Molecular gas in hydrodynamical simulations of galaxy formation

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

von
Alexander Schäbe
aus
Dresden

Bonn, 2021

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

1. Gutachter:	Prof. Dr. Cristiano Porciani
2. Gutachter:	Prof. Dr. Peter Schneider
Tag der Promotion:	30.06.2022
Erscheinungsjahr:	2022

“Solve mentem a molestis
Mentem ad concretum dirige
Unum e jhanas intra
Intra unum e jhanas ”

—from Samadhi - Prelude, *Epica – Design your universe*

Abstract

Observations of nearby galaxies indicate a strong correlation between their molecular gas content and their star formation rate. This leads to the conclusion that molecular hydrogen – the most abundant molecule in galaxies – is the fuel for star formation. Recent results of sub-millimetre facilities like ALMA make it possible to study the molecular gas content in high-redshift systems. This fosters the understanding of the conditions for its formation and its impact at the early phases of galaxy formation – a fundamental building block in the topic of galaxies evolution.

However, there is a subtlety to the hydrogen molecule: its missing dipole moment makes a direct detection impossible. Thus, observations rely on so called tracers. Carbon monoxide is the most prominent tracer, but also neutral and ionized carbon have already been used for the determination of the molecular hydrogen content of a galaxy.

Over the last years, numerical simulations have been used to approach the topic of galaxy formation from a theoretical perspective. Different models allow capturing a variety of effects in this highly complex process. Since these simulations have a limited resolution, they depend on subgrid models which are supposed to compensate for the incapability to include all effects in the interstellar medium such as the chemistry and the abundances of molecules. This becomes even more important for large cosmological volumes. Thus, to study the formation of molecular gas and its impact on star formation and galaxy evolution, subgrid models are essential.

In this thesis, we will compare different recipes for the H_2 formation in galaxies. Chapter 2 shows the results of a suite of cosmological simulations run with the adaptive-mesh-refinement code Ramses where different H_2 formation models are implemented via a subgrid model. One model assumes chemical equilibrium, another one non-equilibrium between the processes of formation and destruction of H_2. Lastly, we apply a semi-analytic model to the simulations which uses scaling relations for galaxies coming from observations in order to assign the molecular hydrogen mass to a dark matter halo at given redshift. We find that all models give similar results. Nevertheless, for morphology studies, running a high resolution simulation with a dynamic H_2 formation model is the best choice. However, we find different results for the H_2 content with respect to the resolution of a simulation.

Applying observational limits to our simulated objects, we find a good agreement with current surveys. However, due to the limitations of observations, our results for the molecular gas budget on cosmological scales imply that observations still miss a significant amount of molecular gas that resides in undetected galaxies.

In chapter 3, we extend the chemistry in our subgrid model used in chapter 2 to be able to trace neutral carbon, ionized carbon and carbon monoxide. For that, a simplified chemical network is embedded into the RAMSES code. This model is applied on-the-fly to a dwarf galaxy up to a reshift of $z = 4$ and in post-processing in a cosmological volume to the objects from chapter 2. In the end, we compare the outcome of the previously mentioned calculations with observational data for the abundance fractions of [C]/[H_2] and [CO]/[H_2] and find good agreement. The spatial distribution of CO coincides well with the H_2 confirming that CO is a good tracer for dense star forming regions.

Overall, the findings can be used in the future to run cosmological simulations that calculate the content

of molecular gas tracers. In a next step, luminosity calculations for C, C^+ and CO can be performed to enable a detailed study of the different dependencies of the conversion factor α_{CO} – i.e. the relationship between molecular hydrogen and CO – as well as the reliability of other tracers.

Contents

CHAPTER 1

Introduction

1.1 A brief history of cosmology

The look of the nightsky always fascinated people at every epoch. Even just with the naked eye, a broad variety of structures and objects is visible, from planets to stars, or even different galaxies. Over millennia, the appearance of the sky was the source of inspiration for a lot of cultures and myths. But the question what our place in the universe is was always one of the most compelling.
One important contribution to this question was made in 1610 by Galileo Gallilei who found out that our galaxy consists of many single stars. Even the ancient Greek already had this idea but Galileo was able to resolve single stars by using a telescope. However, at that time, instruments were not powerful enough to draw the same conclusion for the Andromeda galaxy which was thought to be a nebula due to its diffuse appearance. Thus it was assumed not to have any stellar content at all. In 1755, the German philosopher Immanuel Kant published in his book "Allgemeine Naturgeschichte und Theorie des Himmels" the hypothesis that our galaxy would appear from far away as a nebular structure as well. This thought was not new at all, as Giordano Bruno and Thomas Wright already had that idea nearly 200 years earlier. Further, Kant stated that the universe might be filled with objects like our own galaxy which he called "Welteninseln" ("island universes"). He assumed that the matter in the universe evolves from a gas cloud ("Urwolke") to more complex structures like galaxies and solar systems under the influence of gravity. Almost 40 years later, the French mathematician Laplace published a similar theory. These were the first ideas about our universe and its history. Although at that time, the question of the origin of our universe was more a philosophical question rather than a scientific one, modern physics today still uses the term *cosmology* for the description of the evolution of our universe and the formation of structures (e.g galaxies) therein.

With the beginning of the 20th century, Einstein developed first the theory of *special relativity* which would be published in 1905 (Einstein, 1905a; Einstein, 1905b). Later on, he extended that to the theory of *general relativity*. Between 1908 and 1915, he published several works related to the subject but the correct field equations that provided the mathematical framework for modern cosmology were found later on and published in 1915 (Einstein, 1915). But until then, the nature of the universe and our place therein was not clear at all. This culminated in the *Great Debate* (also *Shapley–Curtis debate*) held on 26 April 1920. It focussed on the question if these nebular structures are either relatively small and close to the Milky Way (Shapley) or large and distant (Curtis), implying that the universe is quite large compared to the scale of our own galaxy (Shapley and Curtis, 1921). Thus, the question of the existence of other galaxies got more attention.
In the year 1923, Hubble was able to resolve single stars in the Andromeda galaxy and to determine the

distance between the Milky Way and the Andromeda galaxy. He realized that Andromeda is not a nebula and – due to the large distance from Earth – that it is not part of our own galaxy. Thus, he provided an answer to the Great Debate.
In 1927, the priest Georges Lemaître published the first paper containing the idea of an expanding universe (Lemaître, 1927) without knowing the work of Alexander Alexandrowitsch Friedmann who found a similar result by solving Einstein's field equations earlier (Friedman, 1922). The concept of an expanding universe led to the question how the universe looked like at its very beginning. In 1929, Edwin Hubble correlated the distance of different galaxies with their velocity (Hubble, 1929) and found a linear relation between both quantities[1]. Hubble extended his study and published the results two years later, confirming his previous findings (Hubble and Humason, 1931), and was even able to quantify it in a formula:

$$v = H_0 D\,. \tag{1.1}$$

This linear relation between velocity v and distance D relative to us on Earth is known as *Hubble's law*. The constant H_0 was found to have a value of $560\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ by Hubble back then. This marks the beginning of observational cosmology. The scaling parameter is also known as the *Hubble constant* H_0 and is still an important parameter in modern cosmology. Summarizing, Friedmann's, Lemaître's and Hubble's findings suggest that the universe is not a static system at all.
In 1933, the first observations of the Coma Cluster by Fritz Zwicky exhibited dynamics that could not be explained by gravitational interaction of the luminous matter. The velocity of the galaxies are too high to keep the cluster together. He concluded that for a stable cluster more than just the luminous mass must be present. His hypothesis was that an unknown form of matter which does not emit any light – thus *dark matter* – accounts for the missing mass. Later, this form of matter turned out to be a key ingredient for the structure formation on large scales in our Universe. In the 1960s, detailed studies on rotation curves of galaxies by Vera Rubin (Rubin and Ford, 1970; Rubin, Ford and Thonnard, 1978) also suggested a lack of mass that could only be explained by an unknown form of matter.
In the meantime in 1946, Ralph Alpher who was a PhD student with George Gamow published the Alpher–Bethe–Gamow paper (or $\alpha\beta\gamma$ paper) (Alpher, Bethe and Gamow, 1948). It provided a calculation from theoretical considerations for the relative abundances of hydrogen and helium after the earliest time the Universe started expanding from the primordial singularity, which is also known as the *Big Bang*[2].
In the same year, Ralph Alpher and Robert Herman published the first estimate for the temperature of the *cosmic microwave background* (Alpher and Herman, 1948). This background radiation consists of remaining photons from a much hotter and denser state of the early universe and they found their energy to be related to a temperature of $T_{\mathrm{CMB}} \approx 5\,\mathrm{K}$. In the following years, Alpher and Gamow did some reestimates. But after this, this concept was almost forgotten for some time – until 1965. Penzias and Wilson (Penzias and Wilson, 1965) discovered by accident the CMB as noise while working with a huge antenna. Dicke (with Peebles, Dicke et al., 1965) interpreted this noise as the cosmic microwave background radiation that had been predicted by Gamow and stated that the CMB spectrum had to be a blackbody spectrum. It was then measured to have a temperature of $T_{\mathrm{CMB}} \approx 3\,\mathrm{K}$.
In the 1970s, the first redshift surveys[3] started with the CfA survey (1977-1982) that was succeeded by

[1] Velocity measurements were already done in 1912 by Vesto Slipher who discovered that some galaxies have redshifted spectra due to their movement away from our Galaxy. Note that he was not aware of the fact that those objects were galaxies and, consequently, his conclusions were different.
[2] Gamow decided to put the name of Hans Bethe to the list of authors although Bethe did not contribute to it at all just to make it sound like the first three letters of the greek alphabet.
[3] The concept of redshift is explained in detail in section 1.2.

many others, e.g. the Sloan Digital Sky Survey (SDSS, since 2000) and Galaxy And Mass Assembly survey (GAMA, since 2008). With these redshift surveys, galaxies at different epochs of the universe could be studied.

Around 1985, numerical simulations of N-body systems became a useful and prominent tool for the research of structure formation and the formation of galaxies[4]. From these simulations it was found by Julio Navarro, Carlos Frenk, and Simon White that collapsed structures of the previously mentioned dark matter (hereafter *dark matter haloes*) seem to have a universal density profile (Navarro, Frenk and White, 1996) that follows the relation

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_S}\left(1 + \frac{r}{R_S}\right)^2} \ . \tag{1.2}$$

This is also known as the famous NFW–profile. In general, computational power increased and numerical simulations became efficient with time allowing further levels of complexity. Starting with dark matter simulations only, baryons[5] were added over the years, incorporating the process of star formation in simulations. This provided a new tool to the community to study the field of galaxy formation and evolution. One highlight of this was the Millenium simulation by Volker Springel (Springel, White et al., 2005).
The improvement of observational tools in synergy with space exploration brought observations to a whole new level by the launch of the Hubble Space Telescope in 1990. Further CMB research was done by the data from the Cosmic Background Explorer Satellite (COBE) from 1989 to 1993. COBE was followed by the Wilkinson Microwave Anisotropy Probe (WMAP, operating from 2001 to 2010) and PLANCK (2009-2013) which both better resolved CMB data as they were equipped with better state-of-the-art technology.
One fundamental finding was made by Madau, Pozzetti and Dickinson (1998). It shows the evolution of star formation in a cosmological volume with redshift. The result was astonishing: The cosmic star formation rate has a peak around a time of 3.3 Gyr after the Big Bang (redshift of $z = 2$, see equation (1.12)). Almost at the same time, the model of an expanding universe was further confirmed by observations of Supernovae Ia performed by Riess, Filippenko et al. (1998) and Perlmutter et al. (1999).
Today, radio telescopes are probing the galaxies at redshift ranges $z > 3$ at (sub-)mm wavelength to detect ionized carbon and carbon monoxide which are indicators for cold gas that is believed to be the fuel for star formation. Impressive examples for this are ALMA (Atacama Large (sub-)Millimeter Array), VLA (Very Large Array), Plateau de Bure Interferometer and its upgraded version NOEMA (Northern Extended Millimeter Array).
The detection of gravitational waves in 2015 (Abbott et al., 2016) by the LIGO collaboration marks the beginning of the new field of *gravitational wave astronomy*. In the future, it is expected to provide further information about the history of the universe.
All those theories and observations flow into the *standard model of cosmology* which is also known as the ΛCDM model. We will discuss its mathematical modelling in the following chapter in more detail.
In the near future, the EUCLID satellite and the James Webb Space Telescope (JWST) will be launched. The first will gather information for the research on dark matter and dark energy, the second will perform infrared observations to better understand the galaxies of the early universe.

[4] We will discuss the evolution of numerical simulations in more detail in section 1.10.1.
[5] The term *baryons* means in the context of cosmology all matter that is not dark matter.

1.2 Mathematical formulation of cosmology

A fundamental ingredient to modern cosmology is the *cosmological principle*. It consists of the following two assumptions:

- Homogeneity: the universe is on large scales (ranges of several hundreds of Mpc) homogeneous. The meaning behind that is that on average on several hundreds of Mpc, the average matter content is roughly constant.
- Isotropy: the universe is isotropic. No matter where an observer is, the universe looks the same in every direction.

Just having basic concepts is not sufficient and we need a mathematical formulation of an expanding universe. General relativity provides here the necessary tools. The starting point in General relativity (GR hereafter) is always the concept of distances between two points. This is defined by the line element $\mathrm{d}s^2$,

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu \,. \tag{1.3}$$

$g_{\mu\nu}$ represents the metric tensor and has to be chosen properly as it changes with the problem one wants to consider. For the description of the universe as a whole, the *Friedman–Lemaître–Robertson–Walker* (short: FRLW) metric is used. It takes the constraints coming from the cosmological principle into account and the idea of a universe that expands with time. The metric (tensor) reads

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right] . \tag{1.4}$$

k is the curvature parameter and has a large influence on the geometry of the universe. $k = -1$ corresponds to the case of a *positively curved*, $k = 0$ to a *flat*, and $k = 1$ to a *negatively curved* universe. This parameter determines the future of the universe: the positively curved and flat universe will expand forever, the negatively curved one will recollapse.

The whole dynamics of the universe is captured by the time-dependent parameter $a(t)$ which is also known as the *expansion factor*. The Einstein field equation can be used to describe the dynamics of the universe and its content. It reads

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi \mathrm{G} T_{\mu\nu} \,. \tag{1.5}$$

The Einstein tensor $G_{\mu\nu}$ is defined as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \,, \tag{1.6}$$

where $R_{\mu\nu}$ represents the Ricci tensor and $R = R^\mu_\mu$ is the Ricci scalar. G is the gravitational constant[6] and assuming a perfect fluid the energy momentum tensor $T_{\mu\nu}$ reads[7]

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & & & \\ 0 & & g_{ij}p & \\ 0 & & & \end{pmatrix} , \tag{1.7}$$

[6] $G = 6.67408 \times 10^{-11}\ \mathrm{m^3\,kg^{-1}\,s^{-2}}$

[7] The indices i and j run from 1 to 3 and contain the spatial components of the involved tensors.

with p being the pressure and ρ the density of the matter content in the universe.

Turning back to equation (1.5), we see that it contains an important piece of information: it relates the matter content of the universe with curved space time. This revolutionized the picture of gravity from the Newtonian era. Instead of the attraction of two bodies, it is the curved spacetime that pulls a mass in a certain direction as it is supposed to move along the geodesics.
Λ represents the cosmological constant which was introduced by Einstein to make the universe static. His ansatz behind that was that a matter dominated universe should collapse under its own gravity and needs an additional component that counteracts this effect. With Hubble's discovery of the expansion of the universe, the idea of the cosmological constant in the context was rejected. Close to the end of the 20th century however, the interest in it increased again due to the discovery of the accelerated expansion of the universe. Today's concept of Λ is different as it is associated with an energy density.
Using the metric tensor, all components of $G_{\mu\nu}$ can be calculated. Specifically, one obtains for the $\mu\nu = 00$ and the $\mu\nu = ij$ components the following equations that are also known as the Friedman equations[8]:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) \tag{1.8}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{1.9}$$

The dot denotes the time derivative. It is often useful to introduce the *Hubble parameter*

$$H(t) = \frac{\dot{a}(t)}{a(t)} \,. \tag{1.10}$$

We already came across H_0 in equation (1.1) and indeed, there is a connection to $H(t)$ as $H_0 = H(t = 0)$. Hence, H_0 represents the present day value of $H(t)$ which is roughly $70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [9].
It is useful to introduce the *density parameter*

$$\Omega = \frac{8\pi G}{3H(t)^2}\rho = \frac{\rho}{\rho_{\mathrm{crit}}} \,, \tag{1.11}$$

with the *critical density* ρ_{crit}. There are different components in the universe, radiation (photons and relativistic particles), dark energy and matter (baryonic and dark matter). To each of them, density parameters are assigned, namely Ω_Λ for the vacuum energy, Ω_r for radiation, and Ω_m for matter. Due to the cosmic expansion, objects move away from the observer. As a consequence, the electromagnetic radiation that is emitted is redshifted. Hence, it is convenient to introduce the *redshift* which combines the information about the age and the distance at the same time:

$$z = \frac{a(t_0)}{a(t_1)} - 1 = \frac{\lambda_{\mathrm{obs}} - \lambda_{\mathrm{rest}}}{\lambda_{\mathrm{rest}}} \,, \tag{1.12}$$

where λ_{obs} is the observed wavelength and λ_{rest} is the wavelength of the photon in the restframe of the observed object. $a(t_1)$ and $a(t_0)$ are the expansion factors of the time when a photon has been emitted and observed, respectively. In summary, the *standard model of cosmology model* or Λ*CDM model* is constrained by six parameters as shown in Tab. 1.1. Additionally to the density parameters and the Hubble constant, we will explain briefly in the following the remaining ones. σ_8 is the normalization of the fluctuations of the power spectrum. The power spectrum is a statistical quantity to characterize the

[8] The Friedman equations can also be determined by a Newtonian approach but we focus on the GR way.

	Ω_m	Ω_b	σ_8	n_s	H_0 [km s^{-1} Mpc^{-1}]	τ	$\Omega_\Lambda = 1 - \Omega_m$
Planck 2013	0.308	0.0481	0.862	0.9608	67.8	0.066	0.692
Planck 2018	0.315	0.0498	0.811	0.965	67.4	0.054	0.685

Table 1.1: Cosmological parameters constrained from the cosmic microwave background from Planck Collaboration (2014) and Planck Collaboration (2018).

clustering of the matter. Initially, for the calculation of the fluctuations, a length scale to average over has to be chosen. To normalize it to unity, $8\,h^{-1}$ Mpc were chosen but as observations got better, the value of σ_8 decreased, since the scale for the average remained unchanged.

τ is the total optical depth for electron scattering at the epoch of reionization, the time when the first galaxies started to ionize the hydrogen that has recombined before. Another cosmological quantity is n_s. At early times, the power spectrum can be approximated by a power law and the exponent of this primordial power spectrum is n_s.
The use of all these parameters will not reproduce our universe, but a universe that has the same statistically properties as ours.

1.3 A *very* brief history of the universe

We will highlight the most important epochs in the history of the universe at specific redshifts.

- Before $z \approx 1100$: Due to the continuous temperature decrease, several high energy processes between fundamental particles (hadrons, gluons, leptons) stall step by step. Lastly, in the process of *nucleosynthesis*, protons, neutrons along with atomic nuclei of helium and small amounts of lithium form.

- $z \approx 1100$ or 380000 years after the Big Bang: The *cosmic microwave background* (CMB) is created by the recombination of electrons and protons to hydrogen. Its temperature fluctuations contain the oldest clustering information about the young universe. The corresponding density perturbations grow during the evolution of the universe. Photons decouple from matter and start to propagate freely. The universe is said to become *transparent* as photons are now allowed to travel without any interaction. That decoupling left an imprint on the clustering signal, the so called *baryon acoustic oscillations*.

- $z \approx 1100 \sim 20$: This epoch is also known as the *Dark ages*. The universe has already cooled down. Photons from the CMB and the 21 cm hyperfine structure-line photons of hydrogen are propagating through the universe.

- From $z \sim 20$: The first stars and galaxies start to form (see section 1.4 for details).

- $z \approx 12 - 8$: As a consequence of galaxy formation, the radiation emitted from galaxies starts ionizing the neutral hydrogen, also known as the *epoch of reionization*.

- $z = 1 \sim 2$: The cosmic star formation rate, after continuously increasing, reaches its maximum. From this time on, fewer stars are formed.

- $z = 0$: today

1.4 Galaxy formation and evolution

The study of galaxy formation and evolution is a very challenging one as we are unable to do experiments on our own. All our knowledge is based on the detailed observation of galaxies at different epochs and their properties. At the beginning of the 20th century, telescopes became powerful enough to identify nearby galaxies. Hubble started to put them in a scheme based on their shape which is still known as the *Hubble tuning fork* or *Hubble sequence* (Hubble, 1927). It was believed that this diagram shows also an evolutionary trend, meaning that a galaxy evolves from left to right, starting with an elliptical shape, towards a spiral pattern (see Fig. 1.1). Today, the classification of *early-type* and *late-type* galaxies is still used, although it is widely accepted that elliptical galaxies represent the old galaxy population.

