derived with M_{lim} as the most massive cluster in those formation epochs that do not contain any clusters, leading to the upper estimate of the SFH. Since M_{lim} increases for older ages due to the internal evolution of clusters, the derived SFR also increases, which does not necessarily reflect the underlying SFH.

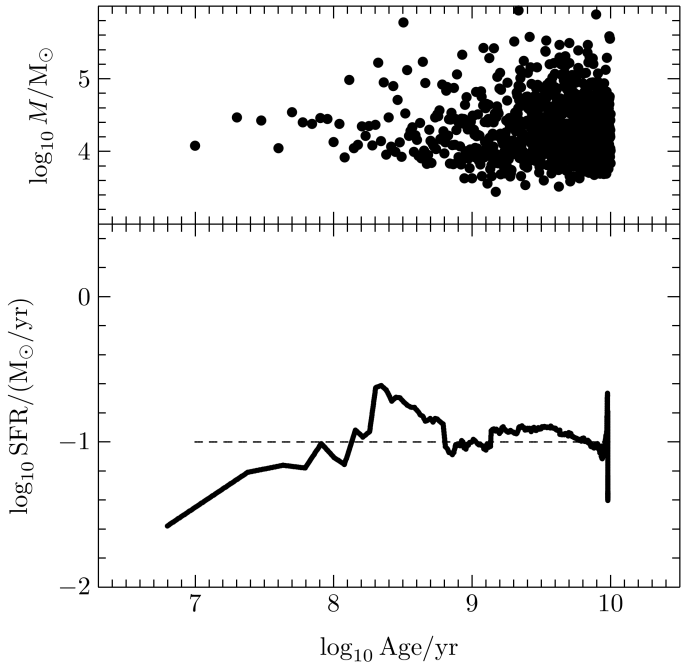


Figure 3.3: Test of the new method for a modelled constant input SFR (dashed line in the lower part). In the upper part the most massive clusters of each 10-Myr formation epoch are shown (dots). The solid line in the lower part is the reconstructed SFH using the method described in the text.

The lower limit of the SFH is given by using $M_{\mathrm{max}} = 0$ in the epochs containing no cluster, corresponding to the assumption that during these epochs no star formation took place at all.

3.3.4 Self-consistency checks

In the next section tests to analyse the systematic errors of our method are made using modelled SFHs. For an observed nearby galaxy the obtained results can be compared to independently determined quantities. The SFH derived using star clusters should be similar to the one obtained using colour-magnitude diagrams. This comparison can be difficult if the used regions of the galaxy differ.

The total mass of stars in a galaxy can be calculated since with our method a SFH for an entire galaxy is derived. For each 10-Myr epoch a SFR has been determined, from which the mass of the formed stars can be calculated. The sum over all epochs gives the mass of the stellar content of the galaxy. When deriving this mass, stellar evolution is not taken into account, i.e. all stars that ever formed are counted regardless of whether they still exist or not. This can be compared to independent determinations of the stellar content of a galaxy.

Furthermore, the cluster formation rate should reflect the structures found in the SFH. For each most massive cluster an appropriate total number of clusters should exist. This comparison of the cluster formation rate (i.e. number of clusters per time) and the SFH can be done e.g. by using a cluster formation rate also derived using a moving time window.

3.4 Modelled star formation histories

To test our method synthetic star cluster populations were generated for different SFHs. The aim is to verify if an input SFH can be re-extracted and to study the effects of the averaging. For this purpose we first consider the simplest case with optimal data, i.e. a constant SFR and no measurement uncertainties for the age. Models with a varying SFR that include the

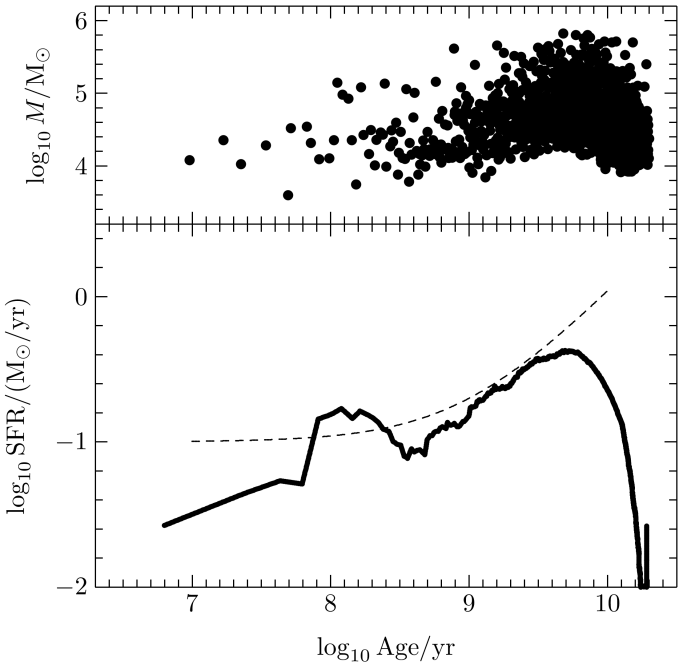


Figure 3.4: Same as Fig. 3.3 but for a slowly linearly decreasing input SFH. Here an age uncertainty with $\sigma_\tau = 0.15$ dex was included, leading to the decline at high ages as discussed in the text.

age uncertainties show the capacity of our method. For clarity and to focus on the effects of the averaging and age uncertainties, cluster evolution is only considered in the last model (Fig. 3.9).

The general procedure of our models is as follows: In time steps of 10 Myr complete cluster populations are generated with a total mass given by eq. 3.4 and the mass limits $M_\mathrm{L} = 5$ M$_\odot$ and $M_\mathrm{U} = \min(M_\mathrm{tot}, M_\mathrm{U,phys})$ with $M_\mathrm{U,phys} = 10^9$ M$_\odot$. Then an age uncertainty is assigned to each cluster with age τ, drawn randomly from a Gaussian in logarithmic age, $\mathcal{N}(\log_{10}\tau'; \mu = \log_{10}\tau, \sigma_\tau)$. The values chosen for the variance σ_τ are 0.15 dex and 0.35 dex, corresponding to the typical uncertainty range of the SED fitting method (de Grijs et al., 2005). Uncertainties larger than 0.5 were rejected and generated again until they are smaller than 0.5. The procedure for synthetic cluster evolution is described in the Section of the respective model. For each 10 Myr interval only the most massive cluster is shown in Figs. 3.3–3.5.

3.4.1 Constant SFR

Figure 3.3 shows the distribution of the most massive clusters for a constant input SFR (dashed line in the lower part) not including age uncertainties. It appears as if the masses of the most massive cluster increase with age, which is a size-of-sample effect. Due to the logarithmic axis 90 % of the clusters are in the age range $9 \leq \log_{10}\tau \leq 10$, therefore the probability to get a very massive cluster by chance is higher in this range. Similarly the probability to sample clusters in the mass range below the average mass of the most massive cluster ($\overline{M_\mathrm{max}}$, eq. 3.6) increases. This leads to the wedge-like shape of the distribution of the most massive clusters.

The derived SFH (solid line) only shows a large variation at young ages, where the large averaging window of 0.5 dex in log age contains only a few 10 Myr formation epochs. Therefore the reconstructed SFH is calculated from too few clusters. For ages older than ≈ 100 Myr the reconstructed SFH agrees within the expected statistical variations (as discussed in Section 3.4.3) with the input SFH. The features in the reconstructed SFH at an age of 10 Gyr are

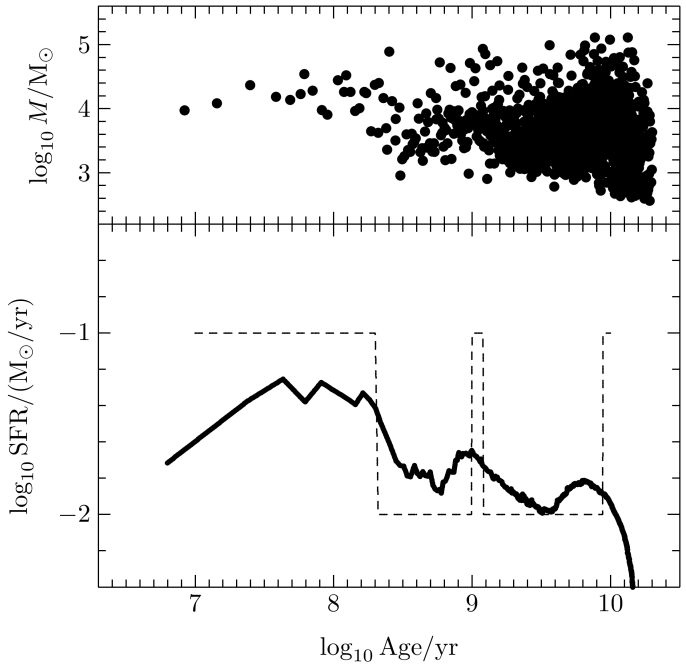


Figure 3.5: Same as Fig. 3.4 but for a SFH with three well-separated bursts.

artefacts due to the shrinking averaging window.

3.4.2 Varying SFR

Figure 3.4 shows the SFH obtained from a linearly decreasing SFR. In this model an age uncertainty with $\sigma_\tau = 0.15$ dex was included as described above. This causes a systematic deviation of the reconstructed SFH towards smaller SFRs at old ages. Although cluster populations were only generated up to an age of 10 Gyr the ages of individual clusters can be allocated up to maximally $\log_{10}\tau = 10 + 0.5$. Thus, for many old formation epochs not the actual most massive cluster corresponding to it is used but the second or third etc. most massive. Therefore the SFR is underestimated. Besides this effect introduced by the age uncertainties the behaviour of the reconstructed SFH is similar to the one of the previous model: The reconstructed SFH follows the slow change of the input SFH within the same degree of deviations. There is a large scatter in the SFH for ages younger than ≈ 100 Myr. Also as in the constant case the artefacts at the oldest ages ($\log_{10}\tau \gtrapprox 10.3$) are visible. This model allows to conclude that our method is capable of reproducing slowly changing SFHs.

The modelled SFH shown in Fig. 3.5 has three well-separated bursts lasting from 0–200 Myr, 1 000–1 200 Myr, and 8 800–10 000 Myr. During the bursts the SFR is increased by a factor of 10. Again the cluster ages were generated with an age uncertainty with $\sigma_\tau = 0.15$ dex. Our method results in an undulating SFH with peaks roughly coinciding with the centres of the bursts. The oldest burst is affected by the decline of the reconstructed SFH due to the way how the age uncertainties are assigned. Thus the maximum is shifted towards younger ages. In the reconstruction the shape of the SFH is much less pronounced than in the original. The age uncertainties and the averaging procedure lead to a larger width of the recovered bursts. As the most massive clusters of a burst are spread over an interval longer than the burst the amplitude of a recovered burst also decreases.

In Figures 3.3 to 3.5 the input SFHs guide the eye to see the appropriate structure in the

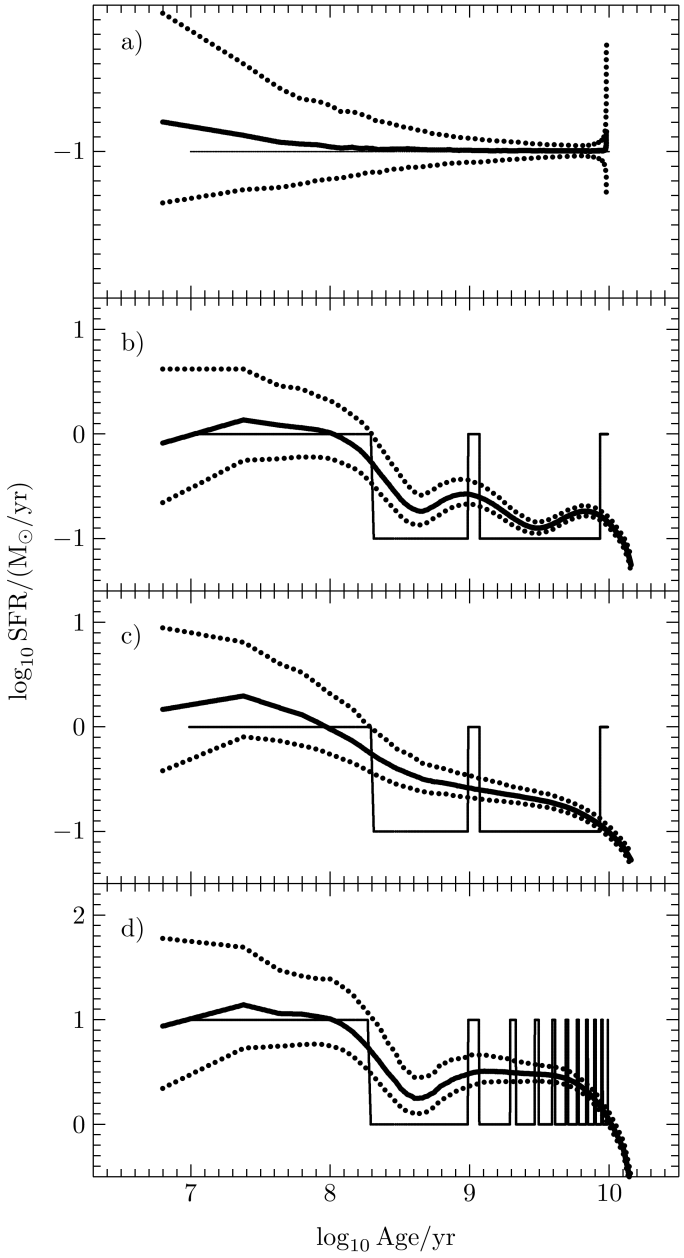


Figure 3.6: Averaged reconstructed SFHs (thick solid lines) of a sample of 1000 synthetic cluster populations for different input SFHs (thin solid lines). Also shown are the corresponding deviations, σ_{SFR+} and σ_{SFR-}, as discussed in the text (dotted lines, eq. 3.13 and 3.14).
a) constant input SFH, no age error.
b) three well separated bursts, age error $\sigma_\tau = 0.15$ dex.
c) three well separated bursts, age error $\sigma_\tau = 0.35$ dex.
d) one burst per Gyr, age error $\sigma_\tau = 0.15$ dex.

SFHs derived from the cluster distributions. The distribution of the most massive clusters, the age uncertainties, and the averaging process in our method lead to a much less distinct structure in the reconstructed SFH than in the initial one. In Section 3.4.4 we present a criterion to decide which features in a reconstructed SFH are caused by variations in the initial SFH and are not merely due to systematic effects. For this purpose it is necessary first to investigate the systematic effects of our method, which is the object of the next Section.

3.4.3 Statistical scatter in the reconstructed SFH

The models presented above show that the input and the reconstructed SFHs differ, especially for ages younger than $\approx$ 100 Myr. Due to the small number of clusters used for the

reconstructed SFH in the respective formation epochs the scatter in the reconstructed SFH increases. To investigate the expected scatter in the reconstructed SFHs, a sample of star cluster populations with the same SFH was created. For each of the input SFHs, 1000 synthetic cluster populations (i.e. 1000 galaxies) were generated with different initial random seeds and the SFH was re-extracted from the clusters. Then the sample average, $\overline{\mathrm{SFR}}(t)$, was calculated for each formation epoch,

$$\overline{\mathrm{SFR}}(t) = \frac{1}{1000}\sum_{i=1}^{1000}\mathrm{SFR}_i(t), \tag{3.12}$$

where $\mathrm{SFR}_i(t)$ is the SFH of an individual cluster population i. To achieve an estimate of the statistical spread we calculated the positive and nevative deviation, $\sigma_{\mathrm{SFR}+}$ and $\sigma_{\mathrm{SFR}-}$, from the average SFR at a given time,

$$\sigma_{\mathrm{SFR}+}(t) = \frac{1}{N_+(t)}\sum_{\mathrm{SFR}_i(t)-\overline{\mathrm{SFR}}(t)>0}|\mathrm{SFR}_i(t)-\overline{\mathrm{SFR}}(t)| \tag{3.13}$$

$$\sigma_{\mathrm{SFR}-}(t) = \frac{1}{N_-(t)}\sum_{\mathrm{SFR}_i(t)-\overline{\mathrm{SFR}}(t)<0}|\mathrm{SFR}_i(t)-\overline{\mathrm{SFR}}(t)|, \tag{3.14}$$

where $N_\pm(t)$ is the number of SFRs larger or smaller than $\overline{\mathrm{SFR}}(t)$ at a given time t. This particular choice of an individual positive and negative average deviation will be discussed after the average SFHs.

Figure 3.6 shows the results of the experiments. For the experiment presented in panel a) a constant SFH without age errors was chosen. The sample average and the input SFH agree well, differences occur only for young ages. This is because of the small number of formation epochs used for reconstructing the SFH at these ages. Since the averaging window contains few epochs the scatter in the reconstructed SFH increases, which is also visible in the progress of $\sigma_{\mathrm{SFR}\pm}$. The average deviations decrease with age because the number of formation epochs used for the reconstructed SFH increases as a consequence of the averaging window which moves in logarithmic time. At the oldest ages the average deviation increases again since the averaging window contains a decreasing number of formation epochs with a cluster, which is the same effect as for young ages. As described in Section 3.4.1 the SFH ends at the oldest ages with an artifact.

Panels b) and c) of Fig. 3.6 show the results for the SFH with three bursts and different age uncertainties. For the smaller age uncertainty, $\sigma_\tau = 0.15$ dex, all three epochs of enhanced star formation can be identified, i.e. the variation of $\overline{\mathrm{SFR}}(t)$ is comparable or larger than $\sigma_{\mathrm{SFR}\pm}(t)$. However, the reconstructed shape of the second and third burst is much wider and less pronounced. The larger age uncertainties ($\sigma_\tau = 0.35$ dex) lead to a reconstructed SFH where the SFR decreases with age and only allow the reconstruction of the youngest burst. Due to the age uncertainties $\overline{\mathrm{SFR}}(t)$ drops at old ages, as discussed in Section 3.4.2. The average deviations $\sigma_{\mathrm{SFR}\pm}$ behave similarly to the constant case.

In the case of one burst every Gyr, lasting for 200 Myr, (Fig. 3.6, panel d) with the small age uncertainty ($\sigma_\tau = 0.15$ dex) only the youngest burst can be detected. Then the reconstructed SFH declines until a minimum between the first and second burst is reached. From the second burst on only a constant SFR with an intermediate value can be recovered. $\sigma_{\mathrm{SFR}\pm}$ has the same features as before.

Table 3.1: Original and reconstructed total masses of the stellar content for the different input SFHs. Output $M_{\rm tot}$ is the ensemble average of the 1000 galaxies with the average deviation.

Model:	input $M_{\rm tot}$ $[\log_{10} M_\odot]$	output $M_{\rm tot}$ $[\log_{10} M_\odot]$
Fig. 3.7 a) Constant SFR no age error		
0.01 $M_\odot$/yr	8.00	$8.02\ ^{+0.02}_{-0.02}$
0.1 $M_\odot$/yr	9.00	$9.01\ ^{+0.03}_{-0.03}$
1 $M_\odot$/yr	10.00	$10.01\ ^{+0.04}_{-0.04}$
10 $M_\odot$/yr	11.00	$11.01\ ^{+0.05}_{-0.04}$
Fig. 3.7 b) Three bursts		
$\sigma_\tau = 0.15$ dex	9.39	$9.40\ ^{+0.05}_{-0.04}$
Fig. 3.7 c) Three bursts		
$\sigma_\tau = 0.35$ dex	9.39	$9.41\ ^{+0.06}_{-0.04}$
Fig. 3.7 d) One burst / Gyr		
$\sigma_\tau = 0.15$ dex	10.45	$10.48\ ^{-0.07}_{-0.05}$

A comparison of the averaged recovered SFHs and the input SFHs of these models show the capabilities of our method. Due to the averaging process structures in the SFH on time scales shorter than the averaging window cannot be reconstructed. In the other cases the sensitivity for structures in the SFH depends on the quality of the age determination. For the situation of a SFH only slowly varying or with bursts that are well separated and small age errors, the shape of the underlying SFH can be extracted with our method. Large age errors of the star clusters or highly variable SFHs do not allow us to recover all of the initial features in the derived SFH. However, even in these cases the absolute value of the recovered SFR is of the same order of magnitude as the actual one.

Thus, the total recovered stellar mass ($M_{\rm tot,rec} = \int {\rm SFR}(t)dt$) corresponds to the underlying SFH, summarised in Table 3.1. The ensemble average of $M_{\rm tot,rec}$ (average of the 1000 galaxies) equals to the input value and no bias is introduced.

The second aim of the models is to investigate the scatter in the reconstructed SFHs. Two properties of the average deviations, $\sigma_{\rm SFR+}$ and $\sigma_{\rm SFR-}$, are noticeable: First, in the log-log diagram the positive and negative deviation have for a certain age approximately the same distance to $\overline{\rm SFR}(t)$. A reason for this effect could lie in the asymmetric distribution of the most massive star cluster (Fig. 3.2). Thus, using the average of $\sigma_{\rm SFR+}$ and $\sigma_{\rm SFR-}$ would lead to a wrong uncertainty estimate. The second property of both average deviations is that their values relative to $\overline{\rm SFR}$ are independent of the value of $\overline{\rm SFR}$. This is visible in Fig. 3.7, where the quantities

$$\widehat{\sigma}_{\rm SFR\pm}(t) \quad = \quad \frac{\sigma_{\rm SFR\pm}(t)}{\overline{\rm SFR}(t)}, \tag{3.15}$$

the relative average deviations, are plotted. To show the independence of $\widehat{\sigma}_{\rm SFR\pm}$ from the shape of a SFH the SFHs from Fig. 3.6, panels b)–d), are used. The model with a constant SFR of 0.1 $M_\odot$/yr (Fig. 3.6, panel a)), and additional models with 0.01 $M_\odot$/yr, 1 $M_\odot$/yr, and 10 $M_\odot$/yr demonstrate the invariance from the absolute value of the SFR. The independence of the relative average deviations from the shape and absolute value of the SFH makes this quantity suitable for estimating the uncertainties of the method. For the implementation in

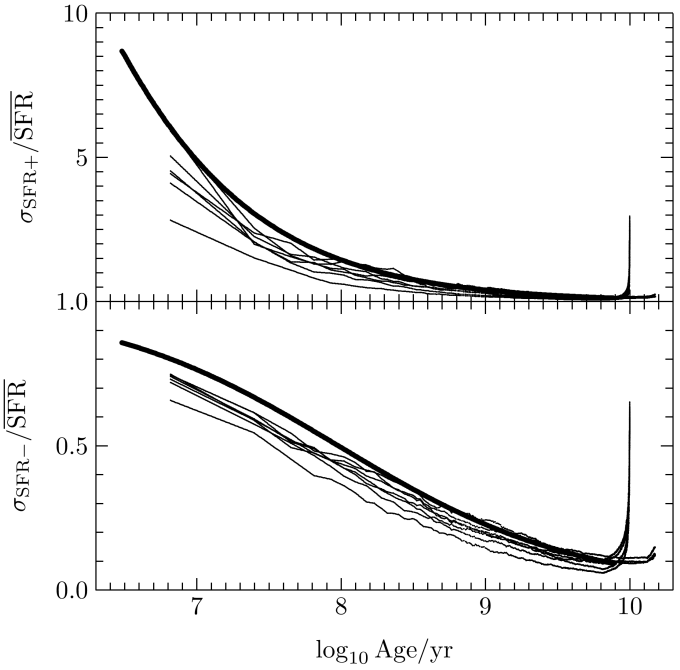


Figure 3.7: Relative errors (thin solid lines) for different constant SFHs with a SFR of 0.01 $M_\odot$/yr, 0.1 $M_\odot$/yr, 1 $M_\odot$/yr and 10 $M_\odot$/yr and for the bursting cases as above, Fig. 3.6. The thick solid line shows an upper envelope of the relative error as described in the text (eqs. 3.16 and 3.17).

our method for deriving SFHs of individual galaxies we used the analytic fitting formulae

$$\widehat{\sigma}_{\rm SFR+}(\tau) \;=\; \frac{45}{1+\exp 1.3(\log_{10}\tau - 5.4)} \tag{3.16}$$

and

$$\widehat{\sigma}_{\rm SFR-}(\tau) \;=\; \frac{1}{1+\exp 1.2(\log_{10}\tau - 8.0)}. \tag{3.17}$$

These fits are a conservative estimate and lie slightly above the experimental data. In the method the absolute average deviation is then calculated by

$$\sigma_{\rm SFR\pm}(\tau) \;=\; \widehat{\sigma}_{\rm SFR\pm}(\tau) \times {\rm SFR}(\tau). \tag{3.18}$$

It will be used in the criterion to detect significant variations in a SFH, discussed next.

3.4.4 A criterion to detect significant variations in a SFH

Due to the probabilistic distribution of the most massive cluster scatter in the reconstructed SFH is expected. As visible in Figs. 3.3 and 3.4 the reconstructed SFH can mimic periods of reduced or enhanced star formation. However, the variations generated in this way are not caused by real events. The average deviations derived in the previous Section can be used to give the region of SFHs compatible with the reconstructed one. But to verify variations in a SFH we suggest to disproof the hypothesis that there are *no* variations. This is done by setting up the null hypothesis of a constant SFH with a SFR equal to the time-averaged SFR, $\langle {\rm SFR} \rangle$, of the reconstructed SFH. If the reconstructed SFH leaves significantly the $1\sigma_{\rm SFR\pm}$-region of the null hypothesis, a constant SFH, can hardly be supported. Thus the actual SFH of the galaxy has to have variations. This is our suggested criterion of significance. Note that this criterion

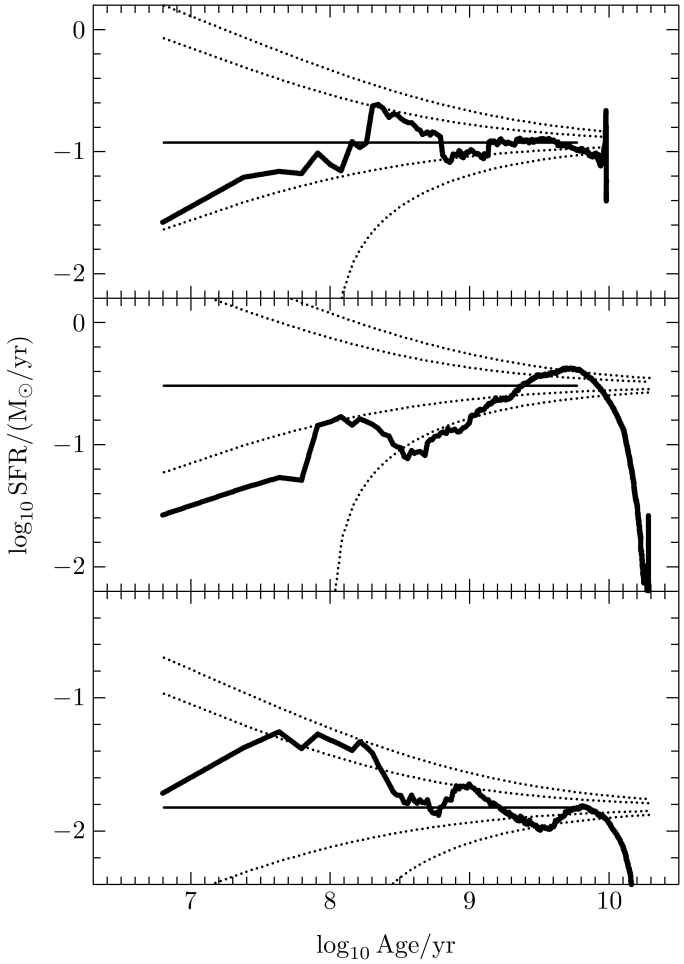


Figure 3.8: Reconstructed SFHs from Figs. 3.3, 3.4, and 3.5 (top to bottom: constant, linearly decreasing, with three bursts; thick lines), shown with the null hypothesis of a constant SFH (thin solid lines, $\langle \mathrm{SFR} \rangle$). For a significant variation of the reconstructed SFH it has to cross the $1\sigma_{\mathrm{SFR}\pm}$ line (inner dotted lines). The outer dotted lines are at $2\sigma_{\mathrm{SFR}\pm}$.

implies that troughs and maxima that result from a truly variable SFR differ by more than $2\sigma_{\mathrm{SFR}\pm}$.

$\langle \mathrm{SFR} \rangle$ is given by the integral over the SFH, divided by the time. With our method two SFHs are reconstructed, the lower limit with gaps where no cluster was observed, and the upper limit with the fading limit mass used in the gaps. Therefore $\langle \mathrm{SFR} \rangle$ has to be calculated as the mean of the average SFR of both limits. In reality these limits start to differ for older ages, caused by incompleteness due to cluster evolution and observational limits. Especially in galaxies with strong cluster evolution the upper limiting SFH can shift $\langle \mathrm{SFR} \rangle$ towards unreasonable values if it is integrated over all ages. To prevent this we integrate only up to the age where the logarithms of the upper and lower limit differ by less than 0.2.

The outcome of this procedure is displayed in Figs. 3.8 and 3.9. Figure 3.8 shows again the SFHs of Figs. 3.3, 3.4, and 3.5, now with the criterion for variations. Since in these models no cluster evolution was incorporated, we integrated up to an age of 6 Gyr (where the thin solid line stops) to obtain $\langle \mathrm{SFR} \rangle$. In the constant case the input SFR (0.1 $\mathrm{M}_\odot$/yr) and the average ($\langle \mathrm{SFR} \rangle = 0.12$ $\mathrm{M}_\odot$/yr) are in good agreement. The extremes of recovered SFH barely exceeds the $1\sigma_{\mathrm{SFR}\pm}$ region. Thus the null hypothesis of a constant SFH cannot be rejected, as is correct in this case.

The reconstructed linearly decreasing SFH lies, during most times, far away from the average. At young ages the recovered SFH lies sometimes even below $2\sigma_{\mathrm{SFR}-}$, and before

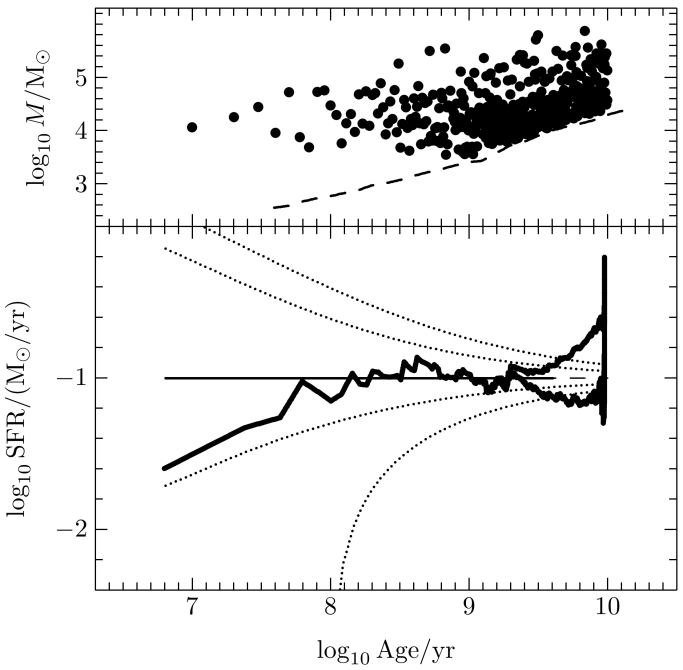


Figure 3.9: Recovered SFH for a LMC-type cluster model that includes dynamical evolution and the missing of clusters due to the observational flux limit (dashed line in the upper part). As previously the upper panel shows the cluster distribution, whereby here each cluster is shown with its current mass after evolving it dynamically, and the lower panel contains the reconstructed SFH (thick lines). Due to missing clusters the reconstructed SFH branches in two parts, as described in Section 3.4.5. The thin solid line is the constant SFH ($\langle\mathrm{SFR}\rangle$)used to detect significant variations of the true SFH and stops when the averaging is stopped. Dotted lines indicate the $1\sigma_{\mathrm{SFR}\pm}$ and $2\sigma_{\mathrm{SFR}\pm}$ regions.

the decline due to the age uncertainties at old ages it rises above $2\sigma_{\mathrm{SFR}+}$. From this can be deduced that the null hypothesis is not consistent with the reconstructed SFH. A clear trend of increase with age is visible.

The bursting case is harder to identify. Only the first burst (0–200 Myr), and the last dip (1.2–8.8 Gyr) leave unambiguous traces, albeit the shape is less pronounced than that of the underlying SFH. The two other bursts and the first dip merely allow an "educated guess" of the shape of the actual SFH. As above a constant SFH is not consistent with the reconstructed SFH. However, the *exact* structure of this SFH cannot be determined with sufficient certainty, although it is visible.

The above cases show that our suggested criterion of significance allows us to distinguish between features in the actual SFH and artefacts due to the method.

3.4.5 Constant SFR including cluster evolution and an observational limit

The models described above do not account for cluster evolution and the observational limiting magnitude for cluster detection. To show the consequences of these effects a galaxy-model with conditions similar to the Large Magellanic Cloud (LMC) was generated. The stellar mass of the LMC was determined by Kim et al. (1998) to be 2.0×10^9 $\mathrm{M}_\odot$. Assuming a constant SFH leads to a SFR of ≈ 0.1 $\mathrm{M}_\odot$/yr, which we used in our model. The cluster-disruption parameter is $t_4 = 7.9 \times 10^9$ yr (Boutloukos & Lamers, 2003) and the flux limit is $\mathscr{M}_{\mathrm{lim}} = -3.5$ mag (Hunter et al., 2003) in the V Band. The generated data and the results are shown in Fig. 3.9.

Because of the weak tidal field clusters evolve only slowly and the high mass-end of the clusters is similar to the case without cluster evolution. The effect of the observational limit is clearly visible as a cut-off in the lower part of the cluster distribution. The minimum observ-

able mass increases with time due to the luminosity evolution of the clusters.

As in the previous models the reconstructed SFH deviates from the input value at young ages because of the small number of formation epochs used for averaging. Then there is a period of good agreement, until some formation epochs contain no clusters any more. Due to the cluster evolution the clusters of these epochs are dissolved or have lost such a large fraction of their stellar content that they cannot be detected. Therefore the reconstructed SFH now shows two branches corresponding to the upper and lower limit. Assuming that no detection of a cluster means that there was no star formation activity leads to the lower limit. Using the detection limit as the mass of the most massive cluster gives the upper limit. The true SFH lies between both limits, which is indeed confirmed: Until the branching of the upper and lower limit the $1\sigma_{\mathrm{SFR}\pm}$-region contains the reconstructed SFH. From this point on only a rough estimate of the SFH can be obtained. From the input SFH a total stellar mass of 1×10^9 $\mathrm{M}_\odot$ was built up. The upper and lower limit reconstructed 0.78×10^9 $\mathrm{M}_\odot$ and 1.74×10^9 $\mathrm{M}_\odot$, embracing the model value. That the reconstructed mass range contains the known input value constitutes a consistency check.

3.5 Conclusions and Summary

Based on the assumption that all stars form in star clusters it is possible to explain the relation between the brightest clusters in a galaxy and the present star formation rate, assuming that the brightest cluster is also the most massive. During a formation epoch lasting for ≈ 10 Myr a complete population of star clusters is formed. The mass of the most massive cluster obeys a distribution function that can account for the scatter in the $M_{\max}(\mathrm{SFR})$ relation. The SFH of a galaxy can then be seen as a sequence of formation epochs. Starting from this we presented a new method to derive SFHs using star clusters taking into account the statistical properties of the most massive cluster as well as their dynamical evolution.

The method was tested for a number of model SFHs for which synthetic cluster populations were created. The tests show that our method is capable of reproducing the modelled SFHs if they are only slowly varying or have bursts which are well separated. To be resolved, the time between two short-time bursts needs to be longer than 0.5 dex, the time over which is averaged. However, the typical uncertainties in the age determination and the need for averaging do not allow a shorter averaging window. Artefacts result from averaging over too few small age bins and from missing data at high ages. The example SFHs show to which degree our method can be used to make confident statements about the SFH of a galaxy.

A model including realistic conditions for observation and cluster evolution also leads to good agreement between the input and the reconstructed SFH. In following contributions we will apply this method to the galaxies LMC (this thesis Chapter 4), SMC, M51 and M101. For the LMC we will compare this new method to the results obtained using the CMD method.

Acknowledgements

We thank Carsten Weidner for useful discussions, and Ylva Schuberth for critical reading of the manuscript. ThM acknowledges financial support by the AIfA.

Chapter 4

The star formation history of the Large Magellanic Cloud as seen by star clusters and stars

Th. Maschberger & P. Kroupa
MNRAS 411:1495–1502 (2011)

In this work we derive the star formation history of the Large Magellanic Cloud from its star cluster population. We follow two methods, either using only the most massive clusters (following Maschberger & Kroupa, 2007, this thesis Chapter 3) or using the whole cluster population, albeit then only for a shorter age span. We compare these results with the star formation history derived from colour-magnitude diagrams and find good overall agreement up to ≈ 1 *Gyr. For later ages, especially the "cluster age gap", there is a deficiency of star clusters in relation to the star formation rate derived from the colour-magnitude diagram. The star formation rates following from the whole cluster population lie a factor of* ≈ 10 *lower than the star formation rates from the colour-magnitude diagram, suggesting that only* $\approx 10\%$ *of all stars form in long-lived bound star clusters.*

4.1 Introduction

The understanding of galaxy evolution is a major goal of astrophysics. Every large-scale event in the life of a galaxy, as e.g. an interaction with another galaxy, has its own pattern of star formation. Since stars can have long lifetimes, the stellar population preserves information of such events, allowing one to re-trace the galaxy's evolution from the present stellar content. In this work we focus on the star formation history, the progression of the star formation rate in time. This study has two main aspects, the comparison of two different methods to obtain a star formation history, from colour-magnitude diagrams and from the star cluster population using the method of Maschberger & Kroupa (2007, this thesis Chapter 3). Furthermore we discuss the star formation history of the Large Magellanic Cloud, which serves as a "guinea

pig" for the comparison.

The common method to obtain a detailed star formation history is to observe the stars in a galaxy (or a part of it). From the distribution of the stars in a colour-magnitude diagram (CMD) the star formation rate at a given time can be derived using modelled tracks of stellar evolution. To get a result which is representative for the whole galaxy it is necessary to observe a significant fraction of the stars in the galaxy, distributed over a large area. This leads to limitations of this method: since individual stars need to be resolved, only nearby galaxies can be examined. Also, a large number of stars and a large area demand a big observational effort. Fortunately, the Large Magellanic Cloud has been extensively observed, so that a set 24 million stars is available from which Harris & Zaritsky (2009) derive the star formation history.

Another approach to infer a star formation history was presented by Maschberger & Kroupa (2007, this thesis Chapter 3). Here the fact is used that practically all stars form in star clusters. The notion of a star cluster is here taken in a wider sense, denoting stellar assemblies from the smallest size, say a dozen stars, up to classical globular clusters, and does not necessarily imply a bound system. Whereas small clusters will disperse their stars rather quickly into the galactic field, massive clusters have lifetimes comparable to a Hubble time. Since the time distribution of massive clusters is related to the star formation rate at their birth, they can be used to find the star formation history of their host galaxy. This approach using the most massive star clusters has been investigated by Maschberger & Kroupa (2007, this thesis Chapter 3) from a theoretical point of view, applying Monte-Carlo models to study how reliable the massive clusters trace the star formation history. The value of this method is that, as (massive) star clusters are observable up to much further distances than individual stars, the star formation history of a wider range and number of galaxies can potentially be obtained.

Additionally the time-sequence star formation occurring in star clusters can be determined by simply tanking all star clusters into account, given that also a large fraction of lower-mass star clusters is observed. Subramaniam (2004) compared in six regions of the Large Magellanic Cloud the cluster formation rates and star formation rates (derived from colour-magnitude diagrams), and found that in general the number of formed clusters follows the "field" star formation rates. Cluster formation rates only have also been derived by Girardi et al. (1995), Pietrzyński & Udalski (2000), Hunter et al. (2003), and de Grijs & Anders (2006) Our approach is slightly different, we use the total mass in clusters per time instead of number of clusters per time. With the diminishing brightness of star clusters as they age the number of older star clusters decreases, so that only a shorter fraction of the galaxy's life-time can be investigated in this way.

The Large Magellanic Cloud has been the target of many research projects and is ideal to compare the CMD and star cluster methods. In the literature there are a number of studies available on the star formation history of the Large Magellanic Cloud, which are based on the CMD approach (e.g. Harris & Zaritsky, 2009; Olsen, 1999; Holtzman et al., 1999; Dolphin, 2000; Smecker-Hane et al., 2002; Subramaniam, 2004; Javiel et al., 2005). Furthermore, the star cluster population of the Large Magellanic Cloud has been investigated, and ages and masses of a large fraction of the star clusters have been determined (Pietrzyński & Udalski, 2000; Hunter et al., 2003; de Grijs & Anders, 2006). This enables us to study the star formation history of a galaxy with two independent methods.

These introductory remarks outline the structure of this work, in summary: After first discussing the star cluster data set, we derive the star formation history of the Large Magellanic

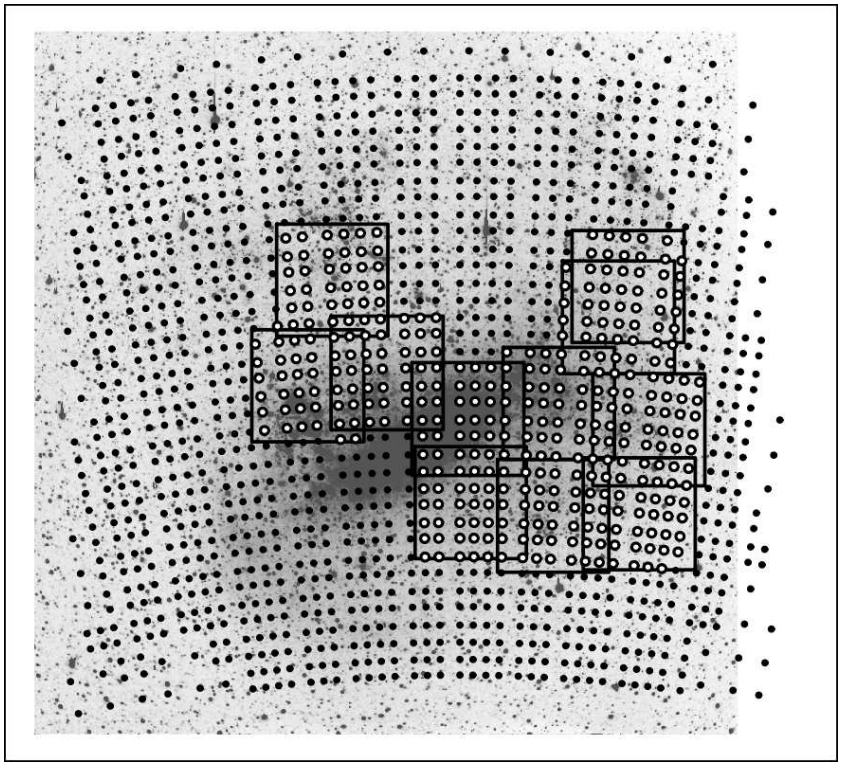


Figure 4.1: Overview of the observed regions in the Large Magellanic Cloud (background image from Bothun & Thompson, 1988, with astrometry by Parker et al., 1998). The dots mark the centres of the fields observed by Harris & Zaritsky (2009) and for which a star formation history was derived using a colour-magnitude diagram. The squares are the boundaries of the regions observed by Massey (2002), in which de Grijs & Anders (2006) derived star cluster ages and masses. The open dots are the fields of Harris & Zaritsky (2009) which we selected for comparison.

Cloud from the most-massive clusters (Sec 4.2.2) and from the total population (Sec. 4.2.3). Then we describe the results obtained from colour magnitude diagrams (Sec. 4.3). We finish with a comparison of the results (Sec. 4.4) and a summary (Sec. 4.5).

4.2 The star formation history of the Large Magellanic Cloud as seen by star clusters

4.2.1 Data

The ages and masses of the star clusters we use for the analysis are taken from de Grijs & Anders (2006), which re-analysed the photometry of Hunter et al. (2003), which itself is based on the observations by Massey (2002). The rectangles in Fig. 4.1 show the spatial coverage of the observed regions with star clusters. The star cluster ages and masses were derived by de Grijs & Anders (2006) from broad-band spectral energy distributions using their AnalySED tool which is based on the GALEV single stellar population models (Kurth et al., 1999; Anders

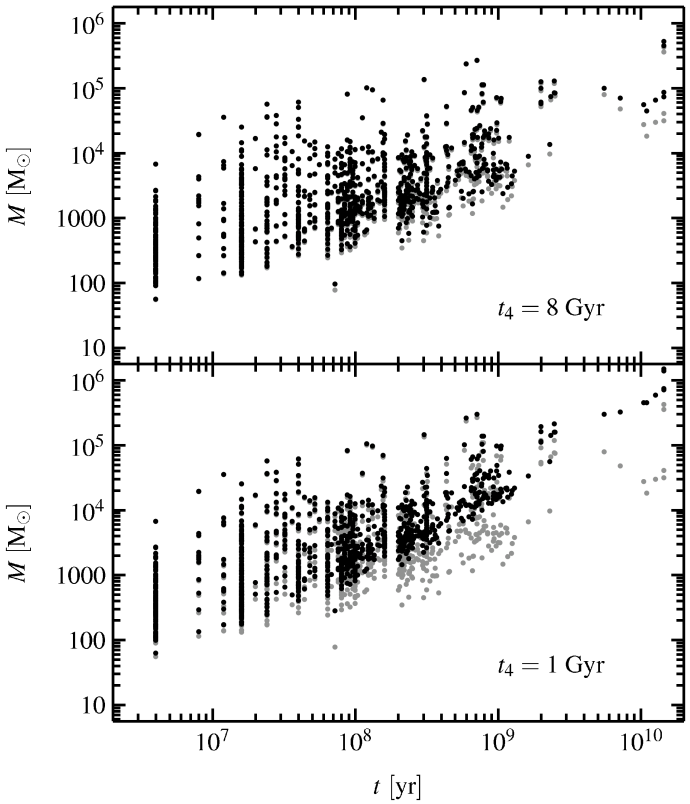


Figure 4.2: Age-Mass diagram of the star clusters in the Large Magellanic Cloud(grey circles original data from de Grijs & Anders, 2006). The cluster masses have been corrected for dynamical evolution using eq. 4.1 with $t_4 = 8$ Gyr (black dots, top panel) and $t_4 = 1$ Gyr (black dots bottom panel). 30Dor is not included in this Figure.

& Fritze-v. Alvensleben, 2003; Anders et al., 2004). The age uncertainties for the 922 star clusters are in the range $\Delta\log_{10}(\tau/\mathrm{yr}) \leq 0.35$ (de Grijs et al., 2005). Due to discrete isochrones the age-mass diagram in Fig. 4.2 shows columns of star clusters of the same age. The lower mass limit for detection of clusters increases with increasing cluster age, as clusters fade due to stellar evolution and dynamical loss of stars, leading to the wedge-like shape of the data in Fig. 4.2.

The AnalySED tool provides "initial" masses of the star clusters which are corrected for mass loss due to stellar evolution. However, the mass of a star cluster diminishes in time also because stars are lost in consequence of dynamical evolution. As we need true initial masses for the star clusters we correct for the dynamical evolution using the formulae of Lamers et al. (2005a). Given the dissolution time of a 10^4 M$_\odot$ star cluster, t_4, the initial star cluster mass of age t is given as

$$M_{\mathrm{ini}} = \left[\left(\frac{M(t)}{\mathrm{M}_\odot}\right)^{0.62} + \frac{0.62 \times t}{(t_4/660)^{1.034}}\right]^{1.61} \tag{4.1}$$

(This follows from combining eqq. 7 and 11 of Lamers et al. (2005a), with $\gamma = 0.62$ and omitting the term for stellar evolution in eq. 11). For the Large Magellanic Cloud Boutloukos & Lamers (2003) found $\log_{10} t_4 = 9.7$ (using a smaller data set) and de Grijs & Anders (2006) gave the slightly larger value of $\log_{10} t_4 = 9.9$ (8 Gyr). Parmentier & de Grijs (2008) carefully performed a reanalysis of the dissolution time and concluded that it is with the current data set only possible to constrain t_4 to be larger than 1 Gyr. Therefore we use two values for t_4, 1 Gyr and 8 Gyr, to correct for dynamical evolution. To visualise the difference between these values we show in Fig. 4.2 the not back-evolved masses as grey circles, and the "true" initial

cluster masses as dots, using $t_4 = 8$ Gyr in the top panel and $t_4 = 1$ Gyr in the bottom panel. For us a larger t_4 seems to be more realistic, as the Small Magellanic Cloud has a similar value ($\log_{10} t_4 = 9.9$, Lamers et al., 2005b), and more massive spiral galaxies with a deeper gravitational potential have smaller values. We will, however, discuss below the implications of both values when determining the star formation history from the most massive clusters.

Further features in the age-mass diagram besides the typical wedge-like shape were pointed out by de Grijs & Anders (2006):
(1) The large densities of clusters at $\log_{10}\tau$ of 6.6 and 7.2: These are caused by the fitting procedure. There are no isochrones for clusters younger than 4 Myr ($\log_{10}\tau = 6.6$), and at $\log_{10}\tau = 7.2$ the isochrones are discrete due to rapid evolution. This does not have a large influence on the determined SFH.
(2) The under-density of data points between $\approx$ 3 Gyr and 13 Gyr ($\approx 9.5 \leq \log_{10}\tau \leq 10.1$), which is the "well-known LMC cluster age-gap".
(3) Overdensities at $7.8 \leq \log_{10}\tau \leq 8.0$, $2.8 \leq \log_{10}(M/\mathrm{M}_\odot) \leq 3.4$ and $8.2 \leq \lg\tau \leq 8.4$, all masses. This feature could be caused by the last encounter between the Large and Small Magellanic Cloud, but this cannot be concluded with sufficient certainty because of the lack of better age resolution and lack of orbital information for the galaxies.

It has also to be noted that this star cluster sample does not contain the 30Dor region, containing the young star cluster R136. It was classified as a newly formed star cluster ("NC") by Bica et al. (1999) and so in a group of objects which were not selected by Hunter et al. (2003). However, R135 is a massive star cluster having a mass of $\approx 5.5 \times 10^4$ $\mathrm{M}_\odot$ (Hunter et al., 1995), and is the most massive star cluster recently formed. The inclusion of this cluster is therefore crucial to the method used in the next Section.

4.2.2 Star Formation History using the most massive star clusters

In Maschberger & Kroupa (2007, this thesis Chapter 3) we presented and tested a method to derive the star formation history of a galaxy using the most massive clusters. This method is based on the observation that the brightness of the brightest young cluster in a galaxy is correlated with the (present) star formation rate (Larsen, 2002; Weidner et al., 2004; Bastian, 2008). This can be understood following the argument of Weidner et al. (2004). Within a certain time span of the galaxy's lifetime, δt, the amount of mass assembled in stellar clusters is proportional to the star formation rate,

$$M_{\mathrm{clusters}} = A\ \mathrm{SFR}\ \delta t \tag{4.2}$$

(A is the proportionality constant). This mass in clusters is related to a number of clusters that have formed,

$$N_{\mathrm{clusters}} = \frac{M_{\mathrm{clusters}}}{\overline{M}}, \tag{4.3}$$

where a universal cluster mass function is assumed to calculate the average mass of a star cluster, $\overline{M}$. Interpreting the star cluster mass function as a probability distribution, this then allows one to calculate the distribution of the most massive star cluster, M_{max}, that would be expected for the given N_{clusters} . From this model follows a relation of the mass of the most massive star cluster with the star formation rate within δt. This can be inverted to $SFR = f(M_{\mathrm{max}})$ which can be used to determine the star formation rate over time, discretised by δt.