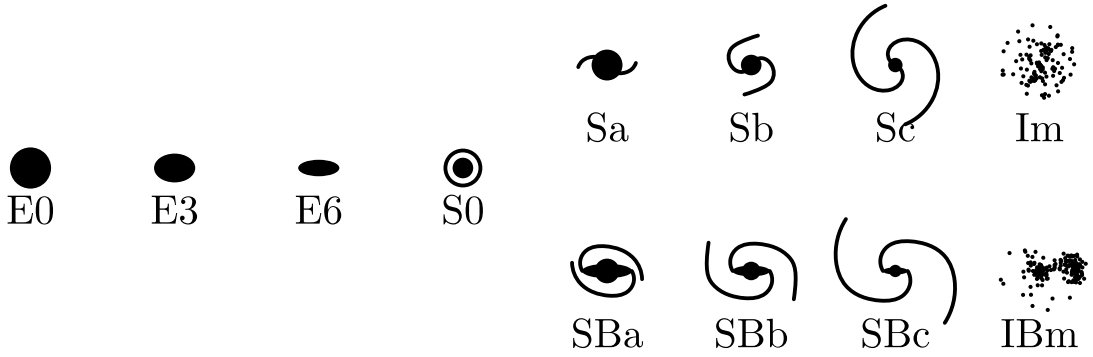

Figure 1.1: The Hubble tuning fork (from Abraham, 1998). The ratio of the major and minor half axis (ellipticity) of the *elliptical galaxies* (E0 to E6) increases from left to right. These galaxies have a rather featureless and uniform light profile. *Lenticular galaxies* (S0) have a spherical central region (*bulge*) which is surrounded by a disk-like structure. The galaxies in the branches of the tuning fork possess spiral arms. Going from left to right, the arms get more pronounced. The difference between the upper and lower branch is the existence of a bar in the central region. Thus, the galaxy types are called *spiral galaxy* (Sa -Sc) and *barred spiral galaxies* (Sba -Sbc).

Rees and J. P. Ostriker (1977) made a more analytical approach and studied the gas dynamics in detail, such as gravity and cooling that are involved for a giant molecular cloud to collapse to a galaxy-sized object. However, these considerations still miss one of the key ingredients to structure and galaxy formation, namely *dark matter*. Dark matter is a hypothetical form of matter that does not emit any electromagnetic radiation. It interacts gravitationally only, providing the necessary potential for the gas to concentrate. The haloes formed by the dark matter are then the sites in which the galaxies reside.
The formation of a halo that provides a gravitational potential strong enough to form stars later on marks the beginning in the life of a galaxy. At this point, the physics of gas becomes relevant. As the gas is accreted by the halo, it is accelerated in this process and thus, gains energy. When the gas arrives in the center of the halo, it is shock-heated: the gravitational energy is converted into internal energy. From both theory and observations, it is clear that hot gas can not form stars; it needs to cool down from temperatures of $\approx 10^4 - 10^7$ K to tens of K. This is done mostly via radiative processes like bremsstrahlung or recombination. The competition of different timescales like for cooling and collapse become relevant in the scenario of hierarchical clustering which has been discussed by White and Rees (1978). Later in the 1980s, e.g. Blumenthal et al. (1984) focused even more on the role of *Cold Dark Matter* in the process of structure formation. *Cold* refers to its speed meaning that it is moving slow compared to the speed of light. This allows structures to form hierarchically in a bottom-up scenario.

Despite the fact that at the early stages of galaxy evolution, smaller companions are absorbed and a galaxy can grow further in stellar and gaseous mass, many galaxies have satellite galaxies moving on stable orbits. Most of them are low-mass objects and thus called *dwarf galaxies*. These galaxies have

masses of several orders less than the main galaxy.
However, the collision of two larger galaxies is not an unlikely scenario. These *mergers* highly influence the fate of a galaxy. The collision of two spiral-shaped galaxies destroys their pattern and the result is an elliptical galaxy. The stars therein move on more random trajectories. Further, star formation can be enhanced by the merging as it leads to gas compression. In turn, stellar feedback increases influencing future generations of stars in their formation process because stars convert hydrogen to heavier elements. At the end of the life of a star, these elements are returned to the interstellar medium. This enhances the radiative cooling of the surrounding gas and consequently supports star formation.
Both the birth of a galaxy and its end of its star forming epoch are topics of current research. After a phase of active star formation, the conditions in the galactic environment or internal processes lead to a decrease of star formation. The color of galaxies changes from blue to red, since their stellar population consists of older stars. Most of the observed galaxies belong to one of these two groups, a histogram of red and blue galaxies exhibits a bimodal distribution (see e.g. Strateva et al., 2001). Thus, galaxies in this transition phase are said to be in the *green valley*.

When a galaxy enters a phase of low star formation rate, most of its evolution is completed. This process is also called *galaxy quenching*. The details of this process are still highly debated. Candidates for possible mechanisms are feedback from active galactic nuclei (AGN-feedback) and missing gas accretion.
The *main sequence of galaxies*, which is a tight correlation between the SFR of a galaxy and its stellar mass, is a tool to classify galaxies. For low-mass galaxies, it exhibits a linear form until it flattens for high-mass galaxies. Outliers that have an unusually high star formation for their stellar mass are called *starburst galaxies* whereas those with a relatively low SFR are called *quenched* or *quiescent* galaxies. The main sequence is also redshift dependent since it changes its slope with redshift (see Förster Schreiber and Wuyts, 2020, for a more detailed overview).
This is an individual view of galaxy evolution and can be summed up in a more cosmological picture. Madau, Pozzetti and Dickinson (1998) discuss the star formation rate history within a cosmic comoving volume. The quintessence was the following: from very high redshift towards $z \sim 2$, the cosmic star formation rate density exhibits an increase. Around $z \sim 1-2$, it reaches its peak and reduces then to present day. Yet it is still unknown why the cosmic star formation rate density shows this behavior. The abundance of molecular gas is considered to be a possible explanation for this shape. The observations of star formation on smaller scales in galaxies indicate that the presence of molecular hydrogen is more significant for star formation than just the gas abundance (see section 1.9 for a more elaborate discussion). Over the last years, the observational data has been updated by dust corrections and modifications on the star formation rate calculations at low redshift (Madau and Dickinson, 2014), but the overall trend of the cosmic star formation rate over redshift stayed the same (see Fig. 1.2).

The possible connection between molecular gas and star formation led to surveys to investigate this issue further. The right panel of Fig. 1.2 shows a summary of the results from observations taken from Lenkić et al. (2020): with decreasing redshift, the cosmic H_2 content increases, reaches its peak at $z \sim 2$ and decreases again towards $z \sim 0$.[10]
In the following, we will discuss the details of star formation. Since this depends on the interplay of many different processes, we will discuss these first and their importance for star formation.

[10] The star formation rate density can be parametrized as $\Psi(z) = 0.015 \frac{(1+z^{2.7})}{1+[(1+z)/2.9]^{5.6}} \mathrm{M}_\odot\ \mathrm{yr}^{-1}\ \mathrm{Mpc}^{-3}$.

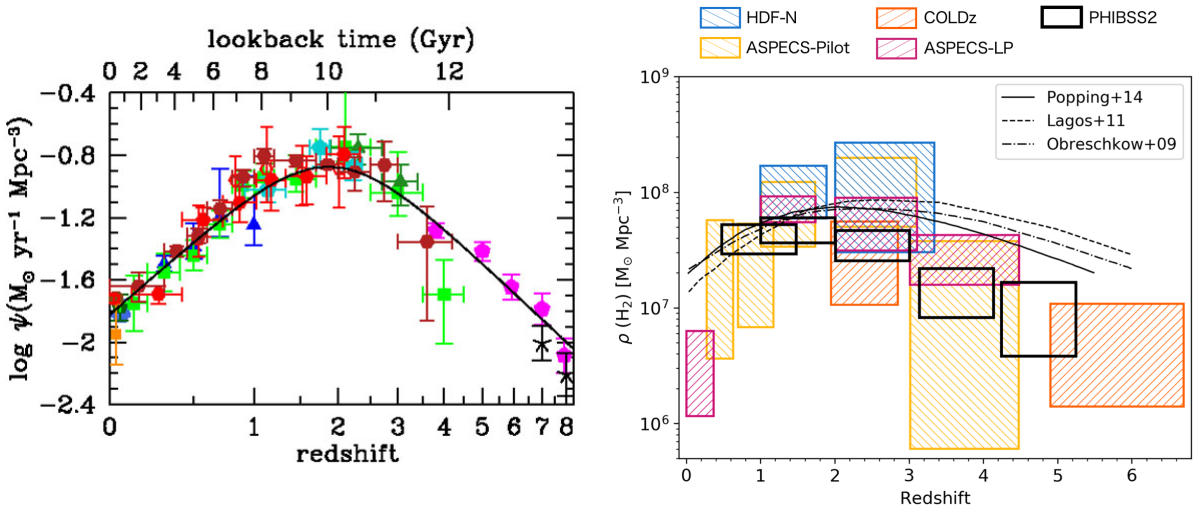

Figure 1.2: Left panel: cosmic star formation history presented in Madau and Dickinson (2014). It exhibits a steady increase from high redshift on, reaching its peak at around $z \approx 2$, and decreases until today. Right panel: the observed molecular gas content in a cosmological volume is shown (from Lenkić et al., 2020). The boxes indicate the constraints from observations whereas the lines show predictions from theoretical models. Both theory and observations show the same trend as noticed for the cosmic star formation rate.

1.5 ISM physics

One key ingredient to understand the process of galaxy formation is understanding the evolution, properties and dynamics of the gas in the so called *interstellar medium (ISM)*. The ISM hosts a lot of different processes on different scales.

The macroscopic process of star formation highly depends on a variety of microscopic events and vice versa – but whether they can take place is influenced by the temperature T and the gas density n_{H}. High gas densities and low temperatures are required to initiate the nuclear fusion in a cloud. Thus, it is not surprising that these quantities are used to subdivide the different phases and components of the ISM (Ferrière, 2001):

- Hot ionized medium (HIM): $T \geq 10^6$ K, $n_{\mathrm{H}} \sim 10^{-3}$ cm^{-3}
- Warm ionized medium (WIM): $T \geq 8000$ K, $n_{\mathrm{H}} \sim 0.2 - 0.5$ cm^{-3}
- Warm neutral medium (WNM): neutral hydrogen of temperatures of several thousand kelvin ($T \geq 6000 - 10000$ K) and $n_{\mathrm{H}} \sim 0.2 - 0.5$ cm^{-3})
- Cold neutral medium (CNM): neutral hydrogen of temperatures $50 - 100$ K, $n_{\mathrm{H}} \sim 20 - 50$ cm^{-3}
- molecular phase: molecular hydrogen of temperatures $T = 10 - 20$ K, $n_{\mathrm{H}} \geq 10^2$ cm^{-3}

Gas does not have to remain in one phase – it can increase or decrease its density and temperature which we will discuss in more detail in section 1.5.1. In fact, the baryonic matter undergoes a continuous cycle: gas that cools undergoes a gravitational collapse and forms stars. At the end of the lifetime of these stars, energy and metals (elements heavier than helium), which are the ingredients for dust, are returned to the

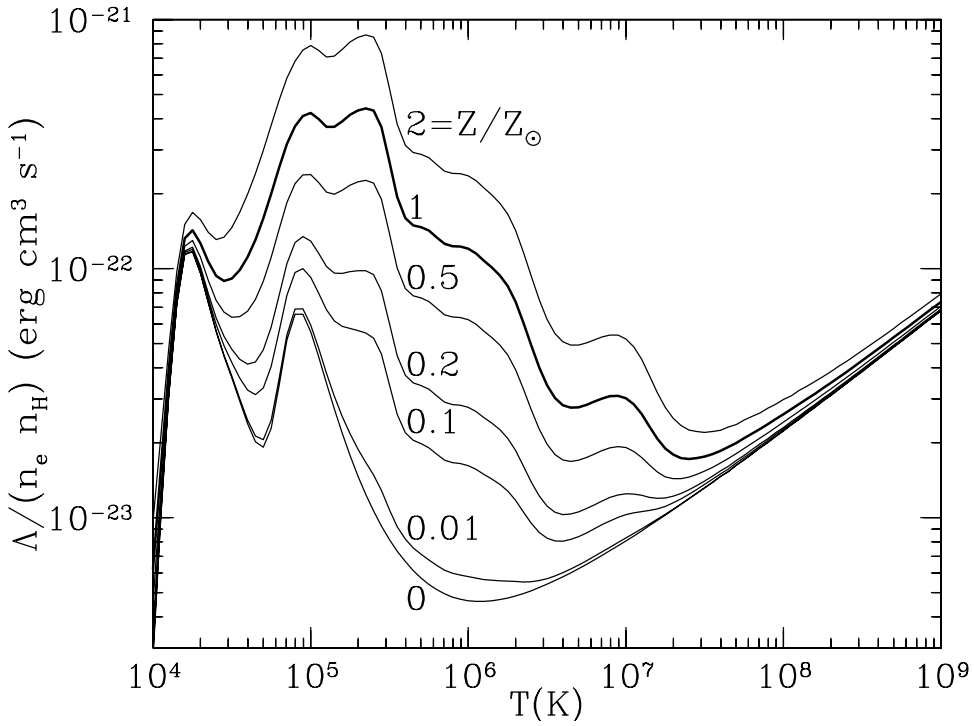

Figure 1.3: Radiative cooling function for different metallicities. The more heavier elements are abundant the higher is the cooling rate. The peak around $T = 2 \times 10^4$ K is the contribution coming from H, the height of the following peaks are influenced mostly by the abundances of He, C and O, but also from Si, N and S. The range of $10^5.8 < T < 10^7.2$ K is shaped by Fe, Ne, Mg and Si. High temperatures are dominated by bremsstrahlung causing the $\Lambda \propto T^{1/2}$ dependency (from Draine, 2011).

diffuse gas in the environment of the star. The dust and metals help to cool the gas down, which enables star formation again. The details of these processes are the subject of this section.

1.5.1 Cooling and heating

To enable gravitational collapse, a gas cloud has to cool down. In the ISM, there are several processes that simply cause the loss of thermal energy of a cloud on the one hand. On the other hand, the interplay of stars and dust with the gas can also cause a temperature increase. We will give a brief overview of the different mechanisms.

Since hydrogen and helium are the most abundant elements in the ISM, their chemistry plays a significant role for the temperature of the gas. The interplay of HI, HII, HeI, HeII, HeIII and the electrons can influence the gas temperature in many different ways. One way that the gas can lose energy is *collisional excitation* where a free electron collides with an atom and excites the bound electron to an excited state. The return to the ground state leads to photon emission and hence, the dissipation of the kinetic energy of the electron. The term *collisional ionization* describes the processes in which the encounter between an electron with an (ionized) atom ionizes the atom (further). The electron that has been removed is carrying some of the atom's kinetic energy away. The *recombination* of ions and free electrons induces that the kinetic energy of the electron is radiated away by the atom. For temperatures of $\gtrsim 10^6$ K, the driving cooling process is *bremsstrahlung*. As the gas is mostly ionized, the electrons are slowed down by passing by an atomic nucleus and lose their kinetic energy by radiating a photon off.

In every phase of the ISM, the transitions of elements (and their ions) and molecules allow the gas to radiate some of the thermal energy away. In the diffuse ionized medium, major coolants are Hα and the fine-structure lines of [OIII] 88 μm, [NIII] 57 μm and of [NII] at wavelengths of 122 μm and 205 μm.

For the neutral gas, cooling happens via the 21 cm hyperfine-structure line of neutral hydrogen, and neutral oxygen [OI] 63 μm and 145 μm. A very important coolant is the fine-structure transition of singly ionized carbon [CII] 157 μm. This line is a key to observations, especially at high redshift to trace star formation as it is a major coolant in the ISM. Especially for primordial gas, the cooling of H_2 is considered to play a key role cool between a range of $100 < T < 1\,000$ K as this is necessary to form the first stars without any metals in the very early universe.

In the components of the dense medium, the *dust continuum* plays an important role in contributing to the cooling. Rotational states of carbon monoxide are excited by collisions and the gained energy is removed from the molecule by photon emission. Also, the forbidden lines of neutral carbon contribute to the cooling mechanism. This highlights the importance of carbon species again: many of them contribute to the cooling and thus, they are useful tracers for star formation.

But there is also a variety of heating processes. In dense clouds, high energy particles called *cosmic rays* pass through the gas. Their kinetic energy is transferred to the gas components. The amount of energy is not only used to ionize the gas particles; it is large enough that the electron from this ionization process has a significant kinetic energy. This energy is then redistributed among the other gas particles which heats up the gas.

Additionally, *photoionization* is caused by the proximity of the gas to the UV radiation from young stars. The gas components can be ionized and the excess energy of the electrons is transferred as thermal energy of the gas.

The interaction of gas components with dust can also heat up the gas as they are coupled to each other. For instance, UV radiation can affect the dust in the ISM in form of ionization or by simply heating it up. Grain assisted chemical reactions can release energy to the gas as well.

Further, interactions between the gas and dust allows an energy transfer. Also mechanical effects are able to increase the temperature of the gas, e.g. *compression* and *shocks*. These processes happen on short time scales which then do not allow to dissipate the energy of the gas.

1.5.2 Feedback mechanisms

Feedback summarizes all effects in which stars return energy or material in any form to the ISM. Relevant for these processes are mainly OB stars. We will briefly discuss the most important mechanisms.

- photoionization: the equilibrium of ionization and recombination of the gas around a star is limited to a spherical region, also known as *Strömgren sphere* (Strömgren, 1939). The size of this region depends on the density of the surrounding gas and the luminosity of the star.

- stellar winds: during its lifetime, a star constantly ejects material from its outer shells. The ejected material contributes to the energy, mass, and metallicity budget. Mass losses are of the order of $10^{-5}\,M_\odot\,yr^{-1}$.

- radiation pressure: the photons emitted from a star transfer their momentum to the gas nearby. This energy is radiated away and is either absorbed by dust or by the gas.

- accretion heating: a protostar accretes the gas it has been born from and the gravitational energy is converted to heat.

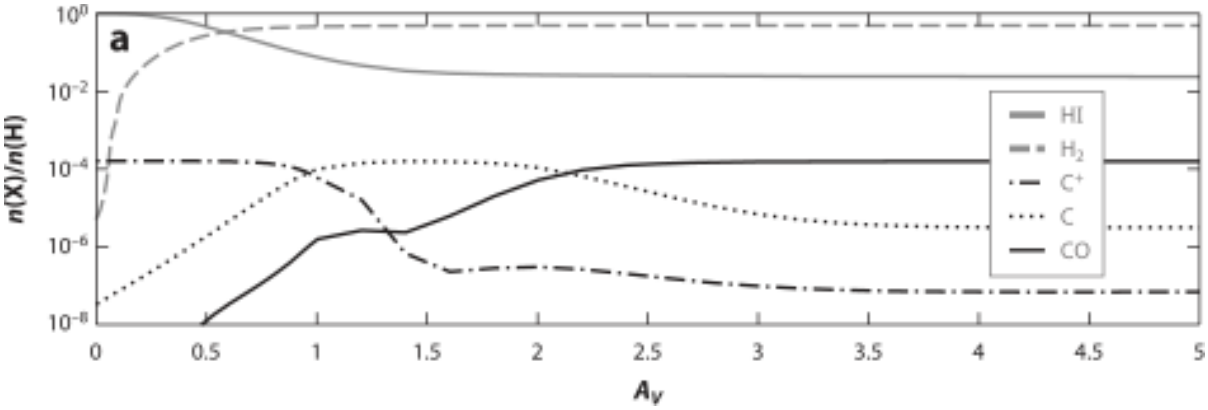

Figure 1.4: Distribution of carbon-based tracers (C^+, C, and CO) and hydrogen species (H and H_2) within a molecular cloud from Bolatto, Wolfire and Leroy (2013). The outer parts of the cloud consist of atomic hydrogen and ionized carbon. As the column density/optical depth increases, the outer layers of the cloud provide enough (dust) shielding against the radiation coming from the environment. This allows hydrogen to exist in molecular form and carbon in atomic form. As CO is very prone to get photodissociated by UV radiation, the very dense part of the cloud provides sufficient shielding.

- jets: stellar jets are bidirectional outflows that form by the accretion of nearby gas. During this process, the matter gains energy. Typical mass outflows are $\sim 10^{-8}$ $M_\odot$ yr^{-1}.
- supernovae: a Type II Supernova represents the end of the life of a massive star. The energy that is injected into the ISM leads to a shock front that propagates through the ISM. Along with that, material locked away in the star previously is released in that explosion. Supernovae Ia occur in binary systems where a white dwarf is accreting mass from its companion until it reaches a mass limit. If it is exceeded, the star explodes, releasing energy and material as well.

1.6 Molecular gas in GMCs and how to trace it

There has been growing evidence that the fuel of star formation is the molecular gas, specifically the molecular form of hydrogen, H_2, as it is the most abundant molecule. However, there is a subtlety to it: its lack of a permanent dipole moment makes it impossible to observe in the cold phase of the interstellar medium with state-of-the-art observations – especially at high redshift. Thus, observers are reliant on so called *tracers*, that coexist with molecular hydrogen but are easier to detect. One of these tracers is carbon monoxide. Since the detection of its rotational transition lines in high redshift galaxies, it has become one of the fundamental tools in the last 20 years for the study of cold gas (Brown and Vanden Bout, 1991), as it gives the ability to infer the molecular gas content from its luminosity. Over the last years, observations made also use of ionized and neutral carbon as molecular gas tracers. In the next sections, we will illustrate how carbon monoxide (including ionized and neutral carbon) helps to detect molecular hydrogen by showing where it is abundant in the ISM (see section 1.6.1) and its spectroscopic properties (see section 1.7).

1.6.1 Carbon-based tracers for H_2

We show in Fig. 1.4 the chemical decomposition of an idealized giant molecular cloud as presented in Bolatto, Wolfire and Leroy (2013) focusing on the abundances of ionized carbon (C^+), neutral carbon

(C), carbon monoxide (CO) and atomic and molecular hydrogen (H and H_2). Carbon at the outer layers of the cloud is ionized by UV radiation that is assumed to surround the GMC. As one moves towards the center of the cloud, the shielding becomes more efficient and blocks out the radiation and allows even the carbon to stay atomic and prevents the photodissociation of molecular hydrogen. Even deeper, CO can form as this is very sensitive to radiation and is not destroyed anymore. Due to its internal structure, CO is an ideal tracer for the very dense and cold regions of the ISM where molecular gas is situated. However, we point out that this is an idealized situation. Observations at high redshift do not resolve the internal structure of the GMCs and one has to consider the processes that take place on a galactic scale (e.g. feedback, turbulence, shocks). The observations of these tracers capture more the abundances on a galactic scale and give information about the distribution on the scale of kpc. Later in section 1.7, we give a more profound insight what makes CO so useful for tracing H_2. Combining the fact that H_2 is the fuel for star formation and the CO luminosity the way to detect molecular hydrogen, the following relation is a crucial tool for state-of-the-art observations:

$$M_{\mathrm{H_2}} = \alpha_{\mathrm{CO}} L'_{\mathrm{CO}}\,, \tag{1.13}$$

where L'_{CO} represents the luminosity of the CO($J = 1 \rightarrow 0$) for its $J = 1 \rightarrow 0$ rotational transition[11]. The luminosity can be further calculated from the flux (Solomon, Downes and Radford, 1992):

$$L'_{\mathrm{CO}} = 3.25 \times 10^7 \times S_{CO}\,\Delta v \frac{D_L^2}{(1+z)^3 \nu_{\mathrm{obs}}^2}\ \mathrm{K\,km\,s^{-1}\,pc^2}\,. \tag{1.14}$$

Here, D_L is the luminosity distance, z the redshift of the observed object and ν_{obs} the observed frequency. $S_{\mathrm{CO}}\,\Delta v$ is the integrated line flux density. The luminosity of an object in the CO transition is measured and then converted via α_{CO} into a mass. α_{CO} is also known as the CO–to–H_2 conversion factor[12] and is of order unity. Depending on the system that is considered, the choice for α_{CO} can still be between 0.8 $\mathrm{M_\odot\,K^{-1}\,s\,pc^{-2}}$ for systems such as for mergers, ULIRGs, starburst systems (Solomon, Downes, Radford and Barrett, 1997; Downes and Solomon, 1998; Tacconi et al., 2008) and the Galactic value of 4.3 $\mathrm{M_\odot\,K^{-1}\,s\,pc^{-2}}$ (Strong and Mattox, 1996; Dame, Hartmann and Thaddeus, 2001; Abdo et al., 2010).

At the current state, there are several uncertainties linked with the conversion factor. One major issue is that its dependency on metallicity is yet unknown and different forms have been proposed. Many of them suggest a power law of the form $\alpha_{\mathrm{CO}} \propto Z^{\beta}$. But there is also the idea to consider two different slopes or even an exponential form. There is also the question of its redshift dependence. Detailed observational studies regarding this are still missing (see Bolatto, Wolfire and Leroy, 2013, for an overview). The high sensitivity of carbon monoxide to UV radiation is also problematic as in some regions of the ISM, hydrogen can be already molecular but CO is photodissociated. Thus, by using L'_{CO}, observations can miss a significant amount of molecular gas which also known in the literature as *CO dark molecular gas*.

Observing the rotational transitions[13] of CO is not easy at all. First of all, the CO transitions are

[11] Note the difference between L'_{CO} and L_{CO}.The first one is the luminosity expressed in units of the areal integrated source brightness temperature, the second one is normalized to the source luminosity in units of $L_\odot$ (see Carilli and Walter, 2013, for details). The units of α_{CO} have to change accordingly.

[12] Alternatively, the H_2 abundance can also be obtained from its column density $N(\mathrm{H_2})$ and the integrated line intensity $W(^{12}\mathrm{C}^{16}\mathrm{O}\ J = 1 \rightarrow 0)$ by using the relation $N(\mathrm{H_2}) = X_{\mathrm{CO}} W(^{12}\mathrm{C}^{16}\mathrm{O}\ J = 1 \rightarrow 0)$. X_{CO} is also known under the designation *conversion factor* and has a value of $X_{\mathrm{CO}} \approx 2 \times 10^{20}\mathrm{cm^{-2}}$.

[13] In the case of ALMA, different rotational transitions are observed depending on the redshift of the observed system. The bands of ALMA cover frequency ranges and the emission frequency of the different rotational transitions of CO are redshifted falling in different frequency ranges. To be able to use equation (1.13), the luminosity for higher transitions is always converted via the line ratio back to the luminosity associated with $J = 1 \rightarrow 0$.

optically thick, i.e. the photons emitted from it are absorbed and scattered easily from the matter between the source and the observer. The luminosity in CO is very sensitive to the conditions along the line of sight. Thus, a very detailed knowledge about the ISM is needed to model all influences accurately.

Knowing the structure of a GMC, one can make use of the other constituents like C^+ and C to trace molecular gas since they both coexist with H_2 in the cloud. Since observations found a correlation between the [CII] and IR luminosities, ionized carbon is also linked to the star formation in galaxies. Additionally, the [CII][14] fine-structure line is an important cooling line and is one of the brightest lines at high redshift. The low ionization energy of 11.3 eV makes carbon a tracer for the cold ISM (hydrogen is ionized at 13.6 eV). Neutral carbon can also be used as an effective tracer for molecular gas (Papadopoulos and Greve, 2004) although it is not as bright as the ionized one. The optically-thin treatment of the [CI] fine-structure transitions makes the analysis of the observations rather simple. The excitation temperatures for the two lowest energy levels are at T =23.6 K and T =62.5 K which makes atomic carbon an ideal tracer for the cold component of the ISM. However, it can be marginally optically-thick under certain conditions in the ISM (e.g. very high densities) (see Krips et al., 2016; Salak et al., 2019, for a detailed discussion) and requires a special treatment. In the case of C and C^+, their luminosities can be converted to a molecular gas mass as well, similar to equation (1.13) but with a different conversion factor (e.g Zanella et al., 2018; Saito et al., 2020).

1.6.2 Other tracers

Further prominent tracers are isotopologues such as ^{13}CO and C^xO (e.g. $C^{18}O$).[15] These isotopologues are used to estimate the molecular gas content, obtain the emission line and isotopologue ratios and check the validity of other tracers (Nishimura et al., 2015; Israel, 2020; Roueff et al., 2021) since their lines are optically thin and their obeserved line properties are not as influenced by the matter content along the line of sight. The disadvantage of this method is that these emission lines of isotopologues are weaker. Additionally, isotope ratios need to be known which might vary (e.g. Q.-H. Tan et al., 2011).

There is also a growing interest to use hydrogen cyanide (HCN) as a tracer for molecular gas (Gao and Solomon, 2004; J. Kauffmann et al., 2017). As an alternative or even a more suitable choice as tracer, HCO^+ is considered for some molecular gas observations in ULIRGs (Graciá-Carpio et al., 2006; Papadopoulos, 2007).

Additionally to HCN, Wu et al. (2010) considers CS as tracer for molecular gas. Along with CO, Goicoechea et al. (2019) used CH^+, HCN, HCO^+ and CH to trace stellar feedback processes in the presence of molecular gas. Another candidate for the determination of molecular gas is the OH molecule which has been found in local observations in the star forming region (Allen, Hogg and Engelke, 2015; Nguyen et al., 2018; Engelke and Allen, 2019). Even polycyclic aromatic hydrocarbons (PAHs) which represent a class of rather complex molecules are discussed in Cortzen et al. (2019) as potential molecular gas tracers. This is especially interesting for high-redshift observations since emissions of PAHs are detected by JWST.

Summarizing, there is a broad variety of molecules that can be used as tracers for molecular gas. However, they can be used only for local observations and not for intensive extragalactic studies since the instruments still have limited resolution and sensitivity.

In the following, we will focus on the internal properties of molecular hydrogen and carbon monoxide and under which conditions and processes it is formed and destroyed.

[14] Square brackets indicate forbidden transitions. For details see section 1.7.
[15] Isotopologues are molecules consisting of isotopes of an element.

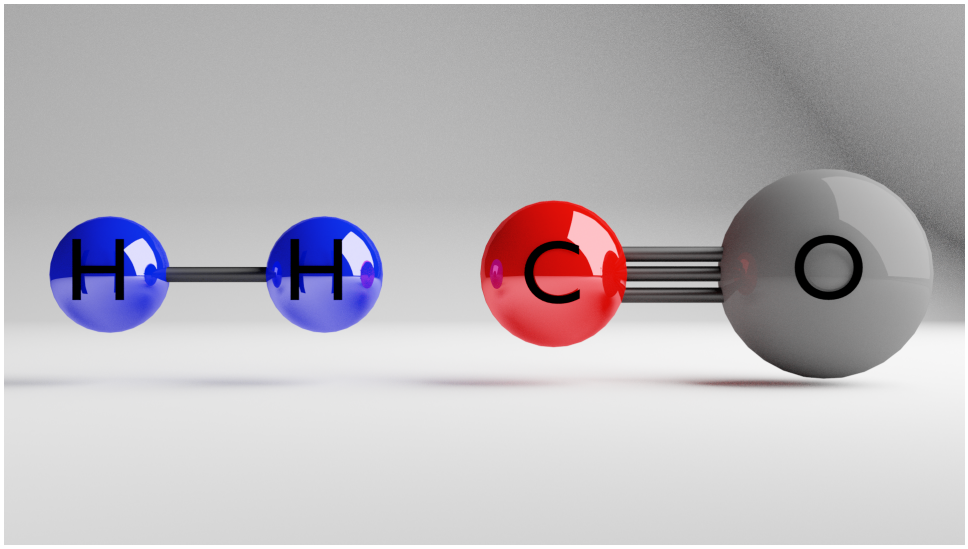

Figure 1.5: Schematic representation of the H_2 and CO molecule. Due to its symmetry, molecular hydrogen has no permanent dipole moment which makes it extremely difficult to observe. In contrast, CO consists of two different constituents giving a dipole moment to it. Note the triple bond in the molecule and the fact that the negative charge is concentrated around the carbon atom.

1.7 Theory of spectra

In the following, we will discuss the basic aspects of energy levels, specifically for the case of the tracers that are relevant for tracing molecular gas. We start with fine-structure transitions as they are relevant for the observations of atomic and ionized carbon.

Fine-structure transitions are caused by the coupling of the angular momentum of the electron with its spin. In theory, this gives rise to an additional Hamiltonian

$$\hat{H}_{LS} \propto \hat{L}\hat{S} = \frac{1}{2}\left[\hat{J}^2 - (\hat{L}+\hat{S})^2\right] \tag{1.15}$$

which results in an energy term of the form

$$E_{LS} \propto \frac{1}{2}\left[j(j+1) - l(l+1) - s(s+1)\right] . \tag{1.16}$$

The resulting energy levels depend on the electron configuration of the atom that is observed. The electron configuration of neutral carbon is $1s^2 2s^2 2p^2$ and for singly ionized carbon one finds $1s^2 2s^2 2p$, consequently. It is also common to use a short-hand notation like $[\mathrm{He}]2s^2 2p^2$ to sum up the already filled shells. The notation convention of the energy levels reads

$$^{2S+1}\mathcal{L}_J . \tag{1.17}$$

Following the rules for the coupling of angular momenta, the allowed values for J for the electron configuration of neutral carbon are 0,1,2, whereas for singly ionized carbon, we find 1/2 and 3/2. The fine-structure transitions are then for [CI] $^3P_2 \rightarrow^3 P_1$ and $^3P_1 \rightarrow^3 P_0$, for [CII] $^2P_{3/2} \rightarrow^2 P_{1/2}$. The square brackets indicate that these transitions are forbidden according to the dipole selections rules. These rules state among other things that a transition comes always along with a change in angular

momentum of 1 which is here not the case.

In contrast to single atoms which can have electronic transitions only, molecules have additional degrees of freedom for excitations, such as vibrational and rotational states. This can be summarized by a formula like the following:

$$\begin{aligned} E &= E_{\mathrm{el}} + E_{\mathrm{rot}} + E_{\mathrm{vib}} \\ &= E_{\mathrm{el}} + \frac{\hbar^2 l(l+1)}{2I} + \hbar\omega\left(n + \frac{1}{2}\right) . \end{aligned} \tag{1.18}$$

The carbon monoxide molecule is composed of an oxygen and a carbon atom which already imposes an asymmetry to the electron distribution within the molecule. In some sense, CO is quite a special molecule since it exhibits a triple bond between the atoms. Due to its asymmetry and the different ability of carbon and oxygen atoms to keep electrons in their proximity, electrons are surprisingly more concentrated around the carbon atom leading to a dipole moment.
The rotational states of CO are used to radiate away the energy it gains from its collisions with H_2. From quantum mechanics, the energies of these states can be calculated by the following Schrödinger equation:

$$\hat{H}\Psi(x) = E\Psi(x) \tag{1.19}$$

with

$$\begin{aligned} \hat{H} &= -\frac{\hbar^2}{2\mu}\Delta \\ &= -\frac{\hbar^2}{2\mu}\left[\frac{\partial}{r^2\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2(\theta)}\frac{\partial^2}{\partial\phi^2}\right] \end{aligned} \tag{1.20}$$

with μ being the reduced mass of the molecule. Since we assume for the molecule a rigid rotor with a fixed distance between both atoms, the radial derivatives can be neglected and one is left just with the angular part. The solution to equation (1.20) is given by the Eigenfunctions for the angular part of the Laplacian, namely the spherical harmonics $Y_{lm}(\theta, \phi)$ and their Eigenvalues. It follows that

$$E = \frac{\hbar^2}{2\mu r_0^2} l(l+1) = \frac{\hbar^2 l(l+1)}{2I} = Bl(l+1) \tag{1.21}$$

with r_0 being the fixed distance between both atoms (equilibrium length) and the rotational constant B combining all constants. Further, it is common to express the vibrational constant in terms of a reciprocal length of the form

$$\bar{B} = \frac{B}{hc} = \frac{h}{8\pi^2 cI} . \tag{1.22}$$

corresponding to a characteristic wave number. For the CO molecule, $\bar{B}$ is found to be 193.13 cm^{-1} and for H_2 it is found to be 60.85 cm^{-1}. Thus, the spectrum for the rotational states and their frequencies can be calculated as shown in Fig. 1.6. From that, the consequences of the weak dipole moment are emphasized as well. The rotational excitations and the transitions of H_2 would trace the phase of the ISM with high temperatures and not the cold neutral medium where star formation takes place.

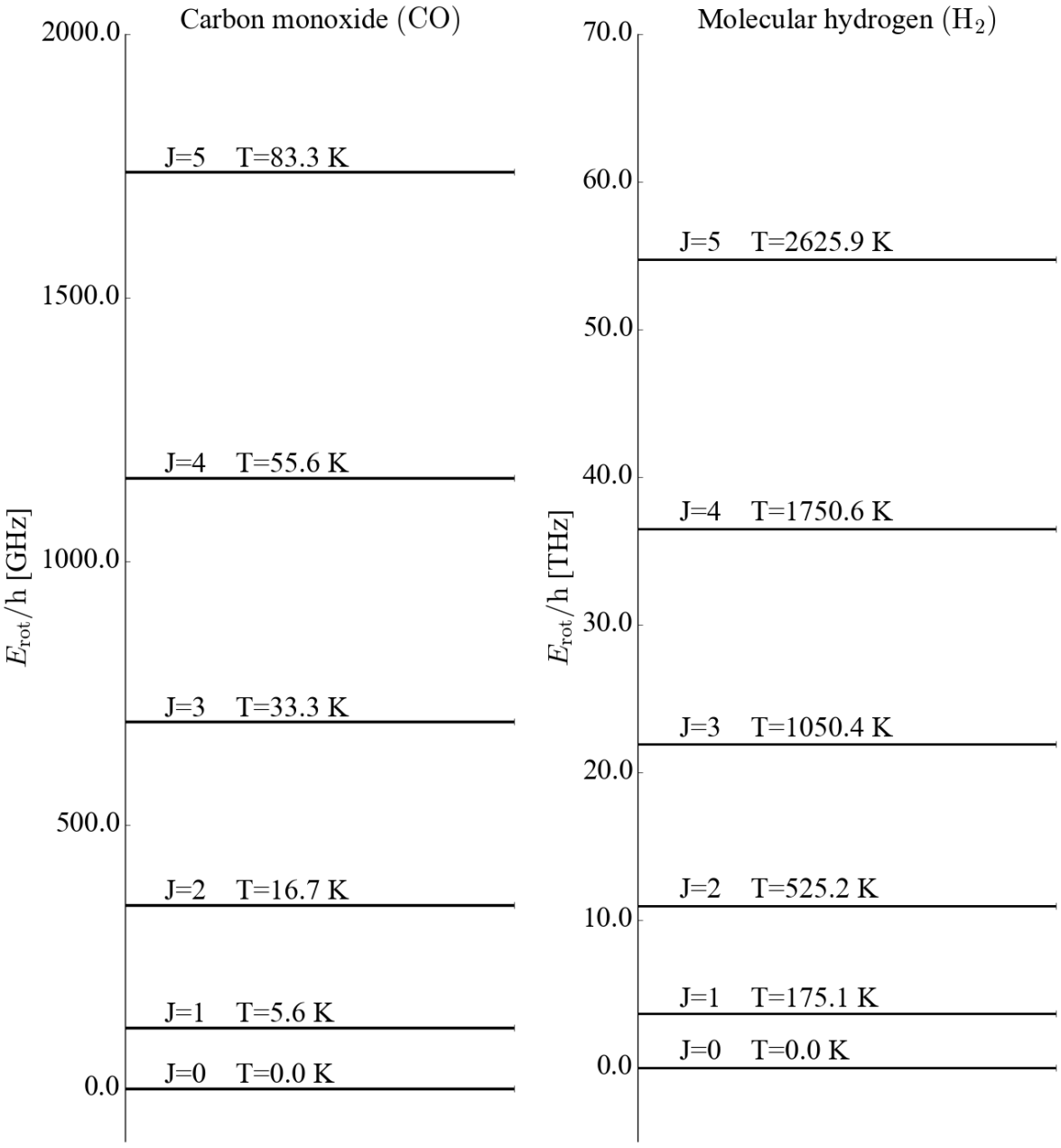

Figure 1.6: Rotational spectra for carbon monoxide (left) and molecular hydrogen (right). One sees that the transitions for CO can take place at much lower temperatures as for molecular hydrogen. That makes carbon monoxide a far better candidate for tracing the cold phase of the ISM.

1.8 Chemistry in the ISM

The ISM hosts various ions, atoms and molecules which interact also on a chemical level. To capture the evolution of the number density of a species n_s, a rate equation of the following form is used:

$$\frac{\mathrm{d}n_s}{\mathrm{d}t} = \sum_{i,j}(\pm k^{ij}_{2-\mathrm{body,s}} n_i n_j) + \sum_i (\pm k^i_{s,\mathrm{gr}} n_i n_\mathrm{H}) + \sum_i (\pm k^i_{s,\mathrm{cr,s}} n_i) + \sum_i (\pm k^i_{s,\gamma,\mathrm{s}} n_i) \tag{1.23}$$

where n_H is the total number density of hydrogen atoms. $k^{ij}_{2-\mathrm{body,s}}$ is the rate coefficient of a two-body reaction between the species i and j that form or destroy – depending on the sign – species s. $k^i_{s,\mathrm{gr}}$ represents the grain assisted reaction rate (for reactions where dust grains act as a catalyser), $k^i_{s,\mathrm{cr,s}}$ and $k^i_{s,\gamma,\mathrm{s}}$ account for interactions with cosmic rays or photons. Many of these coefficients can have dependencies on the UV radiation strength, metallicity or temperature. To obtain the evolution of each species in a chemical network, a system of ordinary coupled differential equations has to be solved. In the following, we give a short insight into the chemistry of H_2 and CO.