In general $M_{\rm max}$ follows a probability distribution, related to the star cluster mass function, which has to be taken into account for the inversion (details of this can be found in Maschberger & Kroupa, 2007, this thesis Chapter 3). To minimise the number of assumptions, especially the exact form (pure power law or Schechter function as suggested by Gieles et al., 2006b) and parameters of the cluster mass function, we use the relation of the mean mass of the most massive cluster and the star formation rate. The $\overline{M}_{\rm max}$–SFR relation can be directly calibrated with the observed relation of the brightest young cluster and the star formation rate in a galaxy (assuming that the brightest cluster is also the most massive one of the most recent time interval, an assumption which is discussed in more detail below). This $\overline{M}_{\rm max}$–SFR relation is then applied to a mean mass of the observed most massive clusters over several δt (choosing the number of used δt such that during the whole time of averaging the star formation rate in the galaxy is not changing significantly). By using a moving averaging window (moved in steps of δt) the time resolution of the obtained star formation history can be increased. The length of the averaging window is essentially constrained by the age uncertainties of the star clusters, which are constant in logarithmic space, so that we keep the averaging window also constant in $\log_{10}$.

By using $\overline{M}_{\rm max}$ and the empirical calibration we have avoided the need of the exact knowledge of the star cluster mass function. However, another crucial ingredient in this method is the formation epoch, δt, which needs more explanation. In this context the often mentioned "size-of-sample" effect has to be discussed. The "size-of-sample" effect is simply the increase of the mass of the most massive cluster with increasing sample size. With the general assumptions of an unchanging cluster mass function and constant cluster formation rate (number per time) a logarithmic age-mass diagram has the characteristic upper envelope of an increasing mass with time. Equally-spaced time intervals in logarithmic space contain more physical time, thus more clusters are formed and subsequently the mass increases. However, this is not the full picture as the cluster (or star-) formation rate can change with time, leading for example to the "age gap" in the Large Magellanic Cloud where barely clusters are found. A mathematically more correct description would be a star cluster mass function depending on both mass *and* time, which is however not very practical. Here the formation epoch, δt, comes into the play: this is the time by which the time evolution of a galaxy is discretised. With a reasonable choice of δt the star formation rate in the galaxy can be assumed to stay constant, simplifying the statistical treatment. The increasing envelope in the log(age)-log(mass) diagram is preserved with using δt (shown in fig. 3, top panel, of Maschberger & Kroupa, 2007, this thesis Fig. 3.3). The difference to the established understanding of the "size-of-sample" effect is that one does not increase the size of a single sample, but one instead increases the number of samples.

The question is now what a reasonable size for δt is. Already mentioned was the need for the star formation rate to be constant over δt, which gives an upper limit for δt of ≈ 100 Myr, the dynamical time of a galaxy. In fact, the star formation rate should be constant over several δt so that it can be averaged over several $M_{\rm max}$. A lower limit is given by the need for a not too meagre number of clusters that form during δt, so that the distribution of $M_{\rm max}$ is not too broad. We follow here the choice of Weidner et al. (2004) and Maschberger & Kroupa (2007, this thesis Chapter 3) of 10 Myr, for the practical reason that most clusters are of this age in the observational $M_{\rm max} - {\rm SFR}$ plot, as the luminosity of a star cluster peaks at about 10 Myr. Thus the brightest cluster is in many cases of this age and at the same time the most massive. With a different choice of δt the brightest cluster would have to be replaced by the actual most massive cluster in the normalisation, i.e. a true $M_{\rm max}$-SFR diagram would have

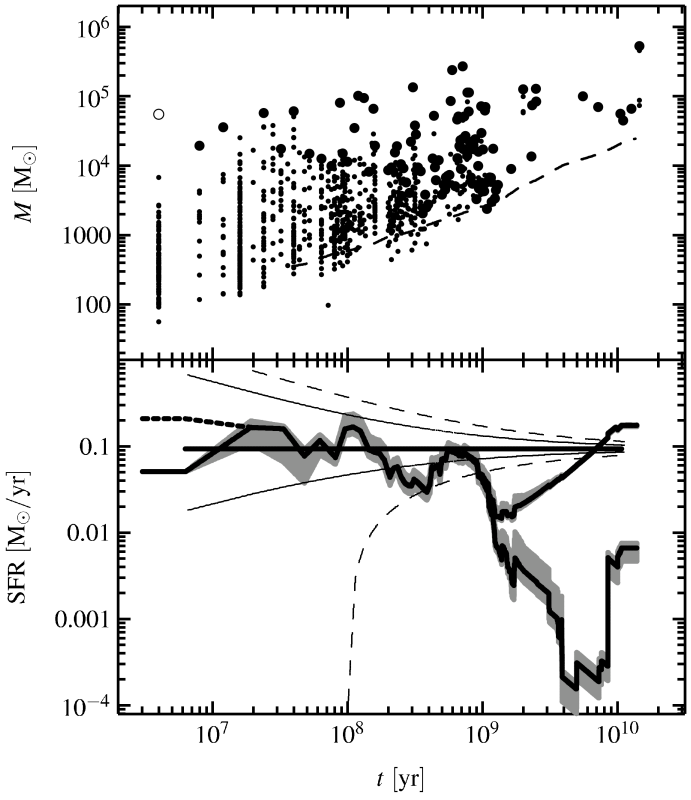


Figure 4.3: Star formation history of the Large Magellanic Cloud derived from the most massive clusters. In the top panel we shows the age-mass diagram of the star clusters (with dynamically back-evolved masses, t_4=8 Gyr), including 30Dor as the open circle and highlighting the most massive clusters of each formation epoch with bigger circles. The dashed line is the fading limit. The lower panel shows the star formation history derived from the most massive clusters (with back-evolved masses and t_4=8 Gyr) as thick solid lines, splitting into two branches, the upper using the fading limit mass in gaps and the lower using $M_{\rm max} = 0$ M$_\odot$. The thick dashed branch at early ages follows with including 30Dor. The gray area is the uncertainty propagated from the uncertainty in the cluster masses. A horizontal line marks the average star formation rate and is enclosed by the 1σ and 2σ curves (solid and dashed) derived from the statistical spread of the $M_{\rm max}$.

to be observed. One possible interpretation of $\delta t = 10$ Myr would be that it is the typical time-scale on which the inter-stellar medium rearranges itself into a coeval population of star clusters that are distributed according to the star-cluster initial mass function (cf. Weidner et al., 2004). Anyway, the comparison of the star cluster results with the CMD results will give an indirect check if our normalisation is correct.

In Fig. 4.3 we show in the upper panel the age-mass diagram of the star clusters, where the clusters identified as $M_{\rm max}$ are the large dots. R136 is shown as an open circle, as it is not contained in the de Grijs & Anders (2006) sample. The dashed line is the fading limit, the mass that a cluster with the lowest observed brightness would have (calculated with the GALEV models). In the lower panel the bold solid curve that splits into two gives the solution

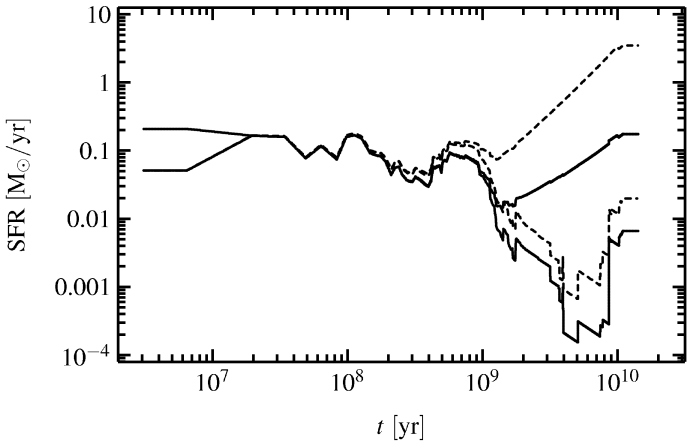


Figure 4.4: Influence of t_4 on the star formation history of the Large Magellanic Cloud using star clusters. For the solid line $t_4 = 8$ Gyr and for the dashed line $t_4 = 1$ Gyr was used.

for the star formation history (calculated with $t_4 = 8$ Gyr). For the youngest ages the dashed branch follows by including R136 in the star cluster sample. Generally the obtained star formation history follows the distribution of the star clusters for about one Gyr, when the number of clusters starts thinning out. As there are formation epochs during which no star cluster is detected we can only estimate a lower and upper limit for the star formation rate, giving the two branches of the bold solid line. The lower limit is simply no star formation at all (SFR $= 0$ M$_\odot$/yr and thus $M_{\rm max} = 0$ M$_\odot$) and the upper limit is given by the fading limit, no cluster more massive than the fading limit has formed (else it would have been observed). The uncertainty in the star formation rate introduced by the uncertainties in the cluster masses is is visualised as a grey region. It is calculated by using the $M_{\rm max}$ values plus/minus their uncertainty in mass.

The peaks in the star formation rate are somewhat displaced when compared to the loci of the massive clusters. This is caused by the way we accommodate for the age uncertainties of the star clusters. The length of the formation epoch δt can become shorter than the age uncertainties, which seems to be an over-interpretation of the data. Therefore we chose an averaging window with the size of the age uncertainties, that is constant in logarithmic time (0.5 dex), at the expense of time resolution and time-offsets.

Due to the probabilistic nature of $M_{\rm max}$ the stochastical scatter in the recovered star formation rate is very large for young ages, as averaging occurs only over a few δt, and decreases with increasing time. To assess the significance in variations of the star formation rate we show a constant star formation rate (the thick solid line at ≈ 0.1 M$_\odot$/yr) which is embraced by the statistical 1σ and 2σ scatter (thin solid lines and dashed lines). The amount of statistical scatter has been determined from Monte-Carlo experiments by Maschberger & Kroupa (2007, sec. 4.3, eqq. (16) and (17), this thesis Chapter 3, Section 3.4.3, eqq. 3.16 and 3.17). For ages between 100 Myr and 1 Gyr the star formation rate shows indeed variations at the $\approx 1\sigma$ level.

We now turn to discuss the effects of the two values for t_4. Figure 4.4 shows this, with the star formation history using $t_4 = 8$ Gyr as the solid line and $t_4 = 1$ Gyr as the dashed line. Our results use only massive clusters, which are significantly affected by dynamical evolution only after a long time. Therefore the two solutions for the star formation history differ only at large ages. Essentially, the shorter t_4 implies a stronger dynamical evolution of the clusters, which consequently had larger initial masses, leading to a higher derived star formation rate. The differences in the star formation histories for ages younger than ≈ 1 Gyr are only small. For larger ages the star formation rates are by a factor of ≈ 10 larger for the smaller t_4.

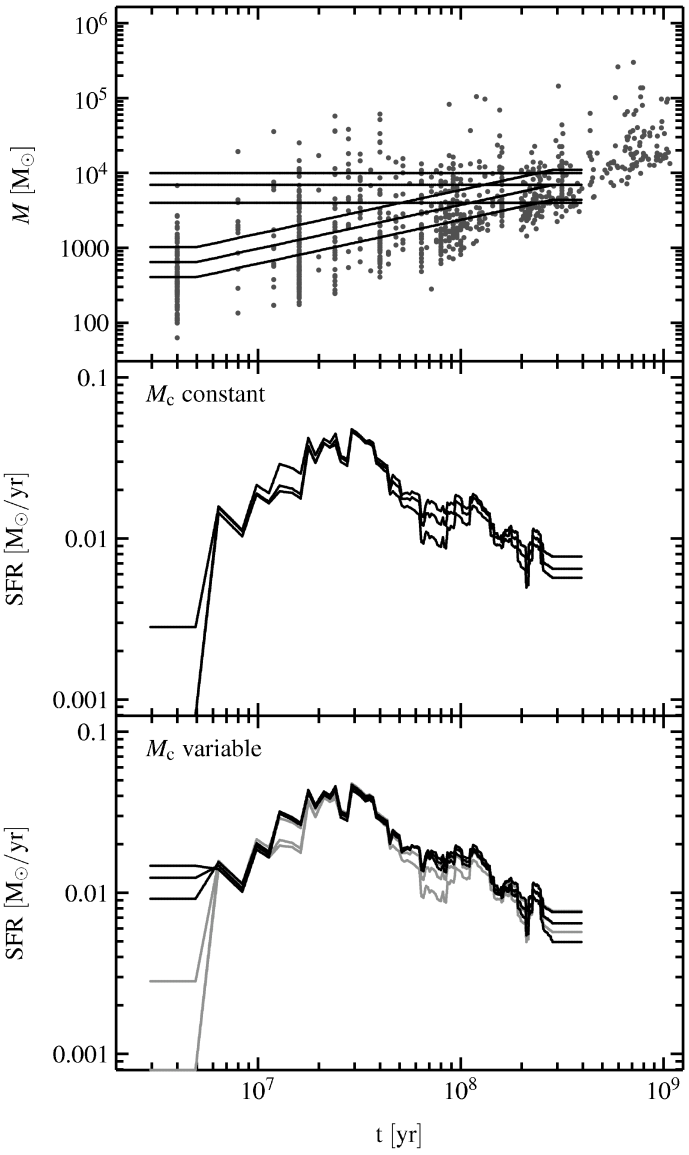


Figure 4.5: Recent history of star formation in star clusters, derived by adding up all cluster masses above a completeness mass, M_c and renormalised as described in the text (Cluster masses were dynamically back-evolved using t_4=8 Gyr). Various values for M_c have been chosen to demonstrate the dependence of the result on the completeness. The top panel shows the location of the constant and time-variable M_c in the cluster age-mass diagram. For the constant M_c the star formation histories are shown in the middle panel. The bottom panel contains the star formation histories for time variable M_c, also showing the results from the middle panel as grey lines for comparison.

4.2.3 Star Formation History using total mass in star clusters

As the Large Magellanic Cloud is very near to the Milky Way not only high-mass but also intermediate-mass clusters (with masses down to a few thousand M$_\odot$) can be detected over an extended time span. This allows us to use not only the most massive clusters to derive star formation rates, but also the whole cluster population. The fraction of star formation in star clusters in a given time interval is simply the ratio of the total mass of star clusters and the length of the interval,

$$\widetilde{\mathrm{SFR}}(t,M_\mathrm{c}) \quad = \quad \frac{1}{\Delta t} \sum_{\substack{M_i > M_\mathrm{c} \\ t(M_i)\in\Delta t}} M_i, \tag{4.4}$$

where M_c is the completeness mass which follows from the detection limit.

We derive the history of star formation in star clusters by moving a time interval of con-

stant logarithmic size (0.35 dex) in steps of 1 Myr until its boundary reaches an age of 400 Myr. For older ages the number of star clusters is too small to reach reasonable results. We use dynamically back-evolved cluster masses with t_4=8 Gyr.

Because the observations do not reach down to the lowest masses which star clusters can have, $\widetilde{\mathrm{SFR}}$ gives only a fraction of the total star formation rate in star clusters. Thus the mass in star clusters has to be extrapolated to a total mass of stars in star clusters by assuming a star cluster mass function, $\zeta \propto M^{-\beta}$, a power law parametrised by an exponent (β) and a lower and upper mass limit (M_{L} and M_{U}). ζ is here normalised as a probability density, i.e. $\int_{M_{\mathrm{L}}}^{M_{\mathrm{U}}} \zeta \mathrm{d}M = 1$. The normalisation factor follows from the observed fraction of star clusters,

$$\frac{M_{\mathrm{obs}}}{M_{\mathrm{tot}}} = \frac{\int_{M_{\mathrm{c}}}^{M_{\mathrm{U}}} M\zeta \mathrm{d}M}{\int_{M_{\mathrm{L}}}^{M_{\mathrm{U}}} M\zeta \mathrm{d}M}, \tag{4.5}$$

as

$$a(M_c) = \frac{\int_{M_{\mathrm{L}}}^{M_{\mathrm{U}}} M\zeta \mathrm{d}M}{\int_{M_{\mathrm{c}}}^{M_{\mathrm{U}}} M\zeta \mathrm{d}M}. \tag{4.6}$$

The correct star formation rate in clusters is then

$$\mathrm{SFR}(t, M_{\mathrm{c}}) = a(M_{\mathrm{c}})\widetilde{SFR}(t, M_{\mathrm{c}}). \tag{4.7}$$

The derived star formation history depends on the chosen completeness mass and parameters of the star cluster mass function.

In order to explore the robustness of the obtained results we turn first to the completeness mass, as it seems not to be too well constrained in our data set (see e.g. the discussion in Parmentier & de Grijs, 2008 and Maschberger & Kroupa, 2009, this thesis Chapter 5) To circumvent this problem we choose various possibilities for M_{c}, shown as lines in the top panel of Fig. 4.5, the age-mass diagram. For a minimal dependence on the parameters of the cluster mass function we choose M_c constant in time, with different values. Therewith one prohibits systematic effects in the shape of the star formation history caused by wrong parameters for the cluster mass function as the normalisation factor is constant for all age bins. The results are shown in the middle panel of Fig. 4.5, for a better comparison re-scaled such that the star formation histories lie near together (actually using in all cases $\beta = 2$, $M_{\mathrm{L}} = 100$ $\mathrm{M}_{\odot}$ and $M_{\mathrm{U}} = 10^6$ $\mathrm{M}_{\odot}$). The general structure is the same for all star formation histories, except for the very youngest ages because of a lack of massive clusters. Further, small variations appear around $\approx$ 15 Myr, $\approx$ 60–90 Myr and after $\approx$ 300 Myr. The first two small variations are caused by an insufficiently small number of clusters more massive than M_c. The discrepancy after 300 Myr is almost certainly caused by a too low mass for the lowest M_c, leading to an incomplete data set at these ages.

For an optimal use of the available data we choose a time-variable M_c, running parallel to the lower envelope of the star clusters in Fig. 4.5, top panel. Now each age bin has an individual normalisation constant, potentially introducing time-dependent systematics. The lower panel of Fig. 4.5 shows the obtained star formation histories with the results for the constant M_{c} (i.e. the results of the middle panel) plotted in grey for comparison. The overall structure of the star formation history is the same as for constant M_c, with the exception that the peak at 90–150-Myr is more like a plateau. The small variations at $\approx$ 15 Myr, $\approx$ 60–90

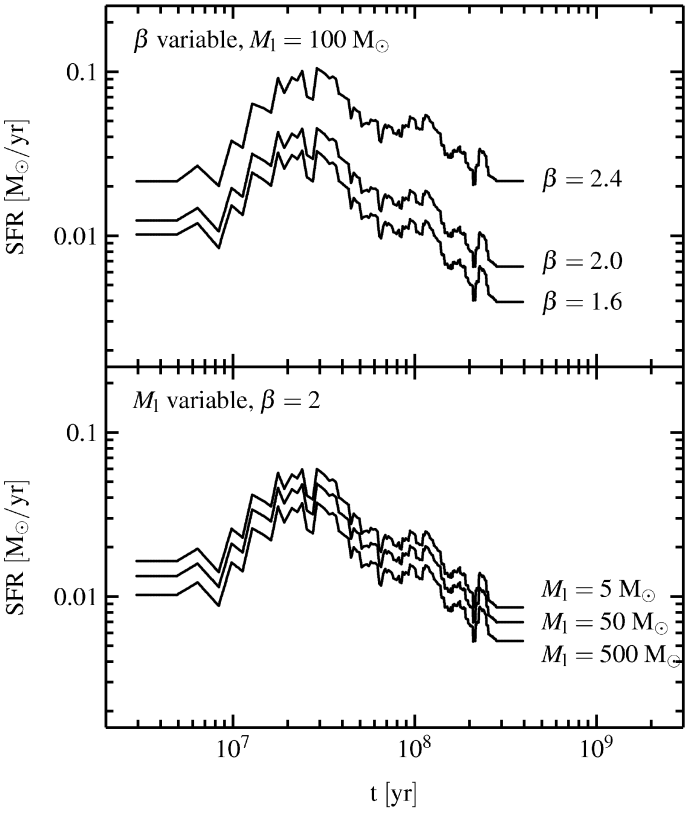


Figure 4.6: Influence of the parameters of the star cluster mass function on the normalisation of the star formation history, derived from the complete cluster population (dynamically back-evolved cluster masses with $t_4 = 8$ Gyr and time-variable completeness mass M_c, the middle in Fig. 4.5, top panel). The top panel shows the changes caused by different exponents β of the cluster mass function, especially on the older-age history. The choice of the lower limit $M_{\rm L}$ (bottom panel) has only minor influence on the absolute value of the star formation rates.

Myr disappear with the larger number of clusters used, but the feature at ≈ 300 Myr is still present. The agreement for the different choices of M_c is better than for constant M_c, which is rather surprising as the influence of the star cluster mass function is changing over time. This indicates that the results are robust and no systematical effects are introduced by the time-variable $M_{\rm c}$.

For the correction of $\widetilde{\rm SFR}$ we used $\beta = 2$, $M_{\rm L} = 100$ M$_\odot$ and $M_{\rm U} = 10^6$ M$_\odot$. These values, especially β, are chosen such that the different $M_{\rm c}$ all lead to the same result. As there are various values reported for β in the literature (e.g. Maschberger & Kroupa, 2009, this thesis Chapter 5, Gieles, 2009 and references therein and Weidner et al., 2004) we show in the top panel of Fig. 4.6 star formation histories corrected with different values of β (1.6, 2.0 and 2.4). $M_{\rm c}$ is variable in time, starting with 260 M$_\odot$(which is the second to lowest line in the top panel of Fig. 4.5). For larger β the fraction of star clusters below $M_{\rm c}$ increases, so that the star formation histories start at higher star formation rates. The increase of the star formation rates for different β is also time-dependent for time-variable $M_{\rm c}$, so that for older ages the amount by which the star formation rates are corrected increases. This leads to the growing difference between the curves in Fig. 4.6, top panel. The overall structure, however, remains the same within our range of β, and no additional features are introduced.

The lower limit of the cluster mass function only has minor influence on the absolute level of star formation histories, as evident in the lower panel of Fig. 4.6. Here we varied $M_{\rm L}$, using $M_{\rm L} = 5$ M$_\odot$, 50 M$_\odot$ and 500 M$_\odot$ ($\beta = 2.0$ and $M_{\rm U} = 10^6$ M$_\odot$). The star formation rates are a factor of 1.6 higher for $M_{\rm L} = 5$ M$_\odot$ compared to their values for $M_{\rm L} = 500$ M$_\odot$.

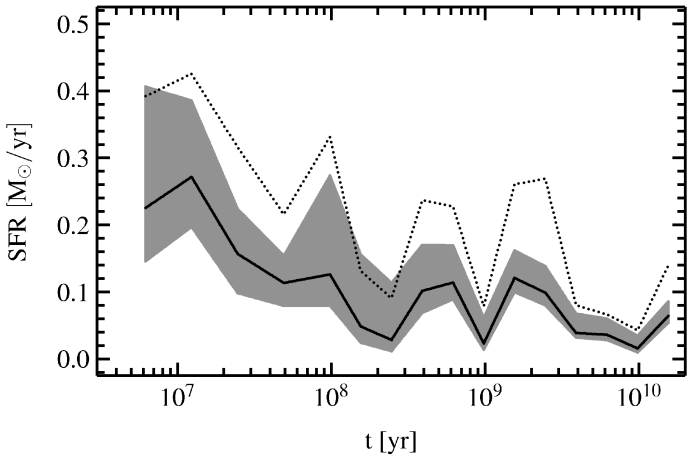


Figure 4.7: Star formation history derived by Harris & Zaritsky (2009) using the colour-magnitude diagram method. The solid line is the SFH for the regions with in the Massey fields only (open circles in Fig. 4.1), with the uncertainty given by the grey region. The dotted line above is for the whole of the Large Magellanic Cloud with the uncertainty.

4.3 The star formation history derived from colour-magnitude diagrams

For the comparison of our star cluster results with the results utilising colour-magnitude diagrams we use the work of Harris & Zaritsky (2009). They presented the star formation history derived with the StarFISH software (Harris & Zaritsky, 2001) in a grid of fields covering the whole Large Magellanic Cloud (the coverage is shown in Fig. 4.1). Their photometric catalogue contains 24 million objects, so that each of the individually analysed fields contains some 10^4 stars. For the synthetic colour magnitude diagrams the isochrones of the Padova group were used (Girardi et al., 2002). The temporal resolution of the star formation history is given by age bins of 0.3 dex size for ages younger than 100 Myr and bins of 0.2 dex for older ages. Solutions for the star formation history were obtained for four metallicities, $Z = 0.001$, $Z = 0.0025$ (interpolated and used only for ages larger than 100 Myr), $Z = 0.004$ and $Z = 0.008$.

As their photometry does not reach the main sequence turnoff point for the old population the extraction of the early star formation history was difficult. Therefore Harris & Zaritsky (2009) restricted StarFISH to fit only a single age bin covering all ages older than 4 Gyr in the bar region. Within the bar region they used the typical star formation history from solutions for the star formation history derived using HST data (Olsen, 1999, Holtzman et al., 1999 and Smecker-Hane et al., 2002, which widely agree with each other).

Figure 4.7 shows the star formation history, derived by Harris & Zaritsky (2009). The star formation history of the fields within the regions of Massey (2002) (open circles in Fig. 4.1) in which star clusters have been observed is shown as the solid line, with the grey area being its uncertainty. This partial star formation history follows the total star formation history (for the whole area of the Large Magellanic Cloud, dotted line) at about half the star formation rate.

4.4 Comparison of the Methods

In Figure 4.8 we summarise the solutions for star formation history of the Large Magellanic Cloud derived from colour-magnitude diagrams, the most massive star cluster or the total star cluster mass. As before in Fig. 4.7 the dashed line within the grey shaded area gives the star formation history with uncertainty derived from colour-magnitude diagrams within the Massey (2002) fields. In the following we refer to the peaks in this star formation history, at

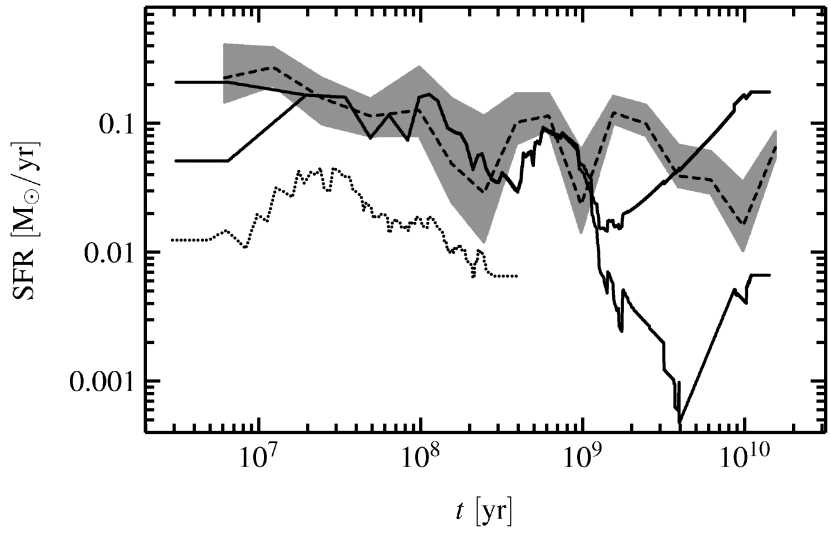


Figure 4.8: Comparison of the results for the star formation history in the Large Magellanic Cloud. The dashed line within the grey shaded region (its uncertainty) is derived from colour magnitude diagrams for the parts lying in the Massey (2002) fields (as in Fig. 4.7, Sec. 4.3). The thick solid lines are the solution using the most massive clusters, with the upper and lower limit for older ages, assuming no cluster formation or the maximum non detectable cluster mass in empty age bins (Fig. 4.3 bottom panel, Sec. 4.2.2). The lowest dotted line is derived from the total star cluster mass (time variable M_c, middle choice from top panel of Fig. 4.5, $M_l = 100$ M$_\odot$, $\beta = 2$, for details see Sec. 4.2.3). Cluster masses have been dynamically back-evolved using t_4 8 Gyr.

$\approx$ 15 Myr, 100 Myr, 400-600 (500) Myr, 2 Gyr and more than 10 Gyr ago.

The thick solid lines give the upper and lower limits for the star formation history derived from the most massive star clusters (Sec. 4.2.2). It branches at young ages into a solution including R136 (upper branch) and omitting R136 as in the original data set (lower branch), and again at ages > 1 Gyr, either assuming the fading limit as upper limit for the star formation rate (upper branch) or no star formation at all (lower branch). The agreement with the CMD solution is reasonable, especially if R136 is included, except for ages older than ≈ 1 Gyr, where the lack of star clusters suggests a significant decrease of the star formation rate. Although there are some clusters at 2 Gyr with $\approx 10^5$ M$_\odot$ (which lead to some kind of "peak" there), many more would be needed to produce a signal. It also seems to be odd that no clusters with lower masses are present, as would be expected for a normal sampling from the cluster mass function assuming the CMD star formation rates, and which would be detected because of the sufficiently low completeness limit. There is currently no plausible explanation for this difference, unless the Large Magellanic Cloud hat a very different mode of star formation at ages older than 1 Gyr, producing field stars without forming the corresponding clusters. That is, the cluster mass function would have been very bottom-heavy, which is not observed in any known galaxy. The peak at 10 Gyr is caused by the single very massive star cluster (10^6 M$_\odot$) and an increased star formation rate at this age is confirmed by the CMD results.

Compared to the CMD solution the most massive cluster solution shows an offset of the peaks in star formation. This is perhaps accounted for by the moving averaging window,

although different isochrone sets could also account for this. If the ages of the star clusters could be more accurately determined the moving window could be reduced, by at the same time including not only the most massive star cluster but also the second, third etc. most massive and thus keeping the sample of data points used large enough for the statistically necessary averaging. An investigation in this direction is beyond the scope of this work.

The absolute value for the star formation rate (up to 1 Gyr), derived from the most massive clusters, is at the level of the CMD star formation rate within the Massey (2002) fields. However, as the most massive cluster-method is intended to give the star formation rate *of an entire galaxy*, and the spatial coverage of the star clusters contains most of the area recently active in star formation, it is perhaps more appropriate to compare to the galaxy-wide star formation rate. In this case the star formation rate would be underestimated by a factor of 2 (compared to the dotted line in Fig. 4.7) and the normalisation in eq. 4.2 would need adjusting. As long as the spatial coverage is incomplete it is unfortunately impossible to disentangle inappropriate normalisation and effects of spatial incompleteness. In any case the absolute values and relative changes of the star formation rate derived from the most massive clusters are comparable, which is very promising and suggests further exploration of the star cluster method (cf. Maschberger & Kroupa, 2007, this thesis Chapter 3).

The star formation rate derived from the total mass in star clusters peaks at 10–40 Myr and again at 90–150 Myr, although the second peak is less clear. These peaks coincide with the peaks of the CMD solution. The fraction of star formation in star clusters, i.e. the ratio between the CMD curve and the dotted curve for all clusters, appears for the whole age range to be at a 10–20% level. If indeed all stars form within star clusters we would expect an increase by up to a factor of 10 in the star formation rate for the very youngest ages, as star clusters in the process of dissolving caused by gas expulsion could still be detected. Our result, however, does not show this feature, rather it shows the opposite. This is likely because very young clusters are in a class of objects not selected by Hunter et al. (2003). Therefore it is only possible to determine the fraction of star formation which leads to (presumably bound) open star clusters, which is $\approx$ 10–20 %. Assuming that indeed all stars form in a clustered way, this can also be stated that star clusters have an infant mortality of 80–90%.

4.5 Summary and Conclusions

We derived the star formation history of the Large Magellanic Cloud from its star cluster population, either following the method of Maschberger & Kroupa (2007, this thesis Chapter 3), using the most massive clusters only, or using the whole mass range of clusters (but then only for the most recent 400 Myr). We found that the results using the most massive clusters, both the absolute value for the star formation rate and the structure of the star formation history, agree well for the first Gyr with the star formation history derived from a colour-magnitude diagram. For older ages the number of detected star clusters is too small compared to expectations from the star formation rate following from CMDs. One possibility to resolve this discrepancy would be that additional clusters are contained in the area which is not observationally covered. Furthermore we derived the star formation history using all available star clusters, which, however, is only feasible for the most recent 400 Myr. As the star cluster data set is likely to be incomplete for young ages we can only state that the fraction of star formation in (presumably bound, open) star clusters after gas expulsion is at a 10–20% level. Alternatively, this means (assuming that all stars form in a clustered way), that star clusters

have an infant mortality of 80–90%.

Our results show that star clusters are a powerful means to investigate the star formation history of a galaxy and invite further investigation in that direction.

Acknowledgements

We wish to thank Richard de Grijs and Peter Anders for providing us with their data of the star clusters and Cathie Clarke for valuable comments on the manuscript. ThM acknowledges funding through CONSTELLATION*, an European Commission FP6 Marie Curie Research Training Network, and the Stellar Populations and Dynamics Research Group at the Argelander-Institut für Astronomie.*

Chapter 5

Estimators for the exponent and upper limit, and goodness-of-fit tests for (truncated) power-law distributions

Th. Maschberger & P. Kroupa
MNRAS 395:931–942 (2009)
A computer program for data analysis is available from
http://www.astro.uni-bonn.de/downloads

Many objects studied in astronomy follow a power law distribution function, for example the masses of stars or star clusters. A still used method by which such data is analysed is to generate a histogram and fit a straight line to it. The parameters obtained in this way can be severely biased, and the properties of the underlying distribution function, such as its shape or a possible upper limit, are difficult to extract. In this work we review techniques available in the literature and present newly developed (effectively) bias-free estimators for the exponent and the upper limit. Furthermore we discuss various graphical representations of the data and powerful goodness-of-fit tests to assess the validity of a power law for describing the distribution of data. As an example, we apply the presented methods to the data set of massive stars in R136 and the young star clusters in the Large Magellanic Cloud. For R136 we confirm the result of Koen (2006) of a truncated power law with a bias-free estimate for the exponent of 2.20 ± 0.78 / 2.87 ± 0.98 *(where the Salpeter-Massey value is* 2.35*) and for the upper limit of* $143 \pm 9\ \mathrm{M}_\odot$ / $163 \pm 9\ \mathrm{M}_\odot$*, depending on the stellar models used. The star clusters in the Large Magellanic Cloud (with ages up to* $10^{7.5}$ *yr) follow a truncated power law distribution with exponent* 1.62 ± 0.06 *and upper limit* $68 \pm 12 \times 10^3\ \mathrm{M}_\odot$*. Using the graphical data representation, a significant change in the form of the mass function below* $10^{2.5}\ \mathrm{M}_\odot$ *can be detected, which is likely caused by incompleteness in the data.*

5.1 Introduction

Many astronomical objects are distributed according to a power law. The probably most prominent example is the mass function of stars more massive than 0.5 $M_\odot$ with the Salpeter-Massey exponent of 2.35. Further examples are the mass functions of young star clusters and of molecular clouds. Modern observational techniques and state-of-the-art models provide data such as stellar masses with unprecedented accuracy. However, the statistical analysis of those data is not yet always optimal. The technique of binning the data suffers from losing a lot of information. The grouping of data into cells instead of using every data point obscures details of the observed distribution. This is an especially serious problem in the upper range, where the bins are only sparsely filled. Furthermore the obtained estimates of the slope can be severely biased (see e.g. Maíz Apellániz & Úbeda, 2005). A method based on a particular graphical display of the data which avoids grouping and allows one an estimate of the upper limit was given by Koen (2006). Another successful approach is to use the Maximum Likelihood method, which has been applied by Jauncey (1967) on extragalactic radio sources. Crawford et al. (1970) derived a Maximum Likelihood estimator for the exponent without grouping the data and including an upper limit.

A further step in data analysis, equally important as estimating the parameters, is the validation of the assumed power law form of the distribution. The simplest way to do this is to look at the histogram of the data in a double logarithmic plot. If this plot appears to be linear then the consistency of the data with a power law is concluded. But the significance of deviations from linearity are hard to state in an objective way by mere visual inspection. A further, more elaborate way is to apply a goodness-of-fit test such as the Kolmogorov-Smirnov test. If the calculated test statistic lies in some acceptance range then also consistency is concluded. But it is possible that the test statistic calculated with data stemming from an alternative hypothesised distribution similar to the power law might as well fall in the acceptance range. The test then fails to produce the right result since it has not enough "power" to discriminate. Therefore the "power" properties of a goodness-of-fit test have likewise to be examined. Such a study is – to our knowledge – not yet available in the astronomical literature.

In this work we describe estimation methods and compare their biases and variances (Section 5.3). Since the data may stem from a truncated power law we focus on estimators which can be used in this case. In the second part (Section 5.4) we investigate the question whether the data are consistent with the assumed power law distribution. As informal aids to answer this we discuss various plotting recipes (Section 5.4.1). For an objective decision we present goodness-of-fit tests with a study of their discrimination power (statistical power) under the hypotheses of a truncated and infinite power law. Finally, in Section 5.5, we will apply the introduced methods on the massive stars in R136 and the young star clusters in the Large Magellanic Cloud.

5.2 General results, definitions and notation

5.2.1 The power law distribution

In this work we only consider power law distribution functions (DF) with a negative exponent, $-\alpha$ ($\alpha > 1$). By convention the sign is separated from the absolute value. Besides the exponent such a distribution is further parametrised by the lower and upper limit, x_L and x_U.

The probability density is then given by

$$p(x;\alpha,x_{\rm L},x_{\rm U}) \;=\; \frac{1-\alpha}{x_{\rm U}^{1-\alpha}-x_{\rm L}^{1-\alpha}}x^{-\alpha}, \tag{5.1}$$

and the cumulative distribution function (DF) is

$$P(x) \;=\; \frac{x^{1-\alpha}-x_{\rm L}^{1-\alpha}}{x_{\rm U}^{1-\alpha}-x_{\rm L}^{1-\alpha}}. \tag{5.2}$$

The family of distributions given by eq. 5.1 includes the "infinite" or "not truncated" distributions with infinite upper limit, $p(x;x_{\rm U}=\infty):=p_\infty(x)$, which is also known in the (non-astronomical) literature as the Pareto distribution. The density function reads then

$$p_\infty(x) \;=\; -\frac{1-\alpha}{x_{\rm L}^{1-\alpha}}x^{-\alpha}, \tag{5.3}$$

and the cumulative distribution is

$$P_\infty(x) \;=\; 1-\left(\frac{x}{x_{\rm L}}\right)^{1-\alpha}. \tag{5.4}$$

An useful property of the power law distribution is its relation to the exponential distribution. By a logarithmic transformation the power law distribution becomes proportional to an exponential, $x^{-\alpha}=e^{-\alpha\log_e x}$. Due to this proportionality it is possible that some techniques for estimation and testing, which were developed for the exponential distribution, can be used for the power law distribution.

5.3 Estimating the parameters

There exist in the literature a variety of methods to estimate the parameters, exponent and upper limit, of a power law distribution. In this first part of the paper we describe them and compare their properties.

5.3.1 Methods

Binning

A particularly simple method is to fit a linear relation to the data grouped in bins of constant size in logarithmic space. As shown by Maíz Apellániz & Úbeda (2005) this method can yield biased results, i.e. results which systematically deviate from the actual parameter, and do not allow one to estimate a possible upper limit.

A solution to avoid biased results was given by Maíz Apellániz & Úbeda (2005), which modified the binning scheme to a constant number of data points per bin and fitted the expected number instead of performing a linear regression. In this way the estimate for the exponent, $\widehat{\alpha}$, can be obtained, together with an estimate of its uncertainty, which is consistent with the sampling variance of $\widehat{\alpha}$. In extension of their work we investigate the properties of

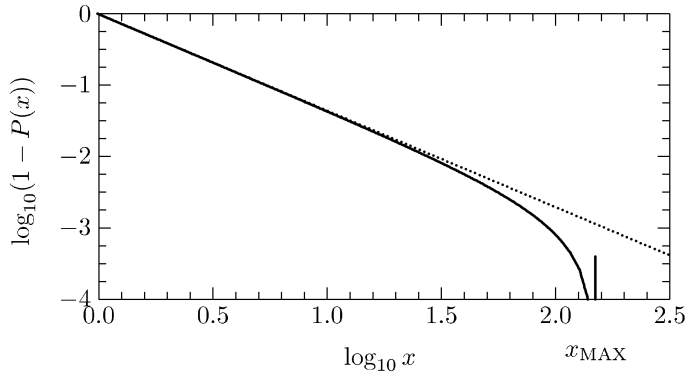


Figure 5.1: Complementary cumulative DF (CCDF) plot for an infinite (dotted line) and truncated (solid line) power-law pdf ($\alpha = 2.35$, $x_{\mathrm{L}} = 1$, $x_{\mathrm{U}} = 150$, shown by the vertical thick bar). For the truncated case a characteristic turn-down appears at the upper end.

the estimate for the upper limit, derived from the normalisation constant of the frequency distribution, $\widehat{k} = n(1-\alpha)/(x_{\mathrm{MAX}}^{1-\alpha} - x_{\mathrm{MIN}}^{1-\alpha})$, which is given by

$$\widehat{x}_{\mathrm{U}} \quad = \quad \left(n\frac{1-\widehat{\alpha}}{\widehat{k}} + X_{(1)}^{1-\widehat{\alpha}} \right)^{\frac{1}{1-\widehat{\alpha}}}, \tag{5.5}$$

where the smallest observation ($X_{(1)}$) is used as an estimate for x_{L}.

Not only the choice of constant-size or variable-size bins has influence on the results of binning, but also the number of bins. D'Agostino & Stephens (1986) give as the optimal number of bins $2n^{2/5}$ for n data points. A smaller number of bins reduces the bias but increases the standard deviation (cf. Table 1 to 3 of Maíz Apellániz & Úbeda, 2005).

Complementary cumulative distribution function plot

Koen (2006) presented a method to estimate both the exponent and the limits of a power law. This method is based on a particular graphical representation of the data, the complementary cumulative DF (CCDF) plot (Fig. 5.1). Data stemming from an infinite power law follow a linear relation with slope $1-\alpha$ in a plot of $\log(1-P_\infty(x))$ versus $\log x$, as can be seen easily by taking the log of eq. 5.4. For a truncated power law, a turn-down appears at the high end. Estimates for the exponent and limits are obtained by fitting $\log(1-P(X_{(i)}))$ (with $P(x)$ from eq. 5.2 and the ordered data $X_{(i)}$) to $\log(1-\frac{i-0.5}{n})$, with using $\frac{i-0.5}{n}$ for the empirical cumulative distribution function.

Beg's estimator

A power-law distribution is closely related to the exponential distribution. Therefore it is possible to apply the uniformly minimum variance unbiased estimators for the slope and limits of a truncated exponential distribution, developed by Beg (1982, 1983) to log-transformed power law data, as shown by Beg (1983). Although these estimators are theoretically an optimal solution, they are only partially practicable, since their computation is numerically difficult and impossible for large data sets. We therefore developed a recursive form which is applicable to arbitrarily large data sets. The original and recursive formulae are given in the Appendix.

Maximum Likelihood (ML) estimator

The Maximum Likelihood estimator for the exponent was given by Crawford et al. (1970), who also included an upper limit. In our case the upper limit is intrinsic to the distribution

function, i.e. the upper limit needs to be estimated simultaneously with the exponent. The likelihood function for a random sample of size n from a truncated power law DF is

$$\mathscr{L} = \prod_{i=1}^{n} p(x_i;\alpha;x_{\mathrm{L}},x_{\mathrm{U}}) \tag{5.6}$$

$$= \left(\frac{1-\alpha}{x_{\mathrm{U}}^{1-\alpha}-x_{\mathrm{L}}^{1-\alpha}}\right)^n \prod_{i=1}^{n} x_i^{-\alpha} \tag{5.7}$$

The estimator for the exponent is obtained by maximising the log-likelihood,

$$\log\mathscr{L} = n\log(1-\alpha)-n\log(x_{\mathrm{U}}^{1-\alpha}-x_{\mathrm{L}}^{1-\alpha})-\alpha\sum_{i=1}^{n}\log x_i. \tag{5.8}$$

The maximisation can be performed by finding the root of the derivative with respect to α of eq. 5.8. The estimator $\widehat{\alpha}_{ML}$ is then the solution of

$$-\frac{n}{1-\widehat{\alpha}_{ML}}+n\frac{Z^{1-\widehat{\alpha}_{ML}}\log Z-Y^{1-\widehat{\alpha}_{ML}}\log Y}{Z^{1-\widehat{\alpha}_{ML}}-Y^{1-\widehat{\alpha}_{ML}}}-T = 0, \tag{5.9}$$

with $Y=\min X_i$, $Z=\max X_i$ and $T=\sum_{i=1}^{n}\log X_i$.

The ML estimates for the upper limit $\widehat{X}_{\mathrm{U}}=\max X_i$ (see e.g. Aban et al., 2006). It is obvious that this estimate will be biased since the upper limit is larger than the largest data point.

Bias-free estimators based on the maximum likelihood estimator

It is possible to construct a minimum variance unbiased estimate of the exponent from the maximum likelihood estimate, as shown by Crawford et al. (1970) or Baxter (1980). For the infinite case the ML estimator for the exponent is given by

$$\widehat{\alpha}-1 = \frac{n}{T-n\log_e Y}, \tag{5.10}$$

with $Y=\min X_i$ or the given lower limit, and $T=\sum_{i=1}^{n}\log_e X_i$. The unbiased estimator is then

$$\widehat{\alpha}'-1 = \frac{n-1}{n}(\widehat{\alpha}-1) \tag{5.11}$$

(if both the exponent and the lower limit should be estimated then $(n-2)/n$ has to be used (Baxter, 1980)).

The simple relation between the ML estimator and the unbiased estimator for an infinite power law suggests a similar relation for the truncated case. However, for a truncated power law a closed form of the ML estimator is not available. This makes a proof of an unbiased estimator very difficult and maybe even impossible, since the distribution of the estimate cannot be calculated analytically. Nevertheless, it is not unreasonable to assume that a simple modification of the ML estimate also leads to unbiased results. A different pre-factor, depending only on the number of data, should give the expected result. We found that

$$\widehat{\alpha}_{\mathrm{MML}}-1 = \frac{n}{n-2}(\widehat{\alpha}_{ML}-1). \tag{5.12}$$

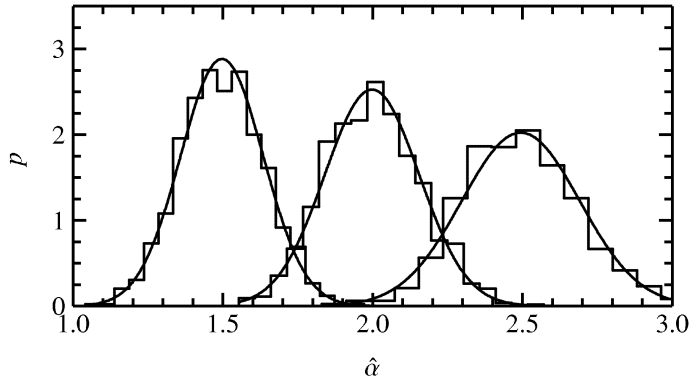


Figure 5.2: Distributions of the modified ML estimates for the exponent (histogram), derived from Monte-Carlo samples of size 1000 for three input values (1.5, 2.0, and 2.5). They follow a Gaussian with mean and variance derived from the samples. For larger exponents the variance increases.

(MML = Modified Maximum Likelihood) provides quasi-bias-free estimates. The pre-factors $\frac{n-2}{n}$ of Baxter (1980) or $\frac{n-3}{n}$ ($n-3$ because there is an additional parameter, the upper limit) lead to biased results. The distribution of the exponents estimated using this method follows a Gaussian, as can be seen in Fig. 5.2, with an increasing variance for an increasing exponent.

The bias of the ML estimate for the upper limit can also be significantly reduced by appropriate modifications. Hannon & Dahiya (1999) developed such a modified estimator for the exponential distribution. This estimator can also be used for the power law distribution and takes then the form (with the ML estimate of the exponent replaced with the bias-reduced form)

$$\widehat{x}_{\rm U} \;=\; X_{(n)} \left(1 + \frac{e^G - 1}{n}\right)^{\frac{1}{1-\widehat{\alpha}_{\rm MML}}}, \tag{5.13}$$

with

$$G \;=\; (1-\widehat{\alpha}_{\rm MML}) \log_e \left(\frac{X_{(n)}}{X_{(1)}}\right), \tag{5.14}$$

where $X_{(1)}$ is the smallest and $X_{(n)}$ is the largest data point. The properties of the modified estimate are discussed in the next Section.

5.3.2 Performance of the estimators

After introducing a number of methods of estimation, we compare their properties. The quality and usability of an estimator is determined by several factors. A main demand is that an estimate is on average equal to the actual parameter, i.e. bias-free. Also, the variance of the estimate should be as small as possible and it should be numerically robust.

To study these properties we carried out a set of Monte-Carlo experiments, each of size 1000, with parameters in the typical range of astronomical applications. The values for the exponent range from 1.6 to 2.85 in steps of 0.25. For each exponent four pairs of limits were used ($\{0.5, 150\}$ and $\{10, 150\}$ corresponding to the stellar mass function, $\{10^3, 10^5\}$ and $\{10^4, 10^6\}$ corresponding to the mass function of young star clusters). The last varied parameter was the number of data (50, 100, 300). For the binning methods the number of bins was chosen according to D'Agostino & Stephens (1986) ($2n^{2/5}$), which gave 9, 12 and 19 bins, respectively.

As diagnostics for the performance in estimating the exponent we choose the average bias

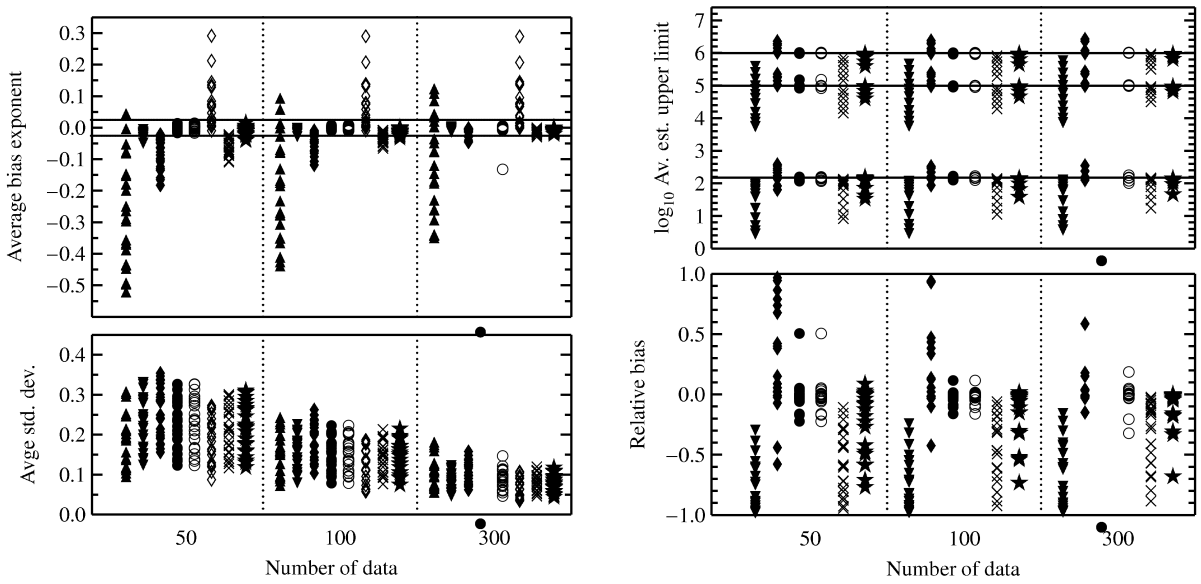


Figure 5.3: In the left panel the results for the exponent are shown, on top the average bias (calculated using eq. 5.15, the horizontal lines mark ± 0.025) and below the average standard deviation (eq. 5.16) of the estimated exponents. In the right panel, average estimates of the upper limit (lines mark the true values) are displayed in the upper part and below the average relative bias is shown (eq. 5.17). The parameter combinations are given in the text, Sec. 5.3.2. The symbols refer to: ▲ constant size binning using linear regression; ▼ variable size binning using χ^2; ♦ CCDF plot fitting. ◊ ML estimator without including truncation; • Beg's estimator; (This estimator starts to fail for $n \geq 150$, dots below the x-axis indicate a failed experiment); ○ Beg's estimator in the recursive form; × ML estimator; ★ Modified ML estimator;

of an estimator for a given parameter set,

$$\overline{B}(\alpha) \quad = \quad \frac{1}{1000}\sum_{i=1}^{1000}\left(\widehat{\alpha}_i - \alpha\right), \tag{5.15}$$

and the standard deviation,

$$\overline{S}(\alpha) \quad = \quad \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}\left(\widehat{\alpha}_i - \overline{\alpha}\right)^2}. \tag{5.16}$$

The left panel of Figure 5.3 summarises the results for the bias. Two horizontal lines at ± 0.025 embrace the region, in which we consider the bias as negligible. The general trend is that the bias decreases with an increasing number of data (except for the ML estimator wich does not include a truncation). The corresponding results for the standard deviation also decrease with larger size of the data set.