1.8.1 Molecular hydrogen

The chemistry of the interstellar medium can be very complicated and there are a lot of reactions that form, destroy or continue processing H_2 but we just want to give a short overview in the following.
There are several ways to form molecular hydrogen in the ISM. Naively, the most obvious way would be

$$\mathrm{H + H \rightarrow H_2 + \gamma} \tag{1.24}$$

which indeed can happen – but on timescales larger than the age of the universe. The problem with this is that two hydrogen atoms have to meet, produce a molecule and radiate the exact amount of energy away, instead of using it to split up again. But also electrons can act as catalyser and initiate a reaction channel that leads to the formation of H_2:

$$\mathrm{H + e^- \rightarrow H^- + \gamma} \tag{1.25}$$

$$\mathrm{H^- + H \rightarrow H_2 + e^-}\,. \tag{1.26}$$

Three-body collisions are also another way for the formation of H_2:

$$\mathrm{H + H + H \rightarrow H_2 + H}\,. \tag{1.27}$$

However, in metal enriched gas, silicates act as a catalyser since they provide a surface hydrogen atoms can stick to. Single hydrogen atoms can move on the surface of these silicates and due to the proximity of one atom to another, they can detach in form of a hydrogen molecule. As a reaction, this reads

$$\mathrm{H + H + grain \rightarrow H_2 + grain}\,. \tag{1.28}$$

In this case, the reaction coefficient has to depend on several variables such as the metal abundance (accounting for the amount of grains), the sticking probability of hydrogen atoms to dust and the temperature. The destruction happens via UV photons of energy between 11.2 eV and 13.6 eV which is also known as the *Lyman-Werner band*:

$$\mathrm{H_2 + \gamma_{LW} \rightarrow H_2^* \rightarrow H + H}\,. \tag{1.29}$$

Also in this case, the reaction rate depends on parameters like the UV field strength, shielding from atomic and molecular hydrogen (self-shielding) and dust. We will give further details for the reaction coefficients in the context of the problem considered in chapter 3.

1.8.2 Carbon monoxide

The second most abundant molecule in the ISM is carbon monoxide. It exists in the cold, dense part of the ISM and – in contrast to molecular hydrogen – has a dipole moment, which makes it a suitable tracer for molecular gas. There is a variety of formation channels. Most of them are initiated by the interaction of a cosmic ray with a hydrogen molecule:

$$\mathrm{H_2 + c.r. \rightarrow H_2^+ + c.r. + e^-} \tag{1.30}$$

$$\mathrm{H_2^+ + H_2 \rightarrow H_3^+ + H\,.} \tag{1.31}$$

The H_3^+ molecule is the basis for a lot of follow-up reactions involving carbon and oxygen intermediate molecules that are required in order to form CO. Most of these reactions do not involve dust grains and are gas-phase reactions. We do not list all reactions in detail, but to give an impression to the reader for the level of complexity of a chemical network that includes CO, we recommend further literature (e.g. Glover, Federrath et al., 2010; Gong, E. C. Ostriker and Wolfire, 2017).

The destruction is mainly driven by UV radiation leading to the photodissociation reaction of the form

$$\mathrm{CO} + \gamma \rightarrow \mathrm{C + O}\,. \tag{1.32}$$

The binding energy of carbon monoxide is 11.1 eV. Thus, each photon with a larger energy can destroy this molecule. Like for the H_2, shielding effects from the CO and H_2 play here an important role.

1.9 Star formation

In this section, we will discuss the basics of the process of star formation. We will give a short overview about the most important aspects. We refer and recommend for further details the reviews of Krumholz (2015) and McKee and E. C. Ostriker (2007).

1.9.1 Star formation in galaxies

One of the most striking features of a galaxy are its stars as this is its most luminous part for our eyes. Additionally, star formation induces a lot of different processes that highly influence the process of galaxy formation and evolution. One of the first hypotheses on star formation came from William Herschel. He believed that stars are produced by the conversion of nebular structures (Herschel, 1811) which is basically still accepted today. Over the years, observations that follow this idea have been made. In 1892, Barnard discovered a nova and the expansion of a gas cloud at the same time (Barnard, 1892). Later in 1908, he noticed objects which he describes as “a number of very black, small, sharply defined spots or holes”. It turned out that these clouds were so dense that the light from stars behind them was completely absorbed. In the 1940s the idea of condensing gaseous structures became more popular (Spitzer, 1941; Whipple, 1946; Bok and Reilly, 1947).

A more quantitative analysis was published in 1959 by Maarten Schmidt (Schmidt, 1959) who investigated the assumption of a correlation of interstellar gas and star formation surface densities (Σ_{gas}

and $\Sigma_{\rm SFR}$) in form of a power law, specifically

$$\Sigma_{\rm SFR} = A\,\Sigma_{\rm gas}^{n}\ . \tag{1.33}$$

Based on the available data at that time, he concluded that the power of that non-linear relation is around $n \approx 2$. From this relation, one reads off the fact that the higher the gas abundance in a galaxy (or a region therein) is, the higher is its ability to form stars.

With better instruments, this study was repeated by Kennicutt (Kennicutt, 1998) and the relation was found to be

$$\Sigma_{\rm SFR} = (2.5 \pm 0.7) \times 10^{-4} {\rm M_\odot yr^{-1} kpc^{-2}} \left(\frac{\Sigma_{\rm gas}}{\rm M_\odot pc^{-2}} \right)^{(1.4\pm0.5)} \tag{1.34}$$

which holds over several orders of magnitude. According to equation (1.34), he found a smaller value of $n \approx 1.4$ from local starburst and disk galaxies compared to Schmidt. This connection of gas and star-formation surface density became known as the *Kennicutt-Schmidt relation* (KS relation hereafter).

It was always of large interest to understand why $\Sigma_{\rm SFR}$ and $\Sigma_{\rm gas}$ are related this way. Many approaches (e.g. Silk, 1997; Krumholz and McKee, 2005) are based on the assumptions that the average star formation rate density in a region is proportional to the ratio of the average gas density over a dynamical timescale $t_{\rm dyn}$ driven by gravitational collapse

$$\dot{\rho}_{\rm SF} = \epsilon \frac{\rho_{\rm gas}}{t_{\rm dyn}} \propto \rho_{\rm gas}^{1.5}\ , \tag{1.35}$$

with ϵ being the star formation efficiency, which is of order 10^{-2} as derived from observations. Depending on the galaxy, variations of the slope and the normalization of KS-relation were observed and until today, the values for the slope and the normalization in formula (1.33) are still revisited for e.g. different galaxy types, different redshifts (de los Reyes and Kennicutt, 2019; Kennicutt and de los Reyes, 2021) meaning that there is no universal KS-relation.

Wong and Blitz (2002) found an excellent correlation between the surface densities of star formation and H_2 by parameterizing the exponent in the Schmidt law in the form of

$$n_{\rm mol} = \frac{{\rm d}\log \Sigma_{\rm SFR}}{{\rm d}\log \Sigma_{\rm mol}}\ . \tag{1.36}$$

It is worth noting that Kennicutt (1998) considered molecular gas already as well, but his $\Sigma_{\rm H_2}$ data suffered from using a single CO-to-H_2 conversion factor for the entire galaxy population. In the following years, spatially resolved CO-observations of nearby galaxies with better telescopes were made on scales of 0.5 – 1 kpc. Those were indicating a significant role of H_2 for star formation due to its strong correlation with the star formation rate density. In the following years, this has been supported by several observations (e.g. Bigiel, Leroy, Walter, Brinks, de Blok, Madore et al., 2008; Kennicutt, Calzetti et al., 2007; Bigiel, Leroy, Walter, Brinks, de Blok, Kramer et al., 2011) and it turned out that the correlation between the molecular gas surface density $\Sigma_{\rm H_2}$ and $\Sigma_{\rm SFR}$ is much tighter compared to the KS-relation. Further, it is much more insensitive to variations in the dust-to-gas ratio and the stellar surface density. Summarizing, this is strong evidence that star formation requires the presence of molecular hydrogen.

Still, it is unclear whether there is a correlation or a causation between star formation and molecular hydrogen. There are studies that do not find any dependence of star formation on H_2 (Krumholz, Leroy and McKee, 2011; Glover and Clark, 2012b; Glover and Clark, 2012c) and point out that the conditions

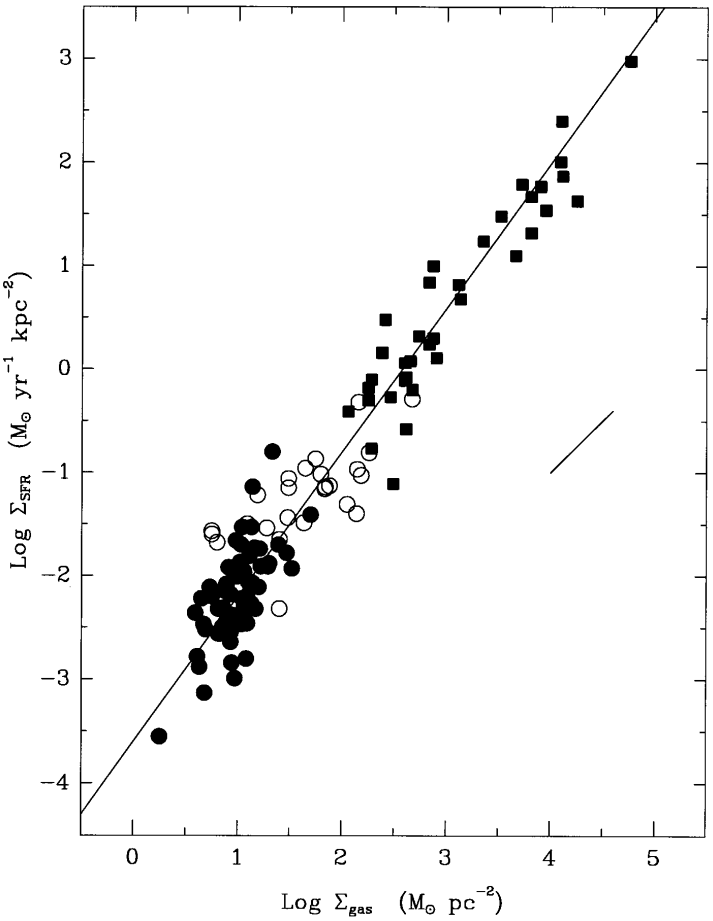

Figure 1.7: The Kennicutt-Schmidt relation from the original publication (Kennicutt, 1998). It directly relates the gas abundance with the SFR of a galaxy. The slope of the linear fit is 1.4 in log space.

for the formation of stars and H_2 are the same. Both form in regions where the gas is dense and cold and well shielded against the UV radiation. Hence, the challenge remains which actual role molecular hydrogen plays in the process of star formation.

1.9.2 Star formation in GMCs

Previously, we considered star formation on a galactic scale but we will now focus on what happens on the level of molecular clouds. Important for the process of star formation is the internal structure of a GMC.

Necessary to understand its properties and dynamics, observations were done by Larson (1981) leading to three main conclusions. The first one was that clouds are supersonically turbulent and their velocity dispersion scales linearly with their size. Further, Larson figured out that these clouds are gravitationally bound objects leading to a relation that connects the velocity dispersion with the cloud mass via a power law. The second one was that the velocity dispersion scales with the cloud mass by a power law as well, supporting the cloud against a gravitational collapse within the free-fall time. The third one is obtained by eliminating the velocity dispersion when combing the first two conclusions: the density of a GMC which is stable against turbulence and gravitational collapse scales inversely with its size. To understand the

collapse of a cloud and its conversion into stars, we now need to have a more detailed look on its energy budget. It consists of several components and keeping its mass fixed, it reads in its full Lagrangian form

$$\frac{1}{2}\ddot{I} = 2(\mathcal{T} - \mathcal{T}_s) + \mathcal{B} + \mathcal{W} \,. \tag{1.37}$$

$\ddot{I}$ is connected to the inertia $I = \int r^2 \mathrm{d}m$ of a cloud and its second time derivative describes the acceleration of the cloud expansion or contraction. $\mathcal{T}$ and $\mathcal{T}_s$ account for the kinetic energy in the cloud and the surface kinetic energy, respectively. $\mathcal{B}$ accounts for the energy coming from magnetic fields (e.g. turbulence from fluctuating magnetic fields) and $\mathcal{W}$ represents the contribution from gravity. The stability of a cloud depends on the balance of kinetic energy and gravity, specifically $\mathcal{W} = -2\mathcal{T}$. The ratio of these two energies is defined as the virial parameter α_{vir}, which is found to be

$$\alpha_{\mathrm{vir}} = \frac{5\sigma^2 R}{GM} \,, \tag{1.38}$$

for $\mathcal{W} = \frac{3GM^2}{5R}$ and $\mathcal{T} = \frac{3}{2}\sigma^2 M$. Provided the internal energy of a cloud cannot counteract gravity anymore, the cloud starts to contract. The time scale for a pressure-free collapse is given by the free fall time t_{ff}:

$$t_{\mathrm{ff}} = \left(\frac{3\pi}{32G\bar{\rho}}\right)^{1/2} = 1.37 \times 10^6 \left(\frac{10^3 \mathrm{cm}^{-3}}{\bar{n}_{\mathrm{H}}}\right)^{1/2} \tag{1.39}$$

with $\bar{\rho}$ and $\bar{n}_{\mathrm{H}}$ being the average densities of the cloud[16]. If this collapse starts, the cloud is going through the process of fragmentation. Jeans (1902) started to model the gravitational instability of a spherical gas cloud and determined a characteristic length scale λ_{J} for the collapsing case:

$$\lambda_{\mathrm{J}} = c_s \frac{\pi}{G\bar{\rho}} \tag{1.40}$$

with the speed of sound $c_s = \left(\frac{k_B T}{m}\right)^{1/2}$. The mass that is enclosed in a sphere of size of this length gives the *Jeans mass*:

$$M_{\mathrm{J}} = \frac{4\pi}{3}\bar{\rho} R_{\mathrm{J}}^3 \,, \tag{1.41}$$

using $R_{\mathrm{J}} = \frac{\lambda_{\mathrm{J}}}{2}$. The *Jeans length* is the critical length scale of a perturbation that leads to an instability of the system. This is followed by the gravitational collapse of the system. At the same time, it marks the smallest size that can collapse. Equivalently, the *Jeans mass* gives the smallest mass that is able to perform a gravitational collapse. The fragmentation of the GMC into smaller gravitationally bound chunks marks the last step, before these cores evolve into stars.

1.10 Numerical simulations of galaxy formation

Galaxy formation and evolution are influenced by many different processes, from the clustering of dark matter on large scales up to the gravitational collapse of gas clouds on ranges of several light years. The interplay of these processes leads to a highly non-linear problem that cannot be treated analytically. The

[16] Numerical values are obtained assuming a He contribution of 10%.

complexity of galaxy formation simulations arises from the combination of the physics from different scales. We will first give an historical overview about the evolution of numerical simulations that were designed in order to capture all these effects before we have a more detailed look on their implementation. To illustrate the synergy between the evolution of numerical methods and computational power, we discuss briefly the history of simulations of galaxy formation.

1.10.1 The art of simulating the formation of galaxies – A historic overview

The very first simulation was set up by Holmberg in 1941 (Holmberg, 1941) and focussed on the dynamics of stars in two galaxies during a merging process. Mathematically, this is an N-body problem that needs to be solved. The force $\mathbf{F}_i$ that acts on the i-th particle coming from all other particles is formulated as follows[17]:

$$\mathbf{F}_i = \sum_{i \neq j}^{N} G \frac{m_i m_j (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} \,. \tag{1.42}$$

Since there were no computers at that time, Holmberg used 37 light bulbs and photocells (galvanometer) to calculate the gravitational potential and updated the positions of the bulbs accordingly. Despite the low number of objects, he was able to conclude that small galaxies merge into a larger one.
Still, the dynamic complexity of galaxies was not fully captured by Holmberg's approach and a more accurate description requires the usage of computers. As the performance of computers increased, Sebastian von Hoerner did the first N-body simulations to investigate the stellar dynamics within a stellar cluster (von Hoerner, 1960; von Hoerner, 1963). In 1963, Sverre Aarseth published the results of simulating clusters of galaxies introducing the *softening length* to keep the gravitational forces convergent at short distances (Aarseth, 1963). The number of particles used reached 100 by then.
With growing evidence of the existence of dark matter throughout the entire universe, the interest in understanding its role in the history of our universe increased as well. Thus, N-body simulations became a useful tool for modelling the gravitational effect of dark matter and its influence on the structure formation of the universe.

But simulating larger structures in a cosmological volume requires the implementation of more particles in order to have a decent resolution for studying the collapse of dark matter haloes.

The first simulations for studying the large-scale structure used a number of particles between 10^3 and 2×10^4 (Gott, Turner and Aarseth, 1979; Efstathiou and Eastwood, 1981). The need to simulate higher numbers of particles on the one hand and the limited computational resources on the other hand led to the invention of different approaches. A lot of these were invented during the 80s like the tree-code method (Barnes and Hut, 1986) or the Fast Multipole Method (Greengard and Rokhlin, 1986) (for reviews, see Hockney and Eastwood (1981) and Sellwood (1987)). The first tree code was then finally released in 1993 (McMillan, 1993).

These new methods and the evolution of computing power allowed to study the clustering effects in a CDM universe (Warren et al., 1992; Gelb and Bertschinger, 1994; Jenkins et al., 1998; Governato et al., 1999) since it was then possible to simulate a cosmological volume.

All these simulations were dedicated to understanding the clustering behavior of dark matter which is the first step in the process of galaxy formation. However, the formation of stars and the gas dynamics must be implemented properly, too. At that point, it became necessary to add a component to the

[17] This can be further simplified, if all the masses are equal which is mostly the case in galaxy formation simulations.

simulations that describes the hydrodynamic behavior of gas accurately. To address this in the framework of astrophysics, Gingold and Monaghan (1977) developed and presented an SPH algorithm (*Smoothed-particle hydrodynamics*). Their approach is used in many other research fields where fluid dynamics is needed. Almost 20 years later, Hultman and Pharasyn (1999) implemented the SPH approach. In parallel, during the 80s, the first grid codes were developed. At that time, some dynamical effects in a fluid were poorly captured in an SPH code (e.g. shocks, Helmholtz instabilities) and a Eulerian treatment of the fluid would be much more advantageous. This description however is in need of a spatial grid which imposes a problem. Due to the constant, limited resolution across the simulation volume and missing computational power, grid codes were unable to resolve the typical high-density regions where star formation occurs. Throughout the 90s, simulations with fixed grid resolution were run (Cen and J. P. Ostriker, 1993; Gnedin and Hui, 1998). But the main problem of these grid codes remained: to have a decent resolution in regions of particular interest, the grid has to consist of a huge amount of cells and in return, a high number of cells increases the computation time. So far isolated objects could be simulated, but very large scales were basically impossible to consider.
In order to avoid the problem of a grid code, AMR (adaptive mesh refinement) codes have been developed. They allow to save computational time, since they increase the grid resolution only in regions of interest. In the context of galaxy formation, AMR codes increase their resolution at the position of the haloes to better capture the gas dynamics which is very important for the galaxy evolution. The most prominent codes are the ENZO code (Bryan and Norman, 1998) and RAMSES (Teyssier, 2002). They both combine an N-body solver with the Eulerian description of the gas, however they differ in their data structure and refinement strategy.
In 2000, Springel, Yoshida and White (2001) released the SPH code GADGET-1. Another milestone in the chapter of computational cosmology was the Millennium Run in 2005 (Springel, White et al., 2005) using a version of the GADGET code. Starting from a redshift of $z = 127$ ($\approx$ 380000 years after the Big Bang) from initial conditions drawn from the CMB, the evolution of $N = 2160^3$ particles within a cubic volume with 500 h^{-1} Mpc boxlength was calculated. To the present day, GADGET has been developed further and its newest version is GADGET-3. This code has been widely used e.g. in the EAGLE (Evolution and Assembly of GaLaxies and their Environments) simulation project (Schaye et al., 2015). Using the AREPO code (Springel, 2010), the Illustris project was carried out by Vogelsberger et al. (2014) as an application. It is also the aim of this project to better understand the process of galaxy formation and evolution. The most recent version of Illustris is IllustrisTNG (The next Generation) (Pillepich et al., 2018). It is an updated version of the Illustris code capturing the ISM physics with state-of-the-art models. There are three different runs called TNG50, TNG100 and TNG300 (for 50 Mpc, 100 Mpc and 300 Mpc comoving boxsize respectively) and all their results are discussed in detail in several publications.

In the following, we will discuss the different types of codes and how they adress the dynamics of gas, stars and dark matter numerically.

1.10.2 Numerical approach to the N-body problem

The force that acts on a single particle in an N-body system can be directly calculated from equation (1.42). Since this needs to be done for all particles, this approach is not suitable for large systems since the computation time for this problem scales with $O(N^2)$ with N being the number of particles.

One possible approach are so called *Tree-codes* which calculate the force within a given range according to equation (1.42). The force coming from all more distant particles is summed up and represented by single effective particles. The trees which contain the information about the interactions among the particles do not need to be calculated at every time step saving a significant amount of computation time.

Other approaches to the N-body problem start from the Poisson equation in order to obtain the potential by calculating the particle density $\rho(\mathbf{x}_{i,j,k})$ on a grid in the simulation volume. $\mathbf{x}_{i,j,k}$ represents the cell centers of the grid. The set of equations to be solved then reads:

$$\Delta\Phi(\mathbf{x}_{i,j,k}) = 4\pi G\rho(\mathbf{x}_{i,j,k}) \tag{1.43}$$

$$\mathbf{F}(\mathbf{x}_{i,j,k}) = -m\nabla\Phi(\mathbf{x}_{i,j,k})\,. \tag{1.44}$$

The difficulty of solving differential equations can be bypassed by going to Fourier space. This reduces the calculation of a gradient or a divergence to a multiplication with the corresponding parameter in Fourier space. It turned out that this treatment of an N-body problem saves a significant amount of time. Traditionally, each particle carries information about its position and velocity throughout the simulation. Depending on the force that acts on it, both quantities are updated. However, some codes do not rely on that principle only. Another way to save computational resources is the *Particle-Mesh* scheme (PM hereafter) where an interpolation scheme is used to calculate a density field based on the positions of *all* particles. For each particle, it is calculated how it is gravitationally affected by that density field.
Particle–Particle–Particle–Mesh (P^3M hereafter) is a hybrid approach that uses a PM scheme for long-range gravitational interaction between particles and the tree scheme for small distances. This decreases the error that is imposed by the limited resolution of the mesh, especially for particles at small separations.
There is still one problem with equation (1.42) that needs to be highlighted. For particles at very small separations, equation (1.42) exhibits a problem as $\mathbf{F}_i$ diverges. To solve this problem, one introduces a softening length a, which is a measure of the spatial resolution of the simulation. In other words, a is the equivalent to Δx in a grid-based code. Formula (1.42) reads in a modified form:

$$\mathbf{F}_i = \sum_{i\neq j}^{N} G\frac{m_i m_j(\mathbf{r}_i - \mathbf{r}_j)}{(|\mathbf{r}_i - \mathbf{r}_j|^2 - a^2)^{3/2}}\,. \tag{1.45}$$

The value for the softening length is chosen to be of order of the mean separation of two particles in the simulation volume. Thus, the more particles are found in a fixed volume, the smaller a is chosen to be. In this case, the short-range interactions are captured in more detail.
The previous discussion is for dark matter and stellar particles only as there is just gravitational interaction. The physical behavior of gas is quite different and the N-body approach is not sufficient. Thus, we will discuss the approaches in more detail in the following.

1.10.3 Numerical approach to hydrodynamics

Dark matter and stars are treated as an N-body problem, but the gas component in a code needs to be treated differently since its physics is very different. For that, the equations of fluid dynamics are used and they read:

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \tag{1.46}$$

$$\frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla\Phi - \frac{1}{\rho}\nabla P \tag{1.47}$$

$$\frac{\partial\epsilon\rho}{\partial t} + \nabla\cdot[\rho\mathbf{v}(\epsilon + P/\rho)] = 0 \tag{1.48}$$

$$\Delta\Phi = 4\pi G\rho\,. \tag{1.49}$$

ρ is the density, $\mathbf{v}$ the velocity, P the pressure, ϵ the energy density of the gas, and Φ the gravitational potential. These equations are known as *Euler equations*, except the last one that is also known as the *Poisson equation*. They describe the conservation of mass, momentum and energy. Solving them is a central aspect of simulations of galaxy formation. We will discuss in the following the different approaches that have been established.

Eulerian codes

In the previous section, we introduced the three Euler equations that are needed to describe the gas dynamics. One can rewrite those equations in a form that is more compact:

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{S} = 0 \tag{1.50}$$

where $\mathbf{U}$ and $\mathbf{S}$ are defined as

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho \mathbf{v} \\ \rho \epsilon \end{pmatrix} \tag{1.51}$$

$$\mathbf{S} = \begin{pmatrix} \rho \mathbf{v} \\ \mathbf{v} \otimes \rho \mathbf{v} + P \mathbb{I} \\ \mathbf{v}(\rho \epsilon + P) \end{pmatrix} . \tag{1.52}$$

Eulerian codes solve the hydrodynamical equations on a grid. This grid consists of small volume elements, called cells. The core task of the simulation code is to calculate the gas flow through the interface of two resolution elements. The mathematical formulation of this issue is also known as *Riemann problem*. One solution scheme is the *Godunov method*. The Godunov scheme starts with the assumption that the value of each hydrodynamic quantity is constant across the whole cell. Further, an approximation of the flux has to be made as the exact treatment of a Riemann problem is computationally too expensive and for that different, schemes can be used, e.g. the Lax-Friedrich flux function. Based on the constraint that the flux is a function of the shape $\mathbf{F} = \mathbf{F}(\mathbf{U}_L, \mathbf{U}_R)$, where $\mathbf{U}_L$ and $\mathbf{U}_R$ are the corresponding vectors of two neighbouring cells, the Godunov scheme allows to write the fluid equations in the following discretized form:

$$\frac{U_i^{n+1} - U_i^n}{\Delta t} + \frac{F_{i+1/2}^{n+1/2} - F_{i-1/2}^{n+1/2}}{\Delta x} = S_i^{n+1/2} . \tag{1.53}$$

For further mathematical details, we suggest Teyssier (2015) to the reader. Another essential component of the Eulerian approach is the shape of the mesh. The simplest geometry is an uniform grid with equally-spaced and -sized resolution elements. But a decent resolution across the whole simulation means that even regions of low interest have a high number of resolution elements. *Moving mesh codes* are exactly designed in order to avoid that and to keep the computational costs low. Based on a condition or a strategy, e.g. each cell has to contain a fixed mass, the grid is refined. Alternatively, Arepo (Springel, 2010) uses an unstructured grid using Voronoi-tesselations to adapt the grid size and to have a better resolution in regions of high gas density. Codes like Flash, Enzo, and Ramses are using the AMR technique where a uniform grid is the starting point but the cells are subdivided into regions where certain (eventually user-set) conditions are met, e.g. the baryonic or the dark matter component exceeds a specific value. Although these codes share the same refinement strategy, there can be differences in the data structures and the hierachy between the different refinement levels.

Lagrangian codes

As an alternative to grid codes, a fluid can be effectively modelled as particles. The fluid equations are then solved in their reference frame which is the basis for *Lagrangian codes*. To each particle, a mass, momentum, pressure, energy, and temperature (or even more properties) are assigned. To represent e.g. a continuous density field, each particle is smoothed out using a *kernel function* $W(\mathbf{r}' - \mathbf{r}, \mathbf{h})$. The local average of a function $g(\mathbf{r})$ then reads

$$\langle g(\mathbf{r}) \rangle = \int g(\mathbf{r}) W(\mathbf{r}' - \mathbf{r}, \mathbf{h}) \mathrm{d}\mathbf{r}' \,. \tag{1.54}$$

In most cases, a spherical symmetry for the kernel is chosen leading to $|\mathbf{r}'| = r'$ and $|\mathbf{h}| = h$. The smoothing length h is set in such a fashion that it reflects the average particle distance. Hence it depends on the particle density within the entire computational volume.
One has to keep in mind that particles are discrete objects and, hence, the integral is replaced by a sum over all particles j in the neighbourhood of a particle denoted by index i:

$$g(\mathbf{r}) = \sum_j m_j \frac{g_j}{\rho_j} W(\mathbf{r}_i - \mathbf{r}_j, \mathbf{h}) \,. \tag{1.55}$$

Specifically, the equation e.g. for the density reads

$$\rho_i = \sum_j m_j W(\mathbf{r}_i - \mathbf{r}_j, \mathbf{h}) \,, \tag{1.56}$$

and this can be applied to other quantities of the fluid in order to rewrite the fluid equations which is not discussed further.

1.10.4 Some final remarks on simulation codes

To have an implementation that captures all physical and chemical processes accurately in a simulation, many simulation codes need further modelling. This is crucial especially for cosmological simulations as they have resolution elements with size of tens of parsecs nowadays. *Subgrid models* provide effective models for the resolved scales and allow to include unresolved processes like e.g chemistry, star formation and feedback.

The different implementations of the hydro codes have advantages and disadvantages which we point out in the following.
A large advantage of *smoothed-particle hydrodynamics* codes is their high stability. Also, the resolution automatically follows those regions where mass is concentrated. Due to their meshless approach, the conservation of angular momentum is assured. One major issue of SPH codes is the shock treatment. Mathematically, a shock is represented by a discontinuity of the fluid equations. But this is exactly the problem for an SPH code, as a shock happens on much smaller spatial scales than those which are usually resolved by the SPH particles. The shock is then poorly resolved, usually by scales a few times the smoothing length. However, introducing an artificial viscosity term to the fluid equation can fix that but this addition may cause other spurious effects.

Since there is a variety of *grid codes*, they all have their own disadvantages. The main issue of a uniform grid is that one has to decide for one resolution across the grid that cannot be that high for computational reasons. Further, in cosmological simulations, the physical resolution gets worse with time due to the expansion of the universe. However, this can be solved by codes that redefine their grid.

Further, in contrast to SPH codes, grid-based codes suffer from losses in the angular momentum which can only be compensated by choosing a symmetry of the grid that reflects the symmetry of the problem, which is not always realistic. Moving mesh codes can suffer from distortions of their grid leading to unknown numerical errors. Apart from those things, grid codes can treat shocks or other fluid instabilities much better without introducing any artificial terms to the fluid equations.

1.11 The Ramses code

Throughout this thesis, the hydrodynamical code Ramses has been used to obtain the presented results. We will dedicate the following section to the details, settings and properties of the code.

The hydrodynamical part of Ramses uses a grid approach which involves the AMR technique to have a decent resolution in the central part of a halo where gas dynamics plays a significant role, and to avoid the necessity to have a highly resolved uniform grid in the entire simulation box. The code uses different criteria to decide if a grid cell is refined. This is the case if a grid cell contains more than 8 DM particles to make sure that the cells after the refinement still contain a particle on average ($N_p \times m_p \times (\Delta x)^{-3l}$ with Δx being the spatial resolution of the base grid, N_p the number and m_p the mass of particles in a cell). For hydrodynamical runs, it depends on the baryonic content whether a cell is refined or not. For this, the gas density of a cell has to exceed $\frac{\Omega_b}{\Omega_m - \Omega_b} \times N_p \times m_p \times (\Delta x)^{-3l}$. Since the resolution of the grid moves along with the flow of the particles, Ramses is said to be a *pseudo-Lagrangian code*.

The dynamics of star formation has already been discussed in section 1.9. A characteristic parameter of this process is the Jeans length. The fact that Ramses might not resolve this length scale has to be treated carefully. In order to avoid *spurious fragmentation*, a thermal pressure floor of the form

$$T_P = T_{2,\mathrm{star}} \left(\frac{n_\mathrm{H}}{n_*} \right)^{g_\mathrm{star}-1} \tag{1.57}$$

is added to each cell. This guarantees that the temperature in a cell does not drop below a value where the Jeans length is not resolved anymore. The numerical inaccuracies would lead to spurious fragmentation and to an overestimation of SF. A resolution criterion in the form of

$$\lambda_\mathrm{J} \geq 4\Delta x \tag{1.58}$$

was found by Truelove et al. (1997) and states that the Jeans length needs to be resolved by at least four resolution elements of size Δx.
However, the resolution has to be limited in order to avoid catastrophic refinement which can occur in fast cooling regions. There, the refinement criterion is always satisfied and the code would start to increase the resolution at early times. To counteract the expansion of the universe and to keep consequently the physical size of the grid roughly constant, the refinement of the grid is triggered at certain times. Further, there are a couple of other things that are used within Ramses that influence the baryonic physics. The cooling module includes several options such as cooling, heating, Compton cooling, Compton heating and metal cooling. Also, an approximate method for the metal cooling has been implemented.
Different feedback mechanisms are implemented in Ramses. The *kinetic feedback* from SNII is taken into account by considering stellar particles exceeding an age threshold. In that, the particle is removed from the simulation and a blast wave is travelling across the grid having its center in the original position of the stellar particle. The blast wave is changing the energy budget in each cell as it is propagating through space leading to an update of all hydro quantities within a certain radius. This includes the update of the density, velocity, pressure, and metallicity.

The *thermal feedback* is considered as well. The effects from SNII, SNIa, and winds are modelled. It takes into account that the stellar particles suffer from a loss of mass and thermal energy and return metals as well during their lifetimes. The temperature of the involved cells is updated along with density and metallicity.
The cooling in Ramses uses a primordial cooling involving the species HI, HII, HeI, HeII, HeIII and e^-. Metal cooling is taken into account by an effective term. Assuming the abundances of the different species are in equilibrium, the cooling and heating rate are calculated.

The initial conditions are generated using the code Music (MUlti-Scale Initial Conditions) (Hahn and Abel, 2011) based on the assumed cosmology.

For this thesis, Ramses is modified for different purposes. The first part of the thesis contains a comparison of H_2 and star formation models, specifically the standard recipe for star formation (see equation 1.35) which is set by default in Ramses. In a second version, this SF law has been removed and exchanged by the H_2–based star formation recipe. This required adding a new separate chemistry routine for a dynamical H_2 formation. All these modifications are discussed in detail in Tomassetti et al. (2015).

For the second part of the thesis, a new version of the Ramses code is used that has been released in 2017. Also, the particle fields were changed. The particles in the first version of the code distinguished stars and dark matter particles by the *epoch*-field. Initial particles had a value of -1, whereas stellar particles found to have a value of $\neq -1$. In the upgraded version of 2017, the particles were distinguished by the new field *family*, that assigns different values to the different families like *DM*, *Star*, *Clouds*, and *Debris*.

Further, for the second part of the thesis, the chemistry has been extended by three carbon species (C, C^+ and CO). The star formation recipe is unchanged with respect to the first part of the thesis. In both versions, the contribution of Lyman-Werner photons to the UV field is calculated and stored.

1.12 Aim of this thesis

From a theoretical perspective, star formation and galaxy evolution is complicated to describe as there are many different processes like chemistry, stellar feedback, and gas dynamics involved. The fact that they take place on different time and length scales is an additional challenge. In the previous sections, we have seen which single ingredients are used to put the big picture together. And yet, after so many years of research, there are still a lot of open questions.
This thesis wants to address two questions that are essential for our understanding of galaxy formation and evolution:

- Which role does molecular gas play on cosmological scales?

The first part of the thesis focuses on this particular question. It incorporates a detailed resolution study which enables us to discuss the numerical effects in simulations. Further, different models for the molecular gas formation are compared among each other and with observations. Also, the molecular gas abundance in a cosmological volume is considered.

However, there is a big difference between the theoretical and observational approach to molecular gas. As previously mentioned, the H_2 content of an observed object is inferred from its luminosity of a molecule (CO), atom (C), or ion (C^+) that is used as tracer. In most simulations within a cosmological box, these tracers are not modelled and, in the end, no luminosity is calculated. Knowing the luminosity of a simulated object would enable a better comparison of simulations and observations. Before starting the calculations for the luminosity, the abundances of the different tracers have to be known. Thus, the second question this thesis will investigate:

- Can we make the first step to bridge observations and theory by calculating the abundances of H_2, C, C^+, and CO at the same time in a simulation with a new subgrid model?

However, the calculation of the luminosity requires further modelling that goes beyond this thesis. There are many more things that need to be considered for this, e.g. the dependencies on the temperature and velocity of the different gas components and dust composition. State-of-the art simulations that include the carbon chemistry need a very high resolution. However, we will develop a subgrid model for simulations with resolutions of tens to hundreds of parsecs to provide a suitable code for a cosmological run in order to study e.g. CO luminosity functions, evolution and dependencies of the CO-to-H_2 conversion factor α_{CO} and $L_{[\mathrm{CII}]}$-SFR-relation in the future.

CHAPTER 2

A comparison of H_2 formation models at high redshift[1]

2.1 Introduction

Observations of molecular gas at high redshift (see e.g. Carilli and Walter, 2013, for a review) are shaping our knowledge of the early phases of galaxy formation. To fully appreciate their implications, it is vital to develop a theoretical framework within which the experimental findings can be interpreted. However, ab initio simulations of galaxy formation generally do not resolve the spatial scales and densities (nor capture the physics) that characterise molecular clouds in the interstellar medium, and therefore fall short of modelling the molecular content of galaxies. The need for more sophisticated models is therefore becoming increasingly important, particularly with the advent of the Atacama Large Millimeter/submillimeter Array (ALMA) which has enabled detections of molecular-gas reservoirs at redshifts as high as $z \sim$ 4-7 (e.g. Riechers, Capak et al., 2010; Capak et al., 2015; Maiolino et al., 2015; Decarli, Walter, Aravena et al., 2016; Bothwell et al., 2017; Santini, Merlin et al., 2019).

Opening a window on the molecular Universe also motivates new theoretical efforts to gain insight into how galaxies grow their stellar component. This requires developing a coherent picture that links molecular gas in the turbulent interstellar medium (ISM) to the various feedback processes that regulate the supply of gas available to form stars. Stellar nurseries in the Milky Way appear to be associated with dusty and dense molecular clouds. Spatially resolved observations of nearby galaxies show that the surface density of star formation (SF) better correlates with the surface density of molecular gas than with the total gas density (e.g. Wong and Blitz, 2002; Kennicutt, Calzetti et al., 2007; Leroy et al., 2008; Bigiel, Leroy, Walter, Brinks, de Blok, Madore et al., 2008). A possible interpretation of these findings is that the presence of molecular material is necessary to trigger SF (Krumholz and McKee, 2005; Elmegreen, 2007; Krumholz, McKee and Tumlinson, 2009), although other viewpoints are also plausible. One possibility, advocated by Krumholz, Leroy and McKee (2011) and Glover and Clark (2012b), is that H_2 and SF are spatially correlated due to the ability of the gas to self shield from interstellar ultraviolet (UV) radiation. That SF primarily takes place in molecular clouds would, in that case, be coincidental rather than a consequence of some fundamental underlying relation between H_2 and SF.

In numerical simulations, the two scenarios generate different galaxies: H_2-regulated SF is delayed in the low-metallicity progenitors of a galaxy where dust and central gas densities are too low to activate an efficient conversion of HI into H_2 (Kuhlen, Krumholz et al., 2012; Kuhlen, Madau and Krumholz, 2013;

[1] This chapter has been published in Schäbe et al. (2020).

Jaacks, Thompson and Nagamine, 2013; Thompson et al., 2014; Tomassetti et al., 2015). The resulting galaxies are thus characterized by lower stellar masses, younger stellar populations, and a smaller number of bright satellites (Tomassetti et al., 2015). In addition, the fact that the energy due to stellar feedback is injected at different locations gives rise to different galaxy morphologies (Tomassetti et al., 2015; Pallottini, Ferrara, Gallerani et al., 2017).

In the ISM, H_2 primarily forms due to the catalytic action of dust grains and is destroyed by resonant absorption of photons in the Lyman and Werner (LW) bands. This is why H_2 is abundant in the densest and coldest regions of the ISM where far-UV radiation is heavily attenuated (Draine, 1978; Hollenbach and McKee, 1979; van Dishoeck and Black, 1986; Black and van Dishoeck, 1987; Draine and Bertoldi, 1996; Sternberg, 1988). The main difficulty in tracking molecular gas within galaxy formation models is the huge dynamic range between the scales that tidally torque galaxies and those that regulate the turbulent ISM and on which SF and stellar feedback take place.

One way to overcome this limitation is to use empirical laws inferred from observations in order to predict the abundance of molecular gas within galaxies. For example, the ratio between the surface densities of molecular and atomic hydrogen is found to scale quasi-linearly with the interstellar gas pressure in the mid-plane of disc galaxies (Wong and Blitz, 2002; Blitz and Rosolowsky, 2004; Blitz and Rosolowsky, 2006; Leroy et al., 2008), a fact that has been exploited to develop semi-analytic models (SAMs) of galaxy formation (Dutton and van den Bosch, 2009; Obreschkow, D. Croton et al., 2009; Obreschkow and Rawlings, 2009a; Fu, Guo et al., 2010; C. D. P. Lagos et al., 2011; Fu, G. Kauffmann et al., 2012; Popping, Somerville and Trager, 2014; Popping, Behroozi and Peeples, 2015; Somerville, Popping and Trager, 2015; Lacey et al., 2016; Stevens, D. J. Croton and Mutch, 2016; C. d. P. Lagos, Tobar et al., 2018) and numerical simulations (Murante, Monaco, Giovalli et al., 2010; Murante, Monaco, Borgani et al., 2015; Diemer et al., 2018) that take H_2 into account.

A second possibility for tracking H_2 is to consider a number of simplifying assumptions under which the coupled problems of radiative transfer and H_2 formation in the ISM can be solved analytically (see, e.g. Krumholz, McKee and Tumlinson, 2008; Krumholz, McKee and Tumlinson, 2009; McKee and Krumholz, 2010). In this case, the processes regulating the molecular gas fraction are assumed to be in local equilibrium and the resulting H_2 abundance depends only on the column density and metallicity of the gas. Model predictions for the HI-H_2 transition profiles appear to be consistent with observations in external galaxies with different metallicities (Fumagalli, Krumholz and Hunt, 2010; Bolatto, Leroy et al., 2011; Wong, Xue et al., 2013). Equilibrium models have been widely used to predict the molecular content of galaxies in the semi-analytic framework (Fu, Guo et al., 2010; C. D. P. Lagos et al., 2011; Fu, G. Kauffmann et al., 2012; Krumholz and Dekel, 2012; Somerville, Popping and Trager, 2015) and in numerical simulations of small-to-intermediate cosmological volumes (Kuhlen, Krumholz et al., 2012; Kuhlen, Madau and Krumholz, 2013; Jaacks, Thompson and Nagamine, 2013; Hopkins et al., 2014; Thompson et al., 2014; C. d. P. Lagos, Crain et al., 2015; Davé, Thompson and Hopkins, 2016). A more complex equilibrium model in which the dust abundance and grain-size distribution evolve with time has been recently employed in simulations of an isolated disc galaxy (Chen et al., 2018).