Variable-size binning gives effectively bias-free exponents for samples having a moderate size or larger. The method of Koen is biased towards lower exponents. The results from Beg's estimator are very good, but the method fails for large data sets, even in the recursive form. A maximum likelihood estimate without considering the truncation can lead to a significantly overestimated exponent. But when the truncation is considered, the bias is small and effectively vanishes if our modified version is used.

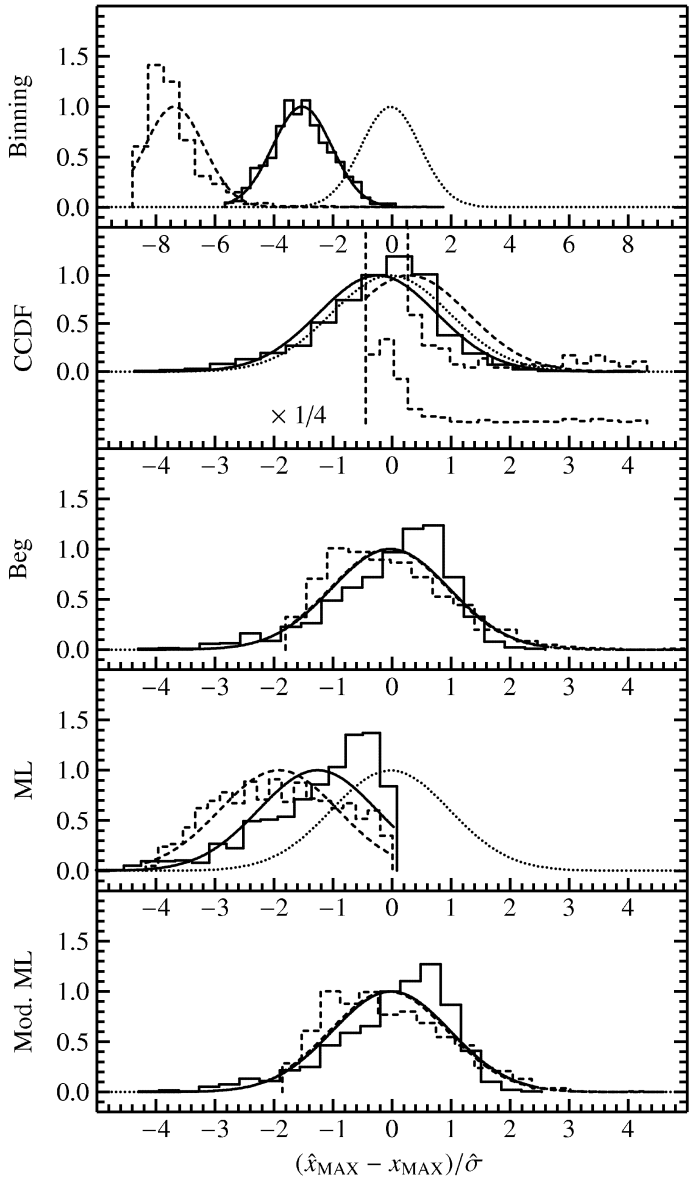


Figure 5.4: Distribution of the estimated upper limits for the different methods (histograms; parameters: $\alpha = 2.0$; $n = 100$; solid: limits $\{10, 150\}$, dashed: limits $\{1000, 100000\}$.) The x-axis has been scaled such that a unbiased estimate should follow a Gaussian of zero mean and unit variance (dotted). Also shown is a Gaussian with mean and variance derived from the estimates.

The results for the estimates of the upper limit are shown in the right part of Fig. 5.3. Generally it can be observed that a larger upper limit leads to larger absolute deviations in the estimate. Because the upper limits used in this study span a wide range of values it is not convenient to compare the absolute biases as for the exponents. The relative bias (also displayed in Fig. 5.3),

$$\widetilde{B}(x_{\mathrm{U}}) \quad = \quad \frac{1}{1000} \sum_{i=1}^{1000} \left(\frac{\widehat{x}_{\mathrm{U},i} - x_{\mathrm{U}}}{x_{\mathrm{U}}} \right), \tag{5.17}$$

is a better measure of trends. Furthermore the normalised distributions of the estimates are shown in Fig. 5.4, for two parameter sets ($\alpha = 2.0$, $n = 100$, limits $\{10, 150\}$ and $\{1000, 100000\}$). The histogram and a Gaussian with mean and variance (σ) calculated from the Monte-Carlo sample are rescaled by $x' = \frac{x - x_{\mathrm{U}}}{\sigma}$ and the y-axis is scaled such that the peak of the Gaussian is

1. If the estimator is not biased and can be approximated by a Gaussian, then the normalised distribution should follow a Gaussian with zero mean and unit variance.

The upper limit is underestimated by using the normalisation constant of the variable-size binning method. The results of fitting the CCDF plot are peaked around the input value, but can have a long tail of very high estimates (for the limits $\{1000, 100000\}$). If in the CCDF plot the data show no strong curvature at the upper end, then the estimated upper limit is very large. The distribution of estimates obtained with Beg's estimator are in reasonable agreement with a Gaussian, but not completely symmetric around the mean. If the largest data point is used (i.e. the direct ML estimate), then the upper limit is underestimated, with a distribution limited by the actual value. With the modification (eq. 5.13) the distribution becomes similar to the one of Beg's estimator, spreading around the true value. Although not completely symmetric around the mean it can be sufficiently approximated by a Gaussian, and has no outliers as fitting the CCDF plot.

In summary, when both the exponent and the upper limit of a truncated power law should be estimated, our modified ML method performs best in terms of bias and stability, being similar to Beg's uniformly minimum variance unbiased estimator, but without the numerical instability. gives the best results in terms of bias and stability.

5.4 Is a power law consistent with the data?

For a thorough analysis of data which are assumed to stem from a power-law distribution it is not sufficient just to estimate the parameters. The parameter estimation answers the question which power-law fits the data best, but leaves open whether the data are originating from a power law *at all*. Or to put it differently: is the (truncated) power law a good parent distribution function of the data? The need to answer this question has already been stressed by Crawford et al. (1970). This question can be addressed by a graphical inspection of the data, which is discussed in the next Section. After the informal visual methods more objective goodness-of-fit techniques are discussed.

5.4.1 Graphical inspection of the data

A common approach to find the parent DF of a data set is to use a histogram as a non-parametric estimate of the form of the parent DF. If in a logarithmic plot the histogram of e.g. stellar masses is a straight line, a power-law is usually assumed as the parent DF. However, a power-law is a heavy-tailed distribution and has only a few counts per bin in the tail. Thus the scatter in a histogram is large in the upper regime and makes deviations from a power-law hard to detect. Alternative, heavy-tailed distributions lead to nearly indistinguishable histograms. It is for example not possible to decide whether the power-law is truncated or not. Therefore a histogram only allows us to roughly determine the parent DF.

A display of the data which avoids grouping them into cells is the probability-probability (or percentile-percenile, PP) plot. For the PP plot the data have first to be sorted in ascending order, $X_{(i)} < X_{(i+1)}$. The x-values then follow as the "theoretical" percentiles, which are the values of the cumulative DF for the i-th data point, $P(X_{(i)}, \widehat{\alpha}, \widehat{x}_{\rm L}, \widehat{x}_{\rm U})$, calculated with the estimated parameters. As y-values the "empirical" percentiles are used, given by $\frac{i-0.5}{n}$. Both axes range from 0 to 1, and when the data lie on the diagonal then they agree with the null-hypothesis. Hypotheses other than the null hypothesis can be shown by plotting the pairs

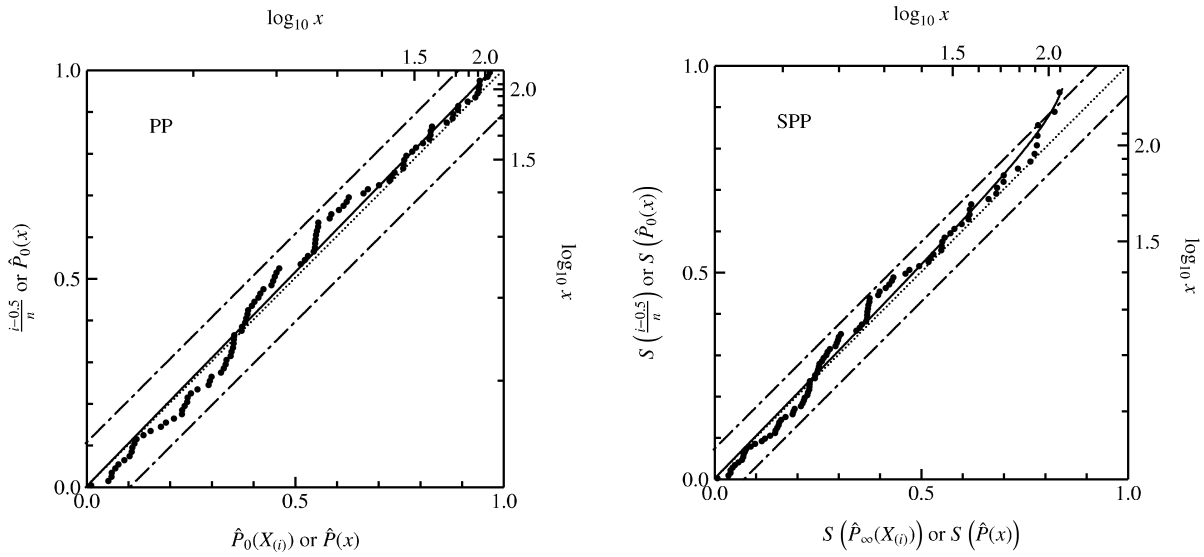


Figure 5.5: Example for a percentile-percentile plot (PP, left) and a stabilised percentile-percentile plot (SPP, right) for the null hypothesis of an infinite power law (=diagonal), using 100 data points sampled from a truncated power law ($\alpha = 2.35$, $x_\mathrm{L} = 1$ and $x_\mathrm{U} = 150$). Also shown are the curves for a truncated power law (solid line, parameters as estimated, $\widehat{\alpha} = 2.35$ and $\widehat{x}_\mathrm{U} = 149$). The acceptance region of the (in the right plot stabilised) Kolmogorov-Smirnov statistic (significance level 5%) is given by the two parallels to the diagonal. The data lie within this region in the PP plot, wherefore from the PP plot the infinite power law cannot be significantly rejected. After stabilisation the KS test is more powerful and thus allows us to detect truncation in contrast to the PP plot.

{x=alternative cumulative DF,y=null cumulative DF}. An example for a PP plot with an infinite power law as null hypothesis (diagonal) is shown in Fig. 5.5, but with data generated from a truncated power law. The curve for the alternative hypothesis of a truncated power law (of the same exponent, solid line) barely deviates from the diagonal and does not allows one to distinguish the infinite and truncated versions. In this plot the acceptance region of the Kolmogorov-Smirnov (KS) test can be directly shown as parallels to the diagonal, in Fig. 5.5 calculated with a significance level of 5%. The data do not exceed this region, not even in the tails, giving evidence for the known insensitivity in the tails of the Kolmogorov-Smirnov statistic.

Before showing a way to improve the insensitivity of the PP plot in the tail we shortly compare it with other possible plots which show all available data (see e.g. Chambers et al. (1983) or Wilk & Gnanadesikan (1968)). A plot which uses e.g. the $X_{(i)}$ as x-values (the "empirical cumulative density plot") has a curved reference line for both an infinite and truncated power law, with only small and not well perceptible differences between them. Following the suggestion of Koen (2006), in a plot of $\log(1 - P(X))$ against $\log X$ (the CCDF plot, see Fig. 5.1 and Sec. 5.3.1) the data should only show a curvature for a truncated power law. But from the results for the estimator based on this plot, the scatter in the highest data points can be large, and a graphical goodness-of-fit criterion would not be very sensitive. A third alternative to the PP plot would be a plot of the inverse cumulative DF ($P^{-1}(\frac{i-0.5}{n})$, giving the expected value for the data point $X_{(i)}$) against the ordered data $X_{(i)}$ ("quantile-quantile" plot). This plot

has a linear reference, the upper tail would be curved when constructed for the infinite null hypothesis but with truncated data. Again, as for the complementary cumulative density plot there is large scatter, and additionally the inverse cumulative DF has to be calculated.

5.4.2 The stabilising transformation and the SPP plot

Goodness-of-fit methods based on the empirical cumulative DF, such as the PP plot or the KS test, have the advantage that their intrinsic properties do not depend on the actual choice of the null hypothesis. A PP-plot for e.g. Gaussian variates looks identical (modulo the random scatter) to a PP plot for power-law data. The reason for this is that by taking the cumulative DF the data are transformed to uniformly distributed variates, if they are following the null hypothesis. Therefore the location and variances of the points in the PP plot are independent of the null hypothesis distribution, as is the distribution of the KS statistic. This transformation reduces the goodness-of-fit task from arbitrary DFs to testing for uniformity. However, the variances of uniform ordered variates are not independent from their position: the scatter in the PP plot is larger in the middle than in the tails. Thus a test which measures the differences between the expectation and the empirical values of uniformly distributed ordered data will be dominated by the points with the larger variances. Hence the insensitivity of the KS-test in the tails.

A way to overcome the unequal variances was introduced by Michael (1983). The stabilising transformation of uniform variates u (= the cumulative DF),

$$S_0(u) \quad = \quad \frac{2}{\pi} \arcsin(\sqrt{u}) \tag{5.18}$$

gives asymptotically equal variances of the transformed ordered variates. In a stabilised PP (SPP) plot every part of the plot has the same weight and no region is particularly emphasised. Although the distribution of the $S_0(u)$ is not uniform any more, tests based on the differences between expectation and empirical value can still be used (as long as they do not use other properties of the uniform distribution). These transformed tests are equally sensitive to every part of the distribution function.

However, for testing the tail-behaviour of a DF it is useful to emphasise the tail. This can be achieved by using only a half transformation, which is possible because S_0 is symmetric around the point $\{0.5, 0.5\}$ and the interval $[0.5, 1]$ is mapped onto $[0.5, 1]$. A one-sided transformation of the percentiles to stabilise a right-tailed distribution consists then of three steps. First, map the interval $[0, 1]$ on $[0.5, 1]$, then use S_0, and lastly map $[0.5, 1]$ back on $[0, 1]$. The formula for this is

$$S(u) \quad = \quad 2S_0(0.5 + 0.5u) - 1. \tag{5.19}$$

We use S instead of S_0 in the SPP plot (Fig. 5.5) and in the goodness-of-fit tests which are related to it, because of the one-tailed power law distribution.

5.4.3 Goodness-of-fit Tests

A goodness-of-fit test provides an objective way to "measure" the agreement of the fit with the data. We follow here the Neyman-Pearson ansatz of hypothesis tests. At first the type I error probability or significance level needs to be specified. This is the rate at which the

test is allowed to falsely declare a data set as too discrepant to be compatible with the null hypothesis (the assumed distribution function), even though it is in reality consistent. There is a value of the distribution of the test statistic, the critical value, which corresponds to this rate. If the value of the test statistic calculated from the data set then exceeds the critical value the null hypothesis is rejected for the data set.

For some tests, the distribution of the test statistic can be calculated analytically for a fully specified null hypothesis, i.e. if no parameter is estimated. If parameters are estimated, the distribution of the test statistic is not universal any more, but depends on the properties of the specific estimator. The distribution of the test statistic can then be obtained using a Monte-Carlo approach, of which follow the critical values. Typically the such derived critical values are larger than for a fully specified hypothesis (cf. Lilliefors, 1967, 1969, for the Kolmogorov-Smirnov test for the normal and exponential distribution). Therefore, if the critical values for the fully specified hypothesis are used when parameters are estimated, the results are conservative with an actually smaller type I error, but also less powerful.

The significance level is not the only quantity characterising a statistical test. It can happen that a data set is not too discrepant to be rejected, but actually does not stem from the null hypothesis, i.e. a type II error occurs. The probability that a type II error does not occur is the (statistical) power of the test. If the power of the test is small then it is not very selective and the alternative hypothesis cannot be strongly excluded. For a given test the power can differ for various alternative hypotheses of the distribution. A demand for a general purpose test is to be powerful against a wide variety of alternative hypotheses.

In order to be able to evaluate the "strength" of a statement concluded from a statistical test, it is therefore necessary to know the type I and type II error rates (the significance level and the power). However, not every test has the same power, therefore we conduct in what follows a power study which has the purpose of finding a powerful goodness-of-fit test to decide between infinite and truncated power laws. The astrophysical motivation for the choice of these hypotheses is the discussion in the literature about an upper mass limit for the distribution of stars in a star cluster (cf. Weidner & Kroupa, 2004; Oey & Clarke, 2005; Koen, 2006) or about an upper limit for the star cluster luminosity function (cf. Gieles et al., 2006b).

5.4.4 Description of the goodness-of-fit test statistics

Goodness-of-fit tests can roughly be classified as tests based on the empirical DF (EDF), based on distance measures (e.g. the KS test) or the correlation coefficient, tests especially developed for a chosen null hypothesis (e.g. the Shapiro-Wilk test for exponentiality), and tests to distinguish between two hypotheses (e.g. the Likelihood Ratio). Below we describe the test statistics used for a comparison.

For the selection of the tests included in the comparison the properties of tests for exponentiality can be used, which can be found in the studies of Stephens (1978), D'Agostino & Stephens (1986) and Gan & Koehler (1990). Gan & Koehler (1990) which use EDF based tests included the alternative hypothesis of a truncated exponential, finding only very low powers. With the stabilising transformation Kimber (1985) finds a larger power for the KS test, but did not include the truncated alternative. We include some EDF based tests in the original and stabilised version, as well as some tests based on tests for exponentiality, and two tests explicitly for truncation.

n	β	D	SD	C^2	SC^2	A^2	r^2	k^2	k_0^2	Sk^2	Sk_0^2	W	T	Λ	X
10. 150.															
33	1.7	54.1	72.0	59.9	69.4	60.4	51.7	7.7	59.9	17.4	65.9	47.4	56.1	64.6	100.0
50	1.7	70.9	88.3	75.0	86.0	78.7	69.9	5.1	75.3	16.3	87.3	68.7	78.6	86.3	100.0
99	1.7	93.8	100.0	96.8	99.8	98.6	98.2	3.7	96.3	40.3	99.9	96.7	98.6	99.9	100.0
33	2.0	23.3	29.6	24.4	28.2	26.0	22.9	4.5	26.2	7.0	28.2	22.2	23.9	29.8	53.4
50	2.0	25.0	58.2	32.8	46.3	36.2	31.3	3.7	32.4	7.8	50.4	38.8	43.5	61.0	100.0
99	2.0	39.1	93.7	50.7	80.4	63.7	58.7	3.1	50.3	11.5	85.2	68.9	74.8	94.9	100.0
33	2.3	7.8	11.5	8.9	10.5	9.2	5.9	5.4	9.7	5.5	10.5	10.7	11.7	11.3	11.4
50	2.3	8.5	18.3	8.8	14.1	9.3	4.0	2.9	9.2	3.7	15.0	14.2	15.0	18.9	22.3
99	2.3	11.9	48.9	11.5	30.4	15.6	4.8	4.2	14.1	4.9	31.1	29.3	33.4	53.5	80.4
10000. 1000000.															
33	1.7	14.3	18.5	15.3	19.1	14.5	18.8	6.6	15.3	7.5	19.3	20.0	17.9	17.6	19.5
50	1.7	12.7	33.3	14.9	25.6	17.8	28.3	5.2	18.4	6.5	26.5	23.8	25.6	31.3	50.9
99	1.7	19.9	77.4	29.1	57.3	40.1	60.1	4.2	30.9	8.0	61.8	52.1	58.0	76.9	100.0
33	2.0	5.3	5.7	4.8	5.6	5.8	5.9	6.5	4.8	5.1	5.9	7.0	6.8	5.4	6.2
50	2.0	6.8	8.4	7.6	7.8	7.2	5.5	5.7	8.3	4.1	7.3	8.2	8.8	8.6	8.7
99	2.0	3.4	14.5	3.4	6.2	3.9	5.3	3.9	3.7	5.0	9.8	12.0	11.7	16.6	12.6
33	2.3	5.4	4.9	4.9	4.4	4.5	3.3	4.7	4.1	4.5	4.2	4.4	4.6	4.9	4.9
50	2.3	4.2	6.0	5.2	5.9	5.3	2.2	3.4	4.4	3.9	5.5	4.6	4.9	6.5	5.5
99	2.3	5.8	7.0	6.8	6.9	7.2	1.1	6.5	6.0	6.0	6.9	8.6	8.6	7.8	6.8

Table 5.1: Results of the power study. Tests are conducted under the null hypothesis of an infinite power law against the alternative hypothesis of a truncated power law with a type I error level of $\alpha_I = 0.05$. The first column gives the size of the data set, n, and the second column the value of the exponent, α. The other columns give the power of the test statistic indicated in the top row in percent. The power is the fraction of the time in which the data drawn from the alternative hypothesis would be rejected (at the 5% level) as coming from the null hypothesis. The numbers on top of each group are the lower and upper limit of the parent distribution function.

EDF statistics based on distance measures

The most prominent goodness-of-fit test is the *Kolmogorov-Smirnov (KS) statistic* (Stephens, 1978; D'Agostino & Stephens, 1986; Gan & Koehler, 1990),

$$D = \max_{1\le i\le n}\left|\frac{i-0.5}{n} - \widehat{P}_{(i)}\right| + \frac{1}{2n}, \tag{5.20}$$

which is the largest vertical distance between the data and the diagonal in the PP plot (Rejection for $D > D_{\rm crit}$). The largest distance can also be measured in the stabilised PP plot, leading to the *stabilized Kolmogorov-Smirnov statistic* (Michael, 1983; Kimber, 1985),

$$SD = \max_{1\le i\le n}\left|S\left(\frac{i-0.5}{n}\right) - S(\widehat{P}_{(i)})\right|, \tag{5.21}$$

Kimber (1985) found that for the exponential distribution these statistics are more powerful than the originals, but used a complete stabilising transformation (for both tails, S_0, eq. 5.18).

Here only a right-tail-stabilising transformation (S, eq. 5.19) is used, since it is more appropriate for the right-tailed power law and gives a better power.

Another "measure of discrepancy" is the sum of the squared distances from the diagonal to the data point in a PP plot, the *Cramér-von Mises statistic* (Anderson & Darling, 1952; Stephens, 1978; D'Agostino & Stephens, 1986; Gan & Koehler, 1990),

$$C^2 = \sum_{i=1}^{n} \left(\widehat{P}_{(i)} - \frac{2i-1}{2n} \right)^2 - \frac{1}{12n}. \tag{5.22}$$

Like the KS statistic, this measure can be used in the stabilised PP plot, yielding the new *stabilised Cramér-von Mises statistic*,

$$SC^2 = \sum_{i=1}^{n} \left(S\widehat{P}_{(i)} - S\left(\frac{2i-1}{2n} \right) \right)^2. \tag{5.23}$$

A modified form of the Cramér-von Mises statistic is the *Anderson-Darling statistic*, (Anderson & Darling, 1952; Stephens, 1978; D'Agostino & Stephens, 1986; Gan & Koehler, 1990)

$$A^2 = -\sum_{i=1}^{n} \frac{(2i-1)}{n} \big(\log_e(\widehat{P}_{(i)}) - \log_e(1 - \widehat{P}_{(n+1-i)}) \big) - n, \tag{5.24}$$

which gives more weight to the tails of the distribution.

EDF statistics based on the correlation coefficient

The correlation coefficient is a measure for linearity, given by

$$R^2(X,Y) = \frac{\left(\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) \right)^2}{\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i-1}^{n} (Y_i - \overline{Y})^2}. \tag{5.25}$$

For perfect linearity R^2 has the value 1 or -1 and for uncorrelated points $R^2 = 0$. If the points $\{X,Y\}$ are always positive as in our cases then R^2 lies in the interval $[0,1]$. The rejection criterion is $R^2 < R^2_{\text{crit}}$.

The correlation coefficient can for example be used in the quantile-quantile plot,

$$r^2 = R^2 \left(X_{(i)}, \widehat{P}^{-1}_{(i)} \right), \tag{5.26}$$

but has, as the quantile-quantile plot, the disadvantage of needing the inverse distribution function.

Another possibility is to use the correlation coefficient in the PP plot (*PP correlation statistic*; Gan & Koehler, 1990),

$$k^2 = R^2 \left(\widehat{P}_{(i)}, \frac{i-0.5}{n} \right), \tag{5.27}$$

or in the stabilised PP plot (*stabilised PP correlation statistic*, first proposed here),

$$Sk^2 = R^2 \left(S\widehat{P}_{(i)}, S\left(\frac{i-0.5}{n} \right) \right). \tag{5.28}$$

The PP correlation statistic can be modified, as suggested by Gan & Koehler (1990), to force the points to go through $\{0.5, 0.5\}$. This is done by replacing in eq. 5.25 $\overline{X}$ and $\overline{Y}$ by 0.5, denoting the modified version R_0^2. Gan & Koehler (1990) found that the *modified PP correlation statistic*,

$$k_0^2 = R_0^2\left(\widehat{P}_{(i)}, \frac{i-0.5}{n}\right), \tag{5.29}$$

is somewhat more powerful than k^2. Again, the analogous procedere is possible in the stabilised PP plot, giving the *stabilised modified PP correlation statistic*,

$$Sk_0^2 = R_0^2\left(S\widehat{P}_{(i)}, S\left(\frac{i-0.5}{n}\right)\right). \tag{5.30}$$

Statistics based on tests for exponentiality

Due to the connection of the power law and exponential distribution, the various tests for exponentiality available in the literature are applicable for our purposes (cf. e.g. Beirlant et al., 2006). Since there is only a proportionality between $p(x)$ and $p_e(\log_e x)$ most of the derived (exponential) null distributions for the following statistics are no longer valid for an infinite power law.

The *Shapiro-Wilk statistic* (Shapiro & Wilk, 1972; Stephens, 1978; D'Agostino & Stephens, 1986),

$$W = \frac{n}{n-1}(\overline{X'} - X'_{(1)})\sum_{i=1}^{n}(X'_i - \overline{X'})^2, \tag{5.31}$$

($X'_i = \log_e X_i$) is originally a two-sided statistic with minimum $(n-1)^{-2}$ and maximum 1 for the exponential case. For the use with a power law the rejection criterion for the alternative hypothesis of a truncated power law distribution is $W > W_{\mathrm{crit}}$ in one-sided use.

The *Jackson statistic* (Jackson, 1967; Stephens, 1978; D'Agostino & Stephens, 1986; Beirlant et al., 2006),

$$T = \frac{\sum_{i=1}^{n} t_{i,n} X''_{(i)}}{\sum_{i=1}^{n} X''_i}, \tag{5.32}$$

with $X''_i = \log_e(X_i/\widehat{x}_{\mathrm{L}})$ and $t_{i,n} = \sum_{j=1}^{i} \frac{1}{n-j+1}$, is primarily the product of the ordered data and their expectation values $\lambda E(X_{(i)}) = t_{i,n}$, comparable to correlation type statistics. The division by $\sum_{i=1}^{n} X_i$ removes the dependence on the scale parameter λ. For a truncated power law alternative in one-sided use the rejection criterion is $-T > -T_{\mathrm{crit}}$.

Other tests for exponentiality are the the statistics of Brain & Shapiro (1983), the Moran statistic (Stephens, 1978; D'Agostino & Stephens, 1986), and the Greenwood statistic (Bartholomew, 1957; Stephens, 1978; D'Agostino & Stephens, 1986). We have also tested their powers when used for a power law, but they are not more powerful than the Shapiro-Wilk or Jackson statistic, and so we do not include details on them here.

5.4.5 Tests for truncation

The above described tests only allow one to distinguish whether the data are described by the null hypothesis or not. When the null hypothesis is rejected the test has to be made again, now with the alternative hypothesis as new null hypothesis. The *Likelihood Ratio* test combines this two-stage procedure into one test, by which can be decided whether of the two hypotheses is favourable. We use here the Likelihood ratio in the same way as the test statistics above. The test statistic is given by

$$\Lambda = \frac{\prod_{i=1}^{n} p_{\infty}(X_i; \widehat{\alpha}_{\infty})}{\prod_{i=1}^{n} p(X_i; \widehat{\alpha}, \widehat{x}_{\mathrm{U}})}. \tag{5.33}$$

For the infinite case we used the ML estimate which does not include a truncation to estimate the exponent and for the truncated case the modified ML for the exponent and upper limit. For numerical reasons the logarithm of eq. 5.33 is evaluated.

An answer to the problem of estimating parameter and simultaneously deciding between hypotheses can also be given in the Bayesian framework of statistics which is, however, not in the scope of this article.

A further specific test for truncation is the *exceedance statistic*,

$$X = \max_{i \leq 1 \leq n} X_i \tag{5.34}$$

(the largest data point), is only designed to test whether the distribution function is truncated or not. It cannot be used to detect a deviation from the power-law assumption. Furthermore it is one-sided with rejection criterion $-X > -X_{\mathrm{crit}}$.

5.4.6 Power comparisons

The power of the various statistics was calculated at a significance/type I error level of $\alpha_I = 0.05$ with parameters for the power law as given in Table 5.1. The critical points for the null hypothesis of an infinite power law were calculated as follows. For each of the parameter combinations, but with $x_{\mathrm{U}} = \infty$, a Monte-Carlo sample containing 1000 data sets was generated. For each data set the parameters were estimated using the modified ML estimator (exponent eq. 5.12, upper limit eq. 5.13). Then the test statistics were calculated using the estimates when necessary. This gives the distribution of the respective test statistic, from which the critical value follows as the 95% quantile.

For the power again a sample of 1000 data sets was generated, but now from a truncated power law. As before the parameters were estimated and the statistics calculated. The power is then the percentage of data sets with a test statistic smaller than the critical value.

The obtained powers are shown in Table 5.1. The exceedance statistic, X, is the most powerful test for truncation. However, it cannot be used for detecting deviations from the power law distribution. Thus it has to be used in conjunction with one or more of the other tests which include a test for the power law family as the parent distribution function.

A general effect appearing for all statistics is that the power decreases with increasing slope and range of the limits. By such changes the truncated distribution becomes – informally speaking – more similar to the infinite distribution and thus harder to discriminate. Above $\alpha = 2$ the performance of the tests drop significantly and therefore strong statements on truncation can barely be made. Unfortunately for the Salpeter value of the slope, $\alpha = 2.35$,

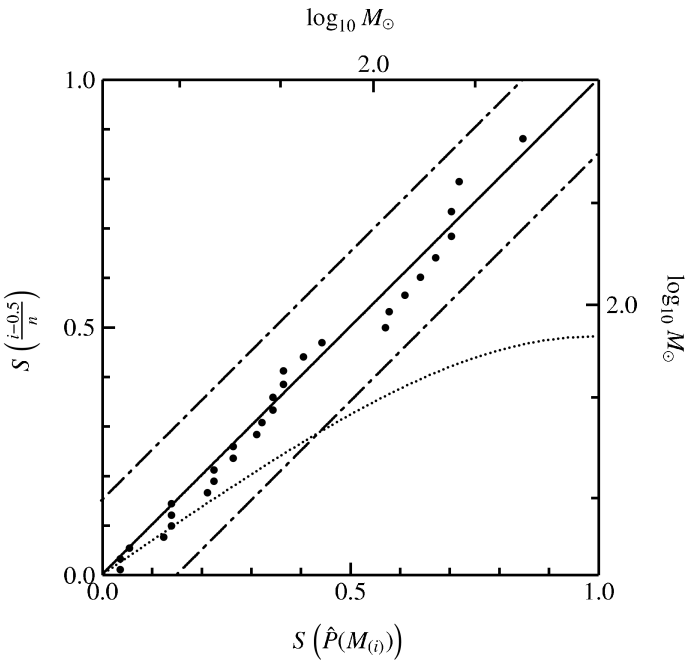


Figure 5.6: Truncated SPP plot of the massive stars in R136 with masses according to the model of Chlebowski & Garmany (1991) and parameters estimated using the modified ML method. Also shown is the curve for a infinite power law (dotted). The parallels to the diagonal limit the acceptance region of the stabilised KS test, null hypothesis of a truncated power law, significance level 5%.

the studied tests are mostly not powerful enough to decide whether an upper truncation is present or not. However, in some not so extreme real cases such as the data set of massive stars in R136 a sufficient power can be achieved. Furthermore, even if a truncation cannot be detected then deviations from a power law might still be discoverable.

Besides the general performance behaviour of the test statistics a further, rather surprising trend exists in the power. The most powerful tests are not necessarily the tests derived especially for the power law distribution from tests for exponentiality. The stabilising transformation (eq. 5.19) strongly enhances the power of general-purpose ECDF or correlation statistics so that they outperform the specialised tests. The Kolmogorov-Smirnov statistic which is known to be not very powerful (cf. Gan & Koehler, 1990) becomes, after stabilisation, more powerful than all other tests except for the exceedance test. In their not stabilised forms the general-purpose tests are, as expected, less powerful than the specialised tests. This enhanced power is a useful property since general-purpose statistics can easily be modified to tests for a different null hypothesis, e.g. a two-part power law.

In summary, the best test for truncation is the exceedance test, X. To confirm the hypothesis of a power law and for better significance this test should be followed by some of the most powerful remaining tests. These are, loosely ordered in descending power, the stabilised Kolmogorov-Smirnov test SD, the stabilised PP correlation test Sk_0^2, the stabilised Cramér-von Mises test SC^2 and the Jackson statistic T.

When a truncation is detected, then the hypothesis of a truncated power law has to be confirmed by again applying the respective statistics with this distribution (the truncated power law) as the null hypothesis.

Data using Chlebowski & Garmany (1991)

Estimate for	Slope $\widehat{\alpha}$	Bias slope	$\widehat{M}_{\mathrm{U}}$ $[\mathrm{M}_\odot]$	Bias $\widehat{M}_{\mathrm{U}}$
Const. Bins, LR	3.38±0.72	0.43		
Var. Bins, χ^2	2.42±0.75	-0.01	134±12	-11
CCDF	2.02±0.88	-0.11	140±9	-1.9
Beg	2.17±0.77	<0.01	142±8	-0.4
Beg, recursive	2.17±0.77	0.01	142±8	-0.4
ML∞	3.51±0.35	1.34		
ML	2.11±0.73	-0.06	136±7	-8
Mod. ML	2.20±0.78	0.02	143±9	< 0.1
Results of Koen (2006)				
CCDF	2.10		143.9	
ML	2.11		136	

Table 5.2: Estimates for the 29 most massive stars in R136. The standard deviations of the estimators were calculated using a Monte-Carlo sample of size 10 000. For the binning methods 5 bins are used. The bias was calculated using the results of the modified ML method as input values.

5.5 Examples

5.5.1 The massive stars in R136

As a first exemplary application of the presented statistical techniques, in particular of the estimators, we chose the data set of massive stars in R136 published by Massey & Hunter (1998). They gave for the 29 most massive stars the masses based on two different stellar models (Chlebowski & Garmany, 1991, with masses ranging from 56 $\mathrm{M}_\odot$ to 136 $\mathrm{M}_\odot$, and Vacca et al., 1996, $75 - 155\ \mathrm{M}_\odot$). The results of the estimators are shown in Table 5.2 where a Monte-Carlo sample of size 10000 was used to calculate the standard deviations.

Beg's estimator and the modified ML method agree well ($\widehat{\alpha} = 2.2$), the ML estimate is slightly smaller ($\widehat{\alpha} = 2.1$). In reasonable agreement with this value are also the results of variable-size binning and fitting the complementary cumulative DF plot. For comparison the results of Koen (2006) are also given in Table 5.2. The ML estimates are equal, only the CCDF result differs, likely due to a different definition of the empirical DF (Koen uses $i/(n+1)$ whereas here $(i-0.5)/n$ is used). The ML method without including an upper limit gives a much larger exponent ($\widehat{\alpha} = 3.5$) which shows the effect of a model mismatch. A comparison of this value only with a constant-size histogram, where a linear regression gives $\widehat{\alpha} = 3.4$, would not give any indication of the mismatch.

The upper limit is determined as $\approx 140\ \mathrm{M}_\odot$ by Beg's estimator, the ML and the CCDF method. The results from variable-size binning are not consistent with the data set, because this upper limit is smaller than the largest data point.

For the goodness-of-fit analysis an SPP plot with a truncated power law as null hypothesis is shown in Fig. 5.6. The curve for the infinite power law is clearly not fitting the data. The stabilised Kolmogorov-Smirnov, Cramér-von Mises and SPP correlation coefficient test all give a strong disagreement of the data with an infinite power law and no disagreement with a truncated power law.

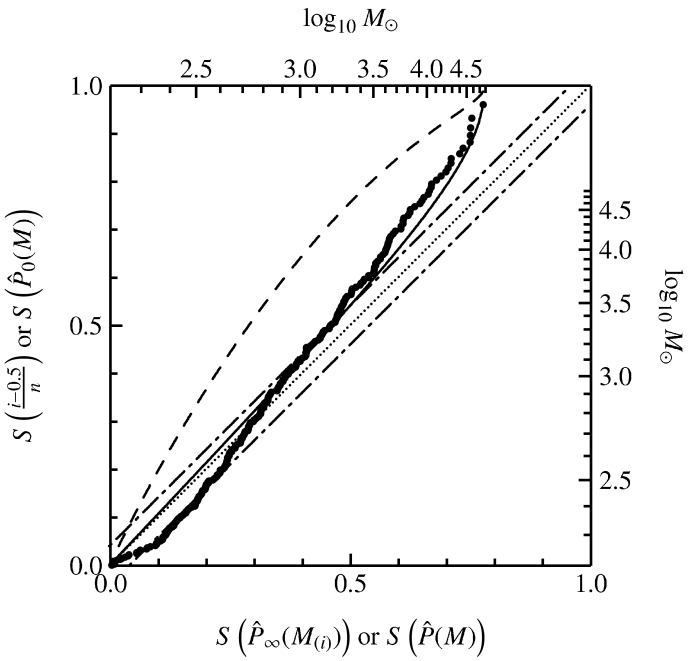


Figure 5.7: Infinite SPP plot of the LMC star clusters (age $< 10^{7.5}$ yr) with the lower mass limit of de Grijs & Anders (2006), $10^{2.2}$ $M_\odot$. (Dotted line: infinite hypothesis, $\widehat{\alpha}_{\mathrm{MML}} = 1.47$; solid curve: truncated hypothesis, parameters as estimated ($\widehat{\alpha} = 1.47$, $\widehat{M}_{\mathrm{U,MML}} = 64200$ $M_\odot$); dashed curve: truncated hypothesis, $\alpha = 2$, $\widehat{M}_{\mathrm{U,MML}} = 64200$ $M_\odot$; dash-dotted lines: limits of the acceptance region of the stabilised Kolmogorov-Smirnov test, significance level 5%).

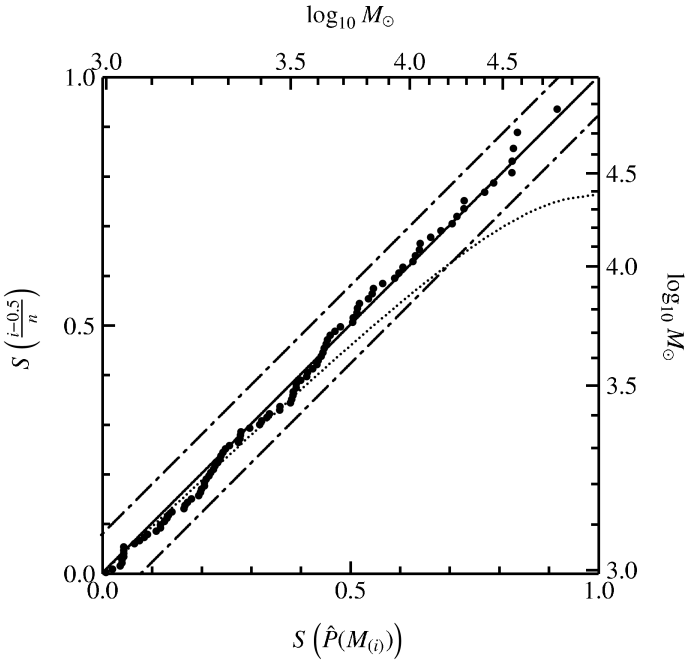


Figure 5.8: Truncated SPP plot of the LMC star clusters (age $< 10^{7.5}$ yr) starting at $10^{2.5}$ $M_\odot$. (Dotted line: infinite hypothesis, $\widehat{\alpha}_{\mathrm{MML}} = 1.62$; solid line: truncated hypothesis, parameters as estimated ($\widehat{\alpha}_{\mathrm{MML}} = 1.62$, $\widehat{M}_{\mathrm{U,MML}} = 68000$ $M_\odot$); dashed: truncated hypothesis, $\alpha = 2$, $\widehat{M}_{\mathrm{U,MML}} = 68000$ $M_\odot$; dash-dotted: limits of the acceptance region of the stabilised Kolmogorov-Smirnov test, significance level 5%).

The modified ML estimates from the data set using the models of Vacca et al., 1996 are $\widehat{\alpha} = 2.87 \pm 0.98$ and $\widehat{M}_{\mathrm{U}} = 163 \pm 9$ $M_\odot$. The goodness of fit tests indicate a truncated power law with high significance too.

5.5.2 The young star clusters in the Large Magellanic Cloud

The second example for the methods presented above, with an emphasis on the advantage of the SPP plot, is the analysis of the mass distribution of young star clusters in the Large

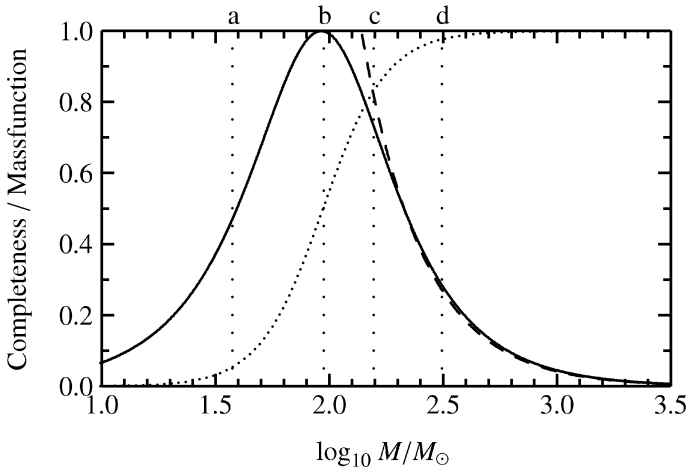


Figure 5.9: Influence of the completeness (dotted line) on the observable mass function (solid line), based on an assumed power law (dashed line) as the underlying distribution function. The parameters were chosen to match the situation for the Large Magellanic Cloud, see text. The vertical lines correspond to a: $10^{1.58}$ $M_\odot$, b: $10^{1.98}$ $M_\odot$, c: $10^{2.2}$ $M_\odot$, and d: $10^{2.5}$ $M_\odot$. The mass function is scaled arbitrarily for better visibility.

Magellanic Cloud. We use a part of the data set given by de Grijs & Anders (2006), the star clusters with ages younger than $10^{7.5}$ yr and masses larger than $10^{2.2}$ $M_\odot$.

Based on an inspection of the shape of a histogram of the data de Grijs & Anders (2006) concluded that there is a significant flattening of the mass function for $M < 10^3$ $M_\odot$ (see their fig. 8). Indeed, also an SPP plot with an infinite null hypothesis, Fig 5.7, shows that the empirical curve of the data is strongly bent in the lower mass range ($M \lesssim 10^{2.5}$ $M_\odot$, estimated exponent ≈ 1.5). In the upper mass range the infinite SPP plot reveals that the data are better described by a truncated power law (solid line). This indicates that above $\approx 10^{2.5}$ $M_\odot$ the data presumably will be consistent with a truncated power law.

With an SPP plot using only the star clusters more massive than $\approx 10^{2.5}$ $M_\odot$ this hypothesis is confirmed (Fig. 5.8), the stabilised Kolmogorov-Smirnov acceptance region is not exceeded. Thus, a change in the slope or shape of the mass function in the mass range $10^{2.5}$–10^3 $M_\odot$, as stated by de Grijs & Anders (2006), cannot be deduced using our techniques. Only the mass range $10^{2.2}$–$10^{2.5}$ $M_\odot$ seems to deviate from the power law. The slope which is derived from the data with masses larger than $10^{2.5}$ $M_\odot$ is $\widehat{\alpha} = 1.6 \pm 0.06$ and an upper limit $68 \pm 11 \times 10^3$ $M_\odot$ is obtained by the modified Maximum Likelihood method. This exponent is smaller than the value determined by de Grijs & Anders (2006), $\alpha = 1.8 \pm 0.1$, who used constant-size binning and star clusters more massive than 10^3 $M_\odot$ (for this mass range the modified maximum Likelihood estimate is $\widehat{\alpha} = 1.63 \pm 0.1$).

The feature in the mass range $10^{2.2}$–$10^{2.5}$ $M_\odot$ could be caused by an actual change of the mass function. However, since it is at the lower mass end, it could also be caused by an incomplete data set. The completeness limit adopted by de Grijs & Anders (2006) was derived by Hunter et al. (2003) from the behaviour of the luminosity function (see fig. 4 of Hunter et al., 2003). They used as the brightness limit the brightness where the luminosity function reaches at the faint side half of its peak value, obtaining $\mathscr{M}_V = -3.5$ mag or $10^{1.58}$ $M_\odot$ (using $M = 10^{6+0.4(-14.55-\mathscr{M}_V)}$ $M_\odot$, Hunter et al., 2003, eq. 1). The mass related to the brightness limit is valid for clusters of an age of 10 Myr. In a similar way Parmentier & de Grijs (2008) derive — starting from the mass distribution of a chosen age interval older than 10 Myr— from the mass which separates the lower 25% from the upper 75% a completeness limit of $\mathscr{M}_V = -4.7$ mag. If we use this value also for younger clusters, then a completeness mass of $10^{2.06}$ $M_\odot$ would result (However, for unknown reasons an application of their method to clusters younger than 10 Myr leads to a different completeness mass of $10^{2.35}$ $M_\odot$, Parmentier, priv. comm.).

A completeness limit derived in such a way coincides approximately with the peak of the observed mass function. But the transition from no detection to complete detection is smooth and has a certain broadness in which only a fraction of all sources is detected which can affect a wide mass range, illustrated in Fig. 5.9. The observable mass function (solid line) is the product of the actual mass function (dashed, exponent $\alpha = 1.6$) and the completeness function (dotted). As the functional form for the completeness function we chose

$$c(M) \quad = \quad 1 - \left(1 + \left(\frac{M}{M_0} \right)^{\phi} \right)^{-1} . \tag{5.35}$$

The parameters of the completeness function were chosen such that the half peak point of the observable mass function is at $10^{1.58}$ $M_\odot$ (or $\mathscr{M}_\mathscr{V} = -3.5$ mag, as Hunter et al., 2003, point a in Fig. 5.9) and the peak mass is $\approx 10^2$ $M_\odot$ ($\mathscr{M}_V \approx -4.5$ mag, point b in Fig. 5.9). It is just a coincidence that for the used parameters ($\log_{10} M_0 = 1.98$ and $\phi = 3.12$) the 50% completeness mass of the completeness function coincides with the peak mass of the observable mass function. With these empirically determined parameters the observable mass function is shallower than the actual power law in the mass range below $\approx 10^{2.5}$ $M_\odot$. This strongly supports the argument that the deviation of the data from the power law in Fig. 5.7 is caused by incompleteness. The distribution of star clusters with ages $< 10^{7.5}$ yr are well consistent with a single power law with $\widehat{\alpha} = 1.6$, starting from $10^{2.5}$ $M_\odot$.

5.6 Summary and conclusions

In this work we compared methods to estimate the exponent and upper limit of a truncated power law distribution. We reviewed graphical methods to represent the data. Finally we studied goodness-of-fit tests, specifically to test for truncation.

Our results are:

1. A generally working estimator for the exponent and upper limit is our modified maximum likelihood method. It performs well with respect to bias and standard deviation.

2. A maximum likelihood estimate of the exponent without considering a truncation can lead to biased results if the data stem from a truncated power law.

3. The estimator of Beg (1983) is also performing well but is numerically not stable. Variable-size binning as introduced by Maíz Apellániz & Úbeda (2005) performs well for the exponent. The estimate for the upper limit based on the normalisation constant is biased.

4. The stabilising transformation introduced by Michael (1983) enhances plots and goodness-of-fit tests. For one-sided distributions only a half transformation should be made to achieve optimal results.

5. The stabilised PP plot is a particular useful display of the data.

6. The stabilised Kolmogorov-Smirnov statistic (SD), the stabilised PP correlation test (Sk^2), the stabilised Cramér-von Mises statistic (SC^2), the Jackson statistic (T) and the QQ correlation (r^2) test are powerful goodness-of-fit tests for the truncated power law.

7. The exceedance statistic (X) is the most powerful test for truncation. Since it does not test for power-law behaviour it has to be used in combination with a powerful goodness-of-fit test for the truncated power law, as the ones mentioned in the previous point.

8. The massive stars in R136 are well described by a truncated power law with $\widehat{\alpha} = 2.20 \pm 0.78$ and $\widehat{M}_{\rm U} = 143 \pm 9$ M$_\odot$, using the Chlebowski & Garmany (1991) stellar models for mass determination, or $\widehat{\alpha} = 2.87 \pm 0.98$ and $\widehat{M}_{\rm U} = 163 \pm 9$ M$_\odot$, using the Vacca et al. (1996) stellar models.

9. The young star clusters in the Large Magellanic Cloud (ages younger than $10^{7.5}$ yr) with masses larger than $10^{2.5}$ M$_\odot$ are well described by a truncated power law with $\widehat{\alpha} = 1.62 \pm 0.06$ and $\widehat{M}_{\rm U} = 68.8 \pm 11.6 \times 10^3$ M$_\odot$.

10. A change in shape of the star cluster mass function in the Large Magellanic cloud in the low mass range $M < 10^3$ M$_\odot$, as reported by de Grijs & Anders (2006), cannot be verified. For $M > 10^{2.5}$ M$_\odot$ the observed distribution follows a truncated power law, a flattening below $10^{2.5}$ M$_\odot$ is most likely caused by an underestimated completeness limit.

Acknowledgements

We thank Cathie Clarke and Douglas Heggie for critical reading of the manuscript and valuable comments. ThM acknowledges financial support by the AIfA.

5.7 Appendix: Beg's estimator

The estimator for the exponent (Beg, 1983) is given in its original form as

$$\widehat{\theta} \;=\; \frac{(n-3)!\sum_{j=0}^{j^\star}(-1)^j\binom{n-2}{j}K_{j+1}^{n-4}}{(n-4)!\sum_{j=0}^{j^\star}(-1)^j\binom{n-2}{j}K_{j+1}^{n-3}}, \tag{5.36}$$

where $\widehat{\theta}=\widehat{\alpha}-1$ and $K_j = T-nY-j(Z-Y)$ with $Y=\log_e X_{(1)}$, $Z=\log_e X_{(n)}$ and $T=\sum_{i=1}^{n}\log_e X_j$. The terminating index of the sum, $j^\star$, is determined by the condition $T-nY-j(Z-Y)>0$, as shown by Beg (1983). The estimate for the exponent is then $\widehat{\alpha}=\widehat{\theta}+1$.

The direct evaluation of eq. 5.36 involves the calculation of $\binom{n-2}{j}$, which is only practicable for less than about 170 data points in double precision arithmetic. This problem can be handled with a recursive implementation of the estimator, feasible for any number of data, as follows.

To abbreviate we introduce $L_j = \left(1-j\frac{Z-Y}{T-nY}\right)$ which leads to

$$K_j^{n-4} \;=\; (T-nY)^{n-4}L_j^{n-4}. \tag{5.37}$$

With changing the limits of the sum and omitting $(n-3)!$ the numerator of eq. 5.36 reads

$$-(T-nY)^{n-4}\sum_{j=1}^{j^\star}(-1)^j\binom{n-2}{j-1}L_j^{n-4}. \tag{5.38}$$

Omitting the prefactor $(T-nY)^{n-4}$, the expanded sum reads

$$-\underbrace{\frac{(n-2)!}{0!(n-2)!}}_{=1}L_1^{n-4}+\frac{(n-2)}{1}L_2^{n-4}-\frac{(n-2)(n-3)}{1\cdot 2}L_3^{n-4}+\ldots. \tag{5.39}$$

Starting with the second term this can be written as

$$\underbrace{\frac{n-2}{1}\left(L_2^{n-4}\,\underbrace{\frac{n-3}{2}\left(L_3^{n-4}--\ldots\right)}_{=:S_3^{(n-4)}}\right)}_{=:S_2^{(n-4)}}. \tag{5.40}$$

The superscript $(n-4)$ should only indicate the exponent and is not used as an exponent in $S_j^{(n-4)}$. From this the recursion can easily be seen:

$$S_{j'-1}^{(n-4)} \;=\; \frac{n-j'}{j'-1}\left(L_{j'}^{n-4}-S_{j'}^{(n-4)}\right), \tag{5.41}$$

where j' descends from $j^\star$ to 2 and $S_{j^\star}=0$. The last step is

$$S^{(n-4)} := S_1^{(n-4)} = L_1^{n-4}-S_2^{(n-4)}. \tag{5.42}$$

The recursion for the denominator in eq. 5.36 is as for the numerator, but replacing the exponent $n-4$ by $n-3$ in equations 5.41 and 5.42. The estimator of θ is then (remembering all omitted factors)

$$\widehat{\theta} \;=\; \frac{n-3}{T-nY}\frac{S^{(n-4)}}{S^{(n-3)}}. \tag{5.43}$$

The estimators for the upper limit in the form of Beg (1983) is

$$\widehat{x}_{\mathrm{U}} \;=\; X_{(n)}\left(1+\frac{1}{n(n-1)}\frac{\sum_{j=0}^{j^\star}(-1)^j\binom{n-1}{j}K_j^{n-2}}{\sum_{j=0}^{j^\star}(-1)^j\binom{n-2}{j}K_{j+1}^{n-3}}\right). \tag{5.44}$$

The recursion formula for the sum in the numerator follows by analogous steps as before with

$$S_j^{'(n-2)} \;=\; \frac{n-j}{j}\left(L_j^{n-2}-S_j^{'(n-2)}\right), \tag{5.45}$$

the last step

$$S^{'(n-2)} \;=\; 1-S_1^{'(n-2)}. \tag{5.46}$$

and

$$\widehat{x}_{\mathrm{U}} \;=\; X_{(n)}\left(1+\frac{(T-nY)}{n(n-1)}\frac{S^{'(n-2)}}{S^{(n-3)}}\right), \tag{5.47}$$

with $S^{(n-3)}$ from the estimator for the exponent above.

Chapter 6

On the mass function of young star clusters in M51

Th. Maschberger, P. Kroupa, R. Scheepmaker & S. Goodwin
MNRAS, to be submitted

We analyse the mass function of young star clusters in M51. For the initial star cluster mass function (ages up to 12 Myr) we test the agreement of the data with the functional forms of an infinite power law, a truncated power law and a Schechter function which all fit the data, although with different parameters. Also, we investigate the dependence of the results on the choice of isochrones (Geneva and Padova), and the choice of metallicity (solar or free). For the power law an exponent of ≈ 2.6 *fits the data, the exponent for a Schechter function is smaller and not well constrained. The turnover mass of the Schechter function,* $M_\star$ *,lies at* $\approx 10^4$ $M_\odot$*. The older clusters (age range 12–25 Myr) have a flatter mass function, which could be caused by infant evolution but is also consistent with the strong dynamical evolution found e.g. by Gieles et al. (2005) for clusters in M51. For older ages (up to 100 Myr) only age/mass fitting with the Padova isochrones provides a sufficient number of clusters for analysis. There the mass function steepens again to reach values similar to the ones of the initial mass function. This could be the sign of cluster revirialisation in a galaxy with dynamical evolution of star clusters as expected for spiral galaxies.*

6.1 Introduction

The galaxy M51 is one of the few galaxies in which a large part of the young star cluster population has been observed and their properties have been subject to a number of studies (Bik et al., 2003a; Bastian et al., 2005; Gieles et al., 2005; Scheepmaker et al., 2007; Haas et al., 2008). We investigate in this work the initial cluster mass function, the mass function of young star clusters which just have ended their embedded phase. Several functions have

been proposed for the form of the initial cluster mass function, an infinite power law, a truncated power law or a Schechter function (a power law with exponential decay at high masses, Gieles et al., 2006a). The shape of the initial cluster mass function can give insights on properties of galaxy-wide star formation. Are, for example, different star formation rates leading to varying parameters, especially upper truncation limits (see e.g. Weidner et al., 2004), or does the range of massive clusters follow purely from random sampling (e.g. Bastian, 2008)? To shed light on this question we estimate the parameters for all propositions and perform goodness-of-fit tests to establish or refute one or the other function using the statistical apparatus of Maschberger & Kroupa (2009).

A further central topic of our work is to investigate in how much results on the cluster mass function depend on the choice of isochrones and metallicities. Scheepmaker et al. (2009) found that especially in the first ≈ 100 Myr the fitted ages and masses are strongly dependent on the isochrone/metallicity choice. Therefore we perform our analysis for two isochrone sets (Padova and Geneva) and for each of these for two metallicities (solar and free).

The time-evolution of the star cluster mass function follows the dynamical evolution of the star clusters. For the earliest times (up to ≈ 30 Myr) the transition from embedded to gas-free star clusters can leave it's imprint on the mass function, especially if this transition is mass-dependent (Kroupa & Boily, 2002; Weidner et al., 2007; Parmentier et al., 2008). For older ages clusters follow their normal path of dynamical evolution via two-body relaxation leading to an increasingly shallower mass function. The degree of flattening is dependent on the galactic environment, e.g. the tidal potential (e.g. Baumgardt & Makino, 2003). A convenient description of the dynamical evolution is given by Lamers et al. (2005b), where the dynamical environment is parametrised by t_4, the life-time of a 10^4 $M_\odot$ cluster. t_4 is typically derived from the total age-mass distribution assuming a cluster (number) formation rate (Gieles et al., 2005, finding $t_4 \approx 10^8$ yr). By considering only a restricted age range but performing an estimate of the exponent the necessity of assuming a cluster formation rate can be avoided when determining t_4.

After describing the data (Sec. 6.2), the functional forms of the cluster mass function (Sec. 6.3) and the methods for data analysis (Sec. 6.4) we first analyse in Section 6.5 the initial cluster mass function and its dependence on isochrones and metallicity. In Section 6.6 we discuss the time-evolution of the cluster mass function with respect to it's relation to cluster infant mortality/evolution and dynamical evolution. We finish with our conclusion and a summary in Sec. 6.7.

6.2 Observations and derivation of ages and masses

For our analysis we utilize the data sets of Scheepmaker et al. (2009) which contain 1850 star clusters of M51. The data set is based on HST observations in the B, V and I bands (described in Scheepmaker et al., 2007) and in the U band (Scheepmaker et al., 2009). The selection criteria for any one of the 5502 resolved sources with photometry to be classified as a star cluster were 1) photometry brighter than the 90% completenes limit of Scheepmaker et al. (2007), 2) photometric accuracy better than 0.2 mag and 3) $U < 22$ mag. Ages and masses were estimated using the GALEV simple stellar population models (Schulz et al., 2002; Anders & Fritze-v. Alvensleben, 2003) together with the ANALYSED tool (Anders et al., 2004). The typical age uncertainty is < 0.4 dex (Anders et al., 2004; de Grijs et al., 2005; de Grijs & Anders, 2006). ANALYSED gives masses for the star clusters that have been corrected for stellar evolution,

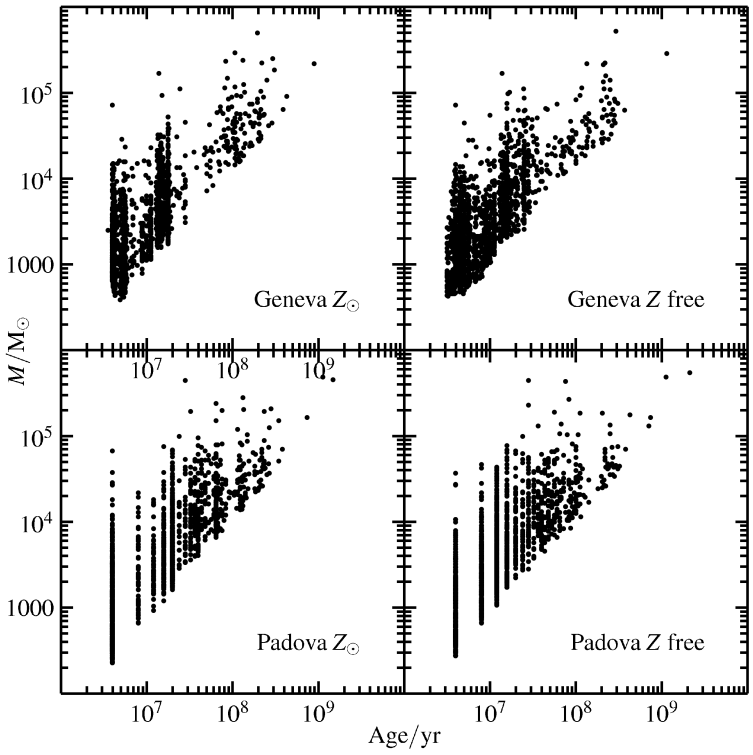


Figure 6.1: Age-mass diagrams for the M51 star clusters where ages and masses have been derived with Geneva or Padova isochrones, either with fixed (solar) or free metallicity. Cluster masses have only be corrected for stellar evolution.

but not for dynamical mass loss. For the fit of the ages and masses isochrones were used either of the Geneva group (Charbonnel et al., 1993; Schaerer et al., 1993) or the Padova group (Bertelli et al., 1994; Girardi et al., 2000). The metallicity was either taken to be constant at solar metallicity or allowed within the range of 0.4–2,5 $Z_\odot$

The age-mass diagrams are given in Fig. 6.1, with the typical lower envelope which follows the observational limiting magnitude. The Padova isochrones have less resolution at young ages which leads to the typical 'chimneys' in the age-mass diagrams. The choice of fixed or free metallicity has a comparably small impact on the age-mass distributions compared to the disturbingly large discrepancies for different isochrone sets. For example, with the Geneva isochrones there is a cluster desert between 20 and 100 Myr, a region which is well populated with Padova isochrones.

The data anlysed by us are

6.3 Models for the star cluster mass function

The most widely adopted functional form of the (initial) star cluster mass function is a power law, ranging up to infinity,

$$p_\infty(M;\beta,M_{\rm L}) \quad = \quad \frac{\beta-1}{M_{\rm L}^{1-\beta}} M^{-\beta}. \tag{6.1}$$

L is the lower mass limit, usually given by observational completeness. We use all mass functions as probability densities, i.e. normalised such that $\int_{M_{\rm L}}^{\infty} p(M){\rm d}M = 1$.

As star clusters of potentially infinite mass are physically not possible the power law can

be truncated at an upper limit U, giving the truncated power law probability density,

$$p_U(M;\beta,M_{\rm L},M_{\rm U}) = \frac{1-\beta}{M_{\rm U}^{1-\beta}-M_{\rm L}^{1-\beta}}M^{-\beta} \quad (6.2)$$

A third possibility is to introduce an exponential truncation near $M_\star$, what leads to the Schechter function (Gieles et al., 2006a),

$$p_\star(M;\beta,M_{\rm L},M_\star) = \frac{1}{\int_{M_{\rm L}}^{\infty} M^{-\beta}\exp(-M/M_\star){\rm d}M}M^{-\beta}e^{-\frac{M}{M_\star}} \quad (6.3)$$

For the Schechter function the integral in the normalising constant has no analytic form and has to be evaluated numerically.

6.4 Methods for data analysis

The analysis of the star cluster data has two aspects: determining the parameters of the distribution function and verifying the assumed functional form of the distribution function.

To estimate the parameters of a truncated power law or a Schechter function we perform Maximum-Likelihood fits. For the infinite power law the exponent is approximately the same as for a truncated power law. To minimise the bias occurring with small samples the Maximum Likelihood method is modified for the truncated power law, following Maschberger & Kroupa (2009, Chapter 5 this thesis). The given uncertainties are the intrinsic statistical scatter of the method (derived from 1000 simulated data sets) and do not account for observational uncertainties. As the likelihood surface does not show a pronounced peak for the Schechter function the usual methods to derive the maximum do not converge well. Therefore we search for the maximum on a grid of β and $M_\star$ values (β ranging from 1.5-3 in steps of 0.025, $M_\star$ from 10^3 $M_\odot$ to 10^7 $M_\odot$ in logarithmic steps of 0.025).

It should be noted that both the estimate for the upper limit of the truncated power law (U) and the turnover mass of the Schechter function ($M_\star$) cannot be determined very well. They both can show a deviation of up to 1 dex, usually lying below the actual value.

To assess informally the agreement of the data with the assumed distribution function we use the stabilised percentile-percentile (PP) plot, examples of which are shown in Fig. 6.2. For this plot an empirical and a theoretical cumulative probability is calculated for each data point (star cluster), which is transformed to a stabilised cumulative probability. The empirical cumulative probability for a cluster of a certain mass is derived from the position of the mass in the sorted set of all masses. In the ascendingly sorted data set of size n the value $P_{\rm E}(M_{(i)}) = (i-0.5)/n$ is assigned to the i-th most massive star cluster, e.g. for the smallest mass $0.5/n$ and for the largest mass $(n-0.5)/n$. The corresponding theoretical cumulative probability for the null hypothesis , $P_{\rm H}$, is calculated using the mass of cluster i and either specified or estimated parameters. In absence of stochastic noise the pairs of $P_{\rm E}(M_{(i)})$ and $P_{\rm H}(M_{(i)})$ have values between zero and unity and lie on the diagonal, the reference line for any null hypothesis in this plot. Alternative hypotheses can also be shown in the SPP plot, e.g. the left panel of Fig. 6.2, constructed with an infinite power law as null hypothesis, shows as solid line a truncated power law. It is clear that in this figure the truncated power law is a better description of the data.

Compared to a histogram a plot of the empirical against the theoretical probability has the advantage of showing every data point individually without a need for grouping, at the

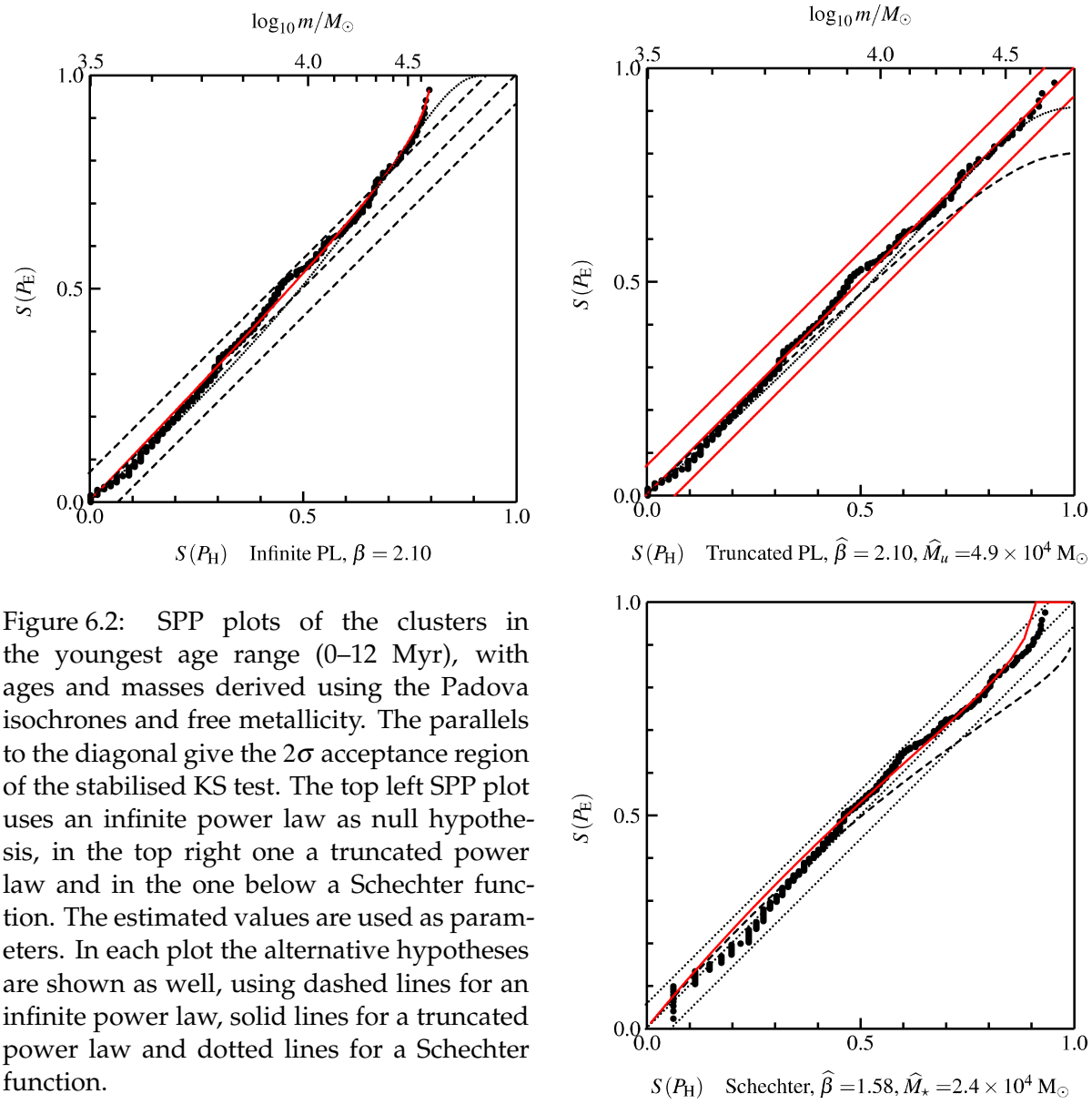


Figure 6.2: SPP plots of the clusters in the youngest age range (0–12 Myr), with ages and masses derived using the Padova isochrones and free metallicity. The parallels to the diagonal give the 2σ acceptance region of the stabilised KS test. The top left SPP plot uses an infinite power law as null hypothesis, in the top right one a truncated power law and in the one below a Schechter function. The estimated values are used as parameters. In each plot the alternative hypotheses are shown as well, using dashed lines for an infinite power law, solid lines for a truncated power law and dotted lines for a Schechter function.

price of a more complex construction. Another convenient property of this plot is that an acceptance region of the Kolmogorov-Smirnov (KS) test can easily be shown as parallels to the diagonal, allowing one to visualise a formal goodness-of-fit criterion. Since it is well known that the KS test is not so significant in the tails we apply a (variance-) stabilising transformation to both the empirical and the theoretical probabilities (see Maschberger & Kroupa, 2009, Chapter 5 this thesis). The appearance of the SPP plot is essentially unchanged by the transformation. The stabilising transformation can be used either as a half transformation, emphasising particularly the upper mass end, or as a full transformation, giving every data point the same weight. To distinguish between an infinite and a truncated power law we use the half transformation, as this has a higher decisive power for these alternatives (see Maschberger & Kroupa, 2009, Chapter 5 this thesis). For the Schechter function we apply the full transformation, as a wide mass range can be affected by the exponential component.

Age range	M_{ll}	$\beta_{\rm PL}$	$M_{\rm U}$	$\beta_{\rm Sch}$	$M_\star$	n	$\beta_{\rm PL}$	$M_{\rm U}$	$\beta_{\rm Sch}$	$M_\star$	n
		Geneva $Z_\odot$					Geneva Z free				
1 – 12 Myr	3.50	2.74±0.13	11.02	1.86	1.73	205	2.54±0.10	9.37	1.70	1.60	247
1 – 12 Myr	3.70	3.09±0.23	13.10	2.30	23.27	101	2.97±0.19	11.13	2.32	22.86	137
12 – 25 Myr	3.70	2.30±0.09	21.58	1.60	3.45	258	2.14±0.09	20.74	1.65	6.03	205
12 – 25 Myr	3.80	2.49±0.12	23.40	1.77	4.58	203	2.25±0.12	21.52	1.74	7.87	163
		Padova $Z_\odot$					Padova Z free				
1 – 12 Myr	3.50	2.43±0.16	9.55	1.90	4.75	114	2.10±0.08	4.90	1.58	2.35	351
1 – 12 Myr	3.70	2.57±0.24	9.94	1.89	19.88	64	2.22±0.13	4.93	1.63	2.61	211
12 – 25 Myr	3.70	2.16±0.10	11.22	1.61	4.13	212	1.89±0.11	10.98	1.60	6.52	151
12 – 25 Myr	3.80	2.23±0.13	11.36	1.62	4.29	162	1.98±0.14	11.11	1.62	6.32	123
12 – 50 Myr	3.70	2.06±0.06	56.45	1.57	6.68	409	1.94±0.06	54.58	1.57	9.03	321
12 – 50 Myr	3.80	2.19±0.07	60.94	1.62	6.88	339	2.08±0.07	58.66	1.58	8.15	276
25 – 50 Myr	3.90	2.22±0.11	73.00	1.76	13.60	157	2.14±0.12	70.81	1.70	14.35	134
25 – 50 Myr	4.00	2.34±0.13	80.52	1.90	25.47	127	2.20±0.14	73.93	1.70	15.47	107
25 – 100 Myr	4.10	2.29±0.11	61.83	1.67	11.80	182	2.24±0.11	61.89	1.88	34.34	159
25 – 100 Myr	4.30	2.70±0.17	77.47	2.22	93.89	120	2.49±0.17	70.54	2.18	209.90	98
50 – 100 Myr	4.20	2.39±0.21	32.66	1.75	16.99	72	2.38±0.21	82.30	2.05	196.23	60
50 – 100 Myr	4.30	2.66±0.26	35.60	2.06	72.41	60	2.46±0.26	85.47	2.09	254.95	47

Table 6.1: Estimates of the parameters of the cluster mass function for the combinations of isochrones and metallicities. The table gives the analysed age range and lower mass limit of the fit (M_{ll} in $\log_1 0$), the estimated exponent and upper limit of the power law ($\beta_{\rm PL}$ and $M_{\rm U}$), and the exponent and turnover mass of the Schechter function ($\beta_{\rm Sch}$ and $M_\star$). $M_{\rm U}$ and $M_\star$ are given in 10^4 $M_\odot$.

In the stabilised PP plot the KS test can analogously be defined as the largest distance between a data point and the diagonal, but now a new null distribution of the test has to be calculated for which we use the Monte-Carlo approach. The null hypothesis can then be rejected when the stabilised KS distance is larger than e.g. the 95% quantile of the null distribution, which corresponds to a significance level of 5%. Using the stabilising transformation the statistical power of the KS test can be significantly increased. It is as effective in deciding between an infinite and truncated power law as a likelihood ratio test, with the added bonus of a graphical representation.

When agreement or disagreement of the data with the functional form of the cluster mass function, evaluated using the estimated parameters, is established, the data analysis is in principle finished. However, as we find below agreement of the data with exponents and turnover masses which are commonly believed to be rather extreme, we extend our analysis. Because the error bars in the estimates can be large it is possible that the "canonical" parameter values are fitting the data as well. Therefore a (parametric) goodness-of-fit test should be made with the "canonical" parameters in order to establish their suitability for the data. For this purpose we performed stabilised KS tests for the infinite power law, the truncated power law and the Schechter function on a grid of exponents, truncation masses and turnover masses. The results are shown in Figs. 6.3 and 6.4, with a dot representing agreement of this parameter combination with the data. The sizes of the dots correspond to different significance levels: large dots for a significance of 32% or 1 σ, medium-sized dots for 5% or 2 σ, and small dots

for 1% or $\approx 2.6\sigma$. The open circle is situated at the location of the maximum likelihood estimates. The plots for the Schechter function also include as grey dots the estimates from 1000 realisations of a data set with the estimates from the data as input values. This is to indicate the typical scatter of the maximum likelihood method, as well as to show the potential bias in the estimates.

6.5 Initial cluster mass function

To investigate the initial cluster mass function we derived the exponents for the age range up to 12 Myr, selecting the only the first three chimneys in the Padova data and their less pronounced analoga in the Geneva data. To be reasonably above the completeness limit we start the analysis with a minimum mass of $10^{3.5}\mathrm{M}_\odot$ for a cluster. The results for the exponents with the different isochrones and metallicities are given in Table 6.1. We give also the results starting from $10^{3.7}\mathrm{M}_\odot$ for a rather informal assessment of their robustness. With larger starting masses the number of available clusters becomes very small so that no good results can be obtained.

We now first turn to the question about the functional form of the initial cluster mass function. As visible from Figs. 6.3 and 6.4 all assumptions about the functional form are consistent with the data, except for the Padova isochrones with free metallicity. For the moment we exclude the Padova/free metallicity data from our discussion and return to them below. The exponent for an infinite or truncated power law is generally much larger than the canonical value of 2. The estimated upper limit of a truncated power law, $\widehat{M}_\mathrm{U}$, lies only a little above the most massive cluster. This is the typical behaviour of the fitting method. However, with the stabilised KS test larger values for M_U, or even an infinite power law cannot be excluded.

Schechter functions with the same exponent as the pure power law and a high $M_\star$ also fit the data, as they are essentially power laws in this case. For the clusters fitted with Geneva isochrones allowing free metallicity, however, the agreement with such parameters is on less than a 2σ level. In analogy to the pure power law essentially any value of $M_\star$ is allowed. The Schechter functions fit as well with a lower exponent and a lower $M_\star$, interestingly for a broad range of β with a narrow range of $M_\star$. This behaviour means that the data are essentially fitted by the exponentially decaying tail, so that the value of β becomes unimportant. The estimated parameters of the Schechter function always lie in this part of the β-$M_\star$ plane, but this can be a systematic effect of the Maximum Likelihood method.

After the general shape of the initial cluster mass function we address the influence of the isochrones and metallicity on the parameters. If a range of metallicities is allowed the exponent decreases compared to fixed solar metallicity, by about 0.1–0.3. This occurs for both isochrone sets With free metallicity some kind of truncation seems to be favoured by the data, especially when the Padova isochrones are used. This truncation either manifests in a sharp truncation or a small $M_\star$ of the Schechter function A Schechter function with the same exponent as the power law and a large $M_\star$ is not fitting the data well.

As the last point of our discussion of the initial cluster mass function we discuss the data set derived with Padova isochrones and free metallicity. In contrast to the other data sets these data do not support an infinite power law. As visible in the age-mass diagram (Fig. 6.1, right panel, the three youngest "chimneys") the cluster masses are densely packed over the whole range and do not thin out with increasing mass. This is first by-eye evidence that some truncation must be present. As visualisation of a more objective way to reject the infinite power

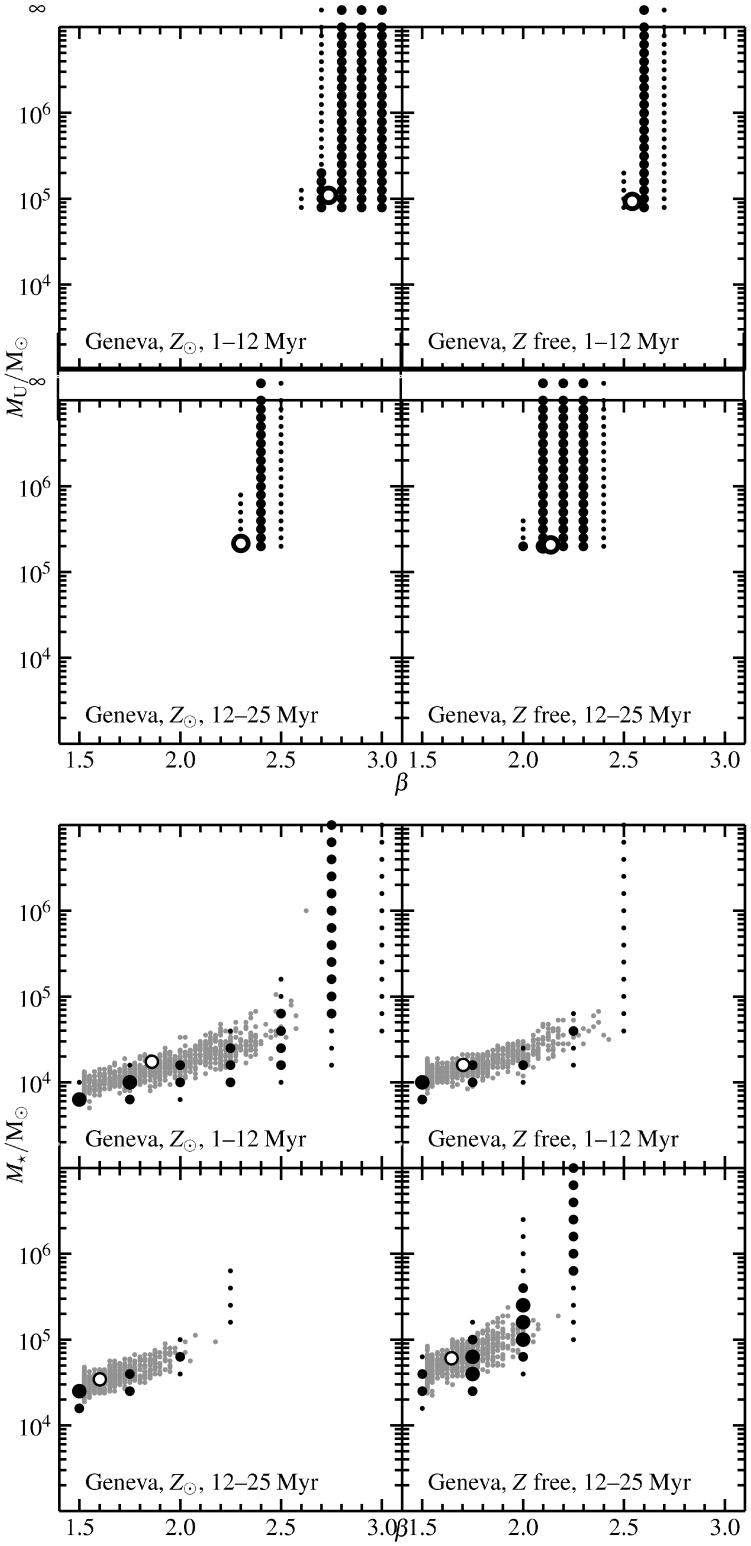


Figure 6.3: Goodness-of-fit plots for the M51 star clusters in the age ranges 0–12 Myr and 12–25 Myr. Shown is the β-U plane for the power law or the β-$M_\star$ plane for the Schechter function. The estimated parameters are marked with an open circle. If a parameter combination is in agreement a dot is shown, with a small size for agreement at the $\approx 2.6\sigma$ significance level, a medium size at 2σ and a large dot at 1σ level. In the plots for the Schechter function the small grey dots show estimates which have been derived from a Monte-Carlo simulation with the estimated values as input parameters, which should visualise the quality of the fitting method.

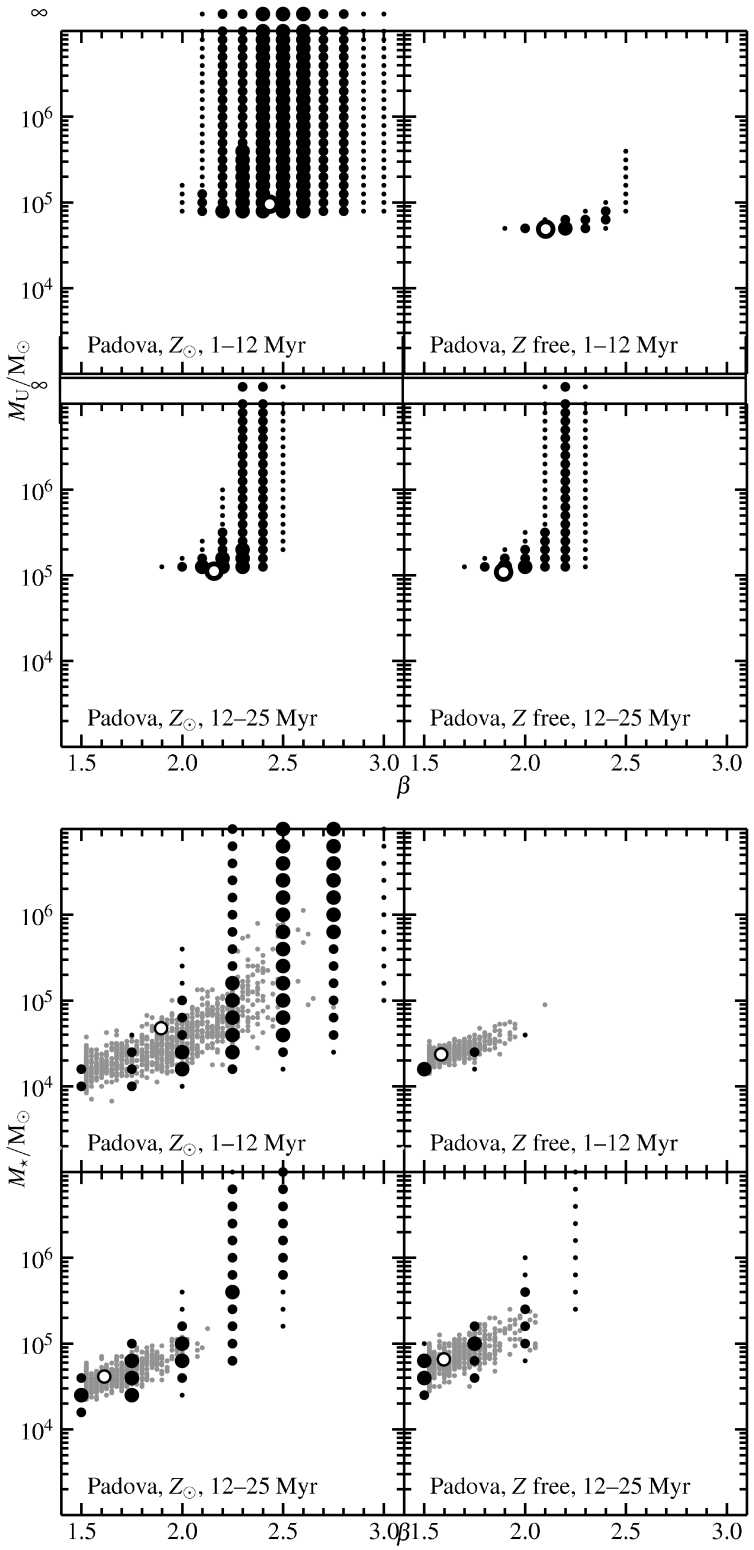


Figure 6.4: Goodness-of-fit plots for the M51 star clusters as Fig. 6.3, but for Padova isochrones.

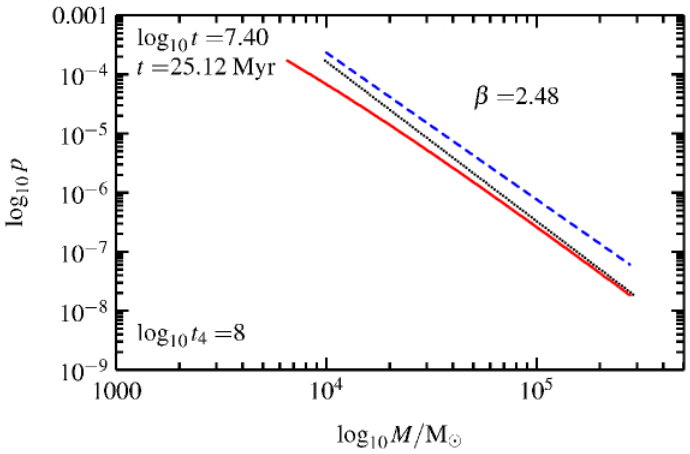


Figure 6.5: Dynamical evolution of the star cluster mass function from an initial power law with $\beta = 2.6$ (dotted line) where $t_4 = 10^8$ yr. After 25 Myr the mass function (solid line) has become shallower and can be fitted by a power law with $\beta \approx 2.5$ (dashed line).

law we show in Fig. 6.2 (left panel) an SPP plot with the infinite power law as null hypothesis (using estimated parameters). Clearly the data bend away from the diagonal and leave the 5% significance region, rather following the truncated power law (solid line). We also show in Fig. 6.2 SPP plots constructed with a truncated power law (middle) and a Schechter function (right), both using estimated parameters. In these two the data do not exceed the acceptance region. By mere visual inspection the truncated power law seems to be in better agreement with the data, as for the Schechter function some trend in the upper half seems to be present. However, as both cannot be rejected the preference of one to the other is more a matter of taste and perhaps practicality.

6.6 Time evolution of the cluster mass function

The star cluster population of M51 is sufficiently rich that our statistical methods can also be applied to older clusters. In the age range $\tau = 12$ Myr to $\tau = 25$ Myr with all isochrone/metallicity combinations the number of clusters is sufficient for an analysis. The lower mass limit of the analysis has to be increased to $10^{3.7}$ $\mathrm{M}_\odot$ to be well above the observational completeness limit. As before we constructed goodness-of-fit plots in the β-M_U or β-$M_\star$ plane, which are shown in the upper parts of Figs. 6.3 and 6.4. The picture with respect to agreement with the functional form is similar to the younger age range, although with less support for Schechter functions with large β and large $M_\star$. The acceptable parameter range is shifted from $\beta = 2.5$–2.7 (1–12 Myr) to $\beta = 2.3$–2.5 (12–25 Myr). For the Schechter function with small exponent $M_\star$ is increased from 10^4 $\mathrm{M}_\odot$ to 10^5 $\mathrm{M}_\odot$, as more clusters of higher mass are present. This shift implies that there is presumably no universal $M_\star$. Again the combination of Padova isochrones and free metallicity falls out of the consensus, with a lower exponent and an infinite power law now being in an agreement with these data.

The natural explanation for the shift in the exponent is the dynamical evolution of the young clusters. The gas that is not converted into stars (perhaps 70% or more of the total initial gas mass) accounts for a large fraction of the potential, so that clusters that expell this gas are left in a dynamically unstable state and loose stars in reaction to this rapid change of the potential. After, say, a few tens of millions of years they reach again a "stable" state and continue with their "normal" dynamical evolution. As the time-scale of the dynamical evolution of a star cluster depends on the number of stars (or total mass), the mass function

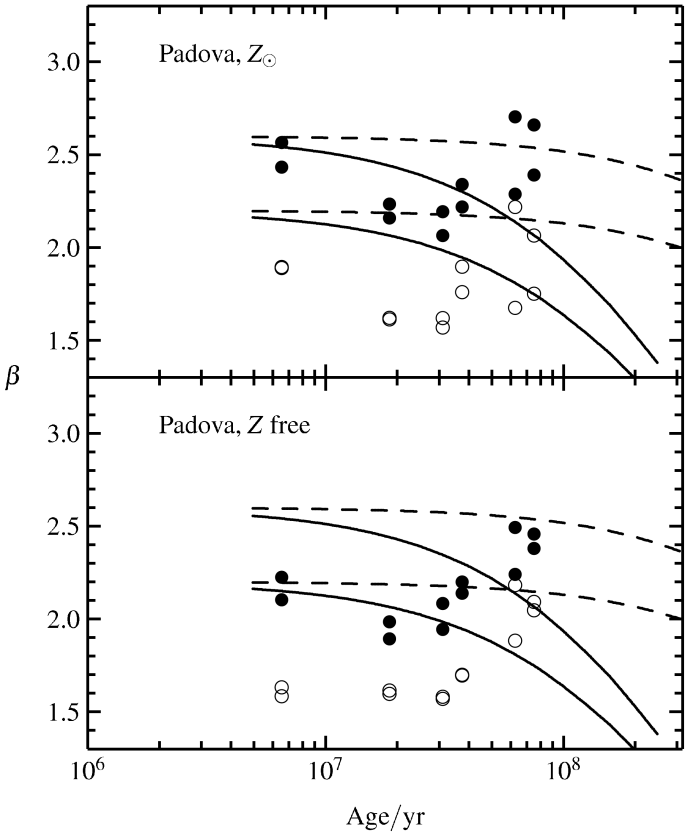


Figure 6.6: Time-evolution of the exponent of the cluster mass function. The dots show the estimates for a power law, the open circles for a Schechter function. Solid lines give the theoretical evolution of the exponent, starting with $\beta = 2.6$ or $\beta = 2.2$ for $t_4 = 10^8$ yr and dashed lines for $t_4 = 10^9$ yr.

of a star cluster system changes in time.

To describe the dynamical evolution we use the analytic models by Lamers et al. (2005a). The age-dependence of a cluster mass is there given by

$$M(t) \quad = \quad M_i \left(1 - \frac{\gamma t}{t_0} \left(\frac{\mathrm{M}_\odot}{M_i}\right)^\gamma\right)^{\frac{1}{\gamma}}, \tag{6.4}$$

with typically $\gamma = 0.62$. Mass loss due to stellar evolution is not included in the above equation as in the mass determination of our data this has been accounted for. t_0 parametrises the influcence of the environment, such as the galactic potential, spiral arms, molecular clouds etc. It is usually transformed into t_4 which is the dissolution time of a 10^4 $\mathrm{M}_\odot$ cluster, using

$$t_4 \quad = \quad 660 \times t_0^{0.967}. \tag{6.5}$$

By using eq. 6.4 and d $t_4 = 1 \times 10^8$ yr as found for M51 by Gieles et al. (2005) we can calculate the time-evolution of the star cluster mass function. Figure 6.5 shows the evolved mass function after 25 Myr. It is somewhat shallower at the low mass end compared to the initial mass function (dotted line, $\beta = 2.6$). With some offset for visibility a power law with an fitted exponent is also shown (dashed line). The fitted exponent, which has been derived using a linear regression in logarithmic space, has a value of ≈ 2.4, comparable to what is found in the observational data. The structure that is introduced in the cluster mass function by dynamical evolution is not sufficiently pronounced to be seen in the data.

The value of 100 Myr for t_4 is much shorter than expected for M51, as it is comparable to the Milky Way. Dynamical models show that the typical dissolution time scale of a cluster

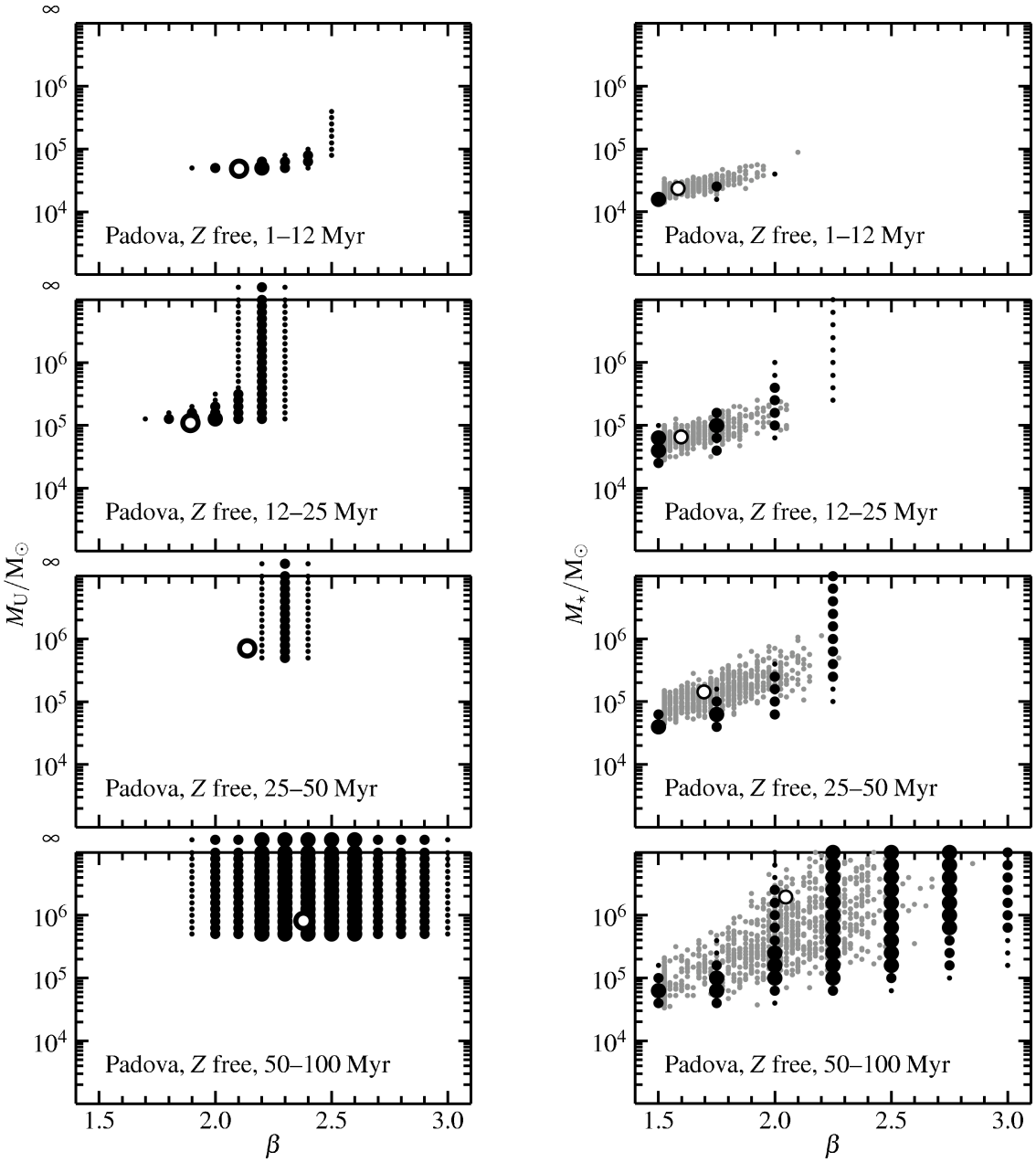


Figure 6.7: Goodness-of-fit plots showing the time-evolution of the cluster mass function, using the Padova isochrones and free metallicity. The left figure is for the infinite and truncated power law and the right figure for a Schechter function. The symbols in the plots are as in Fig. 6.3. A version of this plot with solar metallicity can be found in the Appendix (Fig. 6.8).

containing 10 000 stars is of the order of one to a few Gyr, assuming it resides in a Milky Way-type galaxy (see Baumgardt & Makino, 2003). There have been several explanations for this discrepancy, as for example the collisions with molecular clouds (Gieles et al., 2005).Until now the methods to derive t_4, by using the cluster (number) formation rate (Boutloukos & Lamers,

2003; Gieles et al., 2005) or by using the most massive cluster (Gieles & Bastian, 2008), have both the drawback that a (constant) cluster formation history has to be assumed. However, by the direct analysis of the mass function this serious limitation can be circumvented.

For the Padova isochrones the number of clusters older than 25 Myr is still large enough for an analysis. The parameters and results of the analysis are shown in Table 6.1, where we also included overlapping time bins to assess the stability of the fits. With this we can construct the time-evolution of the exponent, shown in Fig. 6.6 with dots for the power law and open circles for the Schechter function. Surprisingly the exponent does not monotonously decrease with time, but has a minimum at $\approx 20-30$ Myr to increase again to values even slightly higher than for the youngest ages (but consistent within the error bars). Additional to the data we give the expected evolution of the exponent, based on $t_4 = 10^8$ yr (solid) and 10^9 yr (dashed) and starting with $\beta = 2.6$ and $\beta = 2.2$.

Various explanations can be given for this behaviour of the mass function. By taking the error bars of the exponents into account and looking at the goodness-of-fit plots (Fig. 6.7, showing only the free metallicity case. The $Z_\odot$ plots are more homogeneous.) it could be deduced that the evolution of the exponent is less drastic than suggested by the changes from $1-12$ Myr $12-25$ Myr only. M51 would then have a t_4 comparable to what would be expected for a spiral galaxy (10^9 Myr). Alternatively one could argue that, as only for the Padova isochrones this effect appears it might well be possible that the cluster sample for older ages is affected by systematics of the fitting, especially as a increase of the exponent disagrees with the typical physical processes. In this case the results for the older ages would be considered as being not very trustworthy and a small t_4 would be given the preference. A third explanation could be that star clusters pass through a state of low surface brightness as they expand after gas expulsion (see fig. 2 in Kroupa et al., 2001). In this case they would drop out of the observational sample, particularly at the lower mass end, so that for a time an apparent decrease of the exponent would be detected. When the clusters are not completely unbound they would fall back to a higher surface brightness and reappear in the sample again, which gives an apparent increase of the exponent. This middle age range would have to be neglected for determining t_4.

Unfortunately the available data do not allow us to favour or reject any of the possible explanations. However, the main point of this exercise is perhaps not to find exact numbers. Rather, we could show that it is possible to directly analyse the star cluster mass function without any prior assumptions about the cluster formation history. The puzzling behaviour of the exponent shows that the data set has properties not yet understood which were concealed until now. Results obtained from global properties of the cluster mass function (number of clusters per age or most massive cluster) are presumable also affected by these properties. Hopefully new observations and improved age/mass fitting will shed light on this issue in the future.

6.7 Summary and conclusions

We analysed the star cluster population of M51, with ages and masses determined using Geneva and Padova isochrones with solar or free metallicity. The results for the initial cluster mass function show a dependence on the choice of isochrones and metallicities, with the very particular case of Padova isochrones and free metallicity. Generally an inifinite power law, truncated power law or a Schechter function with large $M_\star$ all fit the data, with a rather steep

exponent of $\beta \approx 2.6$. A Schechter function with smaller exponent and a small $M_\star$ fits the data as well, with no strong constraint of the exponent as one fits only with the exponentially decaying part. The value found for $M_\star$, 10^4 $M_\odot$, is by an order of magnitude smaller than found by Gieles (2009), presumably caused by the differently selected set of clusters.

The time-evolution of the cluster mass function is somewhat ambiguous. All combinations of isochrones and metallicities show a decrease in the exponent by ≈ 0.2 in the age range 12–25 Myr. Assuming the generally accepted value for the parameter of star cluster evolution, $t_4 = 10^8$ yr, this decrease is in accordance with normal dynamical evolution of star clusters and not caused by the transition from the embedded to gas-free phase.

For the Padova isochrones the number of star clusters older than 25 Myr is sufficiently large to also analyse older age ranges. Surprisingly, here the exponent does not continue to decrease, but increases again to similar values as at 1–12 Myr. This may be the first direct evidence for cluster revirialisation after residual gas expulsion.

Acknowledgements

ThM acknowledges financial support via the European Union Research Training Network CONSTELLATION.