The third option on the market is to model the out-of-equilibrium evolution of the H_2 abundance. The main motivation for doing this is that the formation of H_2 on dust-grains can be a slow process and the chemical rate equations reach equilibrium only if the ISM presents favourable conditions (e.g. high dust content and long dynamical timescales). As a result, the equilibrium models described above may over-predict the abundance of H_2 in certain scenarios. To overcome this problem, one can directly integrate the system of chemical rate equations without resorting to approximate equilibrium solutions. This approach, however, requires accounting for the complex interplay between velocity and density in a turbulent medium that ultimately determines the column density of the gas and dust. While such a line of attack characterizes state-of-the-art simulations of small ISM patches (see, e.g., Seifried et al., 2017, and

references therein), it cannot yet be fully implemented in cosmological simulations of galaxy formation as they do not yet resolve the relevant length, time and density scales. The simplest approach is to solve the chemical rate equations after coarse-graining them at the level of the single resolution elements and introduce a clumping factor in the H_2 formation rate to account for unresolved density fluctuations. The best possible spatial resolution is then achieved by focusing on idealized (Pelupessy, Papadopoulos and van der Werf, 2006; Robertson and Kravtsov, 2008; Pelupessy and Papadopoulos, 2009; Hu et al., 2016; Richings and Schaye, 2016; Lupi, Bovino et al., 2018) or cosmological simulations of individual galaxies (Gnedin, Tassis and Kravtsov, 2009; Feldmann, Gnedin and Kravtsov, 2011; Christensen et al., 2012; Katz et al., 2017; Pallottini, Ferrara, Gallerani et al., 2017; Nickerson, Teyssier and Rosdahl, 2018; Lupi, Volonteri et al., 2019; Pallottini, Ferrara, Decataldo et al., 2019). Alternatively, physics at the unresolved scales can be dealt with by introducing a sub grid model that takes into account the probability distribution of local densities and the temperature-density relation obtained in high-resolution simulations of the turbulent ISM. In this case, a 1D slab approximation is used to associate an optical depth to each microscopic density (see Tomassetti et al., 2015, for details). Such a model yields H_2 fractions (as a function of the total hydrogen column density) that are in excellent agreement with observations of the Milky Way and the Magellanic Clouds.

Given this variety of techniques, it is worthwhile identifying which regimes, if any, the outputs of simulations based on empirical, equilibrium and non-equilibrium models for the H_2 abundance give consistent results. For instance, Feldmann, Gnedin and Kravtsov (2011) show that a tight relation between the H_2 fraction and the ISM pressure emerges naturally in simulations where the chemical rate equations are integrated without assuming local equilibrium. The slope and the amplitude of the relation depend sensitively on the local ISM properties, in particular on the dust-to-gas ratio. When the conditions of the ISM are tuned to those of the solar neighbourhood, the resulting correlation closely matches that observed in local galaxies. Moreover, Krumholz and Gnedin (2011) find that the equilibrium model presented in Krumholz, McKee and Tumlinson (2009) agrees well with time-dependent calculations for a wide range of UV intensities if the H_2 abundance is coarse-grained on scales of ≈ 100 pc and the ISM metallicity is above $0.01\,Z_\odot$. For lower values of Z, however, the agreement rapidly deteriorates. On the other hand, Mac Low and Glover (2012) find that equilibrium models do not, in general, reproduce the results of simulations of the turbulent, magnetized ISM when coarse-grained on scales of $\sim 1 - 10$ pc. These results suggest that the level of agreement or disagreement between the different approaches depends on the length-scales over which the comparisons are made.

In this paper, we use a suite of cosmological, hydrodynamical simulations to investigate similarities and differences between the three model prescriptions for H_2 formation. We focus on the molecular content of high-redshift galaxies, similar to those that can be detected with ALMA. In previous work, equilibrium H_2 models have been employed in simulations with widely different spatial resolutions, ranging from the parsec to kiloparsec scales. In order to provide a benchmark for future studies, we therefore investigate how predictions for the molecular content of galaxies are influenced by the spatial resolution of the simulations, focusing on both equilibrium and non-equilibrium models. To check the reliability of the H_2 models, we also compare the global properties of our simulated galaxies against observations of high-redshift systems, emphasizing differences between the various H_2-formation schemes.

The paper is organized as follows. In section 2.2, we introduce the semi-empirical, equilibrium and non-equilibrium models used to track the abundance of H_2 in our simulations. Our numerical setup is described in section 2.3 and finite spatial resolution effects are investigated in section 2.4. We compare our numerical results with a series of observational data in section 2.5. Finally, we conclude providing a summary of our main results in section 2.6.

2.2 Modelling molecular hydrogen

As mentioned above, modelling the H_2 chemistry in simulations of galaxy formation is extremely challenging because it requires simultaneously resolving the very disparate temporal and spatial scales relevant for molecular cloud dynamics and galaxy evolution. An exact treatment of all relevant processes is clearly impossible, yet progress continues to be made both in SAMs and hydrodynamical simulations. In this work, we consider three approximate methods that have been previously presented in the literature and are representative of entire classes of models. This section provides an overview of their most important aspects as well as the relevant details of their implementation.

2.2.1 The semi-empirical model (PBP)

As an example of how we can use empirical laws inferred from observations to associate an H_2 mass to a simulated dark-matter halo, we use the method presented by Popping, Behroozi and Peeples (2015, hereafter PBP). The model takes, as input, a halo mass and redshift to which a stellar mass and an instantaneous star-formation rate (SFR) are assigned using subhalo abundance matching (see Behroozi, Wechsler and Conroy, 2013, for details). Stars and gas are assumed to be distributed according to an exponential profile. The scale length of the stellar disc is chosen according to the empirical relation of van der Wel et al. (2014), while the size of the gaseous disc is scaled-up by a factor of 2.6. The relative abundance of HI and H_2 is then determined using an empirical scaling with the mid-plane pressure (Blitz and Rosolowsky, 2006) for an assumed cold gas mass. The latter (and the final results for the H_2 mass) are determined iteratively by requiring that the corresponding H_2 surface density yields a SFR equivalent to that implied by the observed $\Sigma_{H_2} - \Sigma_{SFR}$ relation (as given by Bigiel, Leroy, Walter, Brinks, de Blok, Madore et al., 2008). By construction, this method yields galaxy gas masses that are consistent with observed SFRs.

2.2.2 The equilibrium model (KMT)

Locally, the H_2 abundance is determined by the competing actions of molecule formation on dust-grain surfaces and dissociation due to the absorption of LW photons. Provided certain assumptions are made, the local equilibrium abundance of H_2 can be evaluated analytically (Krumholz, McKee and Tumlinson, 2008; Krumholz, McKee and Tumlinson, 2009; McKee and Krumholz, 2010). The calculation assumes a spherical molecular cloud shrouded by an isotropic radiation field of LW photons, and an ISM that is in a two-phase equilibrium between the cold and warm neutral mediums. The dust abundance is assumed to scale linearly with the gas metallicity. In this scenario, both the UV field intensity and H_2 fraction depend only on the local column density and metallicity of the gas (see McKee and Krumholz, 2010).

Due to its simplicity, this equilibrium model is commonly employed within SAMs to estimate the relative contributions of atomic and molecular hydrogen to gaseous discs (Fu, Guo et al., 2010; C. D. P. Lagos et al., 2011; Fu, G. Kauffmann et al., 2012; Krumholz and Dekel, 2012; Somerville, Popping and Trager, 2015). The model can also be implemented in hydrodynamical simulations where, instead, simple estimates of the instantaneous LW radiation field can be used to deduce the local equilibrium H_2 fraction (Gnedin and Kravtsov, 2011; Kuhlen, Krumholz et al., 2012; Kuhlen, Madau and Krumholz, 2013; Jaacks, Thompson and Nagamine, 2013; Hopkins et al., 2014; Thompson et al., 2014; Baczynski, Glover and Klessen, 2015; C. d. P. Lagos, Crain et al., 2015; Tomassetti et al., 2015; Davé, Thompson and Hopkins, 2016).

2.2.3 The dynamical model (DYN)

Our final model tracks directly the time-dependent formation and destruction of H_2 within the resolution elements of our simulations (see Tomassetti et al., 2015, for further details).

State-of-the-art numerical simulations of galaxy formation typically reach spatial resolutions of the order of 50 to 100 pc, comparable to sizes of giant molecular clouds (GMCs). Numerical and observational studies of the turbulent ISM, however, indicate that GMCs are rich in substructure on much smaller scales. This is often accounted for in cosmological simulations using a gas clumping factor, C_ρ – an approximation that neglects the complexity of substructure as well as their temperature-density correlations, which may alter H_2 formation and destruction rates. For that reason, we model the unresolved sub-grid density distribution using a mass-weighted log-normal probability function (Kainulainen et al., 2009; Schneider et al., 2013), whose parameters can be determined once a clumping factor has been specified. We assume $C_\rho = 10$ in all of our simulations since it has been shown to give reasonable results in simulations involving H_2 (Gnedin, Tassis and Kravtsov, 2009; Christensen et al., 2012).

We adopt a temperature-density relation for unresolved clumps consistent with results from simulations of the turbulent ISM (Glover and Mac Low, 2007a). These simulations suggest that the formation of H_2 primarily takes place in dense regions where temperatures remain $\lesssim 200$ K. In this implementation, we neglect the collisional destruction of H_2. Note that we adopt the same value for C_ρ in simulations with linear spatial resolutions that differ up to a factor of four (see section 2.3.5 for further details). At first sight, this choice might appear to be unphysical as, in the limit of infinite resolution, every small clump should be resolved and $C_\rho \to 1$. Therefore, one expects C_ρ to decrease as the spatial resolution of the simulations increases. In Davé, Thompson and Hopkins (2016), for instance, C_ρ is assumed to scale proportionally to the minimum comoving gravitational softening length. In their implementation of the KMT model, the clumping factor assumes the values of 30, 15 and 7.5 in simulations with softening lengths of 0.5, 0.25 and 0.125 h^{-1} kpc, respectively. Note that, continuing this scaling, C_ρ would approach unity if the softening length is further reduced to $\simeq 25$ pc. However, in molecular clouds, most of the clumping takes place at 'microscopic' scales compared with the size of our simulation cells.

Figure 7 in Micic et al. (2012) shows the time evolution of the clumping factor within a simulated molecular cloud in a box of 20 pc (i.e. smaller than our smallest grid cell): C_ρ is always of order 10 for turbulent rms velocities of a few km s^{-1} and even substantially higher in the presence of compressive forcing. Therefore, using a clumping factor that does not depend on resolution (as we do) corresponds to assuming that density fluctuations are much more prominent on microscopic length-scales than on scales that are comparable to size of our smallest grid cells.

Recently, Lupi, Bovino et al. (2018) followed the evolution of a single galaxy at $z = 3$ for 400 Myr using a spatially and temporally varying clumping factor in a simulation with softening lengths of 80, 4 and (up to) 1 pc, for dark matter, stars and gas, respectively. In this case, C_ρ is linked to the subgrid model for the turbulent ISM and assumes median values around 20.

Our assumptions lead to the following system of coupled differential equations describing the formation and destruction of H_2:

$$\frac{\mathrm{d}\langle n_{\mathrm{H}_2}\rangle}{\mathrm{d}t} = \langle \mathcal{R}_f(T)\, n_{\mathrm{HI}}\, n_{\mathrm{H}}\rangle - \langle G\,\kappa\,\Phi_{\mathrm{s}}\, e^{-\tau}\, n_{\mathrm{H}_2}\rangle\,, \tag{2.1}$$

$$\frac{\mathrm{d}\langle n_{\mathrm{HI}}\rangle}{\mathrm{d}t} = -2\,\frac{\mathrm{d}\langle n_{\mathrm{H}_2}\rangle}{\mathrm{d}t}\,, \tag{2.2}$$

with

$$\langle n_{\mathrm{HI}}\rangle + 2\,\langle n_{\mathrm{H}_2}\rangle = \langle n_{\mathrm{H}}\rangle\,. \tag{2.3}$$

Here, the brackets $\langle \dots \rangle$ indicate averages taken over the substructure present within a single resolution element of the simulations. They are computed by integrating over the mass-weighted probability density function (PDF) and temperature-density relation specified by the sub-grid model for the turbulent ISM. The function $\mathcal{R}_f$ controls the formation rate of H_2 on dust grains. The parameter G is the unshielded interstellar UV radiation flux (in Habing units; not to be confused with Newton's constant); κ is the photo-dissociation rate of H_2; Φ_s is the H_2 self-shielding function, and $\tau = \sigma_d N_H$ is the dust optical depth in the LW band with column density N_H and the photon cross section σ_d, corresponding to a column density $N_H = N_{HI} + 2N_{H_2}$. Dust abundances are assumed to scale linearly with gas metallicity as $M_{dust}/M_{gas} = 0.008\ (Z/Z_\odot)$, where $Z_\odot = 10^{-2}$ (Draine, Dale et al., 2007).

2.3 Numerical methods

2.3.1 Simulation setup

We use the adaptive-mesh-refinement (AMR) code Ramses (Teyssier, 2002) to run a suite of hydrodynamical simulations. The code uses a second-order Godunov scheme to solve the hydrodynamic equations, while trajectories of DM and stellar particles are computed using a multigrid Particle-Mesh solver.

We consider a cubic periodic box with a comoving sidelength of $12\,h^{-1}$ Mpc and assume a cosmological model consistent with the Planck Collaboration (2014) results: $\Omega_\Lambda = 0.692$, $\Omega_m = 0.308$, $\Omega_b = 0.0481$ and a present-day value of the Hubble parameter of $H_0 = 100\,h$ km s^{-1} Mpc^{-1} with $h = 0.678$.

Initial Conditions (ICs) are generated using the Music code (Hahn and Abel, 2011) at different spatial resolutions but use the same phases and amplitudes for mutually-resolved modes. Our high-resolution ICs are imposed on a grid of 512 cells per dimension, corresponding to a comoving Lagrangian spatial resolution of $\approx 23.4\,h^{-1}$ kpc. In all cases, the ICs are set at redshift $z = 99$ and assume a ΛCDM model with primordial spectral index $n_s = 0.9608$ and a linear rms density fluctuation in $8\,h^{-1}$ Mpc spheres of $\sigma_8 = 0.826$.

For each resolution, we carry out two types of simulations: one following the evolution of collisionless DM alone and another following the co-evolution of DM and baryons. The former is used to establish a refinement strategy and to identify the maximum level of refinement achieved during the simulation. In principle, grid refinements are based on the standard 'quasi-Lagrangian' criterion, i.e. they are triggered if the number of DM particles in a cell exceeds eight or if the baryonic mass is $> 8\,m_{DM}\,\Omega_b/(\Omega_m - \Omega_b)$. However, in order to prevent runaway refinements at early times, we demand that new levels are only triggered at certain times as described, e.g., in C. Scannapieco et al. (2012). This ensures that the grid resolution in physical units stays approximately constant (although not continuously but in a series of distinct steps). Moreover, in the hydrodynamic runs, we make sure that the maximum level of refinement for the DM component does not exceed that reached in the DM-only simulations at the same cosmic time. On the other hand, the grid for the gas component is allowed to reach one or two additional levels (see section 2.3.5 for details).

We adopt an equation of state with polytropic index $\gamma = 5/3$ for the gas component. To avoid spurious fragmentation, thermal pressure is added where needed by increasing the gas temperature so that the Jeans length is resolved with at least four grid cells (Truelove et al., 1997; Teyssier, Chapon and Bournaud, 2010).

2.3.2 Star formation and stellar feedback

Simulations of galaxy formation do not resolve the time and length-scales on which SF occurs in the ISM. Therefore, SF needs to be treated in a simplified way on scales comparable with the spatial resolution.

It is reassuring that the combined action of this rather crude modelling and of stellar feedback in the simulations leads to the emergence of regularities on kpc scales e.g. the Kennicutt-Schmidt relation (Schmidt, 1959; Kennicutt, 1989; Kennicutt, 1998) and global gas depletion times that are in good agreement with observations (see e.g. Agertz and Kravtsov, 2016; Orr et al., 2017; Semenov, Kravtsov and Gnedin, 2018, and references therein).

Following a standard procedure, we impose that SF only takes place within gas cells that i) are part of a convergent flow and ii) have a temperature $T < 10^4$ K. However, we use two different approaches to model SF. In the simulations based on the KMT model, we impose that SF only takes place where the number density of hydrogen atoms exceeds $n \sim 1\,\mathrm{cm}^{-3}$. The selected gas elements with mass density $\rho_{\rm gas}$ are then converted into star particles according to a stochastic Poisson process with density

$$\dot{\rho}_{\rm SF} = \epsilon \frac{\rho_{\rm gas}}{t_{\rm ff}} \,, \tag{2.4}$$

where $t_{\rm ff} = \sqrt{3\,\pi/(32\,G_{\rm N}\,\rho_{\rm gas})}$ is the free-fall time of the gas (here $G_{\rm N}$ denotes Newton's gravitational constant) and $\epsilon = 0.05$ is an efficiency parameter.

On the other hand, in the runs carried out with the DYN model, we link SF directly to the local H_2 mass density through the relation

$$\dot{\rho}_{\rm SF} = \epsilon \frac{\rho_{\rm H_2}}{t_*} \,, \tag{2.5}$$

without imposing any criterion on n. Note that, in this case, most SF naturally takes place at high n since H_2 formation is inefficient at low particle densities. For example, in our runs, nearly 96 per cent of SF occurs in cells with $n > 2.5\,\mathrm{cm}^{-3}$ and 55 per cent takes place where $n > 100\,\mathrm{cm}^{-3}$. Nevertheless, some H_2-rich cells inevitably fall short of actual GMC densities ($n \gtrsim 100\,\mathrm{cm}^{-3}$). For this reason, we define t_* as the minimum of a cell's free-fall time and that of a uniform cloud of density $n = 100\ \mathrm{cm}^{-3}$ (see Gnedin, Tassis and Kravtsov, 2009, for details).

All simulations include supernova type II feedback and the associated metal enrichment, as well as cooling from H, He and metals (Rasera and Teyssier, 2006). The impact of cosmic reionization is modelled using the spatially uniform UV background derived in Haardt and Madau (2012). Following Kuhlen, Krumholz et al. (2012) and Kuhlen, Madau and Krumholz (2013) and Tomassetti et al. (2015), we set a metallicity floor of $10^{-3}\ \mathrm{Z}_\odot$ at $z = 9$. This approximately compensates for chemical enrichment from early generations of unresolved SF (e.g. Wise et al., 2012) and seeds the initial formation of H_2. Self-shielding of dense gas is approximated by exponentially suppressing UV heating in cells where the gas density exceeds $n_{\rm H} \simeq 0.014\ \mathrm{cm}^{-3}$ (Tajiri and Umemura, 1998).

2.3.3 Local UV radiation field

The KMT and DYN models need an estimate of the intensity of LW radiation in each resolution element of the simulations. We compute this quantity following the approach of Tomassetti et al. (2015). We model each stellar particle as a population of stars with masses distributed according to a Kroupa (2001) initial mass function and with luminosities consistent with the Starburst99 templates (Leitherer et al., 1999). We then calculate the total UV luminosity (in the LW band) of the stellar particles as a function of their ages. Finally, we propagate the photons isotropically from the stellar particles assuming that the ISM transitions abruptly from optically thin to thick at some characteristic length-scale, $r_{\rm LW}$, which corresponds to the size of a few cells. Note that this includes a geometric dilution following the inverse-square law (more sophisticated but time consuming approaches solve the radiative transfer problem for UV radiation on the fly assuming a reduced speed of light, e.g. Gnedin and Kravtsov, 2011; Lupi, Bovino

et al., 2018).

2.3.4 Haloes and galaxies

In order to identify galaxies and their host haloes in our simulations, we use the Amiga Halo Finder code (AHF; Gill, Knebe and Gibson, 2004; Knollmann and Knebe, 2009). We first locate spherical regions with mean density equal to $200\,\rho_{\rm c}(z)$, where $\rho_{\rm c}(z) \equiv 3\,H^2(z)/(8\,\pi\,G_{\rm N})$ is the critical density of the Universe. We then remove unbound particles by iteratively clipping those whose velocities exceed 1.5 times the local escape speed, $v_{\rm esc} = \sqrt{2|\phi|}$, where ϕ is the local gravitational potential. Note that we consider all matter components to identify the haloes (for instance, the thermal energy of the gas is also taken into account in the unbinding procedure). In what follows, we characterise the haloes based on their position (we associate the halo centre with the densest spot), the total mass of the bound material ($M_{\rm h}$), and the maximum distance of a bound mass element from the halo centre ($R_{\rm h}$).