6.8 Appendix

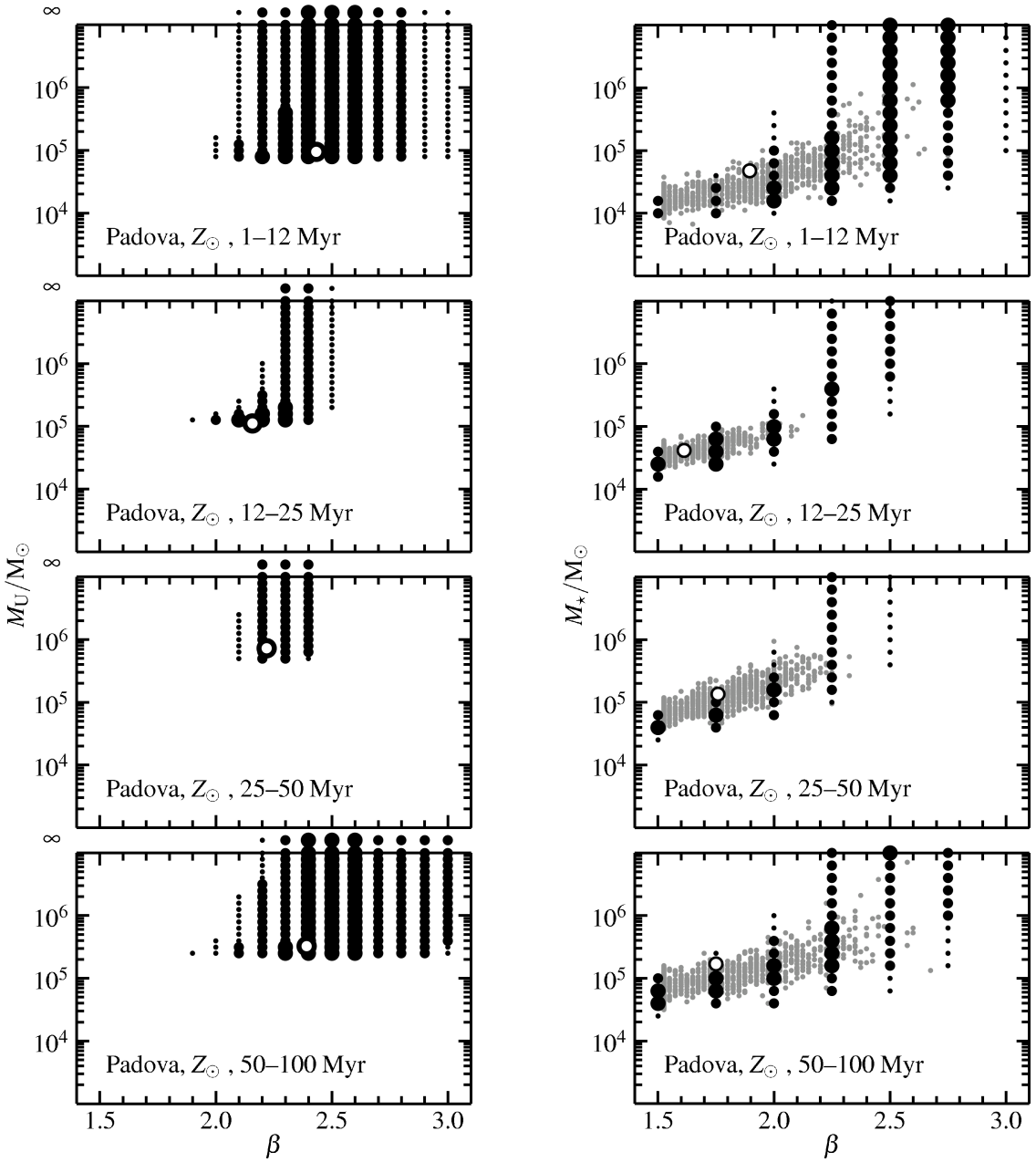


Figure 6.8: Goodness-of-fit plots, as Fig. 6.7, showing the time-evolution of the cluster mass function, using the Padova isochrones and solar metallicity.

Chapter 7

Maximum stellar mass versus cluster membership number revisited

Th. Maschberger & C. J. Clarke
MNRAS 391:711–717 (2008)

We have made a new compilation of observations of maximum stellar mass versus cluster membership number from the literature, which we analyse for consistency with the predictions of a simple random drawing hypothesis for stellar mass selection in clusters. Previously, Weidner and Kroupa have suggested that the maximum stellar mass is lower, in low mass clusters, than would be expected on the basis of random drawing, and have pointed out that this could have important implications for steepening the integrated initial mass function of the Galaxy (the IGIMF) at high masses. Our compilation demonstrates how the observed distribution in the plane of maximum stellar mass versus membership number is affected by the method of target selection; in particular, rather low n clusters with large maximum stellar masses are abundant in observational datasets that specifically seek clusters in the environs of high mass stars. Although we do not consider our compilation to be either complete or unbiased, we discuss the method by which such data should be statistically analysed. Our very provisional conclusion is that the data is not indicating any striking deviation from the expectations of random drawing.

7.1 Introduction

It is well known (following Weidner & Kroupa (2004, 2006) and Oey & Clarke (2005)) that in the case of clusters containing fewer than ≈ 100 OB stars (i.e. those with mass $<$ a few $\times\ 10^4\ \mathrm{M}_\odot$) the maximum stellar mass increases with cluster mass. At higher cluster mass scales, the value of the maximum stellar mass saturates at around 150–200 $\mathrm{M}_\odot$ for reasons that are not entirely clear (see e.g. Zinnecker & Yorke, 2007). In this paper, we restrict ourselves to considering the lower mass regime. In Section 7.2 we review why the statistics of maximum

stellar masses in clusters of various scales can place constraints on high mass star formation in a cluster context and how rather subtle differences in assumed algorithms for cluster building are imprinted on the integrated galactic IMF (the IGIMF). We emphasise that analysis of the statistics of maximum stellar mass versus cluster mass offers the best prospects for an observational determination of whether the IGIMF should be different from that measured in individual clusters (see e.g. Weidner & Kroupa, 2006; Elmegreen, 2006). We also stress that competing algorithms can only be distinguished through proper statistical analysis of the observed distributions and that selecting algorithms according to how they reproduce the *mean* trend can be misleading. In Section 7.3 we present a new (but in all likelihood still incomplete and biased) compilation of observational information on this issue and highlight the sensitivity of the distribution obtained to the method of target selection. In Section 7.4 we discuss the statistical inferences that can be drawn from the current dataset and conclude (Section 7.5) with an appeal for further observational information to be used in future analyses.

7.2 The importance of maximum stellar mass data and its statistical analysis

The simplest interpretation of the fact that the maximum stellar mass is lower, on average, in lower mass clusters is that this just derives from the statistics of random sampling. To take a simple analogy, the average height of the tallest inhabitants of large cities is likely to be greater, on average, than the average height of the tallest individuals in small villages. It would however to be incorrect to infer from this that there is, for example, a nutritional deficiency among village dwellers. On the other hand, a better analogy might be with the wealth of richest individuals in settlements of various sizes, since in this case this might reflect the size of the local economic base. This is the sort of argument used by Weidner & Kroupa, who point out that in the case of cluster formation, the stars acquire their mass directly from the available gas reservoir. Their simple Monte-Carlo simulations build the finite size of the gas reservoir into their algorithms for stellar mass selection and reject any star whose formation causes the total designated cluster mass to be exceeded. This 'rejection' element preferentially affects more massive stars and is chiefly manifest through a statistical lowering of the maximum stellar mass compared with its value in random sampling experiments. Another plausible algorithm was proposed by Elmegreen (2006), motivated by the fact that the fraction of the initial gas in a protocluster that ends up in stars may be significantly less than unity. In this algorithm, therefore, although stellar mass selection is terminated once the total stellar mass exceeds the designated total mass, the last star is only rejected if this takes the total cluster mass over a value equal to the sum of the designated cluster mass and the mass of an additional gas reservoir. Since this is a softer rejection criterion, this algorithm produces results that are closer to random sampling than a strictly mass constrained algorithm.

Since in all these cases the maximum stellar mass follows quite a broad distribution (at fixed cluster scale), these differences cannot be discerned from a single observational datapoint. Instead it is necessary to compare observed distributions (at given cluster scale) with the results of simulations (or, in the case of the simplest, random drawing, hypothesis, with the results of analytic predictions). Some care is needed when considering the best property of the distribution that should be compared with observational data, as is demonstrated in Figure 7.1 which illustrates the probability density function of maximum stellar mass at given cluster mass scale in the case of the random drawing hypothesis. Clearly, this distribution

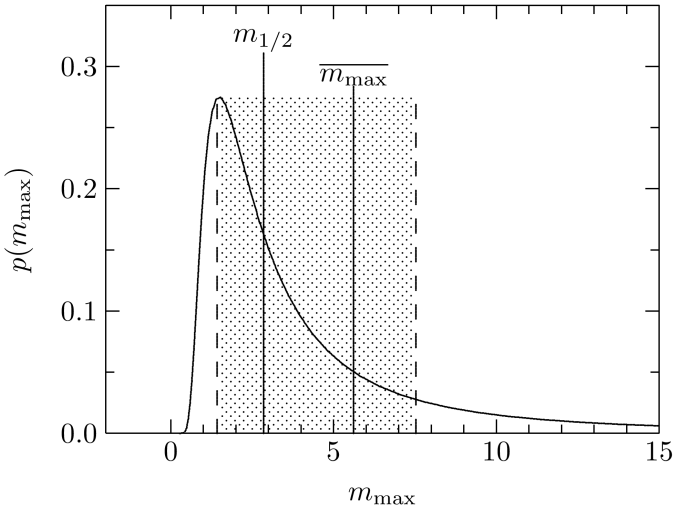


Figure 7.1: Probability density of the most massive star, $p(m_{\rm max})$ (eq. 7.5 in Appendix 7.6) for a star cluster containing $n = 30$ stars. Characteristic quantities are the mean, $\overline{m_{\rm max}}$, and the median, $m_{1/2}$. The 1/6th and 5/6th quantiles limit the shaded region containing 2/3rd of the most massive stars.

is highly asymmetric, with the mean significantly exceeding the median. This means that in the case of sparsely sampled observational data (i.e. not many clusters at given mass scale), the observed mean is likely to be lower than the true mean. A better approach is instead to compute the predicted distribution (as a function of cluster mass) and use a non-parametric method (e.g. a Kolomogorov-Smirnov test) to compare these with the distribution of observational datapoints in the plane of cluster mass versus maximum stellar mass (see Oey & Clarke (2005) and Section 7.3).

Thus far we have discussed the interpretation of these statistics in terms of what light they may shed on cluster formation (obviously the algorithms described above are simple 'toy models' but a clear signature in favour of one of them could be useful, for example, in determining how much stellar mass assignments are shaped by strict limitations in available gas supply). Another implication is purely empirical: we have stressed that the steepening of the upper IMF (and the resulting reduction in maximum stellar mass) in the case of non-random mass selection algorithms are too subtle to be detectable in any given cluster (i.e. each cluster is statistically consistent with being drawn from the input IMF). However, when one combines the results of many clusters (i.e. —on the assumption that the galactic field is composed of dissolved clusters— if one turns an IMF into an IGIMF) the signature of algorithms that preferentially reject high mass stars is seen in a steepening of the IGIMF. This important insight was first discussed in this way by Kroupa & Weidner (2003) (though see Vanbeveren (1982) for an earlier version of the argument). The reason why such star rejection algorithms —which are only important in the lower mass clusters that we discuss here— have a discernible effect on the IGIMF is simply that, given the steepness of the observed cluster mass function, low mass clusters make an important contribution to the galactic field.

Naturally, a steeper IGIMF has implications for how a range of quantities (such as supernova rate or ionising luminosity or chemical enrichment) relate to the galactic star formation rate. Weidner et al. (2004) also extended the argument by positing that similar considerations apply to *star cluster* maximum masses in galaxies of different masses. This means that in lower mass galaxies, the field population would be more dominated by lower mass clusters, and therefore that the IGIMF would be more steepened by the effect described above. Indeed, Pflamm-Altenburg et al. (2007b) have gone on to argue that this has important implications for the mapping between Hα luminosity and star formation rate in dwarf galaxies and would

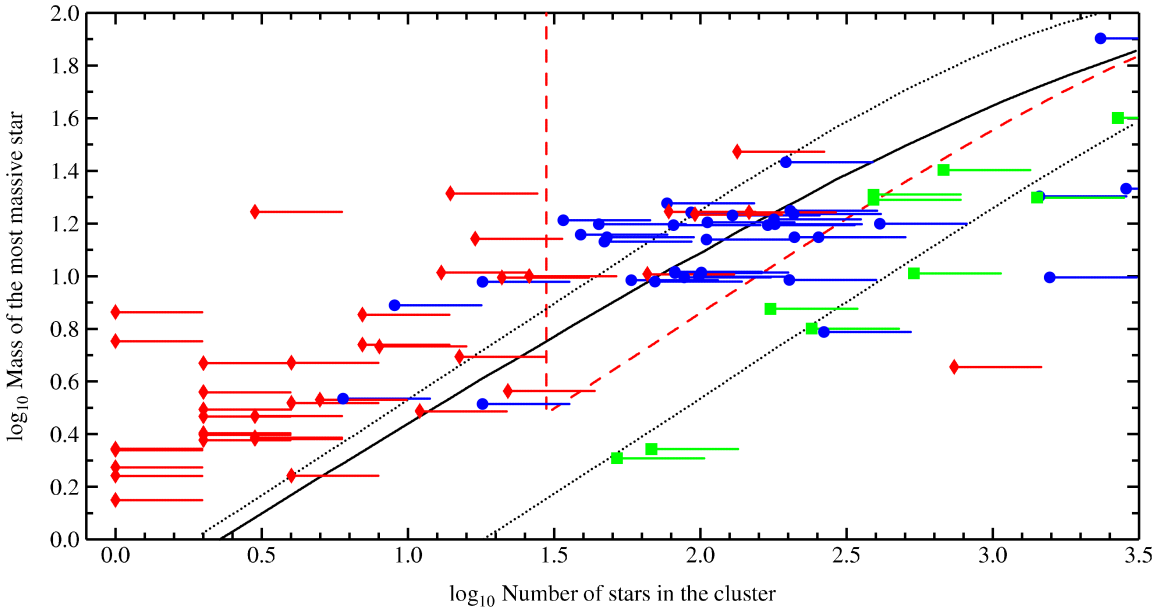


Figure 7.2: Mass of the most massive star versus the number of stars in the cluster (for better visibility, a small random scatter was applied to the (discrete) masses). The data are collected from the literature, with the main sources Testi et al. (♦) and Weidner & Kroupa (■). The references for the other points are given in Appendix 7.7. The solid line is the mean value of $m_{\rm max}$ depending on n. The dotted lines follow the 1/6 and 5/6 quantiles, and should confine 2/3rd of the observed data.

lead to the systematic *under*-estimation of the SFR in dwarfs.

Observationally, opinion is strongly divided as to whether there is good evidence that the IGIMF is steeper than Salpeter (or if it varies between galaxies): see e.g. Elmegreen (2006), Pflamm-Altenburg et al. (2007b), Selman & Melnick (2008) and discussion in Clarke (2008). We may, however, be able to turn this question around: *if* we can use cluster data to determine whether the maximum stellar mass statistics are indeed compatible with random drawing models then we can immediately learn whether the IGIMF should be equal to the IMF (without recourse to any Galaxy-wide or extragalactic data). Although, as we shall see in the following Section, it is not straightforward to achieve an unbiased sample for analysis, it is obviously attractive to be able to use rather simple, local observations to constrain a quantity which is potentially of extragalactic significance.

7.3 Observational data

In what follows, we attempt a simple test and enquire: is observational data on maximum stellar mass as a function of cluster scale compatible with the hypothesis of random drawing from a universal IMF (i.e. the same exponent and mass limits for all clusters)? Incompatibility would have important consequences for the IGIMF, as we have seen above and we would then need to enquire what other algorithms could achieve consistency. On the other hand, compatibility (at whatever desired significance level) would not necessarily imply that ran-

dom drawing is the 'best fit model'; it would however remove much of the motivation for finding more complex alternatives.

The analytical model for random drawing of stellar masses (see Oey & Clarke (2005) or Selman & Melnick (2008), and Appendix 7.6 of this paper that recapitulates the main results) is based on the expected distribution of maximum stellar masses in the case that one makes a given number, n, of selections from a given mass function. It therefore makes sense to make n the independent variable (unlike in the case of mass constrained models where cluster mass is the obvious choice). In order to obtain a homogeneous sample the data are renormalised to a common lower limit because of differing observational lower mass limits for each cluster. Each cluster is designated by the expected number of stars that it would contain down to 0.08 $M_\odot$ (i.e. not including brown dwarfs), given the observed number and mass limit, and assuming an IMF for the missing range.[1] We employ a two-part power law IMF (Kroupa, 2001, 2002) with a scaling as $m^{-2.35}$ (Salpeter, 1955) for stars > 1 $M_\odot$. In order to compare the completeness magnitude of a particular set of observations with the lower mass limit of our analysis, we use the conversion between K magnitude and mass given by Carpenter et al. (1993) ($m = 10^{-0.24 m_K + 0.24}$ or $m = 10^{-0.25 m_H + 0.44}$). All clusters are corrected (at least in a statistical sense) for background or foreground contamination. Furthermore we demanded that the observed region was large enough to contain the whole cluster area.

As far as the maximum stellar mass is concerned, we either use values quoted in the literature or else estimate masses from listed spectral types using the conversion given in Schmidt-Kaler (1982) (the masses for spectral types not contained in the list being interpolated).

Our criterion for including a cluster is only that we have found it to be possible to derive estimates of both n and $m_{\rm max}$ in this way. As we discuss below, it is unlikely to be either a complete or an unbiased sample and this makes any conclusions that we draw from this dataset extremely preliminary. Figure 7.2 compares all the data that we have assembled with the predicted centiles of the random drawing model (i.e.the mean and the 1/6 and 5/6 contours of the cumulative distribution). The data is coded according to source: ■ for the data tabulated in Weidner & Kroupa, ♦ for that obtained by Testi et al and ● for miscellaneous other observations (see Appendix 7.7).

One of the hardest aspects of constructing Figure 7.2 is the assignment of realistic errorbars (in n; errors in $m_{\rm max}$ are negligible by comparison, since we include only clusters which are young enough for their most massive members not to have expired as supernovae). We have drawn one-sided errorbars, on the grounds that we are probably missing stars that are located at large distances from the most massive star where the density of sources on the sky falls below the local background value, either as a result of initial conditions or dynamical evolution. We are interested here in the total population of stars that was formed with the most massive object, irrespective of whether these stars are currently bound to the natal cluster. Dynamical evolution in small n clusters can however cause significant expansion over a few Myr (Bonnell & Clarke, 1999) and this effect increases the likelihood that we may be missing stars at large distances. In order to estimate the possible error introduced in this way, we really need dynamical simulations on a cluster by cluster basis, which limit the range of original configurations (cluster n and size) that are compatible with the present census of background corrected objects. To our knowledge, this exercise has only been undertaken in one cluster (η Cha, Moraux et al., 2007) where the total n lies in the range of 18 (as observed)

[1] Note that the choice of this lower limit is arbitrary, provided that it is self-consistently applied to all the observational data and to the analytic predictions.

to 40 (the maximum number that is compatible with leaving a cluster with the parameters observed). With this in mind, we add one-sided errorbars of a factor of 2 in Figure 7.2, although note that this is pessimistic (i.e. too large) for the larger n clusters where the rate of dynamical dispersion is probably lower.

An obvious feature of Figure 7.2 is that different regions of the $\{n, m_{\rm max}\}$ plane are populated by clusters obtained through different observing strategies. Evidently, the squares are low compared with the centiles, explaining why Weidner & Kroupa found it necessary to invoke a non-random algorithm for stellar mass selection. The data of Testi et al. is apparently discrepant in the opposite direction — i.e. maximum stellar masses are, if anything, rather large, given the number of stars in the clusters. The reason why these two datasets are complementary (and both necessary to the statistical analysis) is simply one of the order in which the properties of the systems were determined. Weidner & Kroupa sought data on regions recognised as 'star clusters' and then found the maximum recorded mass. Testi et al. instead first identified massive stars (including those that are apparently isolated) and then undertook deep infrared imaging of the environs in order to identify any surrounding over-density of low mass stars. Unsurprisingly, the ratio of maximum stellar mass to cluster number is considerably higher in the latter case.

7.4 Analysis of the dataset

The observational data contained in Figure 7.2 is highly incomplete and biased and so great care must be taken in its statistical analysis. In this section we discuss whether a subset of the data can be used to settle whether the results are consistent with the expectations of the random drawing hypothesis. We here remind the reader that acceptance (rejection) of this hypothesis means that the high mass tail of the IGIMF should be identical to (steeper than) the input Salpeter IMF.

If we simply took all these datapoints at face value, we could evaluate a cumulative probability for each datapoint (i.e. evaluate the probability that the maximum mass is less than or equal to the datapoint value, according to the theoretical distribution for that particular value of n): if the observations match theoretical expectations, these probabilies should be uniformly distributed between 0 and 1. We can then use a KS test to compare the probability distribution with a uniform distribution. The probability that data generated through random drawing would be as discrepant from the theoretical prediction as is that observed is 10^{-17} (adopting the membership numbers denoted by symbols in Figure 7.2) and 10^{-8} (if one instead adopts twice these values, i.e. corresponding to the upper end of the errorbars shown in Figure 7.2). At face value, therefore, one would overwhelmingly reject the hypothesis of random drawing. The reason for the discrepancy (as can be seen from Figure 7.2) is that the lower range of the cumulative distribution function is actually *under*-populated by the data.

This conclusion is however highly misleading — the discrepancy is strongly driven by the very large number of datapoints in Figure 7.2 (from Testi et al.) which populate the upper regions of the cumulative distribution. However, it needs to be remembered that we simply do not have complete data.

One way forward is to define a stellar mass, $m_{\rm cpl}$, such that we deem that we have the information on all clusters (of all n), for which the maximum mass is $> m_{\rm cpl}$. By retaining all the data with $m_{\rm max} > m_{\rm cpl}$, however, we then have data which is complete down to different values of the cumulative distribution depending on the value of n, which is impractical for a

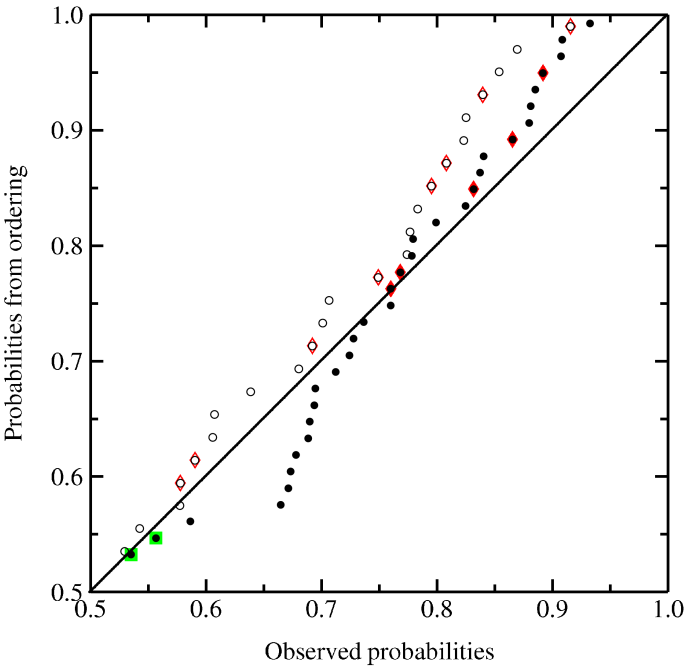


Figure 7.3: A plot of the ordered cumulative probabilities versus the observed cumulative probabilities derived from the distribution of $m_{\rm max}$ for a particular n. The shown data fullfill the selection criterion discussed in Section 7.4, i.e. they lie in the dashed wedge in Fig. 7.2. The filled symbols follow with the observed n, whereas for the open symbols $2n$ was used, corresponding to the right end of the error bars in Fig. 7.2 (■: Weidner & Kroupa; ♦ Testi et al.). If the data were uniformly distributed they should follow the diagonal.

statistical test and its interpretation. A more convenient approach is to select the data such that they are complete down to the same cumulative probability $P_{\rm cut}$. By this criterion all data points are selected which lie above the mass $m_{\rm cut}(n)$ which corresponds to $P_{\rm cut}$ in the cumulative distribution of $m_{\rm max}(n)$.

Furthermore, we do not take into account the very small n data, because of their presumably large error bars. Thus we introduce a minimum number of stars in a cluster, $n_{\rm cpl}$, which are needed to ensure the quality of the selected data. The selection parameter $P_{\rm cut}$ follows then as the cumulative probability of $m_{\rm cpl}$ for the given $n_{\rm cpl}$. The choice for the values of $m_{\rm cpl}$ and $n_{\rm cpl}$ is discussed below.

We demonstrate the selection criterion in Figure 7.2 for the case $m_{\rm cpl} = 3$ M$_\odot$, $n_{\rm cpl} = 30$, where the diagonal dashed line tracks the values of $m_{\rm cut}(n)$ with $P_{\rm cut} = 0.53$. By restricting ourselves to data within the wedge of the diagram above the dashed line, we are obviously not using a lot of the data, but have now defined a sample which is complete down to a fixed point in the cumulative distribution at every value of n included. We can therefore test whether these cumulative probabilities are uniformly distributed in the range $P_{\rm cut}$ to 1.

If we, for the moment, disregard any physical or observational ground for choosing particular values of $m_{\rm cpl}$ and $n_{\rm cpl}$, we can use different values of these parameters so as to explore various aspects of the two dimensional distribution. As expected, the KS probabilities are sensitive to the values of the $\{m_{\rm cpl}, n_{\rm cpl}\}$ adopted, with probabilities being higher if the choice is such (e.g. through high $m_{\rm cpl}$ or either very high or very low $n_{\rm cpl}$) that the number of datapoints retained is small.

We are particularly looking for evidence for a lack of data at the upper end of the cumulative distribution (as this would imply $m_{\rm max}$ was less than that implied by random drawing); we are however finding in general that this is not a striking feature of the distribution. In fact, a significant dearth of high $m_{\rm max}$ values is only recorded if one selects $m_{\rm cpl}$ and $n_{\rm cpl}$ so as to just include the clump of datapoints at around 10 M$_\odot$ and $n \approx 60$, since in this case the very top of the cumulative distribution is slightly under-represented (for example, for $m_{\rm cpl} = 9$ M$_\odot$ and

$n_{\rm cpl} = 60$, there are many more datapoints with cumulative probabilities slighly larger than $P_{\rm cut}$ ($= 0.75$) compared with those at higher values, and this is reflected in a relatively low KS probability of around 2%). In general, however, the feature of the plot that is generally flagged as most discrepant by the KS test is *not* a dearth of the highest values, but instead the lack of systems at around $3\ {\rm M}_\odot < m_{\rm max} < 10\ {\rm M}_\odot$ and $n \approx 50 - 100$, which is readily visible as a data hole in Figure 7.2.

It is obviously unsatisfactory if we tune the values of $m_{\rm cpl}$ and $n_{\rm cpl}$ so as to retroactively highlight a particular aspect of the data. Therefore, as our best guess of plausible parameter values, we adopt $m_{\rm cpl} = 3\ {\rm M}_\odot$ and $n_{\rm cpl} = 30$. The choice of $n_{\rm cpl}$ is motivated by the fact that we consider the errorbars in membership number of smaller clusters to be very high: both because dynamical evolution is more rapid in smaller n systems (Bonnell & Clarke, 1999) and also because the ejection of even a few stars to radii where they cannot be distinguished from the background causes relatively large fractional errors in n. The choice of $m_{\rm cpl}$ reflects the mass of a moderately luminous Herbig Ae star as targeted by Testi et al 1998. We have cautioned above that we must not go to much lower masses, since young stars of close to solar mass and below have not been systematically targeted for surrounding clusters.

The distributions derived from the data, for this choice of $m_{\rm cpl}$ and $n_{\rm cpl}$, are represented graphically in Figure 7.3. For each retained datapoint, we calculate the position in the theoretical cumulative distribution function (plotted on x-axis) and on the y-axis we plot the ranked position of the datapoint. (Note that this latter quantity has been renormalised to lie in the interval $[P_{\rm cut}, 1]$ instead of $[0, 1]$) The filled symbols use membership numbers denoted by the symbols in Figure 7.2, whereas the open symbols correspond to the case when values of n a factor two larger are adopted. Note that the number of datapoints are not the same in the two cases, since shifting n by a factor of two moves data values in and out of the region above the dashed line in Figure 7.2. If the data conforms to the random hypothesis, then the data plotted in Figure 7.3 should be following the diagonal.

A KS test performed on this data yields a KS probability of nearly 20%, implying quite adequate agreement with the random drawing hypothesis [2]. Although both curves in Figure 7.3 are somewhat above the diagonal at the uppermost end (implying a mild deficit of data values at the top of the predicted cumulative distribution function), this discrepancy is not significant in this sample (and, as discussed above, the feature in the filled curve that is most discrepant is actually the mild deficit of data around ≈ 0.6 in the cumulative distribution, corresponding to the 'data hole' at $m_{\rm max} \approx 3 - 10\ {\rm M}_\odot, n \approx 50 - 100$ in Figure 7.2).

7.5 Conclusions

We have high-lighted the difficulty in analysing the data contained in Figure 7.2 owing to difficulties to assigning regions of the diagram where the data is believed to be complete.

Nevertheless, our preliminary conclusion is that we are *not* seeing strong evidence for a systematic suppression in maximum stellar mass in small n clusters in addition to that expected on the basis of the statistics of random drawing (see also the complementary analysis of the statistics of isolated stars by Parker & Goodwin (2007), which reached similar conclusions). Indeed, if anything, the feature of Figure 7.2 that seems to be most discrepant with the

[2] We also tested the sensitivity of our results to the mapping employed between spectral type and mass by noting that if one uses the calibration of Martins et al. (2005), two of the O-stars in our sample are significantly reduced in mass. This adjustment however does not change the reasonable agreement with the null hypothesis.

random drawing model is the data hole in the range $m_{\max} \approx 3-10\ \mathrm{M}_{\odot}, n \approx 50-100$. We are however aware that this might indeed be filled in if we have under-estimated the incompleteness in smaller n clusters (particularly due to the effects of dynamical evolution).

Our conclusion (in support of the random drawing hypothesis) remains provisional. Although we have set out what we believe to be a statistically correct methodology for analysing the problem, we are highly aware of the difficulties of properly quantifying observational selection effects. We therefore seek further input from observers in compiling a good sample for this kind of analysis.

Acknowledgements

We thank Carsten Weidner, Leonardo Testi, Thomas Preibisch, Pavel Kroupa and the referee Hans Zinnecker for useful discussions and helping to collect the data. ThM acknowledges financial support via an EARA-EST fellowship and the European Union Research Training Network CONSTELLATION.

7.6 Appendix: The distribution of $m_{\rm max}$

The IMF used in this work is according to Kroupa (2001, 2002) with a Salpeter exponent for massive stars (Salpeter, 1955; Massey, 1998),

$$\xi(m) \quad \propto \quad \begin{cases} m^{-1.3} & m_{\rm MIN} \leq m < 0.5\ {\rm M}_\odot \\ m^{-2.35} & 0.5\ {\rm M}_\odot \leq m < m_{\rm MAX} \end{cases}, \tag{7.1}$$

where the lower limit $m_{\rm MIN} = 0.08\ {\rm M}_\odot$ and brown dwarfs are not included. Since we consider 'pure' random sampling the upper limit does not depend on the total mass of the star cluster ($m_{\rm MAX} = 150\ {\rm M}_\odot$). We use the IMF as a probability density, i. e. normalised to

$$\int_{m_{\rm MIN}}^{m_{\rm MAX}} \xi(m){\rm d}m \quad = \quad 1. \tag{7.2}$$

In a sample of "identical" star clusters (with the same n) the mass of the most massive star in each cluster will not be the same but follow its own distribution function. For this kind of Monte-Carlo experiment the distribution of the most massive star can analytically be derived. The probability for the most massive star to lie in the mass interval $m_{\rm max}, m_{\rm max} + {\rm d}m$ is

$$P(m \in [m_{\rm max}, m_{\rm max} + {\rm d}m]) = \xi(m_{\rm max}){\rm d}m. \tag{7.3}$$

All other stars must have a mass smaller than m. The probability to pick randomly $n-1$ stars from the mass range $m_{\rm MIN}, m$ is

$$P(m_{1\ldots n-1} \in [m_{\rm MIN}, m_{\rm max}]) \quad = \quad \left(\int_{m_{\rm MIN}}^{m_{\rm max}} \xi(m'){\rm d}m' \right)^{n-1} \tag{7.4}$$

The probability distribution of the most massive star is then the product of eqns. 7.3 and 7.4, multiplied with the factor n because every star could be the most massive star. To obtain the probability distribution the product has to be differentiated with respect to m. This gives

$$p(m_{\rm max}) \quad = \quad n \left(\int_{m_{\rm MIN}}^{m_{\rm max}} \xi(m'){\rm d}m' \right)^{n-1} \xi(m_{\rm max}) \tag{7.5}$$

7.7 Appendix: Data of the used clusters

In the compilation of Weidner & Kroupa (2006) the total mass of a cluster down to 0.01 $M_\odot$ is given. We converted this mass into a membershp number by dividing with the average stellar mass (0.36 $M_\odot$), and subtracted the expected number of brown dwarfs ($\xi(m) \propto m^{-0.3}$ for 0.01 $M_\odot \leq m < 0.08$ $M_\odot$) . In the cases where numbers are given we used them.

From the list of Testi et al. we took the I_C values, which are background-corrected total numbers and range down to the limiting magnitude $\mathscr{M}^c_{\mathscr{K}}$. One particular star/cluster, MWC 300 ($m_{\max} \approx 5$ $M_\odot$), is quite distant (15.5 kpc) and has a very high limiting mass ($\approx$ 2.4 $M_\odot$). Therefore the observed number of 21 stars is corrected to 740 stars. This is the far right outlier in Fig. 7.2, and the correctness of this data point is questionable, but included to be complete.

Name	$m_{\max}$	n	Ref	Name	$m_{\max}$	n	Ref
Taurus-Auriga	2.2	68	1	IRAS 05274+3345	10.0	18	9 10
Ser SVS 2	2.2	52	1	IRAS 05275+3540	16.0	179	9 10
ρ Ophiuchi	8.0	174	1	IRAS 05377+3548	14.0	39	9 10
IC 348	6.0	241	1	IRAS 05490+2658	10.0	58	9 10
IC 348	5.9	265	2	IRAS 05553+1631	10.0	82	9 10
NGC 2024	20.0	1447	3 4	IRAS 06056+2131	10.0	202	9 10
NGC 2024	20.0	392	1	IRAS 06058+2138	10.0	99	9 10
σ Orionis	20.0	392	1	IRAS 06068+2030	16.0	81	9 10
Mon R2	10.0	538	1	IRAS 06073+1249	16.0	412	9 10
Mon R2	10.0	1568	5 6	IRAS 06155+2319	14.0	48	9 10
NGC 2264	25.0	679	1	IRAS 06308+0402	16.0	45	9 10
NGC 6530	20.0	1421	1	MWC 1080	17.5	78	12
NGC 6530	80.0	2337	7	η Cha	3.5	18	13
Ber 86	40.0	2682	1	M20	27.0	196	14
USco	22.0	2859	8	IRAS 01546+6319	14.0	47	11
IRAS 00494+5617	14.0	105	9 10	IRAS 02044+6031	17.5	129	11
IRAS 02575+6017	14.0	210	11	IRAS 02232+6138	16.0	180	11
IRAS 02575+6017	14.0	254	9 10	IRAS 02245+6115	16.0	106	11
IRAS 02593+6016	19.0	77	11	IRAS 02461+6147	10.0	101	11
IRAS 02593+6016	17.5	93	9 10	MWC 137	17.5	96	12 15
IRAS 03064+5638	16.0	34	9 10	MWC 297	21.0	14	12
IRAS 05100+3723	17.5	203	9 10	NGC 7129	10.0	66	12
IRAS 05197+3355	16.0	170	9 10	NGC 7129	10.0	70	16

NGC 7129	10.0	82	17
NGC 7129	10.0	88	18
MaC H12	2.0	4	12
VX Cas	2.9	4	12
RNO 1B	5.0	8	12
XY Per	5.2	7	12 15
MWC 480	2.5	3	12
HD 245185	2.7	3	12
HD 37490	7.6	7	12 15
VY Mon	3.8	22	12
VV Ser	3.5	11	12
LkHα 257	3.8	5	12
HD 216629	10.0	26	12
BD+40 4124	10.0	13	12 15
V645 Cyg	30.0	134	12
IP Per	2.3	3	12
MWC 300	5.0	739	12
AS 310	17.5	147	12
HD 200775	7.6	1	12
LkHα 233	1.6	1	12
Elias 1	1.9	1	12
V1012 Ori	3.5	2	12
MWC 758	2.3	2	12
RR Tau	2.3	1	12
LkHα 208	2.3	2	12
BHJ 71	17.5	3	12
LkHα 215	4.6	4	12
HD 97048	3.4	6	17
BD+46 3474	17.5	209	17
BD+46 3471	7.3	9	17
RNO 6	14.0	17	12 15
HD 52721	10.0	21	12 15
HD 259431	5.9	1	12 15
LkHα 25	4.6	15	12 15
HD 250550	4.6	2	12 15
LkHα 218	3.5	2	12 15
AB Aur	2.9	2	12 15
T Ori	2.3	1	12 15
HK Ori	2.1	2	12 15
BF Ori	1.6	1	12 15

References:
1: Weidner & Kroupa (2006);
2: Luhman et al. (2003);
3: Lada et al. (1991);
4: Bik et al. (2003b);
5: Carpenter et al. (1997);
6: Carpenter (2000);
7: Prisinzano et al. (2005);
8: Preibisch et al. (2002);
9: Carpenter et al. (1993);
10: Carpenter et al. (1990);
11: Carpenter et al. (2000);
12: Testi et al. (1998);
13: Moraux et al. (2007);
14: Rho et al. (2001);
15: Testi et al. (1997);
16: Gutermuth et al. (2004);
17: Wang & Looney (2007);
18: Gutermuth et al. (2005);

Chapter 8

Properties of hierarchically forming star clusters

Th. Maschberger, C.J. Clarke, I.A. Bonnell & P. Kroupa
MNRAS 404:1061–1080 (2010)

We undertake a systematic analysis of the early (< 0.5 Myr) evolution of clustering and the stellar initial mass function in turbulent fragmentation simulations. These large scale simulations produce up to thousands of stars in clusters that can individually contain up to several hundred stars and thus for the first time offer the opportunity for a statistical analysis of IMF variations and correlations between stellar properties and cluster richness.
The typical evolutionary scenario involves star formation in relatively small-n clusters which then progressively merge; the first stars to form are seeds of massive stars and achieve a headstart in mass acquisition. These massive seeds end up in the cores of clusters and a large fraction of new stars of lower mass is formed in the outer parts of the clusters. The resulting clusters are therefore mass segregated at an age of 0.5 Myr, although the signature of mass segregation is weakened during mergers. We find that the resulting IMF has a smaller exponent (α =1.8–2.2) than the Salpeter value ($\alpha = 2.35$). The IMFs in subclusters are truncated at masses only somewhat larger than the most massive stars (which depends on the richness of the cluster) and an universal upper mass limit of 150 $\mathrm{M}_\odot$ is ruled out. We also find that the simulations show signs of the IGIMF effect proposed by Weidner & Kroupa, where the frequency of massive stars is suppressed in the integrated IMF compared to the IMF in individual clusters.
We identify clusters in the simulations through the use of a minimum spanning tree algorithm which is readily applied to observational data and which allows easy comparison between such survey data and the predictions of turbulent fragmentation models. In particular we present quantitative predictions regarding properties such as cluster morphology, degree of mass segregation, upper slope of the IMF and the relation between cluster richness and maximum stellar mass.

8.1 Introduction

Recent years have seen a proliferation of simulations of star and cluster formation involving a range of theoretical assumptions and physical ingredients (e.g. Bonnell et al., 2003; Schmeja & Klessen, 2004; Bonnell et al., 2004; Jappsen et al., 2005; Bate & Bonnell, 2005; Dale et al., 2005; Bonnell et al., 2006a; Dale & Bonnell, 2008; Bate, 2009b). Whereas the choice of model ingredients is set by a mixture of theoretical prejudice and numerical feasibility, it has already proved useful to undertake detailed comparisons between the output of such simulations and observational data. For example, the over-production of brown dwarfs in the original simulations of Bate et al. (2002) pointed to shortcomings in the treatment of gas thermodynamics which appears to have been largely remedied in subsequent simulations incorporating radiative transfer (Bate, 2009b).

The simulations of choice for the analysis of the larger scale clustering properties of stars are however those of Bonnell et al. (2003) and Bonnell et al. (2008) which, at the expense of being able to resolve the formation of the smallest objects, are able to follow the formation of hundreds of stars and track the hierarchical assembly of stellar clusters. Qualitatively, these simulations demonstrated how clusters grow through a combination of merging, the formation of new stars through fragmentation and the accretion of gas onto existing stars during cluster merging. Bonnell et al. (2003, 2004) were thus able to use these simulations in order to take a first look at how the mass of the most massive star in a cluster changes as the cluster grows through successive merger events.

In this paper we return to these simulations and their successors in order to analyse the properties of the resulting clusters and to take a more detailed look at issues such as the relationship between maximum stellar mass and cluster growth (the $m_{\rm max} - n_{\rm tot}$ relation; Weidner & Kroupa, 2004, 2006; Weidner et al., 2010; Maschberger & Clarke, 2008, this thesis Chapter 7), the degree of mass segregation (primordial vs. dynamical, cf. Bonnell & Davies, 1998; McMillan et al., 2007; Allison et al., 2009a) and other cluster diagnostics such as fractal dimension, ellipticity and slope of the upper IMF.

We here have the luxury of simulations which produce large numbers of stars: in particular, the large scale simulation discussed here produces thousands of stars, with individual clusters that contain up to hundreds of members. It thus becomes possible to analyse the *statistical* properties of the resulting ensemble. Apart from the superior statistics offered by the large scale simulation, the main difference between our analysis and the preliminary description given in Bonnell et al. (2004) is that we here identify subclusters through use of a minimum spanning tree technique, in contrast to Bonnell et al. (2004) who instead employed the ad hoc device of identifying a cluster as being all the stars within 0.1 pc of a massive star. The obvious advantage of our present analysis is that the clusters in the simulations are identified in precisely the same way as observers would extract clusters from maps of star forming regions and thus allows a much more direct comparison with observations (indeed, parameters such as cluster morphology, mass segregation and the cluster membership number, n, can only be explored if one has a generalised algorithm for defining clusters). This exercise is particularly timely given the accumulating survey data on stellar distributions in star forming regions (see the two substantial volumes on star forming regions edited by Reipurth, 2008a,b or the recent survey by Gutermuth et al., 2009); in particular, the use of Xray observations (for example of the ONC Getman et al., 2005; Prisinzano et al., 2008 or NGC 6334 Feigelson et al., 2009 and further regions mentioned in Feigelson et al., 2009) allows one to distinguish young stars from foreground/background sources and will this provide a good census of the

clustering properties of stars at birth.

The structure of the paper is as follows. In Section 8.2 we recapitulate the main features of the simulations to be analysed and in Section 8.3 describe the algorithm used for cluster extraction. In the following we describe the results for the cluster assembly history (Sec. 8.4), for the structure and morphology of subclusters (Sec. 8.5), for the locations of newly formed stars and for the initial mass segregation (Sec. 8.6) and finally for the initial mass function (Sec. 8.7).

8.2 Calculations

We analyse the data of two SPH simulations, the 10^3 $M_\odot$ simulation by Bonnell et al. (2003) and the 10^4 $M_\odot$ simulation by Bonnell et al. (2008).

The initial condition for the 10^3 $M_\odot$ simulations (Bonnell et al., 2003) is a uniform-density sphere containing 1000 $M_\odot$ of gas in a diameter of 1 pc at a temperature of 10 K, using 5×10^5 SPH particles. Supersonic turbulent motions are modelled by including an initial divergence-free, random Gaussian velocity field with a power spectrum $P(k) \propto k^{-4}$. The velocities are normalised such that the cloud is marginally unbound, and the thermal energy is initially 1% of the kinetic energy.

Protostars are replaced by sink-particles (Bate et al., 1995) if the densest gas particle and its ≈ 50 neighbours are a self-gravitating system (exceeding the critical density of 1.5×10^{-15} g cm^{-3}), sub-virial and occupy a region smaller than the sink radius of 200 AU. Accretion onto the sink particles occurs i) in the case of gas particles moving within a sink radius (200 au) and being gravitationally bound or ii) in the case of all gas particles moving within the accretion radius of 40 AU. The mass resolution for sink particles is ≈ 0.1 $M_\odot$. Gravitational forces between stars are smoothed at 160 AU.

For the 10^4 $M_\odot$ calculation (Bonnell et al., 2008) 10^4 $M_\odot$ of gas are initially distributed in a cylinder of 10 pc length and 3 pc diameter, with a linear density gradient along the main axis, reaching a maximum of 33% higher than the average density at one end, and 33% lower at the other. For computational reasons a particle-splitting method was employed (Kitsionas & Whitworth, 2002, 2007), which gives an equivalent of 4.5×10^7 SPH particles for the calculation, and a mass resolution of 0.0167 $M_\odot$. Turbulence is modelled using an initial velocity field with power spectrum $P(k) \propto k^{-4}$. For the whole cloud the kinetic energy equals the gravitational energy, which results in one end of the cloud being bound and the other unbound. The gas follows a barotropic equation of state of the form

$$P = k\rho^\gamma \tag{8.1}$$

where

$$\begin{array}{ll} \gamma = 0.75; & \rho \leq \rho_1 \\ \gamma = 1.0; & \rho_1 \leq \rho \leq \rho_2 \\ \gamma = 1.4; & \rho_2 \leq \rho \leq \rho_3 \\ \gamma = 1.0; & \rho \geq \rho_3 \end{array} \tag{8.2}$$

and $\rho_1 = 5.5 \times 10^{-19}$ g cm^{-3}, $\rho_2 = 5.5 \times 10^{-15}$ g cm^{-3} and $\rho_3 = 2 \times 10^{-13}$ g cm^{-3}.

Again, star formation is modelled via sink particles, with a critical density of 6.8×10^{-14} g cm^{-3}, a sink radius of 200 AU and an accretion radius of 40 AU. The smoothing radius for gravitational interactions is 40 AU, a quarter of that for the 10^3 $M_\odot$ calculation.

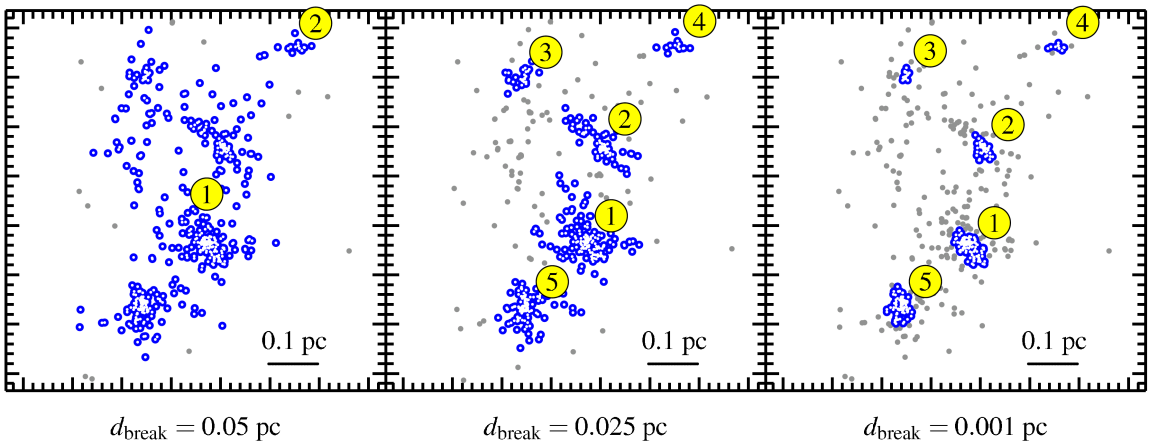


Figure 8.1: Influence of $d_{\rm break}$ on the detected subclusters (large dots), sink particles not in subcluster are shown as small dots. With $d_{\rm break} = 0.025$ pc the five detected subclusters have properties similar to a detection by eye. A too small $d_{\rm break}$ (0.01 pc) cuts off the lower-density outer regions. With a too large $d_{\rm break}$ (0.05 pc) only 2 subclusters are detected, with the larger one being highly substructured.

8.3 Cluster identification

For the identification of subclusters we employ a minimum spanning tree. The minimum spanning tree is a network of connections between points, not containing any closed loops, with the minimum possible total length of the connections (for the relation between the minimum spanning tree and clustering identification, the properties of the minimum spanning tree in general and algorithms for the construction see e.g. Zahn, 1971). The minimum spanning tree and and its properties have previously been used to determine the level of substructure in a star cluster, e.g. the Q measure of structure by Cartwright & Whitworth (2004) or the Λ measure of mass segregation by Allison et al. (2009b). A minimum spanning tree does not only characterise the degree of substructure, but can also be used to identify the sub-clusters themselves. A clustering-algorithm based on the minimum spanning tree has the advantage that the subclusters can have arbitrary shapes, that small-n subclusters can be found and only one parameter, the break distance $d_{\rm break}$, needs to be specified.

Once the minimum spanning tree containing all sinks has been constructed, subclusters can be identified by splitting the global minimum spanning tree into sub-trees by removing all edges which have a length larger than $d_{\rm break}$. The break distance can be related to a minimum density of points per area which is required that groups remain connected. The remaining sub-trees are then identified as a subcluster if they contain more than $n_{\rm min} = 12$ sink particles. Sinks of subtrees with a smaller n are attributed to the "field". To each subcluster we assign an identification number which is unique to the most massive sink particle in it. Sometimes it can occur that another sink particle in the same physical subcluster has accreted so much that it takes over the position as the most massive particle. In this case we assign a new identification number to the cluster.

The clustering algorithm using the minimum spanning tree is not scale-free, as a particular length scale, $d_{\rm break}$, is needed. The choice of $d_{\rm break}$ is somewhat arbitrary, as experiments

following ideas by Zahn (1971) to determine a reasonable $d_{\rm break}$ self-consistently from e.g. the edge length distribution gave no robust scale-independent criteria. Thus we chose $d_{\rm break}$ such that the subclusters found by the clustering algorithm have properties similar to subclusters which are selected by eye. For the effects of various $d_{\rm break}$ we analysed the 10^3 $M_\odot$ data set with $d_{\rm break} = 0.01$ pc, 0.025 pc and 0.05 pc and show a snapshot made at 3×10^5 yr in Fig. 8.1. Clearly one sees that too large a value of $d_{\rm break}$ (0.05 pc) identifies objects as subclusters that themselves contain considerable substructure. On the other hand, a very small $d_{\rm break}$ cuts off low-density regions of the actual subclusters. We found that $d_{\rm break}$ = 0.025 pc gives the best results, as all reasonably rich subclusters are detected with a sufficient quantity of their low-density outskirts, and mergers do not occur prematurely. We however emphasise that the main utility of this approach is that it allows one comparisons with observations that are analysed with the same value of $d_{\rm break}$.

As by observations only a *projection* is available, we also project the simulation data onto two dimensions, for both calculations in the *x*-*y* plane. To exclude projection effects causing artefacts in the results we did all our analyses in other projections as well (*x*-*z* and *y*-*z*). Different choices of the plane of projection do not affect our results qualitatively, and not significantly quantitatively.

8.4 Cluster assembly history

We start our analysis of the simulations by constructing the merging history of the subclusters and the general properties of the simulation, such as the evolution of the total number of sinks and their total mass.

The overall evolution of the two simulations is illustrated by Figure 8.2, showing the projected distributions of the sink particles at different times. The small scale simulation (top row) simply demonstrates a history of hierarchical merging, with the final outcome being the creation of a single merged entity and a smaller population of sinks that are identified as 'field stars' by our clustering algorithm. The bottom row shows the global evolution of the large scale simulation: as is consistent with globally unbound state of this simulation, one sees that merging does not go to completion and that there are instead regions of local merging and a pronounced field population in between. On the other hand, when one homes in on a dense region of this large scale simulation (the box shown in the lower panels) we see (middle row) an evolutionary sequence that is very similar to that shown in the small scale simulation (top row). In general terms we will find in all our subsequent analysis that significant differences between the two simulations all relate to parameters that take into account the dispersed population and the survival of multiple clusters in the larger (unbound) simulation.