It proves useful to track the evolution of particular haloes through simulation snapshots, or to cross-match them between simulations adopting different SF-recipes or H_2 models. This is done using the DM-component only. For simulations with the same initial resolution (e.g. the same $l_{\rm initial}$), this is trivially carried out by matching particle IDs. For simulations with different $l_{\rm initial}$, we associate a lower-resolution counterpart to each highly resolved halo by minimizing an objective function d that depends on the halo positions ($\boldsymbol{x}$), masses and the maximum value of their rotation curves ($v_{\rm max}$):

$$d = \left[(\boldsymbol{x}_{l_1} - \boldsymbol{x}_{l_2})^2/R_{\rm h}^2(l_1)\right]^{2/3} + \log\left(M_{{\rm h},l_1}/M_{{\rm h},l_2}\right)^{2/3} + \log\left(v_{{\rm max},l_1}/v_{{\rm max},l_2}\right)^{2/3}\,, \tag{2.6}$$

where the subscripts l_1 and l_2 refer to the higher and lower values of $l_{\rm initial}$, respectively (Angulo et al., 2017).

We assume that each halo hosts a central galaxy that occupies the spherical region of radius $R_{\rm gal} = 0.1\,R_{\rm h}$ (e.g. C. Scannapieco et al. (2012), but see Stevens, Martig et al. (2014)). The stellar and gas mass of the resulting galaxies are rather insensitive to the precise definition of their outer boundary. Outliers are driven primarily by rare major mergers. In this work, we do not consider satellite galaxies hosted by substructures of the main haloes.

2.3.5 The simulation suite

Table 2.1 summarizes the main characteristics of our hydrodynamic simulations. We adopt a naming convention for the different runs in which the first letter identifies the H_2 model that has been used (K for KMT and D for DYN), followed by a number indicating its Lagrangian refinement level, $l_{\rm initial}$, which varies from 7 for our lowest resolution simulation to 9 for our highest. For two runs, we allow gas cells to refine up to two levels higher than the maximum level attained in the corresponding DM-only simulations. We use the superscript '+' to distinguish these runs from the others.

On top of the simulations listed in Table 2.1, we also build a galaxy catalogue based on the PBP model. This lists the H_2 mass, SFR, and the stellar mass that the semi-empirical model associates to the central galaxies of the haloes extracted from the $D9^+$ run.

Simulation	H_2 model	SF model	$l_{\rm initial}$	$l_{\rm max}$	$\Delta x(z=4)$ pc	$z_{\rm min}$	$M_{\rm DM}$ [$M_\odot$]	M_* [$M_\odot$]
$D9^+$	DYN	H_2	9	16	55	3.6	1.4×10^6	2.0×10^5
D9	DYN	H_2	9	15	110	2.8	1.4×10^6	2.0×10^5
$D8^+$	DYN	H_2	8	15	110	3.8	1.1×10^7	6.3×10^5
D8	DYN	H_2	8	14	220	3.9	1.1×10^7	6.3×10^5
K9	KMT	gas	9	14	110	4.3	1.4×10^6	2.0×10^5
K8	KMT	gas	8	13	220	4.1	1.1×10^7	6.3×10^5
K7	KMT	gas	7	12	440	4.1	8.7×10^7	2.0×10^6

Table 2.1: Main properties of our simulation suite. The first column assigns a name to each run. The names are formed by a letter that denotes the adopted H_2-formation model (second column) and a number that indicates the spatial resolution of the ICs (fifth column). The simulation box of linear size $L_{\rm box}$ (fourth column) is initially divided in $2^{3l_{\rm initial}}$ identical cells. During the evolution, we use the AMR technique to increase the spatial resolution in the high-density regions. The maximum level of refinement achieved by the simulations for the gas component is indicated in the sixth column. Note that a new grid refinement is triggered right after $z = 4$ and therefore the values of $l_{\rm max}$ for the D- and K-series (that have different $z_{\rm min}$) do not seem to match. For this reason, in the seventh column, we report the spatial resolution (in physical units) achieved by the simulations at $z = 4$ (i.e. before the new refinement is triggered). The symbol '+' at the end of the simulation name highlights those runs in which an extra level of refinement is used for the gas component with respect to a normal run. The adopted SF recipe (third column) has a one-to-one association with the H_2 model: the K runs form stars depending on the total gas density while the D ones by the amount of molecular gas (see section 2.3.2 for further details). The minimum redshift reached by the simulations is given in the eighth column. Finally, the minimum masses of the stellar and DM particles are given in the ninth and tenth columns.

2.4 Numerical resolution effects on the H_2 mass

In this section, we investigate how the resulting H_2 mass of the simulated galaxies is affected by the finite spatial resolution of the runs.

2.4.1 Dynamical model

In our dynamical model, the net formation rate of H_2 ultimately depends on the gas density, metallicity and the intensity of UV radiation in each resolution element of the simulations. In Fig. 2.1, we illustrate how the gas cells in the $D9^+$ (blue) and D9 (red) runs populate this space for two particular galaxies at $z = 3.6$. The first one, denoted with the symbol ☆ and displayed in the left block of panels, is the central galaxy hosted by the most massive halo in our simulations at $z = 3.6$, with $M_{\rm h} = 6.1\times10^{11}$ $M_\odot$. The second one, denoted with the symbol ◯ and displayed in the right block of panels, is hosted by a much smaller halo with $M_{\rm h} = 4.4\times10^{10}$ $M_\odot$. For each galaxy, we consider four bins for the UV intensity and we show the scatterplot of the cells in the Z-$n_{\rm gas}$ plane. To improve readability, we only plot one in every 300 cells for the $D9^+$ run and one in every 5 cells for the D9 simulation. The colour and shape of the symbols indicate whether the H_2 fraction in a cell is $f_{H_2} \geq 0.55$ (the dark circles), $0.45 < f_{H_2} < 0.55$ (the diamonds), and $f_{H_2} \leq 0.45$ (the light circles). The green lines represent the loci where the H_2 abundance is in equilibrium (i.e. where the formation rate equals the destruction rate at $f_{H_2} = 0.5$). They are computed assuming the median value of the UV intensity in each panel and a cell size of 59 pc corresponding to the highest refinement level in the $D9^+$ run. Note that they shift towards the right for more intense UV radiation as higher gas densities (at fixed metallicity) are necessary to maintain equilibrium in the presence of an increased destruction rate for the H_2 molecules. For the lowest UV intensities, equilibrium could in principle be reached at relatively small gas densities. However, the

H_2 formation time, $t_{\rm form} = [\mathcal{R}_f(T)\, n]^{-1}$, can be extremely long. In this case, more time is needed to reach equilibrium. For this reason, we use black lines to indicate the loci where $t_{\rm form}$ equals the age of the Universe at $z = 3.6$ (1.7 Gyr). In each panel, we expect to find abundant H_2 only on the right-hand side of both the green and black lines. In other words, the simulations must resolve high-enough densities to produce substantial amounts of H_2. Moreover, the relevant density threshold changes with the local metallicity and UV intensity.

Let us now focus on the gas cells that form the ☆ galaxy. As expected, in both the D9 and D9$^+$ simulations, the numerical resolution elements that contain large H_2 fractions are generally found on the right-hand side of (or around) the green and black lines. In the D9$^+$ simulation, cells with large H_2 fractions are found in all bins of UV intensity. On the other hand, the fact that the ☆ galaxy is less metal enriched in the D9 run (by ~ 0.5 dex) pushes the threshold for copious H_2 formation to higher densities. In consequence, H_2 fractions above 0.5 are almost exclusively found in the densest cells (that typically are also associated with larger UV intensities). To emphasize this difference, in Fig. 2.2 we show the $n_{\rm H}$-T phase diagram of the ISM colour coded by the H_2 fraction. Since the total H_2 content is dominated by the contribution of these dense gas elements, the two simulations give very similar results for the molecular mass of the ☆ galaxy.

The discrepancy between the D9 and D9$^+$ simulations is more extreme for the ◯ galaxy. In this case, the metallicity difference between the two runs is larger ($\sim$ 1 dex) and, even at the largest resolved densities, the D9 simulation contains very few cells in which $f_{\rm H_2} \gtrsim 0.5$ while there are many in the D9$^+$ run. This happens because $t_{\rm form}$ is longer than the age of the Universe for the combinations of densities, metallicities, and UV intensities appearing in the D9 simulation. A substantial difference then appears in the predicted total H_2 mass of the ◯ galaxy in the D9 and D9$^+$ runs.

The two examples discussed above suggest that the earlier onset of SF in higher resolution simulations leads to faster metal enrichment of the ISM which, in turn, boosts the formation rate of H_2. This chain of events is amplified because we link SF to the local H_2 density as explained in section 2.3.2 as no SF and metal enrichment can take place before some H_2 is formed in the first place. Once the metallicity of the ISM is sufficiently large and the timescale for H_2 formation is sufficiently short at the resolved densities, we expect simulations with different resolutions to yield similar H_2 masses. At the high redshifts we are investigating here, this basically implies that the difference in metallicity between the simulations is not too large.

We now proceed to study the impact of the finite spatial resolution on the global galaxy population. The scatter plot in the left panel of Fig. 2.3 displays the relationship between the H_2 mass ($M_{\rm H_2}$) and the total host-halo mass ($M_{\rm h}$) emerging for individual galaxies in the D9$^+$(red), D9 (blue), D8$^+$ (green) and D8 (magenta) simulations. We only consider the central galaxies of haloes with $M_{\rm h} > 10^{10}\, h^{-1}\, {\rm M}_\odot$ at the lowest common redshift of the simulations ($z = 3.9$). To highlight the characteristic trends we represent the running averages (computed in log-log space using a Gaussian filter with a standard deviation of 0.1 dex) with solid lines.

Simulations that achieve the same maximum spatial resolution for the baryonic component (i.e. D8$^+$ and D9) produce very similar results over the entire range of $M_{\rm h}$. The extra refinement level in the ICs of the D9 simulation causes only a minor systematic shift of $M_{\rm H_2}$ towards higher values. On the other hand, the galaxies in the D9$^+$ run contain substantially larger H_2 reservoirs at low and intermediate $M_{\rm h}$. It is only for the most massive haloes (for which we do not plot the running averages as they would be biased low) that the D8$^+$, D9 and D9$^+$ simulations yield consistent values of $M_{\rm H_2}$.

We further explore these trends in the top-right hand panel of Fig. 2.3 where we plot the ratio of the H_2 masses found in the D9 and D9$^+$ simulations for each central galaxy as a function of $M_{\rm h}$ of the host halo (which practically does not change among the various runs). The solid line once again denotes the

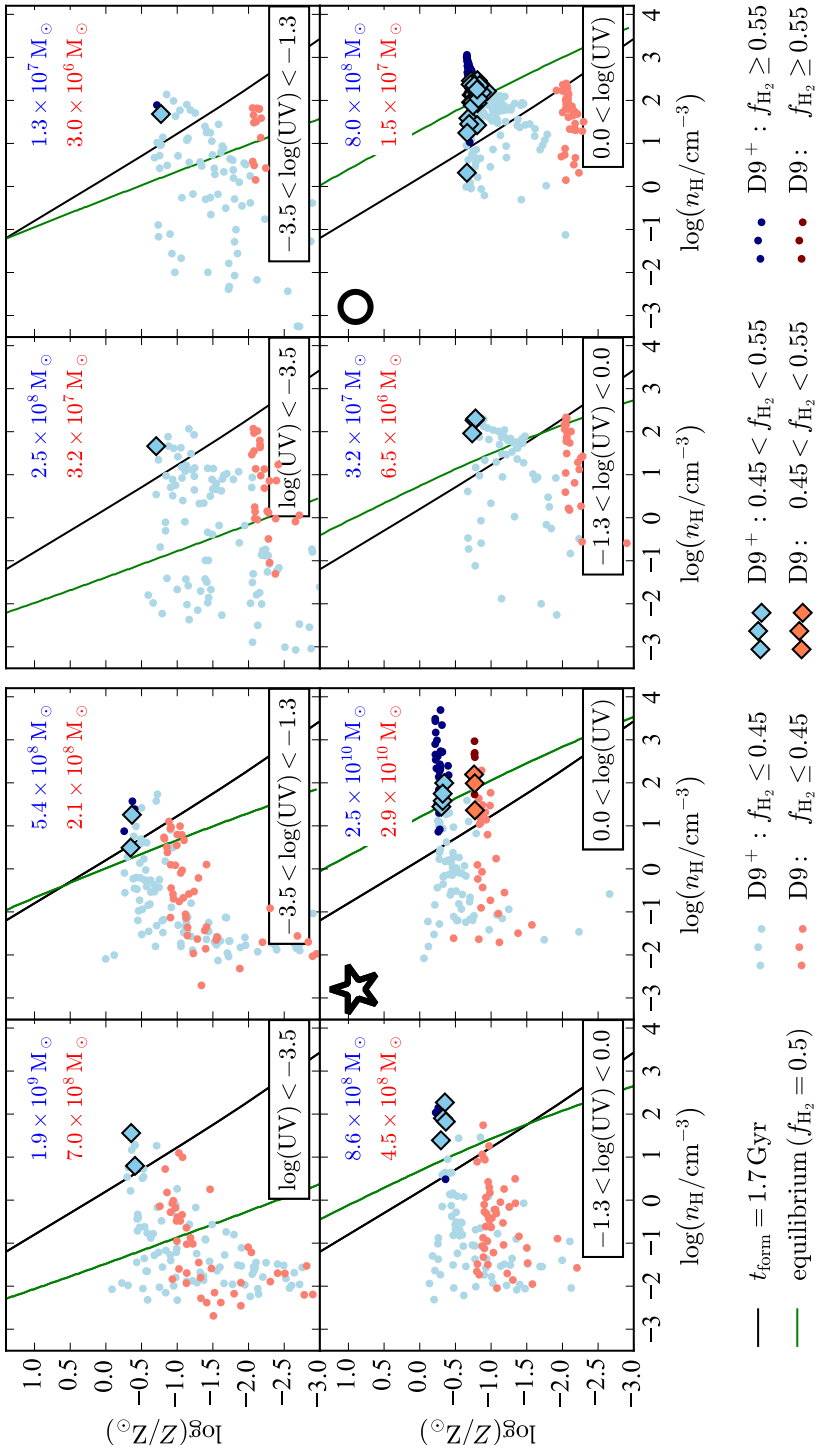

Figure 2.1: Density, metallicity, and intensity of UV radiation for a subsample of the numerical resolution elements found within two simulated galaxies at $z = 3.6$. The four panels on the left-hand side refer to the central galaxy hosted by the most massive halo in our simulations with $M_h = 6.1 \times 10^{11}\,M_\odot$. Those on the right-hand side correspond to the central galaxy hosted by a much smaller halo with $M_h = 4.4 \times 10^{10}\,M_\odot$. The red data points are extracted from the D9 simulation while the blue ones come from the D9$^+$ run. Different symbols indicate the H_2 fraction as described in the legend. The numbers in the top-left corners give the total H_2 mass contained in each panel for the two simulations. The green and black lines mark the loci where the equilibrium H_2 fraction is 0.5 and where the H_2 formation time equals the age of the Universe, respectively (see the main text for further details).

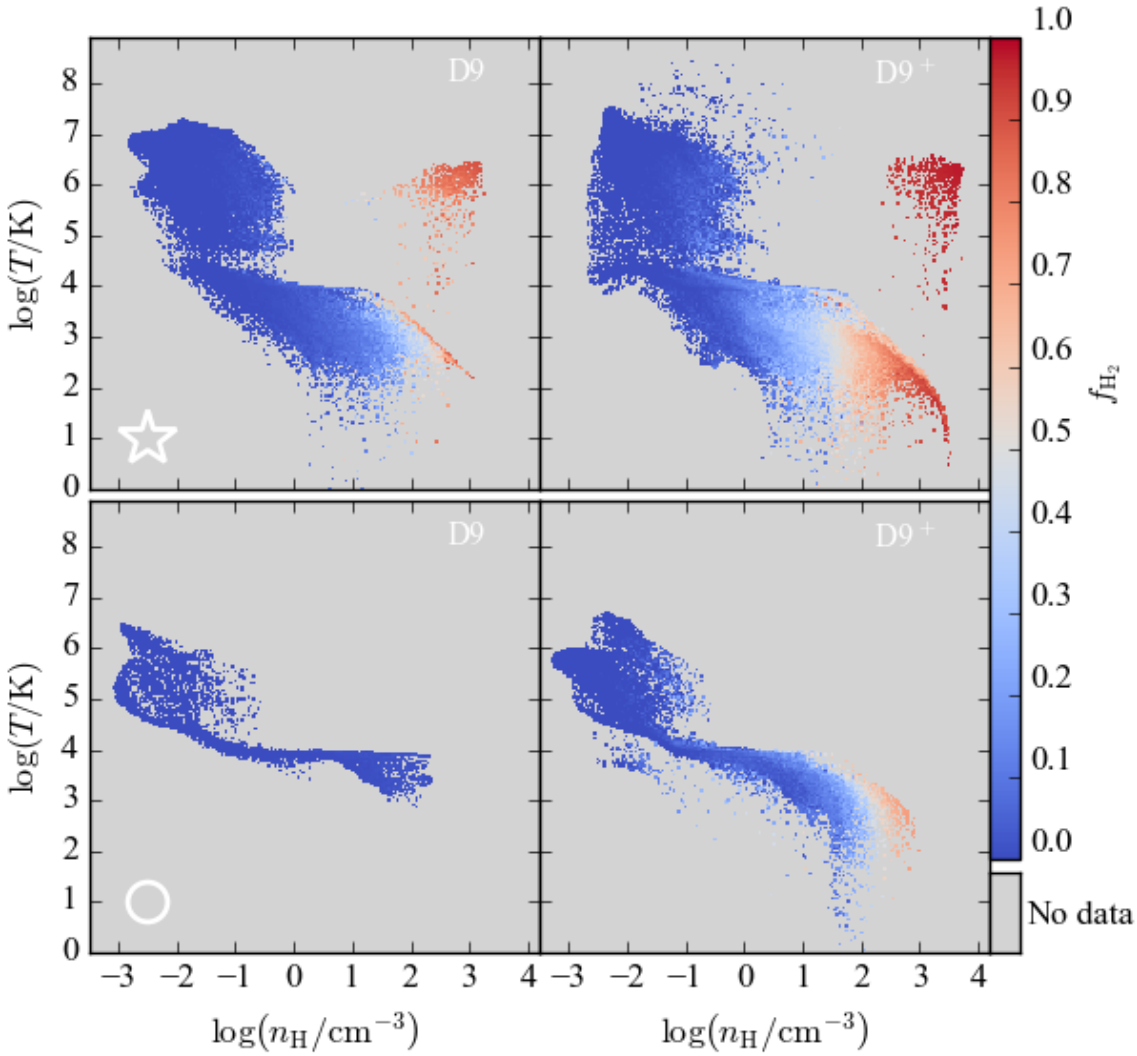

Figure 2.2: Phase diagram in the density-temperature plane colour coded according to the H_2 density for the gas cells of the galaxies considered in Fig. 2.1. In the top panels, the small cloud of H_2-rich gas at high densities and temperatures represents elements that have been recently influenced by stellar feedback and that are located around the centre of the galaxy.