8.4.1 Merging history

Figure 8.3 depicts merger trees for cluster assembly. Each subcluster is denoted by the identification number of its most massive sink particle, with a symbol size corresponding to the number of sinks in the subcluster. The arrows at the end of a lifeline correspond to merger events where the merged subcluster is given the identification number of the subcluster that had previously contained the most massive member of the new combined entity. The upward pointing arrows connecting the end of one lifeline with the start of a new one correspond to cases where the identity of the most massive sink particle changes (i.e. one sink overtakes

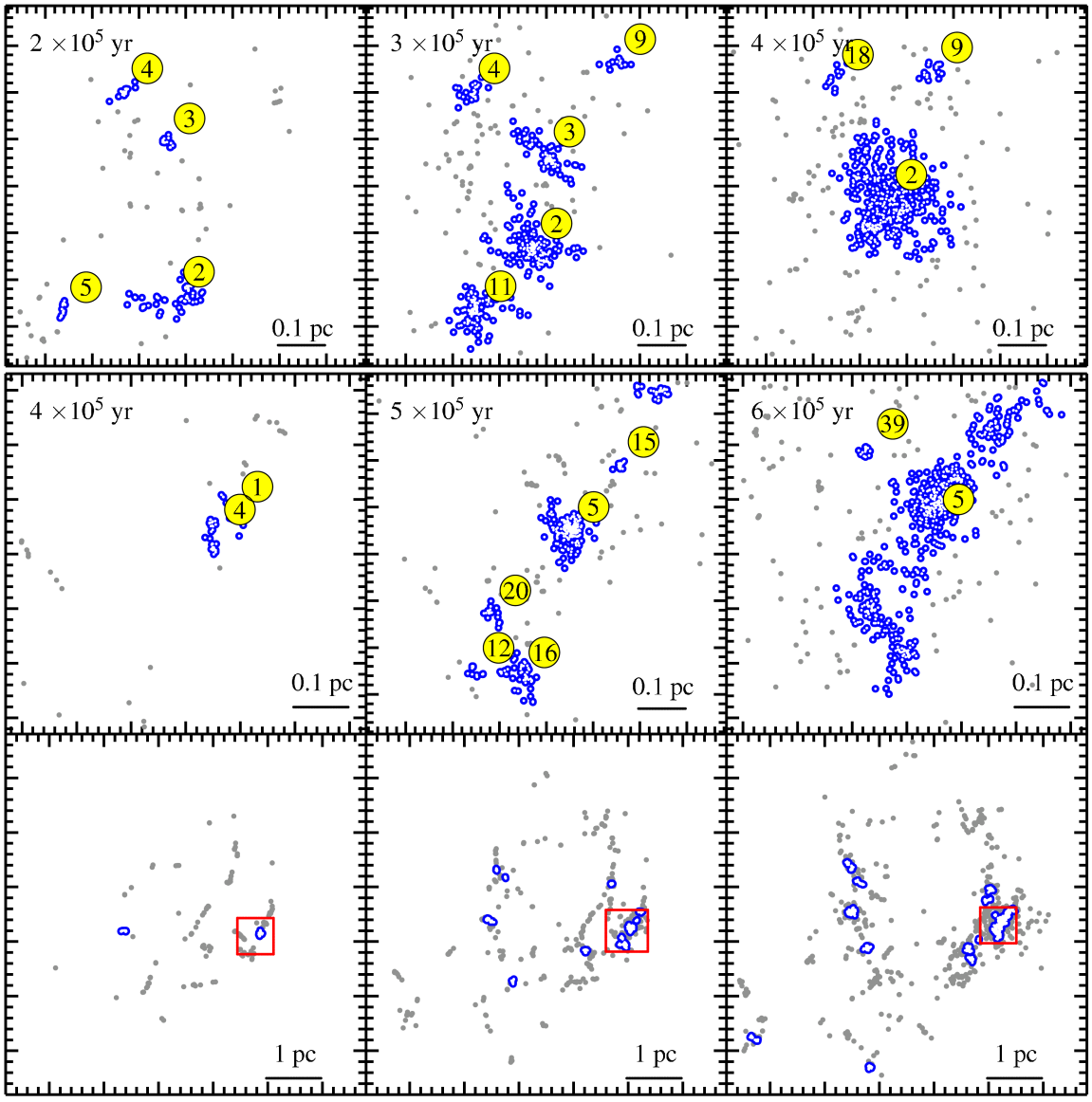


Figure 8.2: Time-evolution of the projected spatial distribution of the sink particles, with large dots representing sinks in subclusters (whose labels correspond to the identification numbers in Fig. 8.3) and small dots for "field" sinks. The snapshots start at different global times of the two calculations, but at similar structures. The top row shows the central 0.6×0.6 pc of the 10^3 $M_\odot$ calculation, the middle row the corresponding section in the 10^4 $M_\odot$ calculation (large ticks = 0.1 pc). In the bottom row displaying the whole area of the 10^4 $M_\odot$ calculation (6×6 pc, large ticks = 1 pc) a box marks the location of the detail section.

another in mass as a result of accretion). The subcluster is then assigned a new identification number (and lifeline), but this is only a re-labelling. Subclusters that are registered as subclusters on less than five occasions do not appear on this plot. We also see occasional gaps in the lifelines of particular subclusters: these are usually small or low density subclusters

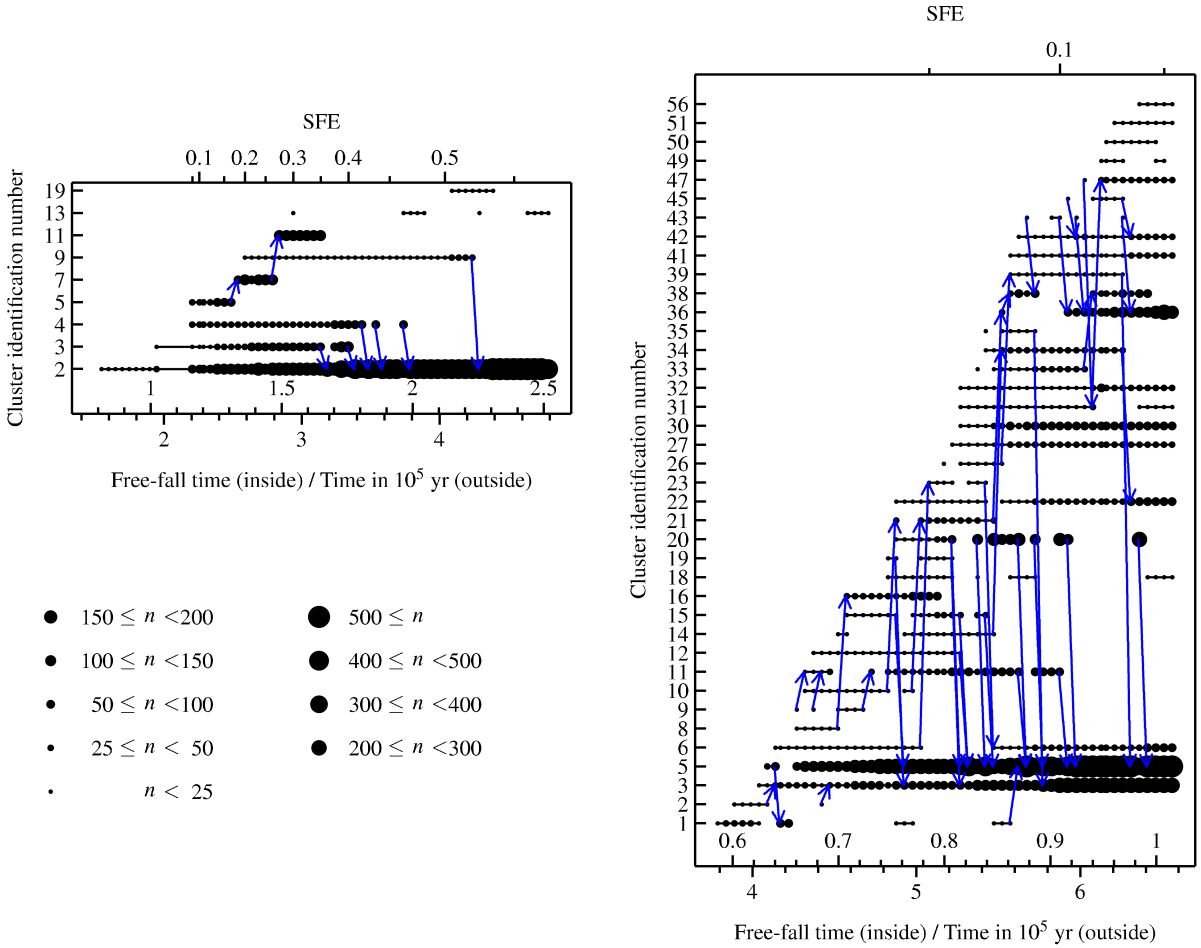


Figure 8.3: Merging history of the subclusters, left for the 10^3 $M_\odot$ and right for the 10^4 $M_\odot$ calculation. Each dot marks the detection of a subcluster, with the size of the dot scaling with the richness of the subcluster (only subclusters which have been detected more than 5 times are shown). Arrows at the end of a lifeline mark mergers of subclusters, or, if they point to the beginning of a new lifeline, a change of the most massive sink particle as the most massive sink is overtaken by another.

where the relatively modest rearrangement of its members due to few body dynamical effects changes whether or not the grouping is classified as a subcluster.

Depending on the size of the subclusters involved, it can take up to $\approx 5 \times 10^4$ yr for a merger to produce a single, stable new structure, as can for example be seen from the sporadic detections of subcluster # 4 during its merger with # 2 in the 10^3 $M_\odot$ simulation. Fellhauer et al. (2009) investigated the time scales for mergers of a spherically symmetric distribution of subclusters embedded in a background potential (typically more than $\approx 5 \times 10^5$ yr for systems comparable to ours). The time scale for mergers we find are perhaps somewhat quicker than theirs, as the subclusters are not distributed isotropically but along filaments, which also direct their motion.

Overall, Figure 8.3 describes a situation of hierarchical merging; in the small simulation the system evolves towards a single merged entity whereas in the large simulation (which is

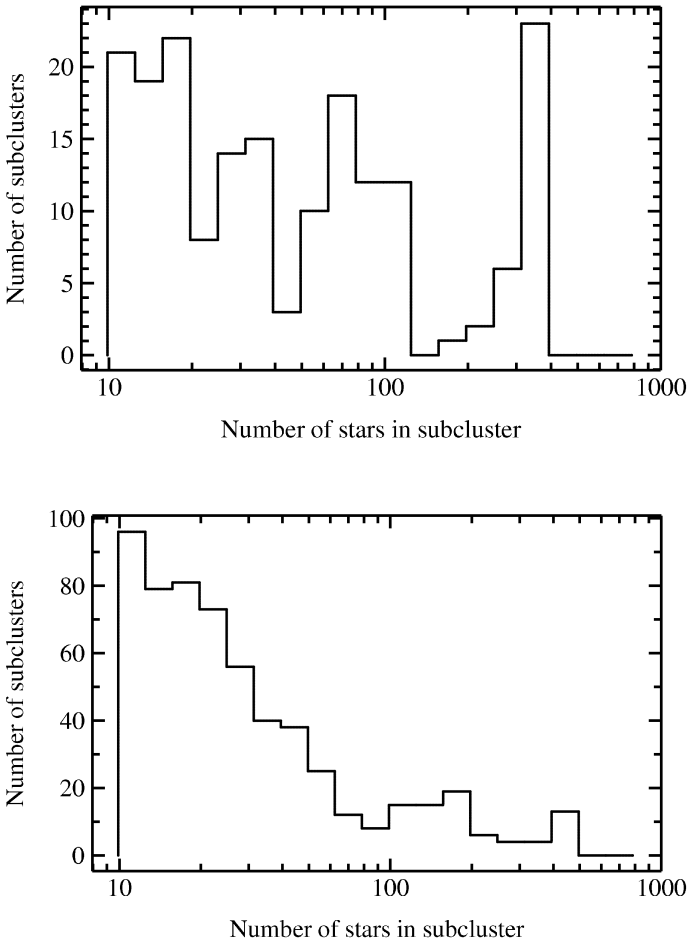


Figure 8.4: Histogram of the subclusters of the composite population (i.e. all subclusters of all time steps) by their number of sinks, for the 10^3 $M_\odot$ (top) and the 10^4 $M_\odot$ (bottom) calculation. The peak at $n = 10^{2.5}$ in the left hand panel corresponds to the formation of a long-lived central cluster in the 10^3 $M_\odot$ simulation, similar to the large-n peaks of the 10^4 $M_\odot$ calculation. Note that these distributions do not correspond to what would be seen in a single snapshot in time, for this see Fig. 8.5.

globally unbound) the system is tending to several merged structures which (from inspection of the simulation) are unlikely to undergo further merging. We note that the change of identity of the most massive sink particle in a cluster occurs relatively frequently. This is rather surprising in the case of a power law mass distribution: in this case the expected spacings in mass between sinks are relatively large and it is not expected that differential accretion would cause one sink to overtake another. In fact, we shall see later that the masses of the most massive sink particles in a cluster are rather well correlated so that relatively minor changes in accretion history can change the identity of the most massive member.

8.4.2 Cluster population

In later Sections we will look at various properties of the subclusters, as for example their shape, mass segregation etc. In an individual time step the number of detected subclusters is not very large, therefore we sometimes use the subclusters from *all time steps together* for the analysis, which we term the *'composite population'*. As they can be at different stages of evolution one has to be careful when interpreting the results.

Figure 8.4 shows a histogram of subclusters in the composite population by their number of sink particles. The composite population is dominated by rather small clusters ($n < 30$–50) which are usually very young subclusters ($< 10^5$ yr since their first detection), or subclusters

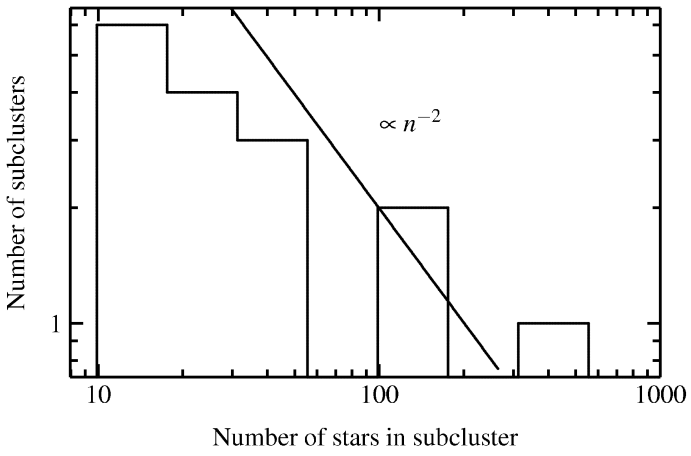


Figure 8.5: Number spectrum of subclusters at the end of the 10^4 $M_\odot$ simulation. We show for comparison a line corresponding to a number spectrum $\propto n^{-2}$.

which have never merged (compare with the merging history, Fig. 8.3, where the symbols' sizes reflect the number of sinks). The large-n peaks in the 10^3 $M_\odot$ histogram is produced by the formation of a cluster of ≈ 300 sinks which, being long-lived, appears in many time steps. We emphasise that the distributions in Fig. 8.4 are provided in order to interpret results based on the composite population and should *not* be interpreted as spectra of cluster richness at a given time.

In order to get an idea of the latter we plot in Fig. 8.5 a histogram of the cluster number spectrum at the end of the 10^4 $M_\odot$ simulation. For comparison we show the n^{-2} spectrum found by Lada & Lada (2003) for the embedded star clusters in the Milky Way (with $m_{\rm cluster}$ between 50-1000 $M_\odot$).

8.4.3 Build-up of stellar number and mass

Figure 8.6 shows that the fraction of all sinks formed by a given time rises more steeply by number (solid curves) than by mass (dotted curves, both normalised to the total number or mass at the end of the simulation). Later on, fewer new sinks are formed but all accrete mass so that the mean stellar mass increases during the simulation (and hence, by implication, the mass function evolves during the simulation). The thin curves in Figure 8.6 refer to the sinks that are classified as being in subclusters at any time (also normalised to the total number or mass at the end of the simulation): they start to increase later than the thick curves (for all sinks) what shows that the classification of clusters is delayed with respect to formation of the first sinks. This can be seen more directly in Figure 8.7, which shows that, after an initial delay, the fraction of sinks in clusters rises to 60–80% (note that the fraction of sinks in clusters is higher in the bound simulation, as expected). The initial delay is comprehensible since we imposed a minimum cluster membership number of 12; the first sinks form in small-n clusters that do not register as clusters until they have acquired enough members by cluster merging. In the 10^3 $M_\odot$ simulation the fraction of sinks in subclusters reaches a maximum and then decreases slightly, which is caused by dynamical evolution.

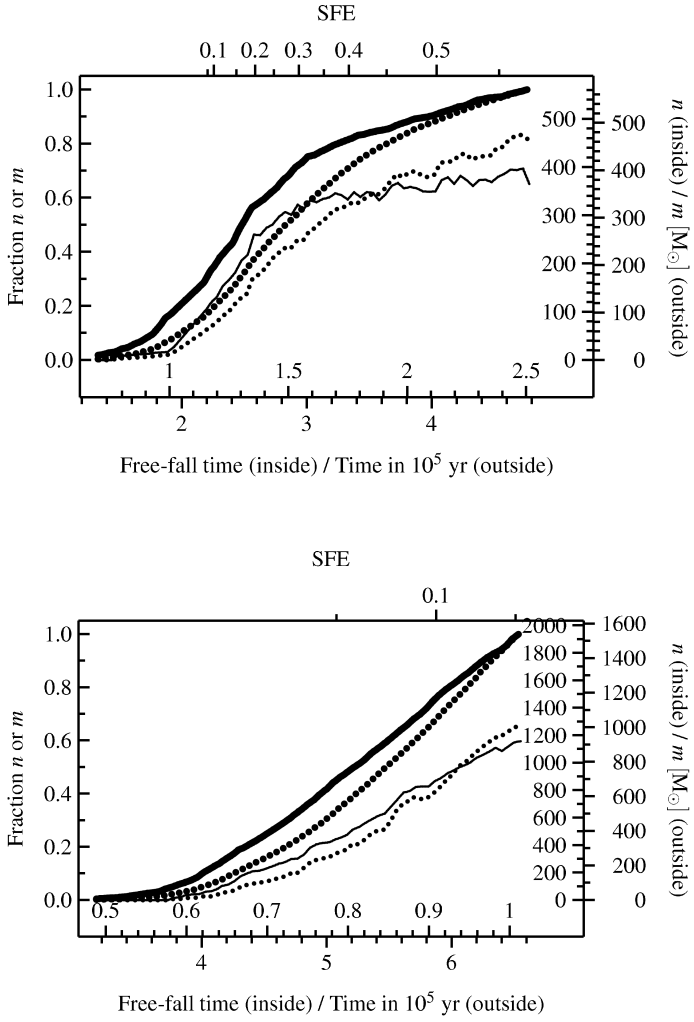


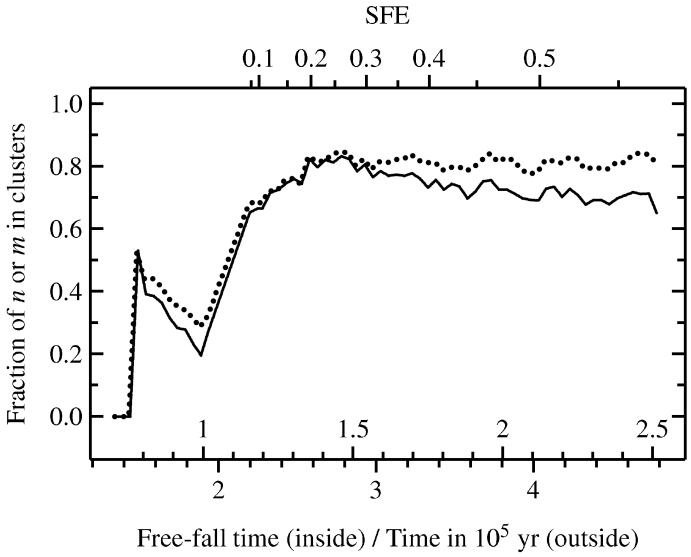
Figure 8.6: Assembly by number (solid) and mass (dotted) for the whole system (thick symbols) and for all sinks in subclusters (thin symbols), respectively, normalised to the total number/total mass of all sinks at the end of the simulation. The top panel is for the 10^3 $M_\odot$, the bottom one for the 10^4 $M_\odot$ calculation.

8.5 Cluster structure and morphology

8.5.1 Structure

Figure 8.8 illustrates the effect of the cluster merging history on a structural parameter of the stellar distribution. Here we use the Q parameter, introduced by Cartwright & Whitworth (2004), which is defined as the ratio of the mean edge length in the minimum spanning tree to the correlation length of the the stellar distribution. Fig. 8.8 shows the time-evolution of the Q parameter for the whole simulation (big dots) and for individual subclusters containing a minimum number of 48 sinks (lines). As discussed by Cartwright & Whitworth (2004) small values of this parameter (< 0.8) correspond to fractally distributed points (the small value reflecting the fact that the existence of multiple nuclei tends to increase the correlation length more than the mean edge length). On the other hand, higher Q values correspond to centrally concentrated distributions, with the Q value rising with the degree of central concentration.[1]

[1]We do not correct our interpretation of Figure 8.8 for the fact that our stellar distributions are not spherically symmetric, since Cartwright & Whitworth (2008) found that such corrections were negligible for aspect ratios less than ≈ 3; we show below that extreme ellipticities are rare in our data. In the normalisation a geometrical factor is implicitly contained by choosing a circle as circumference for the uniform distribution, as in Cartwright



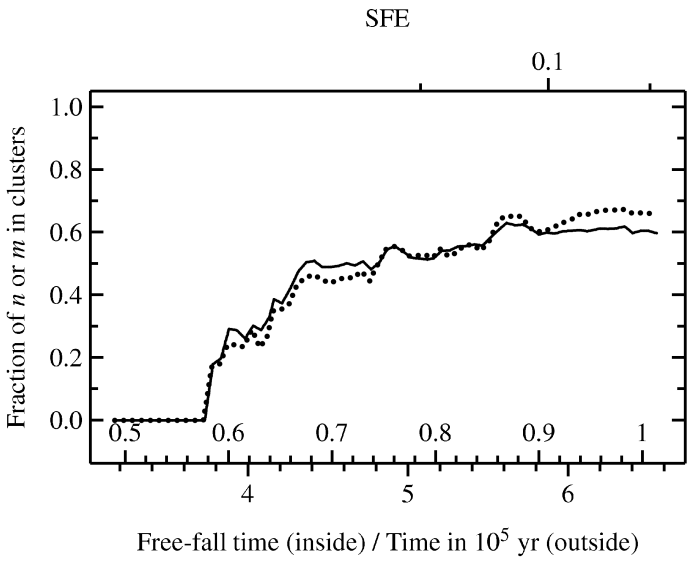


Figure 8.7: Evolution of the fractions of sinks in subclusters at a given time by number (solid) and mass (dotted), for the 10^3 $M_\odot$ calculation (top) and the 10^4 $M_\odot$ calculation (bottom). The fraction in subclusters increases by a mixture of sink formation within the subclusters and the accretion of isolated sinks or small groupings onto the subclusters. The modest decrease in the fraction of sinks in subclusters at late times in the 10^3 $M_\odot$ calculation results from the formation of one large cluster with a low-density halo.

Figure 8.8 shows that in the small simulation, the total stellar distribution is characterised by monotonically increasing Q values, indicating the formation of a single centrally concentrated cluster through hierarchical merging. The recovery from a substructured subcluster to a radially concentrated system occurs over about $0.5–1.0 \times 10^5$ yr, which can be seen as the time for a merger. The large simulation remains in the fractal regime throughout, since (being globally unbound) it retains a multiply clustered structure. In both simulations, the Q values of individual clusters fluctuate, exhibiting periods of increase (as isolated clusters become more centrally concentrated as a result of two body relaxation) followed by abrupt reductions of Q into the fractal regime during episodes of cluster mergers. The range of Q values that we recover from our whole simulations is similar to that found in observations by Cartwright & Whitworth (2004) and Schmeja & Klessen (2006), where fractal dimension as

& Whitworth (2004) (Schmeja & Klessen, 2006 instead use the convex hull of the data set).

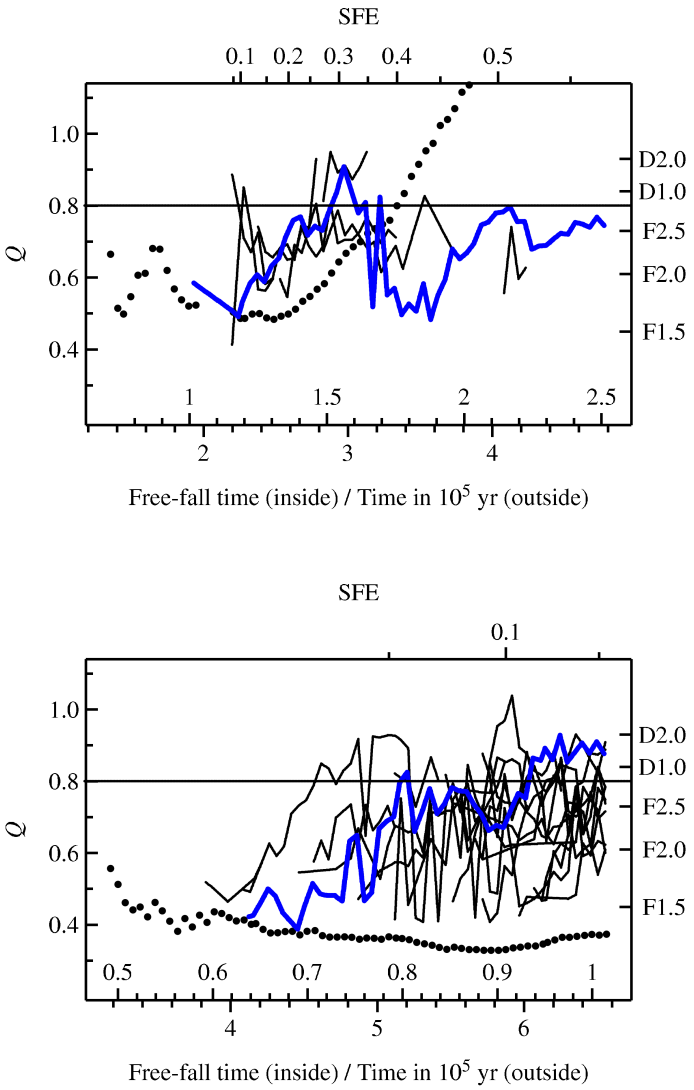


Figure 8.8: Time-evolution of the Q parameter, see Section 8.5.1. The horizontal line marks $Q = 0.8$, which corresponds to a uniform distribution (radial exponent = 0, $D0.0$, or fractal dimension = 3, $F3.0$). Fractally subclustered systems have $Q < 0.8$ and radially concentrated systems $Q > 0.8$. The fractal dimension (F) and the radial exponent (D) can be read off the right axis. The whole system (big dots) starts fractal and evolves towards a centrally concentrated system in the bound 10^3 $M_\odot$ calculation (top panel) and stays fractal in the unbound 10^4 $M_\odot$ calculation (bottom panel). The subclusters (lines) evolve in both calculations towards concentrated systems when they are not disturbed. Mergers lead to the more or less pronounced jumps towards smaller Q. The thick line is for the richest subcluster that is formed in each of the calculations.

low as 1.5 ($Q = 0.47$) are found for Taurus and radial concentrations following $r^{-2.2}$ ($Q = 0.98$) for IC 348. The Orion Nebula Cluster has $Q = 0.82$ (considering only stars; Kumar & Schmeja, 2007). Schmeja et al. (2008, 2009) derived Q in subclusters identified within larger regions (Perseus, Serpens, Ophiuchus and NGC346) and obtained values of $0.59 \leq Q \leq 0.93$.

In the 10^3 $M_\odot$ simulation the Q parameter reaches values of ≈ 1.4 at the end of the calculation, which implies a very steep radial density following r^{-3}, but the central subcluster appears to have a uniform density. In order to resolve this apparent contradiction we investigate the density profile of the whole system at the end of the simulation. For power-law distributed data the cumulative distribution function provides a convenient way of visually assessing all available data without the need of grouping them as in a histogram. The probability density of a power law distribution from l to ∞ is given by $p(x) = -\frac{1-\alpha}{l^{1-\alpha}} x^{-\alpha}$, and the cumulative density is $P(x) = 1 - \frac{x^{1-\alpha}}{l^{1-\alpha}}$. Therefore, a plot of $\log(1 - P(x))$ (the logarithm of the complementary cumulative density) vs. $\log x$ should be a straight line. We show such a plot for the data in Figure 8.9. The radial density distribution does not follow a straight line but falls into three segments, a flat/uniform central region, a main region from 0.1 pc to 1 pc pro-

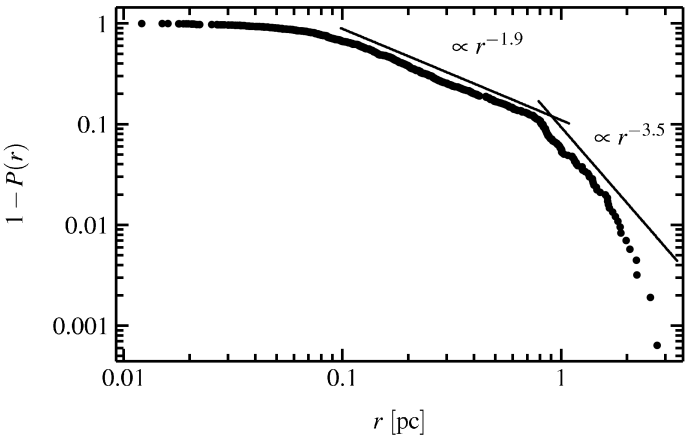


Figure 8.9: Double-logarithmic plot of the complementary cumulative radial density, $1-P(r)$, against distance measured from the geometrical cluster centre containing all sinks at the end of the 10^3 $M_\odot$ calculation. Power-law distributed data follow straight lines in this kind of plot.

portional to $r^{-1.9}$ and an outer halo having $r^{-3.5}$ or even a larger exponent. The halo is formed by low-mass sinks which have left the main region due to dynamical interactions (an effect which is also responsible for the decreasing fraction of sinks in subclusters in Fig. 8.7). Most of the mass in this merged cluster is contained in a region whose density profile is close to the isothermal $\rho \propto r^{-2}$ profile. We thus see that the Q parameter method of estimating the radial exponent is unduly influenced by the steeper distribution in the halo.

8.5.2 Morphology

In Figure 8.10 we plot a histogram of the ratio of the projected major axis to projected minor axis for our clusters. This quantity has been derived by fitting a two-dimensional normal distribution to the projected number density distribution. The eigenvalues of the covariance matrix then give an elliptical contour of equal values of probability density containing $\approx 30\%$ of the sink particles.[2]

We see in Figure 8.10 that most clusters are mildly elongated: the distribution peaks at 1.5 and most clusters have an axis ratio of less than 2. Subclusters form in dense nodes along the filaments of gas, as dense small-n systems which shortly after their formation attain a spherical shape, which gives the peak in Fig. 8.10. One filament can contain several subclusters, so that the distribution of subclusters is elongated, but not the subclusters themselves, as visible in the snapshots in Fig. 8.2. During a merging event the resulting object is naturally elongated, leading to the tail of large ellipticities in Fig. 8.10. An example is the cluster with #5 in the 10^4 $M_\odot$ simulation at 6×10^5yr, which has an ellipticity of 3.86 and $Q = 0.46$ (see the middle right panel in Fig. 8.2 for the projection) and is currently merging with cluster # 20.

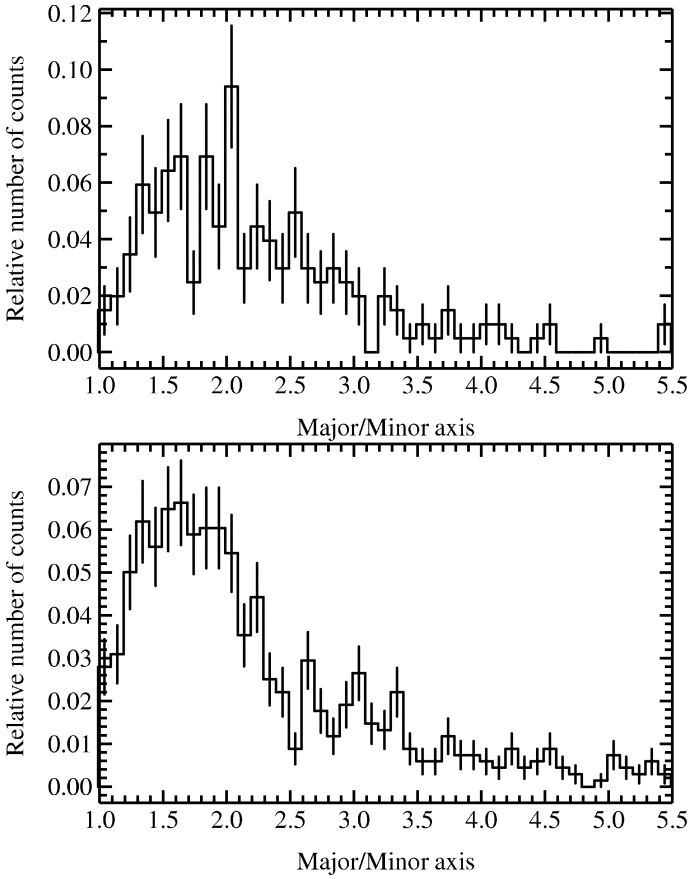


Figure 8.10: Histogram of the ellipticities of the subclusters (derived from fitting a 2D Gaussian distribution), using the composite population of subclusters (from all times). The top panel shows the result for the 10^3 $M_\odot$ calculation and the bottom panel for the 10^4 $M_\odot$ calculation.

8.6 Formation sites of stars and (primordial?) mass segregation

8.6.1 Formation sites of stars

It has already been mentioned in Bonnell et al. (2004) that sinks do not necessarily form close to the centres of existing clusters (with the centre defined using the most massive sink particle, an assumption we test below). With our definition of a subcluster we find that only 50–60% of all sinks form within a subcluster. Within the subclusters the distribution of the formation sites follows the same distribution as existing sinks in the subclusters (with only a very mild concentration towards the inner region), as visible in the histogram of the radial ranking (Figure 8.11). The sinks forming outside of subclusters form either in the immediate neighbourhood of a subcluster or as the centres of new subclusters.

Significantly, we find that the most massive sink particles avoid formation within existing subclusters: indeed virtually no sinks which end up with masses > 1 $M_\odot$ form within the half-number radius of an existing cluster. It is thereby more correct to say that *clusters form around (seeds of) massive stars* than massive stars form in clusters.

[2]Note that - in contrast to some previous algorithms for deriving cluster shapes - we are not unduly sensitive to the locations of the outermost points in the dataset (cf Schmeja & Klessen, 2006; Cartwright & Whitworth, 2008). This can be particularly problematical since the definition of clusters through splitting a minimum spanning tree can lead to 'hairs' at the end of the cluster (sub-trees that reach out of the cluster body and have no branches) and so it is important to avoid an algorithm that gives undue importance to these outlying protrusions.

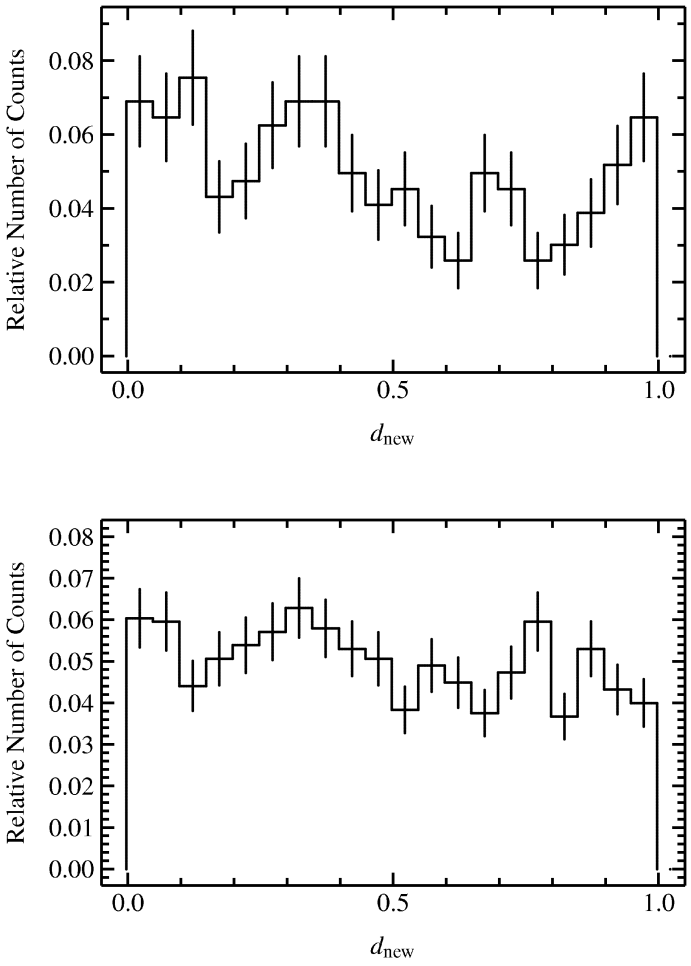


Figure 8.11: Histogram of the fractional radial ranking of newly formed sinks in the subcluster to which they are assigned, measured at the time of formation (top 10^3 $M_\odot$, bottom 10^4 $M_\odot$ calculation). Sinks which are born in the field ($\approx$ 30–40% of all sinks) are not included.

8.6.2 Development of mass segregation

We now turn to the question of where stars of various masses end up within the subclusters (as opposed to where they form). We emphasise that since the entirety of the simulations correspond to the deeply embedded phase (age < 0.5 Myr) then even the *final* state of the simulations can be used to assess what is usually termed primordial mass segregation.

We have looked at a variety of mass segregation diagnostics and find that mass segregation usually applies to the ten to fifty most massive sinks. For example, cumulative radial distributions within clusters for stars in different mass bins rarely reveal consistent evidence for mass segregation apart from its existence in some clusters which are spherically symmetric. Bate (2009a) finds no mass segregation in his data using cumulative distributions whereas Moeckel & Bonnell (2009) using their (non-parametric) technique find mass segregation in the same data.

We use the Λ measure of Allison et al. (2009b) which is based on the minimum spanning tree and allows one to detect mass segregation also if only a few stars are involved. For the ith most massive star it is defined as

$$\Lambda_{(i)} = \frac{\overline{l_i}}{l_{(i)}} \pm \frac{\overline{\sigma_i}}{l_{(i)}}. \tag{8.3}$$

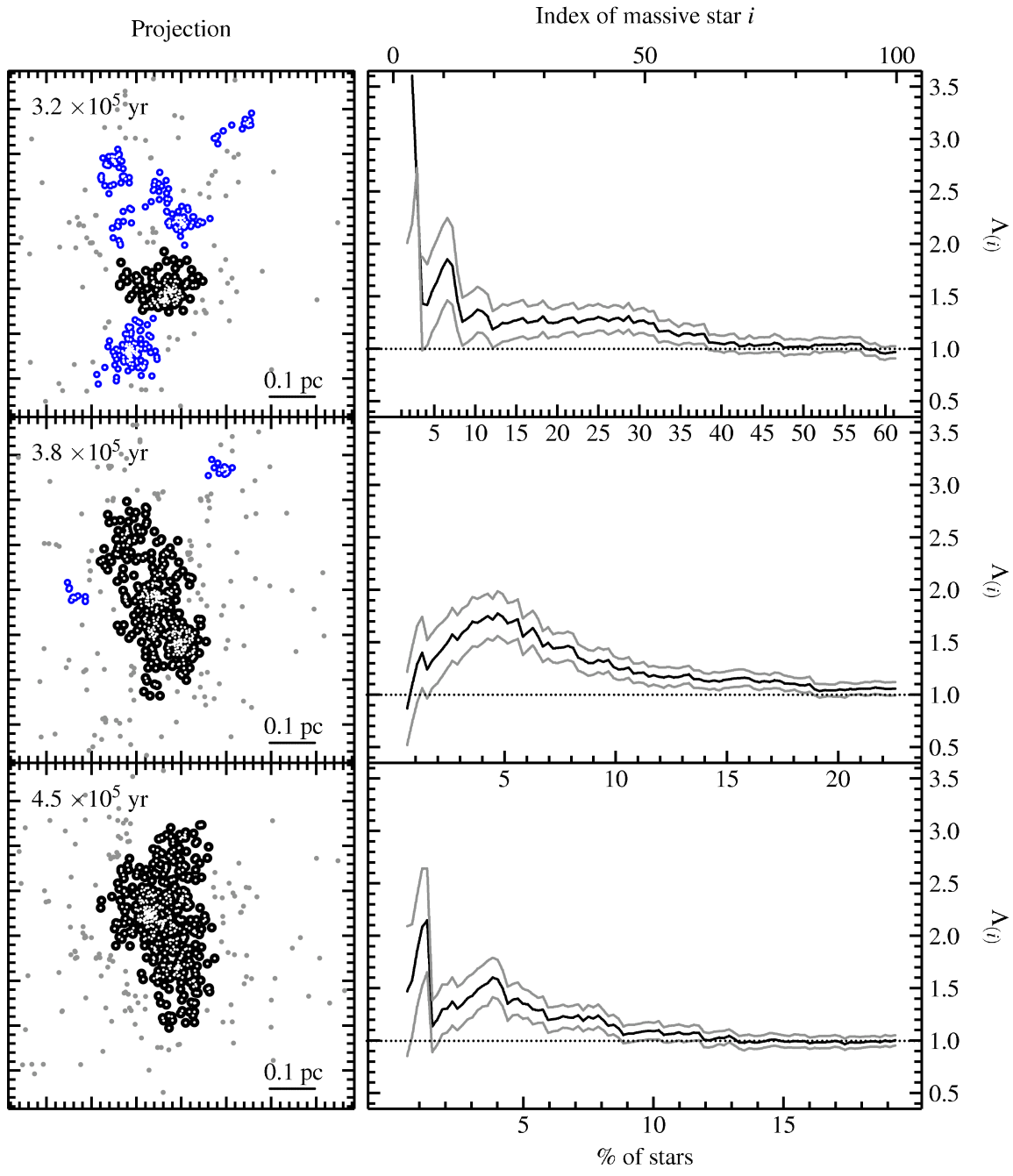


Figure 8.12: Evolution of mass segregation for a particular subcluster during a merging event (# 2 in the 10^3 $M_\odot$ calculation). The left panels show the projection distribution of the sink particles at the snapshots with the analysed subcluster marked with black dots. The right panels display the Λ measure (eq. 8.3, Allison et al., 2009b) for the 100 most massive sinks. The index of the sinks can be read off the top axis of the uppermost panel and is the same throughout. As the subcluster grows in number we show the percentages for the massive sinks of the total number at the bottom axis of each panel. Before the merger (top row) the subcluster is already mass segregated, Λ is larger than unity for the ≈ 60 most massive sinks (40%). During the merger (middle panel) the ≈ 15 most massive sink particles are not mass segregated as they are still in the centres of the merging subclusters, but not randomly distributed (Λ exceeds unity). After the merger (bottom row) the ≈ 10 most massive sinks quickly reach a state of strong central concentration (large Λ) and general mass segregation is at a 10% level.

$\overline{l_i}$ and $\overline{\sigma_i}$ are the mean length and its standard deviation of a minimum spanning tree constructed from a sample of i stars which are randomly drawn from the total sample of stars in the subcluster. $l_{(i)}$ is the length of the minimum spanning tree containing the i most massive stars. $\Lambda_{(i)} = 1$ means that the i most massive stars are distributed as the other stars and there is no mass segregation. Mass segregation is detected if Λ is significantly larger than unity (in terms of standard deviations), and the absolute value of Λ reflects the degree of spatial concentration (i.e. the larger Λ the more spatially concentrated). Λ has the big advantage of being non-parametric, i.e. knowledge about the shape or density profile is not necessary.

The typical states of mass segregation in a rich subcluster are shown in Fig. 8.12, which follows the time evolution of mass segregation in an individual cluster during a merging event. The left panel shows the projected spatial distribution and the right panel Λ. A subcluster that has never undergone a merging event or had a merging event a long time ago (top panel) shows a monotonic decrease of Λ extending over a large fraction of the massive sinks: in our example about 40 per cent of all sinks (by number) are significantly segregated. The snapshot is taken just before a number of subclusters will merge into the analysed subcluster. During the merger (middle panel) the merging clusters are gradually dissolved and incorporated in the merger product, so that for some time the detected subcluster actually has multiple centres. These centres still hold the massive sinks, so that they are spatially more widely distributed than a random sample of sinks, which will contain mostly sinks from the richest previous cluster. However, as soon as with an increasing random sample size sinks are also chosen from the other centres, the massive sinks show a concentration within these centres. This explains the typical behaviour of Λ during a merger, which is increasing from unity for the ≈ 10 most massive sinks until it reaches a maximum, in our example at 5% of the sinks, from which it gradually decreases again. The total percentage of sinks that are mass segregated is smaller compared to before the merger. When the merged subcluster has settled down to a system with a single centre (bottom panel), the ≈ 10 most massive sinks quickly form a close, concentrated system in the centre, leading to large values of Λ. The less massive sinks are more randomly distributed so that in total a smaller fraction of the sinks is mass segregated ($\approx 10\%$).

This quick development of mass segregation after a merger has already been found in nbody simulations of merging subclusters by McMillan et al. (2007) and Allison et al. (2009a). The feature of mass segregation (i.e. that it involves of the order of ten stars shortly after a merger) is the same as found by Moeckel & Bonnell (2009) in the simulation of Bate (2009a). Allison et al. (2009a) analysed the evolution of mass segregation in a cluster evolving from fractal initial conditions to a centrally concentrated system, but without mass segregation of the subclusters. At an age of $\approx 500\,000$ yr they find values of Λ for the whole cluster which are comparable to the values we derived. For the Orion Nebula Cluster (analysed by Allison et al., 2009b) only the nine most massive stars are mass segregated which is comparable to the post-merger state we find.

In the previous paragraph we gave examples of rich subclusters that are mass segregated if they have not undergone a merger recently. In order to establish what is the observational norm we turn to the composite population of the 10^4 $M_\odot$ calculation (Sec. 8.4.2 and Fig. 8.4) and have split our sample between subclusters according to their richness ($n \leq 30$, $30 < n \leq 50$, $50 < n \leq 100$ and $n > 100$). As subclusters gain new sinks during their evolution this sequence of increasing richness can also be seen as a sequence in time. In Figure 8.13 we plot histograms of the fractional radial rankings of the most massive, second and third most massive sinks. In the absence of mass segregation these histograms should be flat, which is

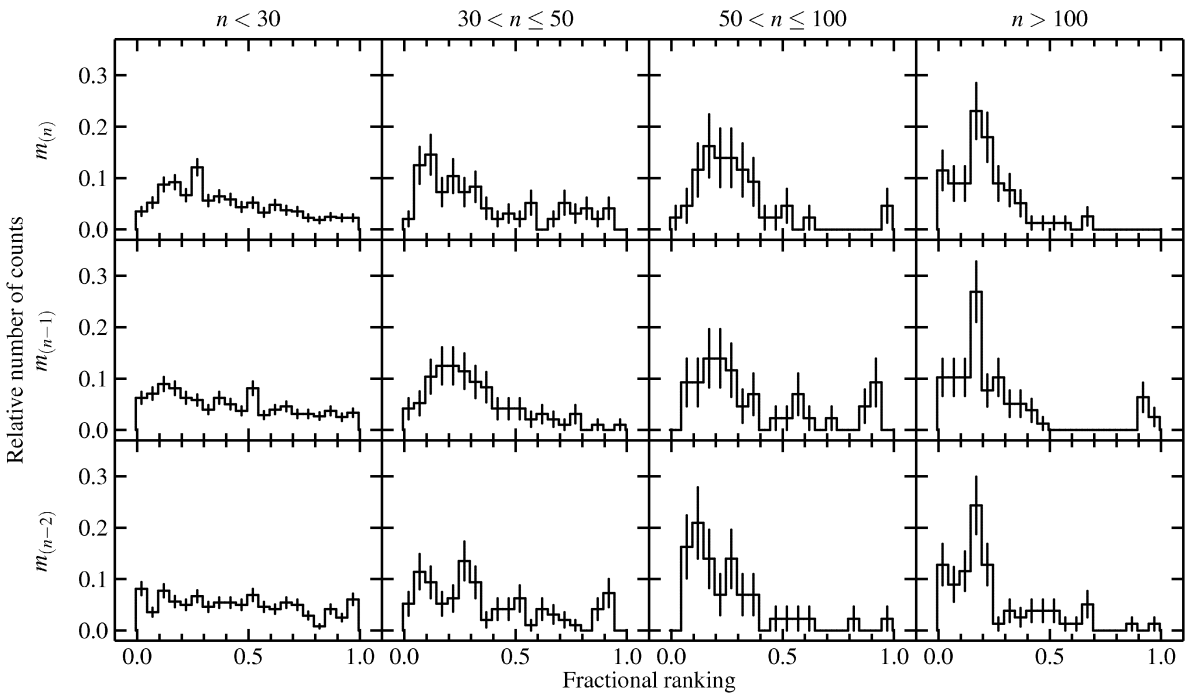


Figure 8.13: Histogram of the fractional radial ranking of the most massive (top), second most massive (middle) and third most massive (bottom) sink particle in its associated subcluster, split up by the number of sinks in the subcluster. The composite population of the 10^4 $M_\odot$ calculation is used to make the histograms. In the absence of mass segregation the histogram would be flat: the peak at small values shows that the massive sinks are preferentially found near the cluster centre. The second peak with a ranking of ≈ 1, especially for the second and third most massive sink, is due to mergers, where two centres are still present.

roughly the case for the very small clusters ($n < 30$), although already for them a weak trend of central concentration is present. These systems already contain the seeds of massive sinks (they have a large average stellar mass, see Fig. 8.17, and will become the central parts of richer subclusters. For the larger clusters there is clear evidence that the most massive sink particle is concentrated towards small radii, being rarely located beyond the inner 25% of sinks (we emphasise that this radial ranking is based on distance from the geometrical cluster centre, rather than centre of mass). The second (and also third) most massive sink particle is also frequently found in the inner regions of populous subclusters, but there is a second peak in the upper quartile, corresponding to the case where the second most massive sink is located in the nucleus of a subcluster that is in the process of merging.

Over all, therefore, we conclude that the most massive sinks are indeed segregated towards the centres of populous ($n \geq 30$) subclusters. We will also see that the most massive sinks are preferentially located in subclusters as opposed to the field as evidenced by the steeper slope of the upper tail of the IMF for the entire population as opposed to the total population contained in subclusters (see Fig. 8.16).

8.7 Evolution of the sink particle mass function

The mass distribution of the stars describes the end product of the star formation process. In this Section we analyse the sink particle mass distribution as proxy for the stellar mass function at each time step, with a focus on the high-mass tail of the mass distribution. We would like to stress that the results presented in this Section are not directly comparable to the observed stellar IMF, as a complete modelling of the star formation process is computationally not possible at the present time. Thus the *actual* mass of a star formed is not the sink particle mass, but lower because of simplifications in the computations. Firstly, star formation is modelled by sink particles with radii larger than the proto-stellar radii, so that fragmentation could also occur within the sinks (formation of close binaries). The sink particle mass function is closer to being a system mass function since the observed distribution of binary separations implies that most binary companions would be located within the sink radius (200 au). (Weidner et al., 2009) found that the system and individual mass function have only slightly different exponents (difference < 0.2). Furthermore, feedback by stellar winds or radiation is not included in the model, so that accretion is not hindered or stopped. These (zero-feedback) calculations thus overestimate system masses. Also, the gravitational force between sink particles is softened on a scale of a few sink radii, so that close encounters and binary formation is suppressed, which could influence the accretion history of the sink particles involved. Thus, the actual mass function of individual stars will have a smaller upper mass limit.

Our reference hypothesis for the stellar mass distribution to compare with the sink particle mass function is the two-part power law parametrisation of the mass function by Kroupa (2001, 2002),

$$\xi(m) \quad \propto \quad \begin{cases} m^{-\alpha_{\rm body}}; & \alpha_{\rm body} = 1.3; \quad 0.08 \leq m/{\rm M}_\odot < 0.5 \\ m^{-\alpha_{\rm tail}}; & \alpha_{\rm tail} = 2.35; \quad 0.5 \leq m/{\rm M}_\odot < 150. \end{cases} \tag{8.4}$$

As upper limit or truncation mass for the IMF, valid for all clusters unless estimated, we adopt the physical upper limit for stellar masses, above which stars do not appear to exist ($m_u = 150{\rm M}_\odot$, Weidner & Kroupa, 2004; Oey & Clarke, 2005; Koen, 2006) We use the stellar mass function as a probability density, i.e. normalised such that $\int_{m_l}^{m_u} \xi(m) dm = 1$. The choice of methods for the analysis of the mass function depends on the number of data points. If the dataset contains a sufficiently large number of data ($n \gtrapprox 100$) direct methods can be applied, i.e. parameters can be estimated and goodness-of-fit tests can be carried out. For meagre datasets one has to rely on indirect methods, which are usually comparisons of quantities derived using the data with expectations derived using a hypothesis for the distribution, fully specified with all parameters.

The most detailed information about the high-mass tail of the stellar mass distribution can be obtained at the end of the calculation, when the dataset has the largest number of data points. Thus we start at this point with our analysis of the mass function and proceed then to the time-evolution, which due to the small sample size can only be studied via more indirect methods. The findings from the final state will facilitate the interpretation of the time evolution.

n	$n_{\rm tail}$	$\widehat{\alpha}_{\rm tail}$	$\widehat{m}_u$	$m_{(n)}$
10^3 M$_\odot$ **calculation, richest subcluster:**				
372	110	1.67±0.10	23±2 M$_\odot$	21 M$_\odot$
10^4 M$_\odot$ **calculation, richest subcluster:**				
476	98	1.93±0.11	39±8 M$_\odot$	30 M$_\odot$
10^4 M$_\odot$ **calculation, second richest subcluster:**				
174	31	1.69±0.36	19±6 M$_\odot$	15 M$_\odot$
10^3 M$_\odot$ **calculation, all sinks:**				
563	148	1.79±0.11	24±2 M$_\odot$	21 M$_\odot$
10^4 M$_\odot$ **calculation, all sinks:**				
1945	459	2.18±0.08	33±4 M$_\odot$	30 M$_\odot$
10^4 M$_\odot$ **calculation, all sinks in all subclusters:**				
1645	267	1.92±0.07	34±4 M$_\odot$	30 M$_\odot$
10^4 M$_\odot$ **calculation, all sinks not in subclusters: ("field stars"):**				
890	202	2.55±0.14	9±1 M$_\odot$	8 M$_\odot$

Table 8.1: Estimated parameters of the mass functions for sinks in the high mass tail. n is the total number of sinks in the object, $n_{\rm tail}$ the number with $m > 0.8$M$_\odot$. $\widehat{\alpha}_{\rm tail}$ and $\widehat{m}_u$ are the estimated exponent and truncation mass, respectively. $m_{(n)}$ is the mass of the most massive sink particle, given for comparison. The estimates were derived at the end of the simulations, with $\tau = 2.5\ t_{\rm ff}$ and $\tau = 1.0\ t_{\rm ff}$ for the 10^3 M$_\odot$ and 10^4 M$_\odot$ calculation, respectively.

8.7.1 Final mass function

For the analysis of the final mass distribution we just assume for the high-mass tail ($m >$ 0.8 M$_\odot$) that the mass distribution is following a power law truncated at some value, not imposing any assumption about the exponent or the truncation mass. To estimate the exponent, $\widehat{\alpha}_{\rm tail}$, and truncation mass, $\widehat{m}_u$ we use the bias-corrected maximum likelihood method of Maschberger & Kroupa (2009, this thesis Chapter 5). The results are given in Table 8.1. In the most populous subclusters we find $\widehat{\alpha}_{\rm tail}$ in the range from $\approx$ 1.7–1.9. These are much smaller values than the Salpeter value, $\alpha_{\rm tail} = 2.35$, which can be explained by the preference of massive sinks to be in subclusters. With only three estimates and considering the size of the error bars it is not unreasonable to assume a universal exponent $\alpha_{\rm tail} \approx 1.8$, valid within the dense subclusters. There is no apparent dependence of the exponent on the number of sinks in the tail.

The estimated truncation masses are only marginally higher (up to $\approx$ 10 M$_\odot$) than the most massive sink particles in the clusters (15–30 M$_\odot$), see also Table 8.1. $\widehat{m}_u$ increases with increasing (total) number of sinks in the subcluster. This could indicate that the truncation mass of the mass function increases as the number of sinks increases. The truncation mass of a power-law distribution is difficult to estimate, and it is possible that despite the bias correction the "true" truncation mass can be underestimated by up to 50%.

Using a graphical goodness-of-fit technique, the SPP plot (stabilised probability-probability plot) described in Maschberger & Kroupa (2009, this thesis Chapter 5), it can be assessed whether the data could be consistent with alternative hypotheses of a larger exponent or a larger truncation mass, and also if the data are obeying the assumed null hypothesis (in our case the power law with the estimated exponent and estimated truncation mass). The SPP plot is constructed by first sorting the data ascending in mass and then calculating for each data point the empirical cumulative density and the hypothetical cumulative density. The empirical cumulative density is given by $P_{\rm E}(m_{(i)}) = \frac{i-0.5}{n}$, where i is the rank of the data point in the ordered sample and n the sample size. The cumulative density for the null hypothesis,

$P_{\rm H0}(m_{(i)})$, is in our case simply the cumulative density of a truncated power law, where the estimated values are used for the parameters. For a data set perfectly obeying the null hypothesis the pairs $\{P_{\rm H0}(m_{(i)}), P_{\rm E}(m_{(i)})\}$ would in a plot exactly lie on the $\{0,0\}-\{1,1\}$ diagonal. An additional bonus of this plot is that the Kolmogorov-Smirnov test has the direct graphical interpretation as parallels to the diagonal with their distance depending on the KS probability. However, a direct plot of $P_{\rm H0}$ and $P_{\rm E}$ is not the best display of the data because the main emphasis lies in the middle region of the plot. But if the cumulative densities are transformed using the stabilising transformation of Maschberger & Kroupa (2009, this thesis Chapter 5) this disadvantage can be overcome, and a transformed version of the KS test can be overplotted. This significantly reduces the likelihood of wrongly classifying data stemming from an alternative hypothesis as being from the null hypothesis.

The SPP plots, using a truncated power law as the null hypothesis (= diagonal in the plot), are shown for the most massive subclusters in Fig. 8.14, using the estimated exponent and truncation mass. For all massive clusters the data are following the diagonal and show no systematic trends. They do not exceed the 95% acceptance region of the stabilised version of the KS test, so that indeed a truncated power law describes the data well. As the truncation mass could be underestimated, we also show the alternative hypothesis of a power law with the same, estimated exponent, but with a truncation mass of 150 $\rm M_\odot$ (solid line). The data show no trend to bend in the same direction so that an underestimate of the truncation mass is not likely; instead the mass distribution is indeed truncated only slightly above the most massive sink particle. A power law with $\alpha_{\rm tail} = 2.35$ and $m_u = 150$ $\rm M_\odot$ gives the dotted line in Fig. 8.14, which has a curvature completely in disagreement with the data. The standard parameters (eq. 8.4) can therefore be excluded for our data.

The SPP plots for the whole systems are shown in Fig. 8.15. For the 10^3 $\rm M_\odot$ calculation the estimated parameters are $\widehat{\alpha}_{\rm tail} = 1.79 \pm 0.11$ and $\widehat{m}_u = 23.5 \pm 2.1$ $\rm M_\odot$. Bonnell et al. (2003), analysing the same simulation, already mention that the tail of the mass distribution could be fitted with either an overall exponent of $\alpha_{\rm tail} = 2.0$, or with a smaller slope in the intermediate-mass range and a steeper slope in the high-mass range. A strong truncation of the mass function can mimic in a histogram a two-part power-law behaviour of the data. From Fig. 8.15 we find that a single power law fits the data well and signs of a two-part power law are not present. Compared to the largest central subcluster the exponent of all sinks is somewhat larger, which means that the sinks in the "field" and the other subclusters (containing < 12 sinks) contribute mostly to the low-mass end of the tail and the massive sinks are preferentially found in the central region. Thus the steeper of the mass function for the whole system is a sign of mass segregation.

In the 10^4 $\rm M_\odot$ calculation we estimated for all sinks $\widehat{\alpha}_{\rm tail} = 2.18 \pm 0.08$ and $\widehat{m}_u = 33.0 \pm 3.7$, which is again steeper than for the subclusters. Here the data deviate from the assumed truncated power law in a sense that implies a gradual steepening of the mass function at the high mass end. We shall discuss this behaviour in Section 8.7.4 as a possible manifestation of the IGIMF effect.

Finally, we draw attention to the fact that all our IMFs are too flat compared to observed distributions, i.e. high-mass ($m > 0.8$ $\rm M_\odot$) are over-abundant. Internal fragmentation within the sink particles will decrease the number of massive sinks and increase the number of lower-mass sinks. Also, feedback from a massive sink could diminish the amount of accretion of sinks in it's surroundings, so reducing the relative masses of massive sinks. Both fragmentation and feedback can lead to a steeper exponent, so that the agreement with the Salpeter exponent can be reached. Those effects do not alter our conclusion that a strong truncation is

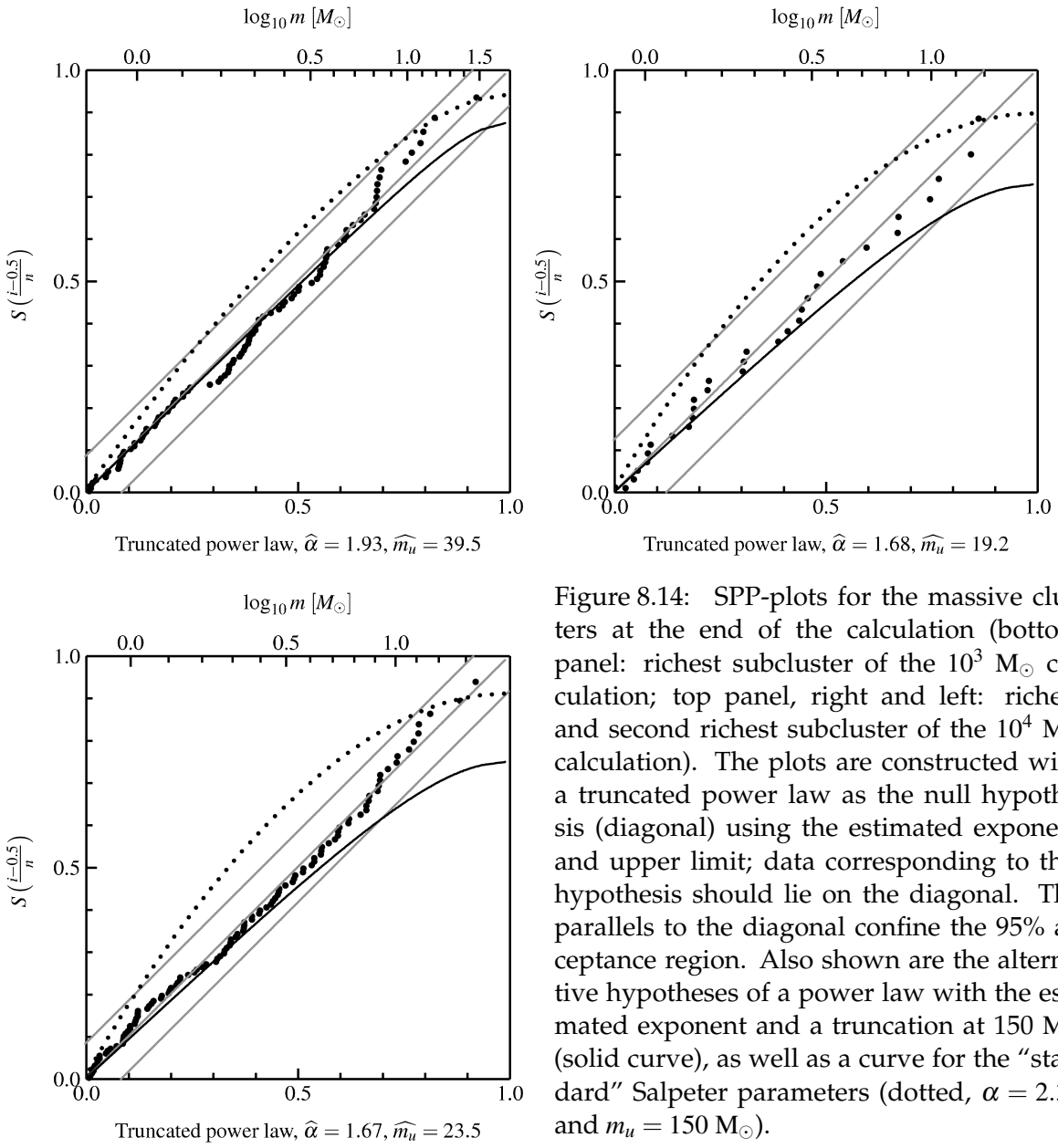


Figure 8.14: SPP-plots for the massive clusters at the end of the calculation (bottom panel: richest subcluster of the 10^3 $M_\odot$ calculation; top panel, right and left: richest and second richest subcluster of the 10^4 $M_\odot$ calculation). The plots are constructed with a truncated power law as the null hypothesis (diagonal) using the estimated exponent and upper limit; data corresponding to this hypothesis should lie on the diagonal. The parallels to the diagonal confine the 95% acceptance region. Also shown are the alternative hypotheses of a power law with the estimated exponent and a truncation at 150 $M_\odot$ (solid curve), as well as a curve for the "standard" Salpeter parameters (dotted, $\alpha = 2.35$ and $m_u = 150$ $M_\odot$).

needed as internal fragmentation and feedback will push the truncation masses even lower. They also do not affect our finding that the mass function is steeper in the 10^4 $M_\odot$simulation, in which regions of the initial gas are unbound. This change of initial conditions prevents cluster merging from going to completion and prevents the over-production of massive sinks.

8.7.2 Time-evolution of the exponent

After the detailed discussion of the mass function at the end of the simulations we now turn to the dependence of the mass function on time and the number of sinks. We first look at the time-evolution of the exponent starting with the larger clusters, which allow us to estimate the parameters, shown in Fig. 8.16. The estimates are made if more than 24 sinks with $m >$

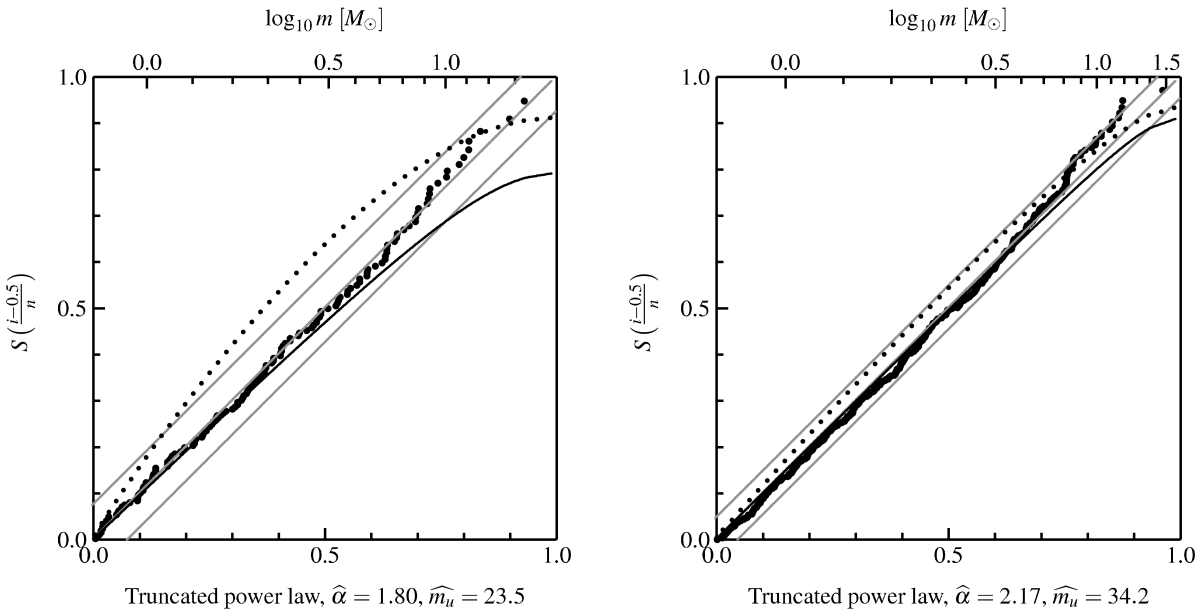


Figure 8.15: SPP plots as in Fig. 8.14 for all sinks at the end of the simulations (10^3 $M_\odot$ left and 10^4 $M_\odot$ right), with a truncated power law as null hypothesis (diagonal) using the estimated exponent and truncation mass. Also shown are the alternative hypotheses of a power law with the estimated exponent and a truncation at 150 $M_\odot$ (solid curve) and with the "standard" Salpeter parameters (dotted, $\alpha = 2.35$ and $m_u = 150$ $M_\odot$).

0.8 $M_\odot$ are present. The small filled symbols are for the entire sample and the open points for the individual subclusters. For the whole system the exponent is initially relatively large ($\alpha_{\rm tail} > 2.5$) consistent with the lack of time available for sinks to grow much by accretion. As sinks gain mass by accretion, the slope rapidly declines over about 5×10^4 years, and then stabilises at about 1.8 in the small simulation and 2.2 in the large simulation. The subclusters only appear when the stable part of the evolution is reached, and their $\alpha_{\rm tail}$ stays roughly constant with similar values in both simulations. The small symbols denote the values of $\alpha_{\rm tail}$ for the whole population of sinks in subclusters together. The fact that these values are smaller (i.e. a flatter IMF) than for the whole population, including the field, is a sign of mass segregation. In addition we note that in the 10^4 $M_\odot$ simulation the values of $\alpha_{\rm tail}$ for individual clusters lie below that for the aggregate cluster population. We however emphasise that the open symbols in Fig. 8.16 are not independent data points and actually only correspond to one (10^3 $M_\odot$ simulation) and up to three (10^4 $M_\odot$ simulation) clusters. Thus whereas the fact that they lie below the solid symbols is interestingly suggestive of a flatter IMF within individual clusters the result is compromised by small number statistics. We return to this in our discussion of possible IGIMF effects in Section 8.7.4 below.

When the number of sinks does not suffice to estimate the exponent, the mean stellar mass, $\overline{m}$, can be used. In Figure 8.17 we show $\overline{m}$ as a function of n. The value derived from the reference mass function is the horizontal thin line with the expected scatter for random sampling (thin lines at the 1/6th and 5/6th quantiles). For the total stellar sample (thick line) the mean mass increases (by a factor 2–3) over the duration of the simulation as a result of accretion, already deduced from Fig. 8.6. It only falls out of the 1/6-5/6 region for larger n,

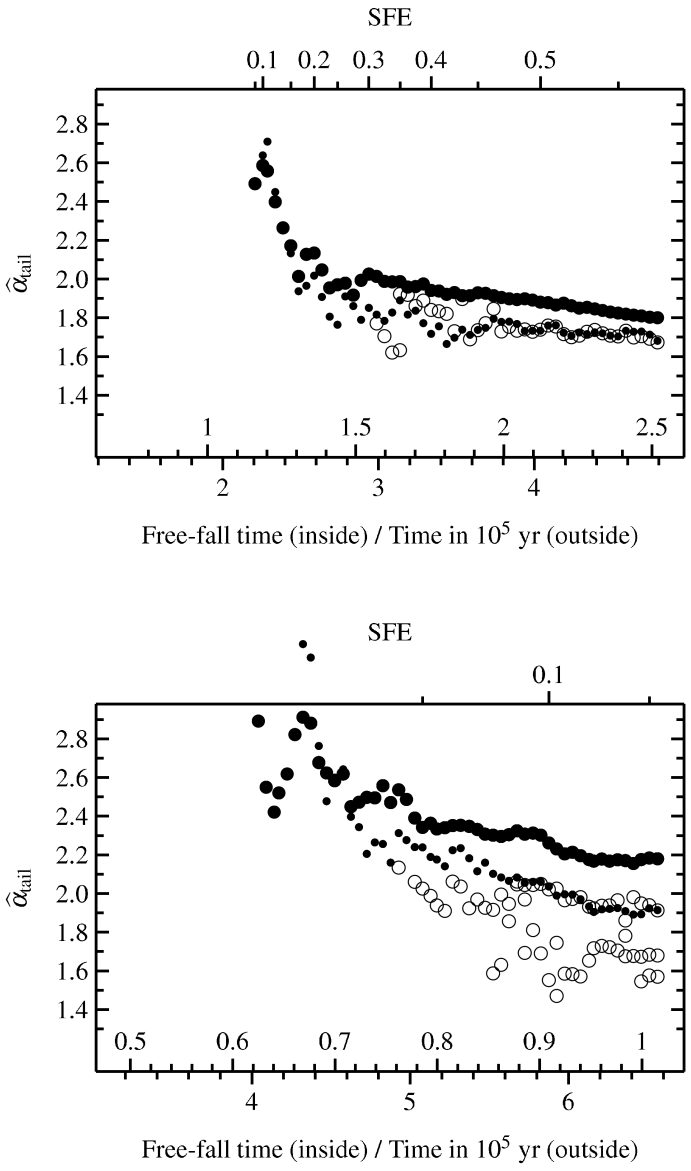


Figure 8.16: Time evolution of the exponent in the tail ($m > 0.8 M_\odot$), estimated when more than 24 sinks are in the sample (10^3 $M_\odot$ calculation top, 10^4 $M_\odot$ calculation bottom). The exponent was estimated for the whole systems (subclusters and field, big filled dots) and the individual subclusters (big open dots). The small filled symbols show the exponent estimated from all sinks in subclusters together (without the field).

interestingly at the point in the simulations at which the maximum stellar mass is around 10 $M_\odot$. It is again tempting to speculate that stellar feedback associated with the steep increase in the ultraviolet output of stars at around 10 $M_\odot$ could remedy this situation.

An increasing $\overline{m}$ is also compatible with the decreasing exponent of the tail that is found in Fig. 8.16 for the larger-n subclusters. A similar trend of an increasing heaviness of the tail is present if only all sinks in subclusters are considered, shown as blue line. The mean mass of the total population in clusters is generally higher than for the whole system, which is a consequence of mass segregation. The data points are instantaneous values for individual clusters and demonstrate that the mean values are not at all consistent with the expectations of random sampling from an invariant reference mass function. Even the smallest clusters can often have large mean stellar masses as would be expected in a scenario where subclusters form around massive sinks.

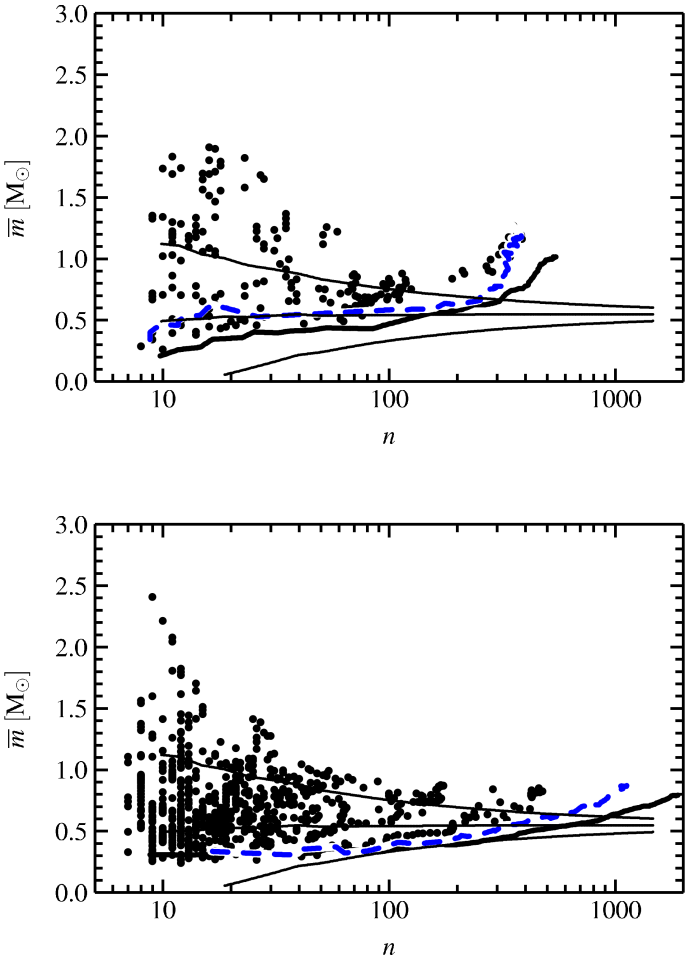


Figure 8.17: Mean mass of sinks in the calculations (10^3 M$_\odot$ calculation top, 10^4 M$_\odot$ calculation bottom), derived for the whole system (thick line), all sinks in subclusters (dashed line) and the individual subclusters (dots). The thin lines are the expected mean from the reference IMF (eq. 8.4), and the 1/6th and 5/6th quantiles for random sampling.

8.7.3 Evolution of the truncation mass

The analysis of the three subclusters at the end of the simulations already gave a tentative indication that the truncation mass of the mass function depends on the number of sinks in the subcluster. Figure 8.18 illustrates this further by showing the estimated truncation mass as a function of the cluster richness, again with the solid dots for the whole systems and open symbols for the subclusters. The actual most massive sink particle in each cluster is also plotted. As above, the deduced truncation mass is always only marginally larger than the largest datapoint, so that a much larger truncation mass is not likely. The points for the whole system are shifted to the right, as it contains many more sinks. We see clear evidence that the truncation mass is a systematic function of the cluster membership number.

We can further test whether the mass functions within individual subclusters are truncated, by examining the distribution of the most massive, second most massive and third most massive sink particle within each subcluster. These three quantities are plotted in the three panels of Figure 8.19 as a function of cluster membership number. These data show the qualitative trend (increasing maximum stellar mass with cluster richness, together with a large scatter in maximum stellar mass at a given cluster n) that is seen in observational data (Weidner & Kroupa, 2004, 2006, Maschberger & Clarke, 2008, this thesis Chapter 7, Weidner et al., 2010) and which is predicted by the statistics of random drawing. The solid and dot-

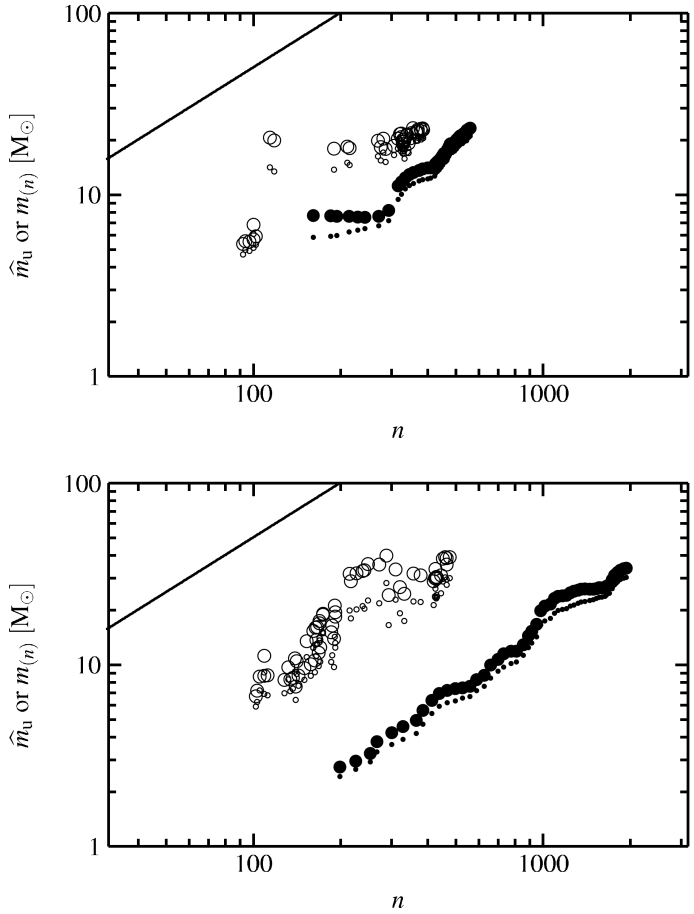


Figure 8.18: Estimated truncation mass as a function of the number of sinks (top 10^3 M$_\odot$ calculation, bottom 10^4 M$_\odot$ calculation), on a cluster by cluster basis (large open dots) and for the whole population (large filled dots). The small symbols are the corresponding values of the actual maximum stellar mass in each cluster (small open) or in the population as a whole (small filled). The line is an estimate of the total stellar mass as a function of n (i.e. $m = \overline{m}n$, with the mean stellar mass $\overline{m} = 0.54$M$_\odot$ as implied by the reference IMF eq. 8.4). The number of sinks can serve as a proxy for time.

ted lines on the plot correspond to the mean and 1/6th and 5/6th contours in the cumulative distribution that is predicted by random sampling from the reference IMF, eq. 8.4.

We see that the simulation data lie progressively higher with respect to the theoretical quantiles as one proceeds from most massive to second and third most massive members: in other words, the masses of the three most massive sink particles are more bunched together than one expects from the models. We illustrate how the form of the IMF affects the *relative* distributions of the most massive three cluster members in Figure 8.20 where we plot the expectation values of the mass of the three most massive members in the case of three 'toy' IMF models. The solid and dotted lines correspond, respectively, to power law distributions with slopes of 1.8 and 2.35 which are truncated at a mass of 150 M$_\odot$. As expected, the flatter power law implies higher means of all three quantities at a given n, but the relative spacing between the most massive and the second and third most massive members is not very different in the two cases. In both cases, the three lines would start to converge only for much richer clusters where the expected masses of the three most massive sink particles approached the cut-off at 150 M$_\odot$. The dashed curves, which are provided for purely illustrative purposes, correspond to an input distribution with a slope of 1.8 but where the upper limit is a function of cluster richness,

$$m_u(n) \quad = \quad \frac{1}{5} n \, \mathrm{M}_\odot. \tag{8.5}$$

In this case, truncation is important in all clusters and the effect of this is to make the three

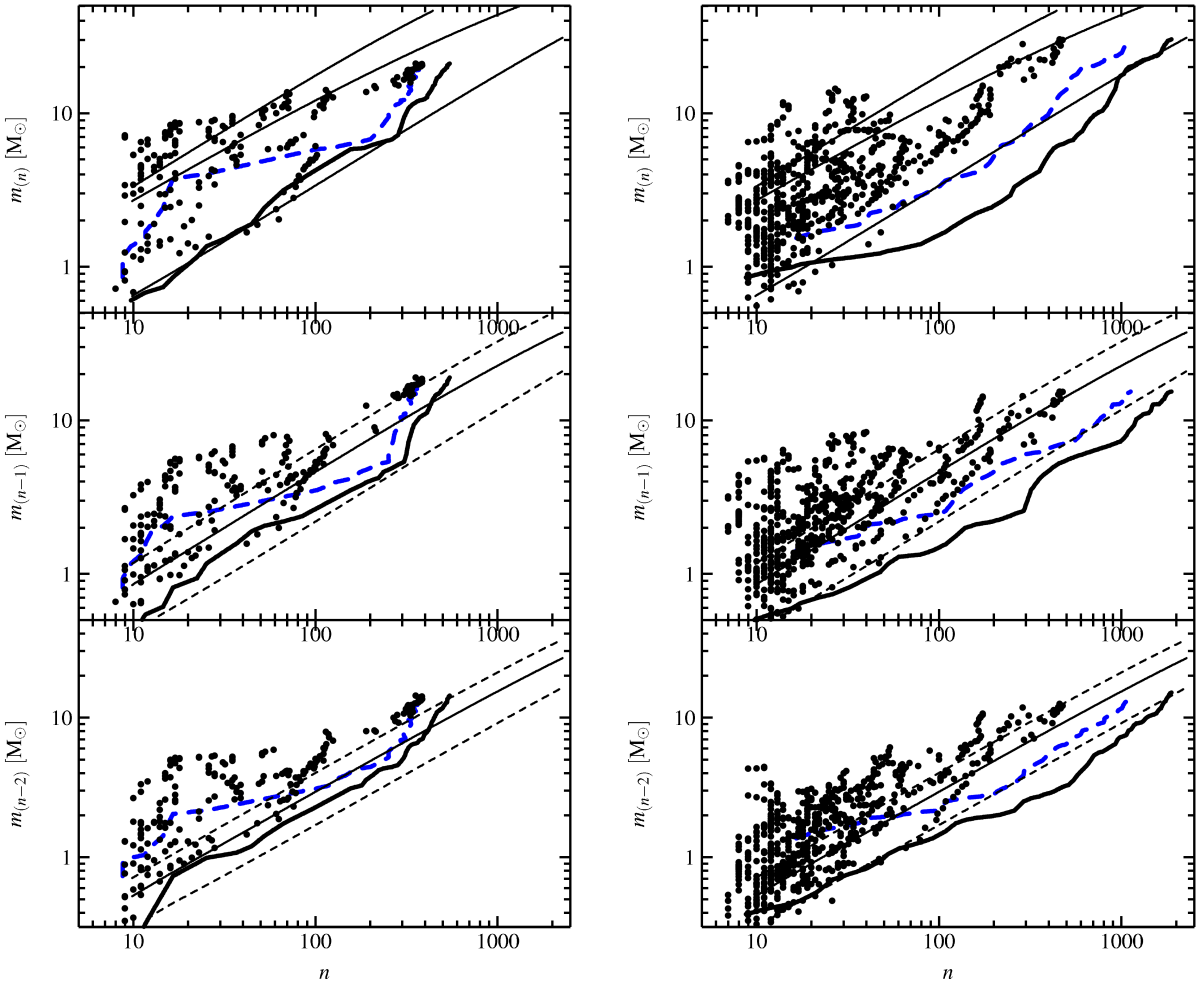


Figure 8.19: Evolution of the mass of the most massive, second and third most massive sink particle (top to bottom) as a function of the total number of sinks (left 10^3 M$_\odot$calculation, right 10^4 M$_\odot$calculation). The thick solid line shows the evolution of the total system (all sinks), the thick dashed line is the track for all sinks in subclusters and the individual subclusters are represented by dots. The solid and dashed lines represent the predicted expectation value of the mass of the nth ranked sink along with the 1/6th and 5/6th quantiles for random sampling from the IMF given in eq. 8.4. Note that the simulation data sits progressively higher with respect to the predicted quantiles, as one proceeds from first to second to third most massive sink particle down the page.

dashed lines much closer together than for the other (fixed truncation) cases. We therefore deduce, at a qualitative level, that the effect seen in Figure 8.19 (whereby the difference in mass between the three most massive sinks is unexpectedly small) may be a hint that the mass functions are truncated even in the lower-n subclusters.

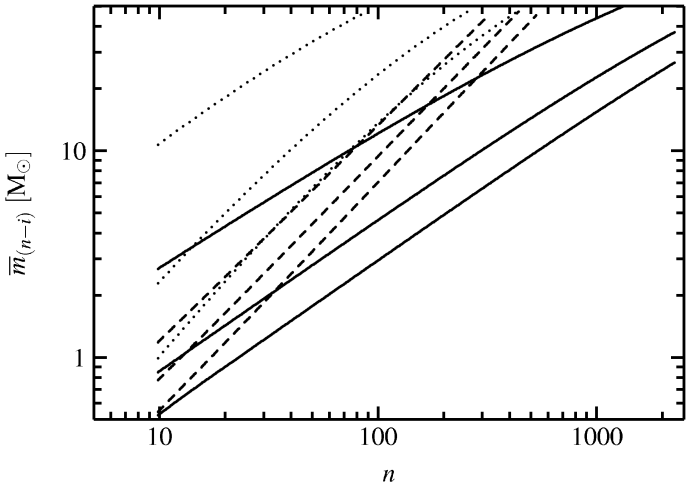


Figure 8.20: Location of the mean mass for the most massive, second and third most massive star for different parameters of the mass function (from top to bottom in each group of lines). The solid lines use $\alpha_{\rm tail} = 2.35$ and $m_u = 150$ M$_\odot$, the dotted lines $\alpha_{\rm tail} = 1.8$ and $m_u = 150$ M$_\odot$. For the dashed lines again $\alpha_{\rm tail} = 1.8$ is used, but the truncation mass is a function number of stars, eq. 8.5.

8.7.4 An IGIMF effect?

The IGIMF (integrated galactic IMF) is a concept introduced by Kroupa & Weidner (2003) and further developed in Weidner & Kroupa (2005) (a similar notion is already present in Vanbeveren, 1982, 1983). If the truncation mass of the IMF in star forming regions (i.e. star clusters) depends on the richness of the region (by number or mass), then the IMFs in the regions are not completely identical any more, and thus the stars of all star forming regions in a galaxy together can have a distribution function, the IGIMF, that differs from the IMF within individual clusters. The IMF (here defined as IMF within an individual star forming region) and IGIMF disagree only in the high-mass tail. For example, if there are 1000 M$_\odot$ in stars of many small star forming regions, with a truncation mass of, say, 10 M$_\odot$, and 1000 M$_\odot$ from star forming regions with $m_u = 100$ M$_\odot$, then the combined sample of 2000 M$_\odot$ will have a deficiency of stars between 10–100 M$_\odot$, compared to a sample of 2000M$_\odot$ with $m_u = 100$ M$_\odot$. For a more realistic case the general trend is that the IGIMF is steeper than the IMF in the high mass tail ($\alpha_{\rm IGIMF} > \alpha_{\rm IMF}$) where the exact relationship depends on the spectrum of cluster masses. This effect can influence for example the relation between the star formation rate and the Hα flux of galaxies (Pflamm-Altenburg et al., 2007a; Pflamm-Altenburg & Kroupa, 2008; Pflamm-Altenburg et al., 2009) and the metallicity of a galaxy (Köppen et al., 2007).

The 10^4 M$_\odot$ calculation covers a region that is sufficiently large and massive that it produces not just a single cluster but a population of objects which may evolve into individual star clusters. To our surprise we found that the mass function of this calculation shows signs of the IGIMF effect, as already mentioned in the Sections above. In the SPP plot containing all sinks of the simulation (Fig. 8.15) the high-mass end of the data bends upwards away from the diagonal, implying a steepening of the mass function. This effect is not only due to the fact that the entire population contains extra (field) sinks that are not included in the cluster and which (due to mass segregation) are of lower mass. Fig. 8.21 is an SPP plot for the aggregate population of sinks in subclusters and here again the upward curvature is a hallmark of a progressive steepening of the IMF. As noted above we expect to see this effect since we have already seen evidence that the IMFs in individual clusters are truncated. Although the observational reality of such IGIMF effects is controversial (e.g. Elmegreen, 2009a and Parker & Goodwin, 2007 on the theoretical side, or Parker et al., 1998; Chandar et al., 2005; Hoversten & Glazebrook, 2008; Meurer et al., 2009, further discussed in Elmegreen, 2009b), it is interesting

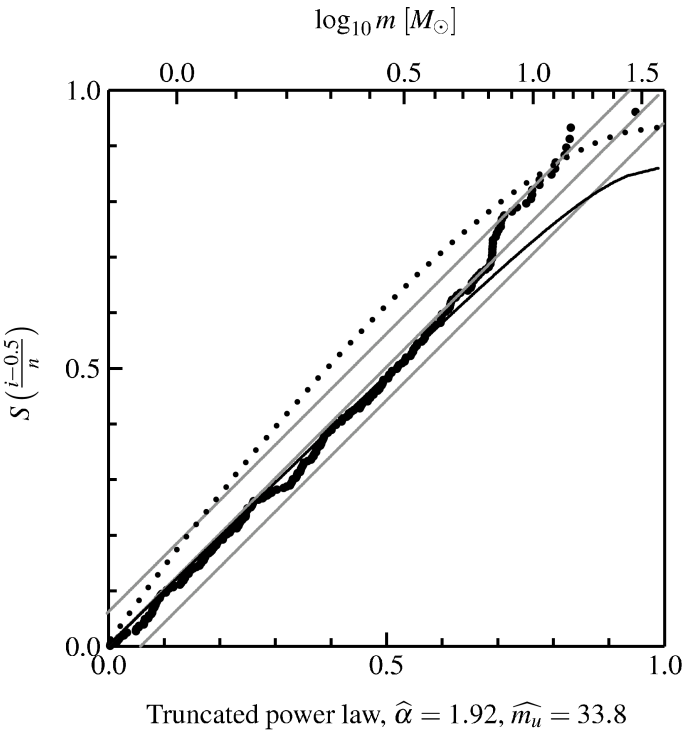


Figure 8.21: SPP plots using all sinks in subclusters together at the end of the 10^4 $M_\odot$ simulation. Also shown are the alternative hypotheses of a power law with the estimated exponent and no upper truncation (dashed curve) and a truncation at 150 $M_\odot$ (solid curve), as well as curve for the "standard" Salpeter parameters (dotted, $\alpha = 2.35$ and $m_u = 150 M_\odot$). The data show a curvature that implies a suppression of high masses.

that the large simulation indeed appears to manifest this behaviour.

Although we stress that the process by which the stellar mass function is built up cannot be seen *physically* as a random drawing experiment, the net effect of the cluster assembly process is to produce clusters that are *mathematically* describable as follows: random drawing from a mass function with an upper cut-off that depends on cluster richness. In this sense, the simulations show a behaviour that is qualitatively similar to the Monte-Carlo simulations of Weidner & Kroupa (2006), who constructed model clusters under a similar assumption. The reason, in the case of the simulations, that the upper truncation increases with cluster richness is because the first sinks to form not only tend to attain the largest masses but also have the greatest opportunities to undergo cluster mergers and hence end up in the largest clusters.

8.8 Discussion

It is often stated that the majority of stars form in clusters and indeed in the simulations we find that by an age of half a Myr $60-80\%$ of sinks are located in clusters[3]. We also find that by this stage the clusters are strongly mass segregated, that more massive sinks are, in a statistical sense, associated with richer clusters and that massive sinks are under-represented in field regions compared with clusters. We find that in the simulations a sink 'forms' (i.e. the mass of bound gas within a radius of ≈ 200 AU increases as a result of infall from the environment) over a variable period which can be as long as the duration of the simulation

[3]Note that in common with observers we here define clusters in terms of association on the sky and do not imply by this that such clusters are necessarily bound or long lived. Obviously the fraction in clusters depends on the choice of $d_{\rm break}$.

(of order half a Myr). Given the ambient gas densities, this period is of order a free fall time and sinks can thus move significant distances (several tenths of a parsec) over this period and experience considerable evolution in clustering properties in the process. Indeed, around half of sinks of all masses do not start to form in the central regions of populous clusters, but in their outskirts or in separate small groupings (with $n < 12$). The more massive sink particles are however those that start to form earlier and are more likely to have undergone mergers into successively larger entities.

A consequence of this cluster formation pattern is that sinks that form together in a small-n group tend to stay together and experience similar accretion histories as they merge into larger entities. This is particularly true of sinks that form early and thus acquire a headstart in mass acquisition, since these tend to end up in the cluster core during cluster merging. Thus the mass distribution of sinks in a given cluster often contains a group of massive sinks of similar mass (see Figure 8.19). In terms of a mathematical description of the resulting IMF on a cluster by cluster basis, this is best represented by a power law upper IMF which is truncated at a stellar mass that depends on the cluster richness. As pointed out by Weidner & Kroupa (2006), a consequence of such behaviour is that in the integrated IMF (i.e. the IGIMF, being that composed of the summed total of a sample of clusters) the massive sinks are underrepresented, which leads to a steeper slope in the power law for a large sample. The 10^4 $M_\odot$ simulation produces several clusters, and indeed when all sinks of them are combined the mass function deviates from a power law. Because of the small number of clusters we do not find a general steepening of the slope but a lack of massive sinks at the high-mass end, which is the IGIMF effect for a small sample of clusters.

While a lot of our analysis has been devoted to understanding the reason that the simulations produce particular observational characteristics, observers can also of course simply use these results as an empirical test of the correctness of the physical ingredients in the simulations. It is of course important for proper comparison that clusters are extracted from spatial distributions on the sky through use of a minimal spanning tree, as here, and that parameters (such as ellipticity) are also derived in the same way.

Apart from the issue of IMF slope described above, we here draw attention to two properties that are particularly suitable for observational comparison. First of all, the ellipticity histogram (Figure 8.10) demonstrates that the clusters are somewhat flattened, typically with an axis ratio of $< 2 : 1$; this moderate flattening is a combined consequence of the filamentary morphology of the gas and the effects of relaxation that tend to sphericalise the inner regions. It is an easy matter to compare the ellipticity distribution of an ensemble of clusters and decide whether this is statistically consistent with the distribution shown in Figure 8.10. Secondly, one may readily compare the degree of mass segregation in an observed cluster ensemble with these simulations through construction of a diagram like Figure 8.13. This diagram involves only a scale free quantity and makes no assumption about the radial density profile or cluster morphology: all that is required in order to construct such a diagram is that one can count sources on the sky and can identify the most massive star in the cluster. We note that upcoming Xray surveys, which offer the potential to identify large numbers of low mass pre-main sequence stars in regions that are heavily embedded, offer an excellent opportunity to test the diagnostics presented in this paper.

Acknowledgements

We thank Simon Goodwin, the referee, for valuable comments that have clarified and improved this article. ThM acknowledges funding by CONSTELLATION*, an European Commission FP6 Marie Curie Research Training Network.*

8.9 Appendix: Dependence of the results on the projection plane

As we analyse the simulation data in projection on a 2D plane the results could be compromised by the choice of the projection plane (in the main text the *x*-*y* plane). To demonstrate the robustness of our results we show here examples for how a different choice of the projection plane would influence them. The following plots are made by analysing the data with exactly the same parameters but different projection planes (*x*-*z* and *z*-*y*). In general the effects of a different projection plane are small and not distinguishable from the already present statistical scatter.

The merging history (Fig. 8.22) shows the influence on the general detection and classification of clusters (number of clusters and their richness). The total number of clusters chances slightly, because small-*n* clusters may not be detected, but all larger subclusters are present with a nearly identical growth history. The ellipticity histogram (Fig. 8.24) and the Cartwright *Q* parameter (Fig. 8.23) should be very sensitive to projection effects, but also here the differences are only minor. Finally, to asses the influence on the Section about the mass function we show in Fig. 8.25 the most massive, second and third most massive sink particle against number of sinks. Again, no major discrepancy that could compromise our results is present.

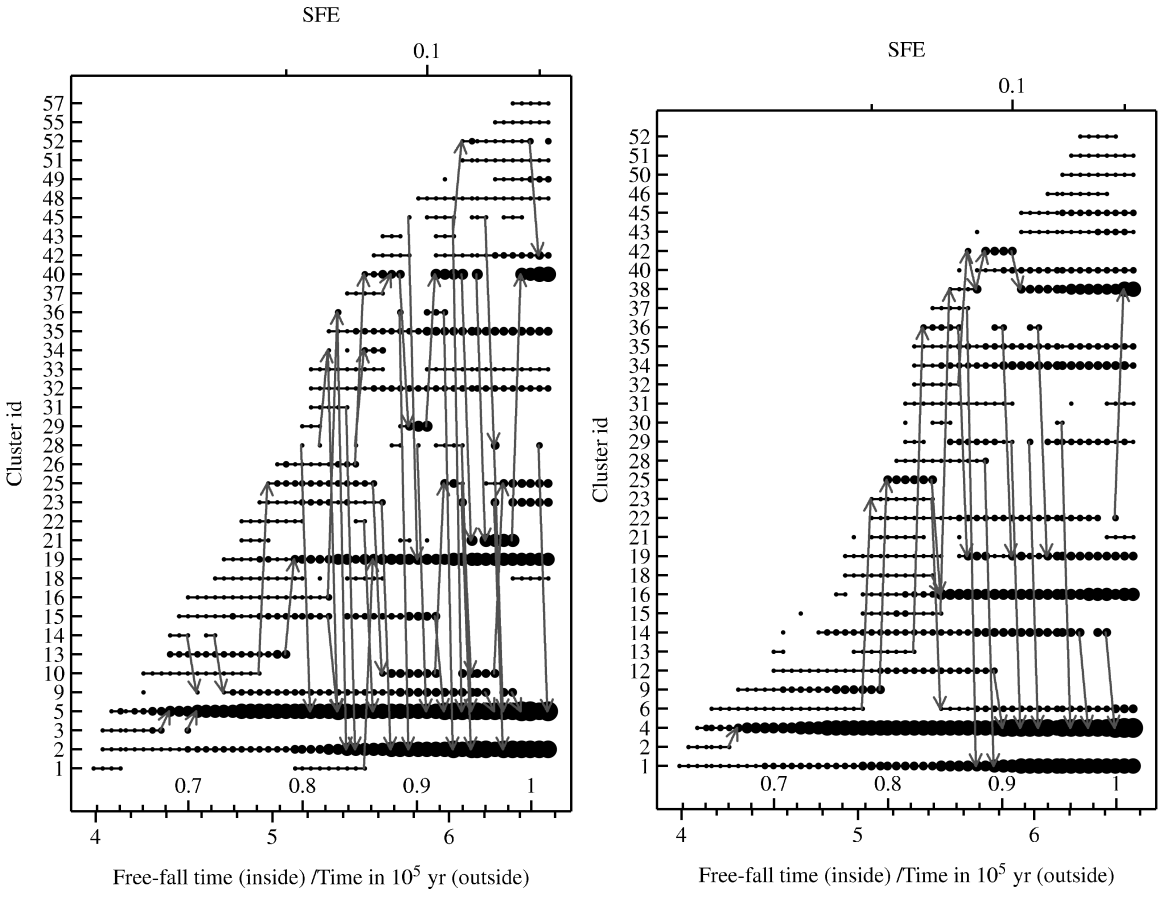


Figure 8.22: Merging history of the subclusters (as Fig. 8.3), derived in *x*-*z* (left) and *y*-*z* projection, large simulation.

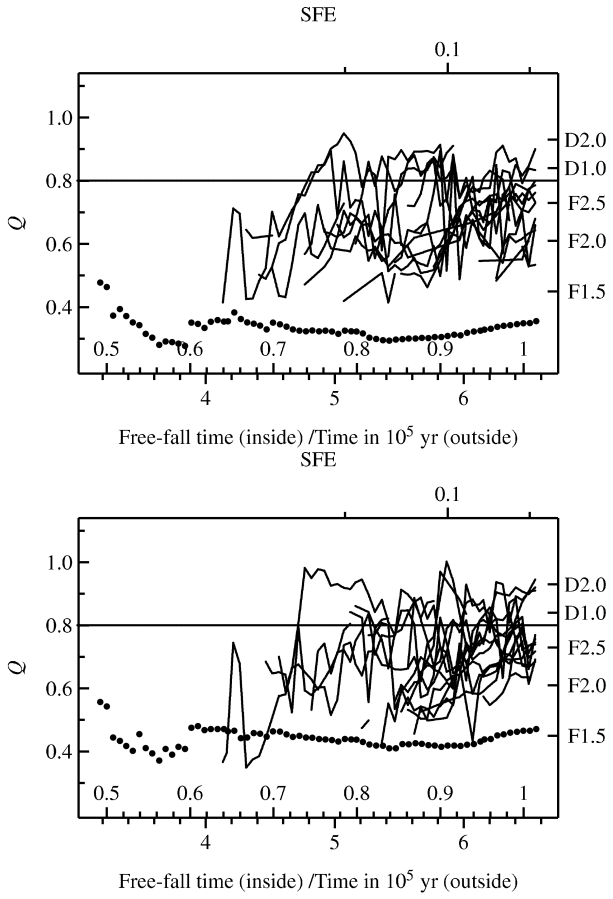


Figure 8.23: Time evolution of Q (lines as in Fig. 8.8 calculated in x-z (left) and y-z projection, large simulation.

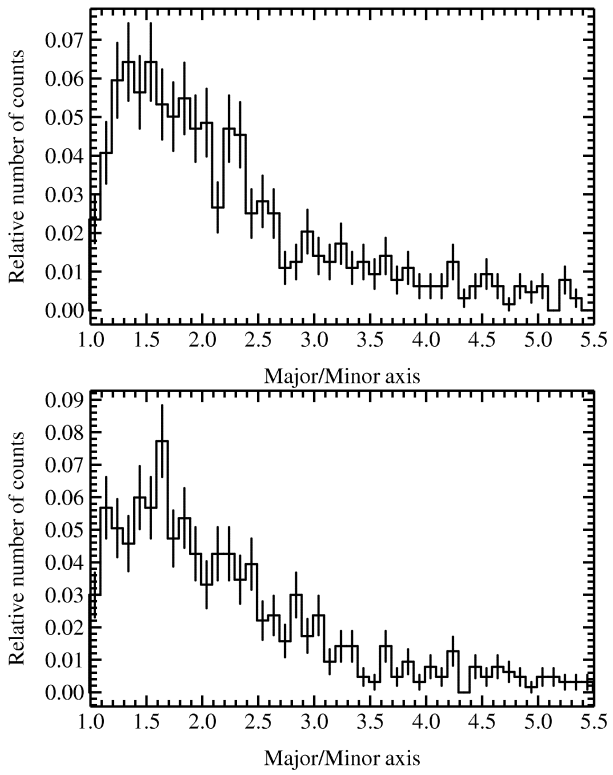


Figure 8.24: Histograms of subcluster ellipticities (as Fig. 8.10) derived in x-z (left) and y-z (right) projection, large simulation.

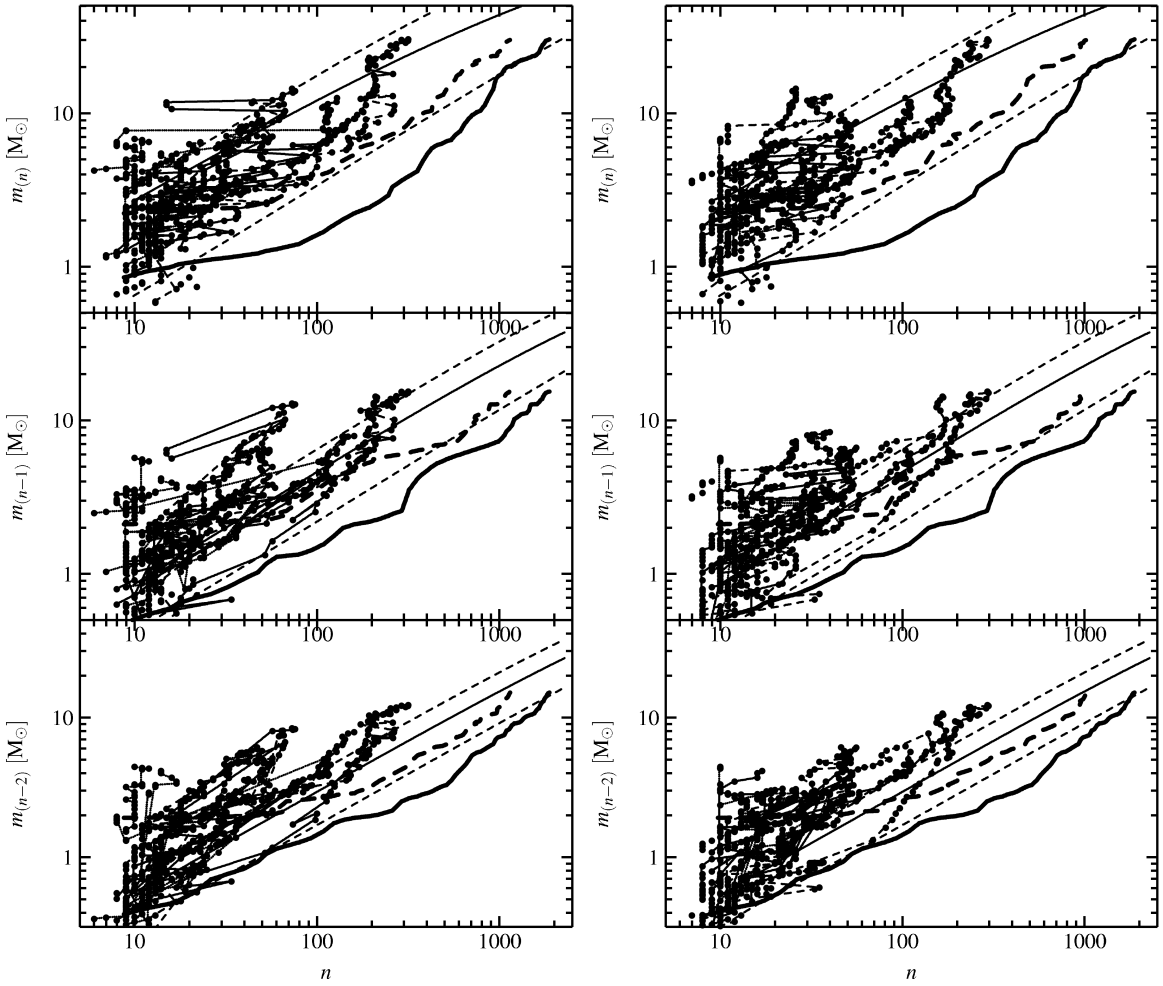


Figure 8.25: Evolution of the most massive, second and third most massive sink particle (as Fig. 8.19), determined in *x*-*z* (left) and *y*-*z* (right) projection, large simulation.

Chapter 9

Outlook

In the preceding Chapters we touched many aspects of star formation, from galaxy scales ($\approx$ 10 000 pc) to sub-cluster scales ($\approx$ 0.1 pc). The following Sections sketch projects which follow up some of these aspects.

9.1 Star cluster mass function

A prominent relation in this thesis was the scaling of the brightest cluster with the star formation rate in a galaxy. Two concepts can account for this, the size-of-sample effect and a variable upper limit of the cluster mass function. For a more quantitative analysis of this relation it is more convenient to go from the observational quantities to "theoretical" quantities, i.e. to analyse the relation between the mass of the most massive cluster and the total mass/number of the cluster population. For this purpose a time interval has to be introduced. Problems that arise here are e.g. the choice of the length of the interval is somewhat arbitrary, and the choice of the brightest cluster does not guarantee that it has an age within the interval.

With the new data from Larsen (2009) it is possible to avoid these problems as he gives masses *and* ages for the clusters. By using the age of the brightest cluster as the length of the interval the problem of the brightest cluster not being the most massive one is solved (clusters become only fainter in time, so that the brightest cluster is the most massive *up to its age*), and by construction the mass of the most massive cluster is paired with the correct number of formed clusters. Therefore it is possible to construct more strictly a physical $M_{\rm max}$–$N_{\rm tot}$ relation from observational data. With this it should be possible to shed light on the size-of-sample vs. truncation question, and perhaps also on the steep exponent for the cluster mass function that is needed to explain the brighest cluster – star formation rate relation.

9.2 Stellar initial mass function

The relation between the upper limit of the stellar mass function and the number of stars in a star cluster has been extensively discussed in Chapters 7 and 8. For an observational confirmation the direct estimation and goodness-of-fit tests as presented in Chapter 5 are superior to the analysis of the $m_{\rm max}$-$n_{\rm tot}$ relation. With modern observational surveys large, homogeneous photometric data sets of star clusters and star forming regions are available. The reason for the choice of photometry over spectroscopy is that large spectroscopic surveys

are too observationally expensive. Photometry has the drawback of reduced precision of the mass determination, but is inevitable in order to arrive at a sufficiently large sample of stars and star clusters for statistically sound conclusions.

During a summer project in 2009 in Cambridge, A. Maharaj investigated with C. Clarke, S. Hodgkin, N. Bastian and myself the possibility to obtain robust stellar masses from multiband photometry. The target was the Orion Nebula Cluster, as it has the largest sample of spectroscopically determined masses for comparison. We were able to pinpoint several issues with the mass determination, as for example the choice of isochrones, the choice of photometric bands, allowing for age spreads etc. The current status of this project is promising, the photometric and spectroscopic mass agree for stars more massive than 5 $M_{\odot}$. Below this mass there is only a weak correlation. There are several issues which still have to be cleared to achieve a conclusive picture of photometric mass determination, as for example the use of homogeneous photometric data sets for both photometric and spectroscopic mass determinations or the explanation of the large scatter in the photometrically and spectroscopically determined redddening. Also, the effects of variability and protostellar disks will have to be considered for the mass determination. Nevertheless, the preliminary results are very promising and need to be pursued to further refine the method.

9.3 Stellar accretion rates

A side project to the analysis of the hydrodynamical star formation simulations (Chapter 8) was the analysis of accretion rates of matter onto the sink particles. A preliminary finding is here that, on average, the accretion rate decreases in time. This is understandable in the smaller simulation, as there a high star formation efficiency, up to 55%, is reached quickly so that the available gas for accretion becomes increasingly scarcer. The same behaviour is appearing in the large simulation with only 15% of the gas transformed into stars at the end. The upper envelope in a plot of accretion rate vs. time stays, however, constant. Perhaps the destruction of accretion flows by near encounters of sink particles is responsible for this behaviour.

Another interesting feature of the accretion rates is their scaling with the (simultaneous) mass of the sink particle. From the theory of Bondi-Hoyle-Lyttleton accretion the accretion rate scales with m^2, but in the simulations this relation is shallower. It is an interesting question what causes this behaviour. An answer could be found by revisiting the mass evolution of a sink particle in the simulation with respect to the local gas environment and dynamics.

List of Figures

List of Tables

Bibliography

Aban, I. B., Meerschaert, M. M., & Panorska, A. K. (2006). *Parameter Estimation for the Truncated Pareto Distribution*. Journal of the American Statistical Association, 101:270–277.

Allison, R. J., Goodwin, S. P., Parker, R. J., de Grijs, R., Portegies Zwart, S. F., & Kouwenhoven, M. B. N. (2009a). *Dynamical Mass Segregation on a Very Short Timescale*. ApJ, 700:L99–L103.

Allison, R. J., Goodwin, S. P., Parker, R. J., Portegies Zwart, S. F., de Grijs, R., & Kouwenhoven, M. B. N. (2009b). *Using the minimum spanning tree to trace mass segregation*. MNRAS, 395:1449–1454.

Ambartsumian, V. A. (1960). *On the Evolution of Stellar Systems (George Darwin Lecture)*. QJRAS, 1:152–163.

Anders, P., Bissantz, N., Fritze-v. Alvensleben, U., & de Grijs, R. (2004). *Analysing observed star cluster SEDs with evolutionary synthesis models: systematic uncertainties*. MNRAS, 347:196–212.

Anders, P. & Fritze-v. Alvensleben, U. (2003). *Spectral and photometric evolution of young stellar populations: The impact of gaseous emission at various metallicities*. A&A, 401:1063–1070.

Anderson, T. W. & Darling, D. A. (1952). *Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes*. Annals of Mathematical Statistics, 23:193–212.

Bartholomew, D. J. (1957). *Testing for departure from the exponential distribution*. Biometrika, 44:253–257.

Bastian, N. (2008). *On the star formation rate - brightest cluster relation: estimating the peak star formation rate in post-merger galaxies*. MNRAS, 390:759–768.

Bastian, N., Gieles, M., Lamers, H. J. G. L. M., Scheepmaker, R. A., & de Grijs, R. (2005). *The star cluster population in M51 II. Age distribution and relations among the derived parameters*. A&A, 431:905–924.

Bate, M. R. (2009a). *Stellar, brown dwarf and multiple star properties from hydrodynamical simulations of star cluster formation*. MNRAS, 392:590–616.

Bate, M. R. (2009b). *The importance of radiative feedback for the stellar initial mass function*. MNRAS, 392:1363–1380.

Bate, M. R. & Bonnell, I. A. (2005). *The origin of the initial mass function and its dependence on the mean Jeans mass in molecular clouds*. MNRAS, 356:1201–1221.

Bate, M. R., Bonnell, I. A., & Bromm, V. (2002). *The formation mechanism of brown dwarfs*. MNRAS, 332:L65–L68.

Bate, M. R., Bonnell, I. A., & Price, N. M. (1995). *Modelling accretion in protobinary systems*. MNRAS, 277:362–376.

Baumgardt, H. & Makino, J. (2003). *Dynamical evolution of star clusters in tidal fields*. MNRAS, 340:227–246.

Baxter, M. A. (1980). *Minimum Variance Unbiased Estimation of the Parameters of the Pareto Distribution*. Metrika, 27:133–138.

Becklin, E. E. & Neugebauer, G. (1967). *Observations of an Infrared Star in the Orion Nebula*. ApJ, 147:799–802.

Beg, M. A. (1982). *Optimal Tests and Extimators for Truncated Exponential Families*. Metrika, 29:103–113.

Beg, M. A. (1983). *Unbiased estimators and tests for truncation and scale parameters*. American Journal of Mathematical and Management Sciences, 3:251–274.

Beirlant, J., de Wet, T., & Goegebeur, Y. (2006). *A goodness-of-fit statistic for the Pareto-type behaviour*. Journal of Computational and Applied Mathematics, 186:99–116.

Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. (1994). *Theoretical isochrones from models with new radiative opacities*. A&AS, 106:275–302.

Bianchi, L., Thilker, D. A., Burgarella, D., Friedman, P. G., Hoopes, C. G., Boissier, S., Gil de Paz, A., Barlow, T. A., Byun, Y.-I., Donas, J., Forster, K., Heckman, T. M., Jelinsky, P. N., Lee, Y.-W., Madore, B. F., Malina, R. F., & Martin, D. C. (2005). *Recent Star Formation in Nearby Galaxies from Galaxy Evolution Explorer Imaging: M101 and M51*. ApJ, 619:L71–L74.

Bica, E. L. D., Schmitt, H. R., Dutra, C. M., & Oliveira, H. L. (1999). *A Revised and Extended Catalog of Magellanic System Clusters, Associations, and Emission Nebulae. II. The Large Magellanic Cloud*. AJ, 117:238–246.

Bigourdan, G. (1916). *La découverte de la nébuleuse d' Orion (N.G.C. 1976) par Peiresc*. Comptes Rendus, 162:489–490.

Bik, A., Lamers, H. J. G. L. M., Bastian, N., Panagia, N., & Romaniello, M. (2003a). *Clusters in the inner spiral arms of M51: The cluster IMF and the formation history*. A&A, 397:473–486.

Bik, A., Lenorzer, A., Kaper, L., Comerón, F., Waters, L. B. F. M., de Koter, A., & Hanson, M. M. (2003b). *Identification of the ionizing source of NGC 2024*. A&A, 404:249–254.

Billett, O. H., Hunter, D. A., & Elmegreen, B. G. (2002). *Compact star clusters in nearby dwarf irregular galaxies*. AJ, 123:1454–1475.

Bonnell, I. A., Bate, M. R., & Vine, S. G. (2003). *The hierarchical formation of a stellar cluster*. MNRAS, 343:413–418.

Bonnell, I. A., Clark, P., & Bate, M. R. (2008). *Gravitational fragmentation and the formation of brown dwarfs in stellar clusters*. MNRAS, 389:1556–1562.

Bonnell, I. A. & Clarke, C. J. (1999). *Clustering around Herbig AeBe stars*. MNRAS, 309:461–464.

Bonnell, I. A., Clarke, C. J., & Bate, M. R. (2006a). *The Jeans mass and the origin of the knee in the IMF*. MNRAS, 368:1296–1300.

Bonnell, I. A. & Davies, M. B. (1998). *Mass segregation in young stellar clusters*. MNRAS, 295:691–698.

Bonnell, I. A., Dobbs, C. L., Robitaille, T. P., & Pringle, J. E. (2006b). *Spiral shocks, triggering of star formation and the velocity dispersion in giant molecular clouds*. MNRAS, 365:37–45.

Bonnell, I. A., Vine, S. G., & Bate, M. R. (2004). *Massive star formation: nurture, not nature*. MNRAS, 349:735–741.

Bothun, G. D. & Thompson, I. B. (1988). *Observations with the Parking Lot Camera. I. Surface photometry and color distribution of the Magellanic Clouds*. AJ, 96:877–883.

Boutloukos, S. G. & Lamers, H. J. G. L. M. (2003). *Star clusters formation and disruption time scales - I. An empirical determination of the disruption time of star clusters in four galaxies*. MNRAS, 338:717–732.

Brain, C. W. & Shapiro, S. S. (1983). *A Regression Test for Exponentiality: Censored and Complete Samples*. Technometrics, 25:69–76.

Briceño, C., Luhman, K. L., Hartmann, L., Stauffer, J. R., & Kirkpatrick, J. D. (2002). *The Initial Mass Function in the Taurus Star-forming Region*. ApJ, 580:317–335.

Carpenter, J. M. (2000). *2MASS Observations of the Perseus, Orion A, Orion B, and Monoceros R2 Molecular Clouds*. AJ, 120:3139–3161.

Carpenter, J. M., Heyer, M. H., & Snell, R. L. (2000). *Embedded Stellar Clusters in the W3/W4/W5 Molecular Cloud Complex*. ApJS, 130:381–402.

Carpenter, J. M., Meyer, M. R., Dougados, C., Strom, S. E., & Hillenbrand, L. A. (1997). *Properties of the Monoceros R2 Stellar Cluster*. AJ, 114:198–221.

Carpenter, J. M., Snell, R. L., & Schloerb, F. P. (1990). *Molecular clouds associated with luminous far-infrared sources in the outer Galaxy*. ApJ, 362:147–164.

Carpenter, J. M., Snell, R. L., Schloerb, F. P., & Skrutskie, M. F. (1993). *Embedded star clusters associated with luminous IRAS point sources*. ApJ, 407:657–679.

Cartwright, A. & Whitworth, A. P. (2004). *The statistical analysis of star clusters*. MNRAS, 348:589–598.

Cartwright, A. & Whitworth, A. P. (2008). *The directional analysis of star clusters*. MNRAS, 390:807–813.

Chambers, J., Cleveland, W., Kleiner, B., & Tukey, P. (1983). *Graphical methods for Data Analysis*. Wadsworth, Belmont.

Chandar, R., Leitherer, C., Tremonti, C. A., Calzetti, D., Aloisi, A., Meurer, G. R., & de Mello, D. (2005). *The Stellar Content of Nearby Star-forming Galaxies. III. Unravelling the Nature of the Diffuse Ultraviolet Light*. ApJ, 628:210–230.

Charbonnel, C., Meynet, G., Maeder, A., Schaller, G., & Schaerer, D. (1993). *Grids of Stellar Models - Part Three - from 0.8 to 120-SOLAR-MASSES at Z=0.004*. A&AS, 101:415–+.

Chlebowski, T. & Garmany, C. D. (1991). *On winds and X-rays of O-type stars*. ApJ, 368:241–251.

Clarke, C. J. (2008). *The IMF in clusters: theoretical and observational perspectives*. In Perez, E., de Grijs, R., & Gonzalez Delgado, R. M., editors, *Young massive star clusters - Initial conditions and environments*. Springer. in press.

Crawford, D. F., Jauncey, D. L., & Murdoch, H. S. (1970). *Maximum-Likelihood Estimation of the Slope from Number-Flux Counts of Radio Sources*. ApJ, 162:405–410.

Cysat, J. B. (1619). *Mathemata Astronomica de loco, motu, magnitudine et causis cometae qui sub finem anni 1618 et initium anni 1619 in coelo fulsit (etc.)*. Ex Typographeo Ederiano, apud Elisabetham Angermariam, Viduam, Ingolstadt.

D'Agostino, R. B. & Stephens, M. A., editors (1986). *Goodness-of-Fit Techniques*. Marcel Dekker, New York.

Dale, J. E. & Bonnell, I. A. (2008). *The effect of stellar winds on the formation of a protocluster*. MNRAS, 391:2–13.

Dale, J. E., Bonnell, I. A., Clarke, C. J., & Bate, M. R. (2005). *Photoionizing feedback in star cluster formation*. MNRAS, 358:291–304.

de Grijs, R. & Anders, P. (2006). *How well do we know the age and mass distribution of the star cluster system in the Large Magellanic Cloud?* MNRAS, 366:295–307.

de Grijs, R., Anders, P., Lamers, H. J. G. L. M., Bastian, N., Fritze-v. Alvensleben, U., Parmentier, G., Sharina, M. E., & Yi, S. (2005). *Systematic uncertainties in the analysis of star cluster parameters based on broad-band imaging observations*. MNRAS, 359:874–894.

Dolphin, A. E. (2000). *The star formation histories of two northern LMC fields*. MNRAS, 313:281–290.

Dreyer, J. L. E. (1888). *A New General Catalogue of Nebulae and Clusters of Stars, being the Catalogue of the late Sir John F. W. Herschel, Bart., revised, corrected and enlarged*. Memoirs of the Royal Astronomical Society, 49:1–237.

Dunlop, J. (1828). *A Catalogue of Nebulae and Cluster of Stars in the Southern Hemisphere. Observed at Paramatta in New South Wales*. Philosophical Transactions, 118:113–151.

Egusa, F., Sofue, Y., & Nakanishi, H. (2004). *Offsets between Hα and CO Arms of a Spiral Galaxy, NGC 4254: A New Method for Determining the Pattern Speed of Spiral Galaxies*. PASJ, 56:L45–L48.

Elmegreen, B. G. (2006). *On the Similarity between Cluster and Galactic Stellar Initial Mass Functions*. ApJ, 648:572–579.

Elmegreen, B. G. (2009a). *The Initial Mass Function in Clusters*. In Livio, M. & Villaver, E., editors, *Massive stars: from Pop II and GRBs to the Milky Way, Proceedings of the Space Telescope Science Institute Symposium, Baltimore, May 8-11 2006*, p. 93ff. Cambridge University Press.

Elmegreen, B. G. (2009b). *The Stellar Initial Mass Function in 2007: A Year for Discovering Variations*. In Sheth, K., Noriega-Crespo, A., Ingalls, J., & Paladini, R., editors, *The Evolving ISM in the Milky Way and Nearby Galaxies*. online only at `http://ssc.spitzer.caltech.edu/mtgs/ismevol`.

Feigelson, E. D., Martin, A. L., McNeill, C. J., Broos, P. S., & Garmire, G. P. (2009). *Stellar Clusters in the NGC 6334 Star-Forming Complex*. AJ, 138:227–239.

Fellhauer, M., Wilkinson, M. I., & Kroupa, P. (2009). *Merging time-scales of stellar subclumps in young star-forming regions*. MNRAS, 397:954–962.

Figer, D. F. (2005). *An upper limit to the masses of stars*. Nature, 434:192–194.

Foderà Serio, G., Indorato, L., & Nastasi, P. (1985). *G. B. Hodierna's observations of nebulae and his cosmology*. Journal for the History of Astronomy, 16:1–36.

Galilei, G. (1610). *Sidereus Nuncius (etc.)*. Apud Thomas Baglionum, Venetiis (Venice).

Gan, F. F. & Koehler, K. J. (1990). *Goodness-of-Fit Tests Based on P-P Probability Plots*. Technometrics, 32:289–303.

Getman, K. V., Flaccomio, E., Broos, P. S., Grosso, N., Tsujimoto, M., & Townsley, L. e. a. (2005). *Chandra Orion Ultradeep Project: Observations and Source Lists*. ApJS, 160:319–352.

Gieles, M. (2009). *The early evolution of the star cluster mass function*. MNRAS, 394:2113–2126.

Gieles, M. & Bastian, N. (2008). *An alternative method to study star cluster disruption*. A&A, 482:165–171.

Gieles, M., Bastian, N., Lamers, H. G. J. M., & Mout, J. N. (2005). *The Star Cluster Population of M51: III. Cluster disruption and formation history*. A&A, 441:949–960.

Gieles, M., Larsen, S. S., Bastian, N., & Stein, I. T. (2006a). *The luminosity function of young star clusters: implications for the maximum mass and luminosity of clusters*. A&A, 450:129–145.

Gieles, M., Larsen, S. S., Scheepmaker, R. A., Bastian, N., Haas, M. R., & Lamers, H. J. G. L. M. (2006b). *Observational evidence for a truncation of the star cluster initial mass function at the high mass end*. A&A, 446:L9–L12.

Gingerich, O. (1987). *The Mysterious Nebulae - 1610-1924*. JRASC, 81:113–128.

Girardi, L., Bertelli, G., Bressan, A., Chiosi, C., Groenewegen, M. A. T., Marigo, P., Salasnich, B., & Weiss, A. (2002). *Theoretical isochrones in several photometric systems. I. Johnson-Cousins-Glass, HST/WFPC2, HST/NICMOS, Washington, and ESO Imaging Survey filter sets*. A&A, 391:195–212.

Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. (2000). *Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 $M_{\odot}$, and from Z=0.0004 to 0.03*. A&AS, 141:371–383.

Girardi, L., Chiosi, C., Bertelli, G., & Bressan, A. (1995). *Age distribution of LMC clusters from their integrated UBV colors: history of star formation*. A&A, 298:87–106.

Gutermuth, R. A., Megeath, S. T., Muzerolle, J., Allen, L. E., Pipher, J. L., Myers, P. C., & Fazio, G. G. (2004). *The NGC 7129 Young Stellar Cluster: A Combined Spitzer, MMT, and Two Micron All Sky Survey Census of Disks, Protostars, and Outflows*. ApJS, 154:374–378.

Gutermuth, R. A., Megeath, S. T., Myers, P. C., Allen, L. E., Pipher, J. L., & Fazio, G. G. (2009). *A Spitzer Survey of Young Stellar Clusters Within One Kiloparsec of the Sun: Cluster Core Extraction and Basic Structural Analysis*. ApJS, 184:18–83.

Gutermuth, R. A., Megeath, S. T., Pipher, J. L., Williams, J. P., Allen, L. E., Myers, P. C., & Raines, S. N. (2005). *The Initial Configuration of Young Stellar Clusters: A K-Band Number Counts Analysis of the Surface Density of Stars*. ApJ, 632:397–420.

Haas, M. R., Gieles, M., Scheepmaker, R. A., Larsen, S. S., & Lamers, H. J. G. L. M. (2008). *ACS imaging of star clusters in M 51. II. The luminosity function and mass function across the disk*. A&A, 487:937–949.

Halley, E. (1716). *An Account of Several Nebulae or Lucid Spots Like Clouds, Lately Discovered among the Fixt Stars by Help of the Telescope*. Philosophical Transactions, 29:390–392 (vol 29 covers 1714–1716, published 1717).

Hannon, P. M. & Dahiya, R. C. (1999). *Estimation of parameters for the truncated exponential distribution*. Communications in Statistics: Theory and Methods, 28:2591–2612.

Harris, J. & Zaritsky, D. (2001). *A method for determining the star formation history of a mixed stellar population*. ApJS, 136:25–40.

Harris, J. & Zaritsky, D. (2009). *The Star Formation History of the Large Magellanic Cloud*. AJ, 138:1243–1260.

Hayashi, C. (1961). *Stellar evolution in early phases of gravitational contraction.* PASJ, 13:450–452.

Henyey, L. G., Lelevier, R., & Levée, R. D. (1955). *The Early Phases of Stellar Evolution*. PASP, 67:154–160.

Herschel, J. F. W. (1847). *Results of astronomical observations made during the years 1834, 5, 6, 7, 8, at the Cape of Good Hope*. Smith, Elder and Co., London.

Herschel, J. F. W. (1864). *Catalogue of Nebulae and Clusters of Stars*. Philosophical Transactions, 154:1–137.

Herschel, W. (1786). *Catalogue of One Thousand New Nebulae and Clusters of Stars*. Philosophical Transactions, 76:457–499.

Herschel, W. (1789). *Catalogue of a Second Thousand of New Nebulae and Clusters of Stars; With a Few Introductory Remarks on the Construction fo the Heavens*. Philosophical Transactions, 79:212–255.

Herschel, W. (1791). *On Nebulous Stars, properly so called.* Philosophical Transactions, 81:71–88.

Herschel, W. (1802). *Catalogue of 500 New Nebulae, Nebulous Stars, Planetary Nebulae and Clusters of Stars; With Remarks on the Construction of the Heavens*. Philosophical Transactions, 92:477–528.

Hodge, P. (1988). *A catalog of 255 new clusters in the Large Magellanic Cloud*. PASP, 100:1051–1070.

Hodierna, J. B. (1654). *Deque admirandis coeli characteribus opuscula duo, ... in secundo vero quid, quales, quoque sind stellae luminosae, nebulosae, necnon, et occultae manifestantur*. Nicolaus Bua, Panormi (Palermo).

Holtzman, J. A., Gallagher III, J. S., Cole, A. A., Mould, J. R., Grillmair, C. J., Ballester, G. E., Burrows, C. J., Clarke, J. T., Crisp, D., Evans, R. W., Griffiths, R. E., Hester, J. J., Hoessel, J. G., Scowen, P. A., Stapelfeldt, K. R., Trauger, J. T., & Watson, A. M. (1999). *Observations and implications of the star formation history of the Large Magellanic Cloud*. AJ, 118:2262–2279.

Hoskin, M. (1990). *Rosse, Robinson , and the Resolution of the Nebulae*. Journal for the History of Astronomy, 21:331–344.

Hoversten, E. A. & Glazebrook, K. (2008). *Evidence for a Nonuniversal Stellar Initial Mass Function from the Integrated Properties of SDSS Galaxies*. ApJ, 675:163–187.

Huggins, W. (1864). *On the Spectra of some of the Nebulae*. Philosophical Transactions, 154:437–444.

Huggins, W. (1865). *On the Spectrum of the Great Nebula in the Sword-Handle of Orion*. Proceedings of the Royal Society of London, 14:39–42.

Humboldt, A. v. (1850). *Kosmos. Entwurf einer physischen Weltbeschreibung*, vol. 3. J. G. Cotta'scher Verlag, Stuttgart und Tübingen.

Hunter, D. A., Elmegreen, B. G., Dupuy, T. J., & Mortonson, M. (2003). *Cluster mass functions in the Large Magellanic Clouds: Fading and size-of-sample effects*. AJ, 126:1836–1848.

Hunter, D. A., Shaya, E. J., Holtzman, J. A., Light, R. M., O'Neil, Jr., E. J., & Lynds, R. (1995). *The Intermediate Stellar Mass Population in R136 Determined from Hubble Space Telescope Planetary Camera 2 Images*. ApJ, 448:179–194.

Huyghens, C. (1659). *Systema Saturnium, sive de causis mirandorum Saturni Phaenomenon, et Comite ejus Planeta Novo*. Ex Typographia Adriani Vlacq, Hagae-Comitits (The Hague).

Jackson, O. A. Y. (1967). *An Analysis of Departures from the Exponential Distribution*. Journal of the Royal Statistical Society B, 29:540–549.

Jappsen, A.-K., Klessen, R. S., Larson, R. B., Li, Y., & Mac Low, M.-M. (2005). *The stellar mass spectrum from non-isothermal gravoturbulent fragmentation*. A&A, 435:611–623.

Jauncey, D. L. (1967). *Re-examination of the Source Counts for the 3C Revised Catalogue*. Nature, 216:877–878.

Javiel, S. C., Santiago, B. X., & Kerber, L. O. (2005). *Constraints on the star formation history of the Large Magellanic Cloud*. A&A, 431:73–85.

Jeans, J. H. (1902). *The Stability of a Spherical Nebula*. Philosophical Transactions Series A, 199:1–53.

Kant, I. (1755). *Allgemeine Naturgeschichte und Theorie des Himmels, oder Versuch von der Verfassung und dem mechanischen Ursprunge des ganzen Weltgebäudes nach Newtonischen Grundsätzen abgehandelt*. Johann Friederich Petersen, Königsberg und Leipzig.

Kelson, D. D., Illingworth, G. D., Freedman, W. F., Graham, J. A., Hill, R., Madore, B. F.,

Saha, A., Stetson, P. B., Kennicutt, Jr., R. C., Mould, J. R., Hughes, S. M., Ferrarese, L., Phelps, R., Turner, A., Cook, K. H., Ford, H., Hoessel, J. G., & Huchra, J. (1996). *The Extragalactic Distance Scale Key Project. III. The Discovery of Cepheids and a New Distance to M101 Using the Hubble Space Telescope*. ApJ, 463:26–59.

Kepler, J. (1609). *Astronomia nova αιτιολογητοσ, etc.* (Gotthard Vögelin), (Heidelberg).

Kim, S., Staveley-Smith, L., Dopita, M. A., Freeman, K. C., Sault, R. J., Kesteven, M. J., & McConnell, D. (1998). *An H I aperture synthesis mosaic of the Large Magellanic Cloud*. ApJ, 503:674–688.

Kimber, A. C. (1985). *Tests for the exponential, Weibull and Gumbel distributions based on the stabilized probability plot*. Biometrika, 72:661–663.

Kitsionas, S. & Whitworth, A. P. (2002). *Smoothed Particle Hydrodynamics with particle splitting, applied to self-gravitating collapse*. MNRAS, 330:129–136.

Kitsionas, S. & Whitworth, A. P. (2007). *High-resolution simulations of clump-clump collisions using SPH with particle splitting*. MNRAS, 378:507–524.

Kleinmann, D. E. & Low, F. J. (1967). *Discovery of an Infrared Nebula in Orion*. ApJ, 149:L1–L4.

Koen, C. (2006). *On the upper limit on stellar masses in the Large Magellanic Cloud cluster R136*. MNRAS, 365:590 – 594.

Köppen, J., Weidner, C., & Kroupa, P. (2007). *A possible origin of the mass-metallicity relation of galaxies*. MNRAS, 375:673–684.

Kroupa, P. (2001). *On the variation of the initial mass function*. MNRAS, 322:231–246.

Kroupa, P. (2002). *The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems*. Science, 295:82–91.

Kroupa, P. (2005). *The Fundamental Building Blocks of Galaxies*. In Turon, C., O'Flaherty, K. S., & Perryman, M. A. C., editors, *ESA SP-576: The Three-Dimensional Universe with Gaia*.

Kroupa, P., Aarseth, S., & Hurley, J. (2001). *The formation of a bound star cluster: from the Orion nebula cluster to the Pleiades*. MNRAS, 321:699–712.

Kroupa, P. & Boily, C. M. (2002). *On the mass function of star clusters*. MNRAS, 336:1188–1194.

Kroupa, P. & Weidner, C. (2003). *Galactic-Field Initial Mass Functions of Massive Stars*. ApJ, 598:1076–1078.

Kumar, M. S. N. & Schmeja, S. (2007). *The spatial distribution of substellar objects in IC 348 and the Orion Trapezium cluster*. A&A, 471:L33–L36.

Kurth, O. M., Fritze-v. Alvensleben, U., & Fricke, K. J. (1999). *Evolutionary synthesis of simple stellar populations. Colours and indices*. A&AS, 138:19–30.

La Caille, N. L. (1755). *Sur les étoiles nébuleuses du ciel austral*. Mémoires de l'Académie Royale des Sciences, pp. 194–199.

Lada, C. J. & Lada, E. A. (2003). *Embedded Clusters in Molecular Clouds*. Annu. Rev. Astron. Astrophys., 41:57–115.

Lada, E. A., Depoy, D. L., Evans, II, N. J., & Gatley, I. (1991). *A 2.2 micron survey in the L1630 molecular cloud*. ApJ, 371:171–182.

Lamers, H. J. G. L. M., Gieles, M., Bastian, N., Baumgardt, H., Karchenko, N. V., & Portegies Zwart, S. (2005a). *An analytical description of the disruption of star clusters in tidal fields with an application to Galactic open clusters*. A&A, 441:117–129.

Lamers, H. J. G. L. M., Gieles, M., & Portegies Zwart, S. F. (2005b). *Disruption time scales of star clusters in different galaxies*. A&A, 429:173–179.

Laplace, P. S. d. (1796). *Exposition du systéme du monde*. Imprimerie du Cercle-Social, Paris.

Larsen, S. S. (2002). *The luminosity function of star clusters in spiral galaxies*. AJ, 124:1393–1409.

Larsen, S. S. (2009). *The mass function of young star clusters in spiral galaxies*. A&A, 494:539–551.

Layzer, D. (1964). *The Formation of Stars and Galaxies: Unified Hypotheses*. ARA&A, 2:341–362.

Lilliefors, H. W. (1967). *On the Kolmogorov-Smirnov Test for Normality with mean and variance unknown*. Journal of the American Statistical Association, 62:399–402.

Lilliefors, H. W. (1969). *On the Kolmogorov-Smirnov Test for the exponential distribution with mean unknown*. Journal of the American Statistical Association, 64:387–389.

Luhman, K. L., Stauffer, J. R., Muench, A. A., Rieke, G. H., Lada, E. A., Bouvier, J., & Lada, C. J. (2003). *A Census of the Young Cluster IC 348*. ApJ, 593:1093–1115.

Maíz Apellániz, J. & Úbeda, L. (2005). *Numerical biases on initiall mass function determinations created by binning*. ApJ, 629:873–880.

Martini, P. & Ho, L. C. (2004). *A Population of Massive Globular Clusters in NGC 5128*. ApJ, 610:233–246.

Martins, F., Schaerer, D., & Hillier, D. J. (2005). *A new calibration of stellar parameters of Galactic O stars*. A&A, 436:1049–1065.

Maschberger, T. & Clarke, C. J. (2008). *Maximum stellar mass versus cluster membership number revisited*. MNRAS, 391:711–717.

Maschberger, T. & Kroupa, P. (2007). *A new method to derive star formation histories of galaxies from their star cluster distributions*. MNRAS, 379:34–42.

Maschberger, T. & Kroupa, P. (2009). *Estimators for the exponent and upper limit, and goodness-of-fit tests for (truncated) power-law distributions*. MNRAS, 395:931–942.

Massey, P. (1998). *The Initial Mass Function of Massive Stars in the Local Group*. In Gilmore, G. & Howell, D., editors, *The Stellar Initial Mass Function (38th Herstmonceux Conference)*, vol. 142 of *Astronomical Society of the Pacific Conference Series*, pp. 17–44.

Massey, P. (2002). *A UBVR CCD survey or the Magellanic Clouds*. ApJS, 141:81–122.

Massey, P. & Hunter, D. A. (1998). *Star Formation in R136: A Cluster of O3 Stars Revealed by Hubble Space Telescope Spectroscopy*. ApJ, 493:180–194.

McCrea, W. H. (1960). *The Origin of the Solar System*. Royal Society of London Proceedings Series A, 256:245–266.

McMillan, S. L. W., Vesperini, E., & Portegies Zwart, S. F. (2007). *A Dynamical Origin for Early Mass Segregation in Young Star Clusters*. ApJ, 655:L45–L49.

McNally, D. (1971). *Theories of star formation*. Reports on Progress in Physics, 34:71–108.

Messier, C. (1771). *Catalouges des nébuleuses et des amas d'étoiles*. Mémoires de l'Académie Royale des Sciences, pp. 435–461 (published 1774).

Messier, C. (1784). *Catalogue des nébuleuses et des amas d'étoiles*. Connoissance des Temps, Pour l'Année commune 1784, pp. 117–269 (published in 1781).

Messier, C. (1801). *Observationes Astronomiques, 1770-1774*. Connaissance des Tems, a l'usage des astronomes et

des navigateurs pour l'an IX [1801], pp. 434–465 (published Frucidor An VI. [1798]).

Meurer, G. R., Wong, O. I., Kim, J. H., Hanish, D. J., Heckman, T. M., Werk, J., Bland-Hawthorn, J., et al. (2009). *Evidence for a Nonuniform Initial Mass Function in the Local Universe*. ApJ, 695:765–780.

Michael, J. R. (1983). *The stabilized probability plot*. Biometrika, 70:11–17.

Mieske, S., Hilker, M., & Infante, L. (2002). *Ultra compact objects in the Fornax cluster of galaxies: Globular clusters or dwarf galaxies?* A&A, 383:823–837.

Moeckel, N. & Bonnell, I. A. (2009). *Does subcluster merging accelerate mass segregation in local clusters?* MNRAS, 400:657–664.

Moraux, E., Lawson, W. A., & Clarke, C. (2007). η *Chamaeleontis: abnormal initial mass function or dynamical evolution?* A&A, 473:163–170.

Murdin, P. & Penston, M. V. (1977). *The Lambda Orionis association*. MNRAS, 181:657–665.

Oey, M. S. & Clarke, C. J. (2005). *Statistical confirmation of a stellar upper mass limit*. ApJ, 620:L43–L46.

Olsen, K. A. G. (1999). *Star formation histories from Hubble Space Telescope color-magnitude diagrams of six fields of the Large Magellanic Cloud*. AJ, 117:2244–2267.

Parenago, P. P. (1954). *Issledovaniia zvezd v oblasti tumannosti Oriona*. Trudy Gosudarstvennogo Astronomicheskogo Instituta, 25:3–538.

Parenago, P. P. & Sharov, A. S. (1961). *On the photometric catalogue of stars in the region of the Orion Nebula*. AJ, 66:103–+.

Parker, J. W., Hill, J. K., Cornett, R. H., Hollis, J., Zamkoff, E., Bohlin, R. C., O'Connell, R. W., Neff, S. G., Roberts, M. S., Smith, A. M., & Stecher, T. P. (1998). *Ultraviolet Imaging Telescope Observations of the Magellanic Clouds*. AJ, 116:180–208.

Parker, R. J. & Goodwin, S. P. (2007). *Do O-stars form in isolation?* MNRAS, 380:1271–1275.

Parmentier, G. & de Grijs, R. (2008). *The poorly constrained cluster disruption time-scale in the Large Magellanic Cloud*. MNRAS, 383:1103–1120.

Parmentier, G., Goodwin, S. P., Kroupa, P., & Baumgardt, H. (2008). *The Shape of the Initial Cluster Mass Function: What It Tells Us about the Local Star Formation Efficiency*. ApJ, 678:347–352.

Pflamm-Altenburg, J. & Kroupa, P. (2008). *Clustered star formation as a natural explanation for the Hα cut-off in disk galaxies*. Nature, 455:641–643.

Pflamm-Altenburg, J., Weidner, C., & Kroupa, P. (2007a). *Converting Hα Luminosities into Star Formation Rates*. ApJ, 671:1550–1558.

Pflamm-Altenburg, J., Weidner, C., & Kroupa, P. (2007b). *Converting Hα Luminosities into Star Formation Rates*. ApJ, 671:1550–1558.

Pflamm-Altenburg, J., Weidner, C., & Kroupa, P. (2009). *Diverging UV and Hα fluxes of star-forming galaxies predicted by the IGIMF theory*. MNRAS, 395:394–400.

Pietrzyński, G. & Udalski, A. (2000). *The Optical Graviational Lensing Experiment. Ages of about 600 Star Clusters from the LMC*. Acta Astonomica, 50:337–354.

Preibisch, T., Brown, A. G. A., Bridges, T., Guenther, E., & Zinnecker, H. (2002). *Exploring the Full Stellar Population of the Upper Scorpius OB Association*. AJ, 124:404–416.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996a). *Numerical Recipes in Fortran 77*. Cambridge University Press, 2. edition.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1996b). *Numerical Recipes in Fortran 90*. Cambridge University Press, 2. edition.

Prisinzano, L., Damiani, F., Micela, G., & Sciortino, S. (2005). *The star formation region NGC 6530: Distance, ages and initial mass function*. A&A, 430:941–957.

Prisinzano, L., Micela, G., Flaccomio, E., Stauffer, J. R., Megeath, T., Rebull, L., Robberto, M., Smith, K., Feigelson, E. D., Grosso, N., & Wolk, S. (2008). *X-Ray Properties of Protostars in the Orion Nebula*. ApJ, 677:401–424.

Ptolemaius, C. (1816). *Κλαυδου Πτολεμαιου Μαθημτικη Συνταισις. Composition Mathematique de Claude Ptolémée, ... Traduite ... par M. l' abbé Halma; et suivie des notes de M. Delambre*, vol. 2. J.-M. Eberhart, Imprim. du College Royal de France, Paris.

Reipurth, B., editor (2008a). *Handbook of Star Forming Regions, Volume I: The Northern Sky*. Astronomical Society of the Pacific, San Francisco.

Reipurth, B., editor (2008b). *Handbook of Star Forming Regions, Volume II: The Southern Sky*. Astronomical Society of the Pacific, San Francisco.

Rho, J., Corcoran, M. F., Chu, Y.-H., & Reach, W. T. (2001). *X-Rays and Protostars in the Trifid Nebula*. ApJ, 562:446–455.

Salpeter, E. E. (1955). *The luminosity function and stellar evolution*. ApJ, 121:161–167.

Schaerer, D., Meynet, G., Maeder, A., & Schaller, G. (1993). *Grids of stellar models. II - From 0.8 to 120 solar masses at Z = 0.008*. A&AS, 98:523–527.

Scheepmaker, R. A., Haas, M. R., Gieles, M., Bastian, N., Larsen, S. S., & Lamers, H. J. G. L. M. (2007). *ACS imaging of star clusters in M 51. I. Identification and radius distribution*. A&A, 469:925–940.

Scheepmaker, R. A., Lamers, H. J. G. L. M., Anders, P., & Larsen, S. S. (2009). *The spatial distribution of star and cluster formation in M 51*. A&A, 494:81–93.

Schmeja, S., Gouliermis, D. A., & Klessen, R. S. (2009). *The Clustering Behavior of Pre-Main-Sequence Stars in NGC 346 in the Small Magellanic Cloud*. ApJ, 694:367–375.

Schmeja, S. & Klessen, R. S. (2004). *Protostellar mass accretion rates from gravoturbulent fragmentation*. A&A, 419:405–417.

Schmeja, S. & Klessen, R. S. (2006). *Evolving structures of star-forming clusters*. A&A, 449:151–159.

Schmeja, S., Kumar, M. S. N., & Ferreira, B. (2008). *The structures of embedded clusters in the Perseus, Serpens and Ophiuchus molecular clouds*. MNRAS, 389:1209–1217.

Schmidt-Kaler, T. (1982). *The stars*. In Schaifers, K. & Voigt, H. H., editors, *Sterne und Sternhaufen*, vol. VI/2b of *Landolt-Börnstein Zahlenwerte und Funktionen aus Naturwissenschaft und Technik, Astronomie und Astrophysik*, ch. 4. Springer-Verlag.

Schulz, J., Fritze-v. Alvensleben, U., Möller, C. S., & Fricke, K. J. (2002). *Spectral and photometric evolution of simple stellar populations at varios metallicities*. A&A, 391:1–11.

Selman, F. J. & Melnick, J. (2008). *The Scale-Free Character of the Cluster Mass Function and the Universality of the Stellar Initial Mass Function*. ApJ, 689:816–824.

Shapiro, S. S. & Wilk, M. B. (1972). *An Analysis of Variance Tests for the Exponential Distribution (Complete Samples)*. Technometrics, 14:355–370.

Siebert, H. (2009). *Peirescs Nebel im Sternbild Orion — eine neue Textgrundlage für die*

Geschichte von M42. Annals of Science, 66:231–246.

Smecker-Hane, T. A., Cole, A. A., Gallagher III, J. S., & Stetson, P. B. (2002). *The star formation history of the Large Magellanic Cloud*. ApJ, 566:239–244.

Stephens, M. A. (1978). *Goodness of fit tests with special reference to tests for exponentiality*. Technical Report 262, Department of Statistics, Stanford University.

Subramaniam, A. (2004). *Cluster formation versus star formation rates around six regions in the Large Magellanic Cloud*. A&A, 425:837–847.

Takáts, K. & Vinkó, J. (2006). *Distance estimate and progenitor characteristics of SN 2005cs in M51*. MNRAS, 372:1735–1740.

Testi, L., Palla, F., & Natta, A. (1998). *A search for clustering around Herbig Ae/Be stars. II. Atlas of the observed sources*. A&AS, 133:81–121.

Testi, L., Palla, F., Prusti, T., Natta, A., & Maltagliati, S. (1997). *A search for clustering around Herbig Ae/Be stars.* A&A, 320:159–166.

Vacca, W. D., Garmany, C. D., & Shull, J. M. (1996). *The Lyman-Continuum Fluxes and Stellar Parameters of O and Early B-Type Stars*. ApJ, 460:914–931.

Vanbeveren, D. (1982). *On the difference between the initial mass function of single stars and of primaries of binaries*. A&A, 115:65–68.

Vanbeveren, D. (1983). *Theoretical evolution of massive stellar aggregates*. A&A, 124:71–76.

von Weizsäcker, C. F. (1951). *The Evolution of Galaxies and Stars.* ApJ, 114:165–186.

Walker, M. F. (1956). *Studies of Extremely Young cluster. I. NGC 2264.* ApJS, 2:365–387.

Wang, S. & Looney, L. W. (2007). *Young Stellar Groups around Herbig Ae/Be Stars: A Low-Mass YSO Census*. ApJ, 659:1360–1372.

Weidner, C. & Kroupa, P. (2004). *Evidence for a fundamental stellar upper mass limit from clustered star formation*. MNRAS, 348:187–191.

Weidner, C. & Kroupa, P. (2005). *The variation of integrated star initial mass mass functions among galaxies*. ApJ, 625:754–762.

Weidner, C. & Kroupa, P. (2006). *The maximum stellar mass, star-cluster formation and composite stellar populations*. MNRAS, 365:1333–1347.

Weidner, C., Kroupa, P., & Bonnell, I. A. D. (2010). *The relation between the most-massive star and its parental star cluster mass*. MNRAS, 401:275–293.

Weidner, C., Kroupa, P., & Larsen, S. S. (2004). *Implications for the formation of star clusters from extragalactic star formation rates*. MNRAS, 350:1503–1510.

Weidner, C., Kroupa, P., & Maschberger, T. (2009). *The influence of multiple stars on the high-mass stellar initial mass function and age dating of young massive star clusters*. MNRAS, 393:663–680.

Weidner, C., Kroupa, P., Nürnberger, D. E. A., & Sterzik, M. F. (2007). *On the infant weight loss of low- to intermediate-mass star clusters*. MNRAS, 376:1879–1885.

Wilk, M. B. & Gnanadesikan, R. (1968). *Probability plotting methods for the analysis of data*. Biometrika, 55:1–17.

Zahn, C. T. (1971). *Graph-theoretical methods for detecting and desrcibing gestalt clusters*. IEEE Transactions on Computers, 20(1):68–86.

Zinnecker, H. & Yorke, H. W. (2007). *Toward Understanding Massive Star Formation*. ARA&A, 45:481–563.

This work made use of the ADS and NED databases. Programs were written in Fortran 95 using the freely available g95 compiler (www.g95.org) and algorithms fromPress et al. (1996a) and Press et al. (1996b). This text was typeset using LaTeX.

Acknowledgements

"If I have seen further it is by standing on ye shoulders of ..." — Prof. P. Kroupa and Prof. C. Clarke, my supervisors. Thank you for guiding me through the exciting topics of this thesis.
I would like to thank also my collaborators and co-authors Ian Bonnell (Chapter 8), Remco Scheepmaker and Simon Goodwin (both Chapter 6).
Of course, the friendly and congenial environments of the Stellar Populations and Dynamics group in Bonn and the Institute of Astronomy in Cambridge have their share as well in helping me to finish this thesis.
With respect to individual chapters I thank Carsten Weidner for useful discussions, and Ylva Schuberth for critical reading of the Chapter on the method to derive star formation histores (Chapter 3); Richard de Grijs and Peter Anders for kindly providing me with the age/mass data for the star clusters in the Large Magellanic Cloud (used in Chapter 4); Douglas Heggie for critically reading the Chapter on statistical methods (Chapter 5); Carsten Weidner, Leonardo Testi, Thomas Preibisch, and the referee Hans Zinnecker for useful discussions and helping to collect the data for Chapter 7; Simon Goodwin refereeing the Chapter on star-cluster formation (8) and his valuable comments.
I appreciate funding by the Stellar Populations and Dynamics group at the AIfA and by the European Union via an EARA-EST fellowship to visit Cambridge for six months and via the Research Training Network "Constellation" for the time I spent in Cambridge since the EARA-EST fellowship.

Finally, I would like to thank my mother for her support during all my studies.