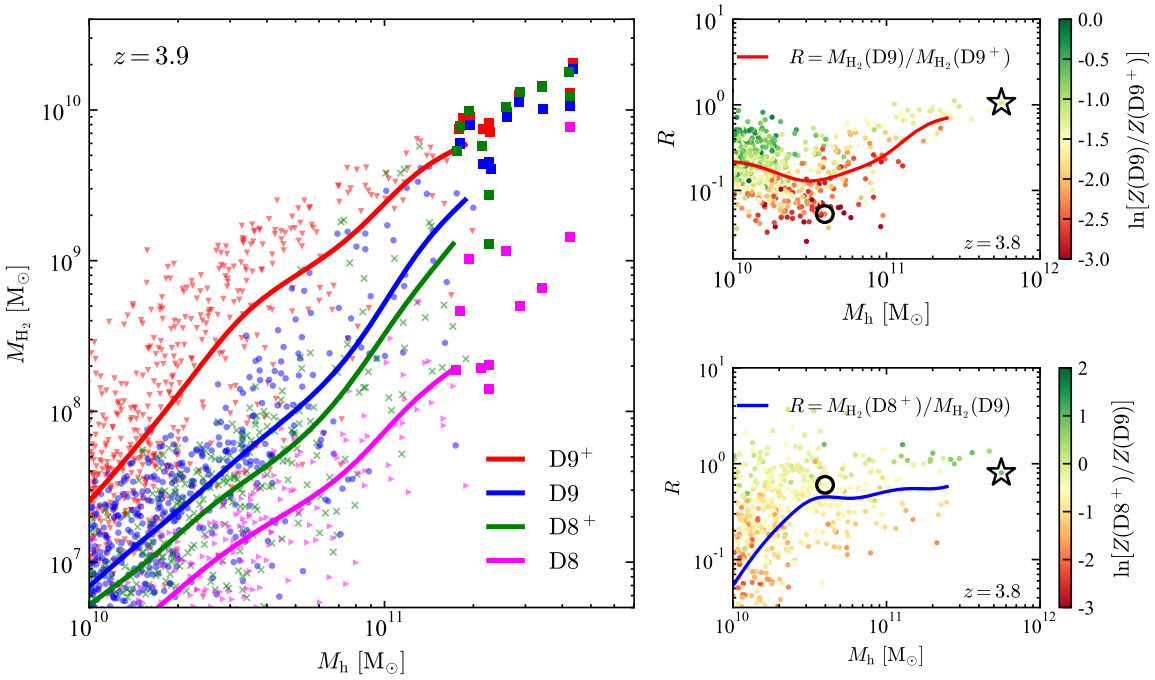

Figure 2.3: Dependence of the H_2 mass of a galaxy on halo mass and on the maximum refinement level achieved in the DYN simulations. Left: Each symbol shows the H_2 mass of a central galaxy in one of the "D" runs as a function of the mass of the corresponding host halo at $z = 3.9$. The colours distinguish simulations with different maximum spatial resolutions as indicated by the labels. The solid lines represent the running average of the points computed in log-log space using a Gaussian kernel with a standard deviation of 0.1 dex. The square symbols highlight the 11 galaxies with the highest M_{H_2} in the D9$^+$ simulation and their counterparts in the other runs. These objects will be further discussed in Fig. 2.4 and section 2.5. Right: The dots represent the ratio between the H_2 masses of individual objects cross-matched between the D9 and D9$^+$ simulations (top) and the D8$^+$ and D9 runs (bottom) at $z = 3.8$ as a function of the host-halo mass. All points are colour coded based on the (natural) logarithmic difference in metal abundance between the lower- and higher-resolution runs. The outsized star and circle highlight the galaxies discussed in Fig. 2.1. The solid lines show running averages taken as in the left panel.

Gaussian-weighted running average. In order to connect this statistical study with the detailed discussion, we have presented for the ☆ and ◯ galaxies in Fig. 2.1; we make sure that the colour of each data point reflects the ratio between the median mass-weighted metallicity of the ISM in the two runs. We also highlight the ☆ and ◯ objects themselves with the corresponding symbols. The plot clearly shows that galaxies hosted by haloes with $M_h < 4 \times 10^{10}\, h^{-1}\, M_\odot$ in the D9 simulation tend to contain nearly an order of magnitude less H_2 than in the D9$^+$ run. However, the scatter is large and strongly correlates with the ratio in the metal content of the galaxies in the two simulations.

As expected from our discussion of Fig. 2.1, galaxies that produce many more metals in D9$^+$ (reddish data points) also show a large difference in the H_2 content between the simulations. For these objects, the onset of SF in their progenitors takes place at earlier times in the D9$^+$ run (which is able to resolve higher densities) than in D9 (see also Kuhlen, Madau and Krumholz, 2013; Tomassetti et al., 2015) and this ultimately leads to more metal- and H_2-rich galaxies.

It is interesting to comment also regarding the cloud of greenish data points that appear on the left-hand side of the plot. They correspond to galaxies that have experienced little SF in both simulations and thus show similar levels of metal enrichment. The higher densities resolved in the D9$^+$ run, though, are

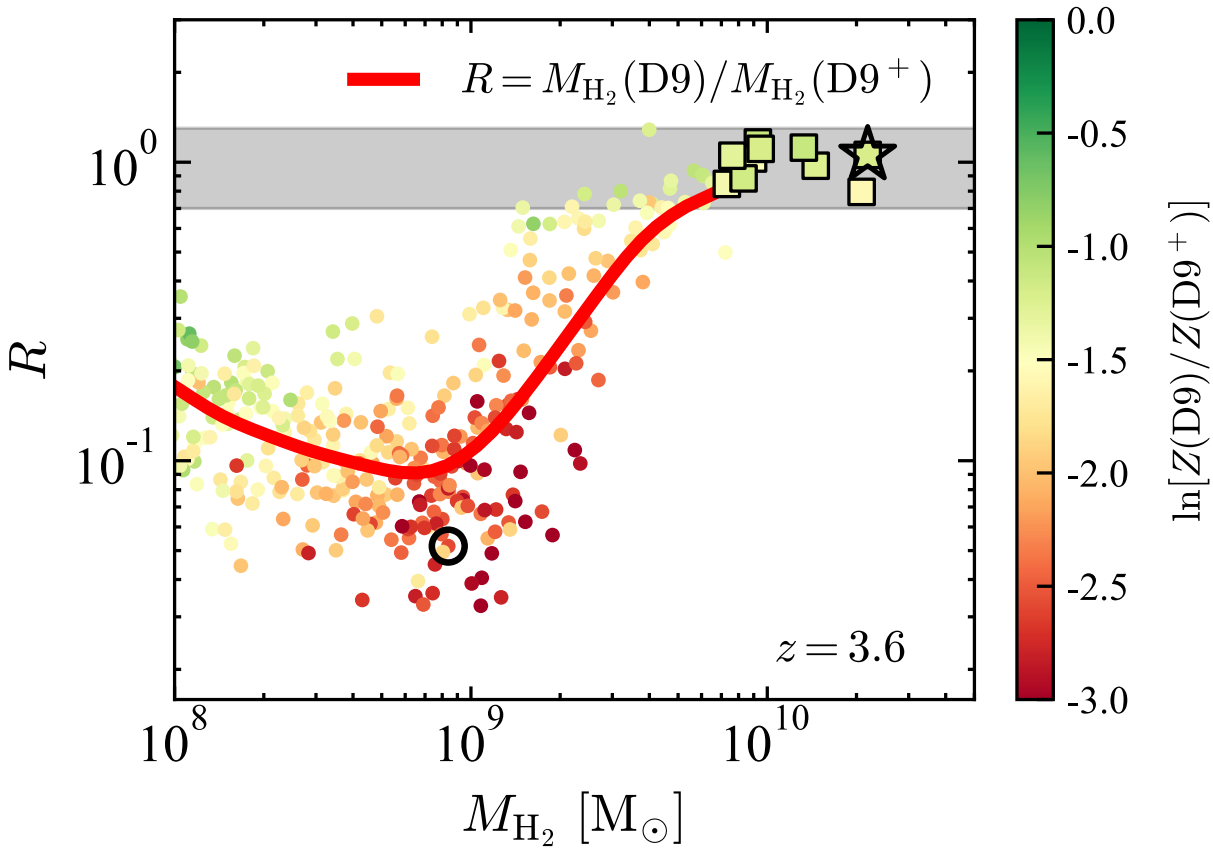

Figure 2.4: As in the top-right panel of Fig. 2.3 but at $z = 3.6$ and using the galaxy H_2 masses instead of the host-halo masses. The grey band highlights the region $0.7 < R < 1.3$. The square symbols mark the galaxies with the highest H_2 content that lie within the grey band. We will use these 11 objects in section 2.5.

enough to yield slightly larger H_2 masses.

Finally, in the bottom-right panel of Fig. 2.3, we compare the H_2 content and the metallicity of the galaxies produced in the D8$^+$ and D9 runs at $z = 3.8$. These simulations achieve the same maximum level of refinement for the gas although the dark-matter distribution is discretized using particles of different masses. Also in this case, we find that the H_2 mass ratio strongly correlates with the relative metallicity. However, the overall trend is different than in the top-right panel. Here, there is a sizeable subpopulation of galaxies for which the two simulations give consistent results at all halo masses. In parallel, there is a second subset whose elements generate substantially less metals and H_2 in the D8$^+$ run. The dichotomy is produced by the absence or presence of a time delay between the epochs in which the progenitors of the galaxies start forming H_2 and stars in the two simulations.

To compare our simulations with observational data and make predictions for forthcoming surveys, we isolate a set of galaxies whose H_2 content does not appear to be strongly affected by spatial resolution effects. In practice, we fix a threshold in M_{H_2} such that, for all galaxies above it, the H_2 masses found in the D9 and D9$^+$ runs differ by less than 30 per cent at $z = 3.6$. As shown in Fig. 2.4, following this procedure, we end up selecting 11 central galaxies with $M_{H_2} \geq 10^{9.8}$ $M_\odot$. We will use this subsample in section 2.5.

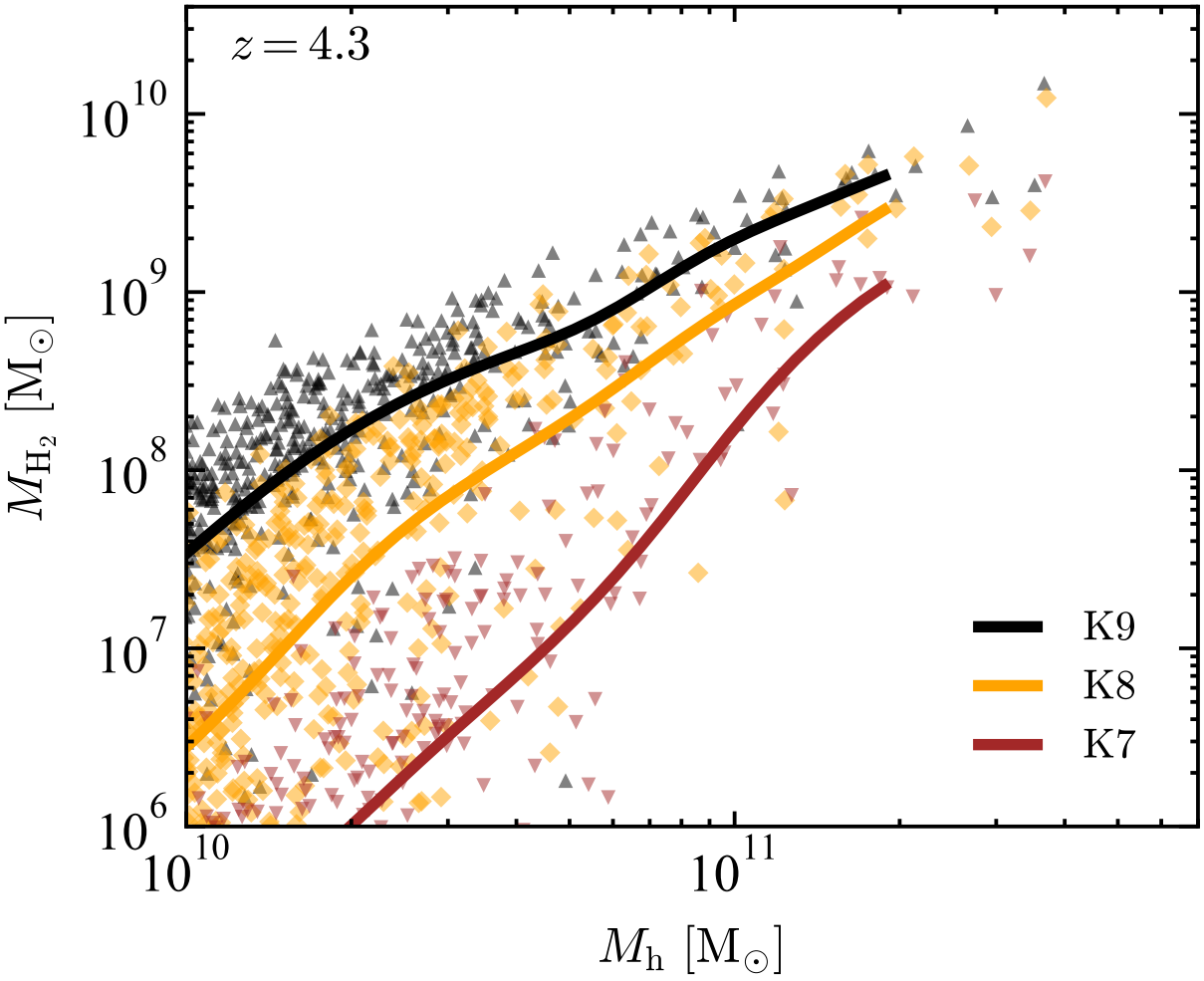

Figure 2.5: Dependence of the H_2 mass of a galaxy on halo mass, for different maximum refinement levels achieved in simulations based on the equilibrium model. The plot is as in the left panel of Fig. 2.3 but using the K7, K8 and K9 simulations at their lowest common redshift, $z = 4.3$.

2.4.2 Equilibrium model

The $M_{H_2} - M_h$ relation emerging in the K7, K8 and K9 runs is shown in Fig. 2.5 at the lowest common redshift of the simulations, $z = 4.3$. The central galaxies hosted by haloes with $M_h \simeq 10^{10}$ M$_\odot$ in the K9 simulation contain nearly a factor of 10 (100) more H_2 than in the K8 (K7) run. This systematic discrepancy decreases with increasing the halo mass. For instance, the difference between K9 and K8 reduces to a factor of $\sim 2-3$ for $M_h \sim 3\times10^{10}$ M$_\odot$ and becomes very small for $M_h > 2\times10^{11}$ M$_\odot$. This trend and its underlying explanation are very similar to those discussed in Fig. 2.3.

2.4.3 Model comparison and H_2 maps

It is interesting to compare the different models at the highest resolution available. In Fig. 2.6, we present the $M_{H_2} - M_h$ relation emerging at $z = 4.3$ in the D9$^+$, K9 and PBP simulations. On a statistical basis, the results from the D9$^+$ and K9 simulations agree very well. while the PBP model predicts substantially lower H_2 masses. If the comparison is performed object by object, the D9$^+$ run always gives slightly

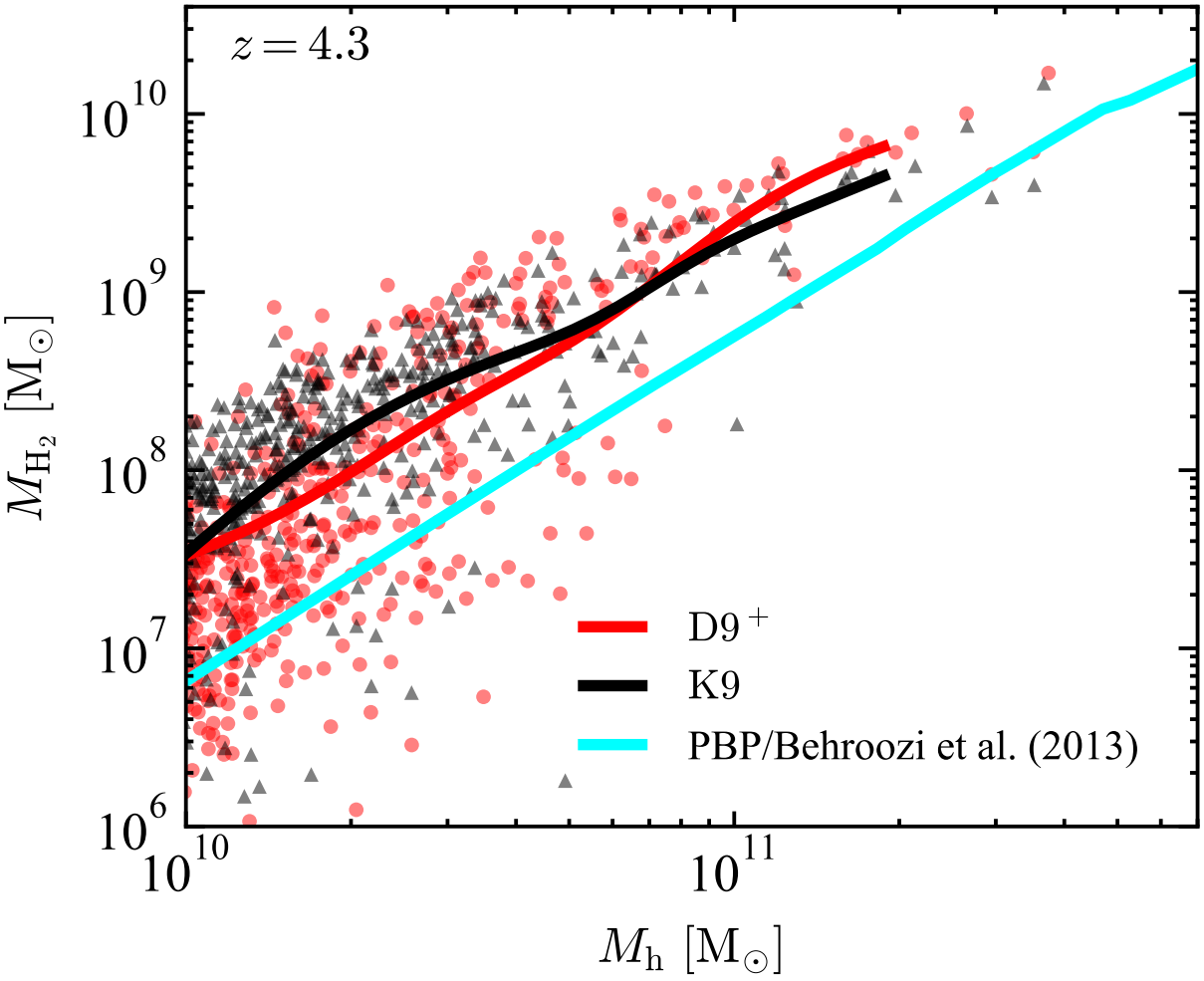

Figure 2.6: M_{H_2}-M_h relation at $z = 4.3$ in simulations based on different H_2-formation models. For each algorithm, we use the highest-resolution run available. Note that the H_2 mass in the PBP model only depends on the host-halo mass of the galaxies and so there is no scatter.

larger molecular masses for $M_h > 10^{11}$ M$_\odot$ and tends to yield smaller M_{H_2} for $M_h < 5 \times 10^{10}$ M$_\odot$.

As an illustrative example, in the top panels of Fig. 2.7 we show H_2 maps of the ☆ galaxy obtained at $z = 4.3$ with different models and by varying the numerical resolution. Each galaxy has been independently rotated using the direction of its stellar angular momentum to obtain a face-on view. Shown is the maximum value of the H_2 density along each line of sight. The name of the simulations and the base-10 logarithm of the total H_2 mass are indicated in each thumbnail. The white bar in the top-left corner corresponds to two physical kpc. For completeness, in the bottom panels of Fig. 2.7, we also show the projected stellar density and the base-10 logarithm of the stellar mass.

A few things are worth noticing. First, the total H_2 and stellar masses come out to be in the same ball park for all runs. In particular, the same value of M_{H_2} is consistently found in all simulations that have a spatial resolution $\Delta x(z = 4.3) \sim 100$ pc or better (i.e. D8$^+$, D9, D9$^+$, K9). Secondly, the detailed morphological structure of the galaxy depends substantially on the maximum spatial resolution and on the adopted SF law that changes the way stellar feedback influences the gas. In the simulations with the lowest resolution (D8, K7, K8), the ☆ galaxy takes the form of a featureless disc. On the other hand, a

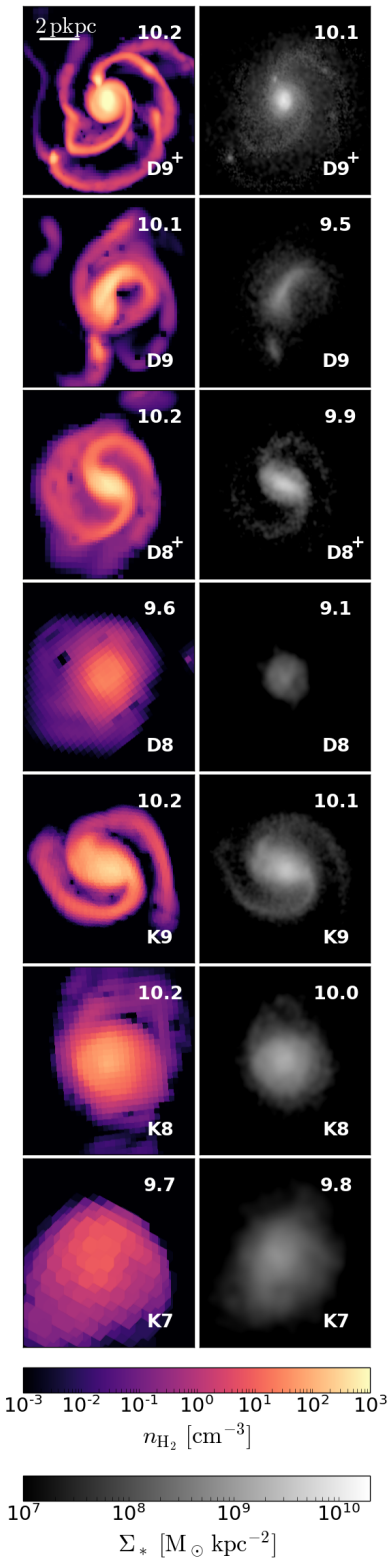

Figure 2.7: H_2 (left) and stellar (right) face-on maps for the ☆ galaxy in the D9$^+$, D9, D8$^+$, D8, K9, K8 and K7 simulations at $z = 4.3$. The images show the maximum H_2 density achieved along the line of sight and the projected stellar density. The white bar in the top-left panel corresponds to a physical length (i.e. not comoving) of 2 kpc. The numbers in the top-right corners give $\log(M_{H_2}/M_\odot)$ and $\log(M_*/M_\odot)$.

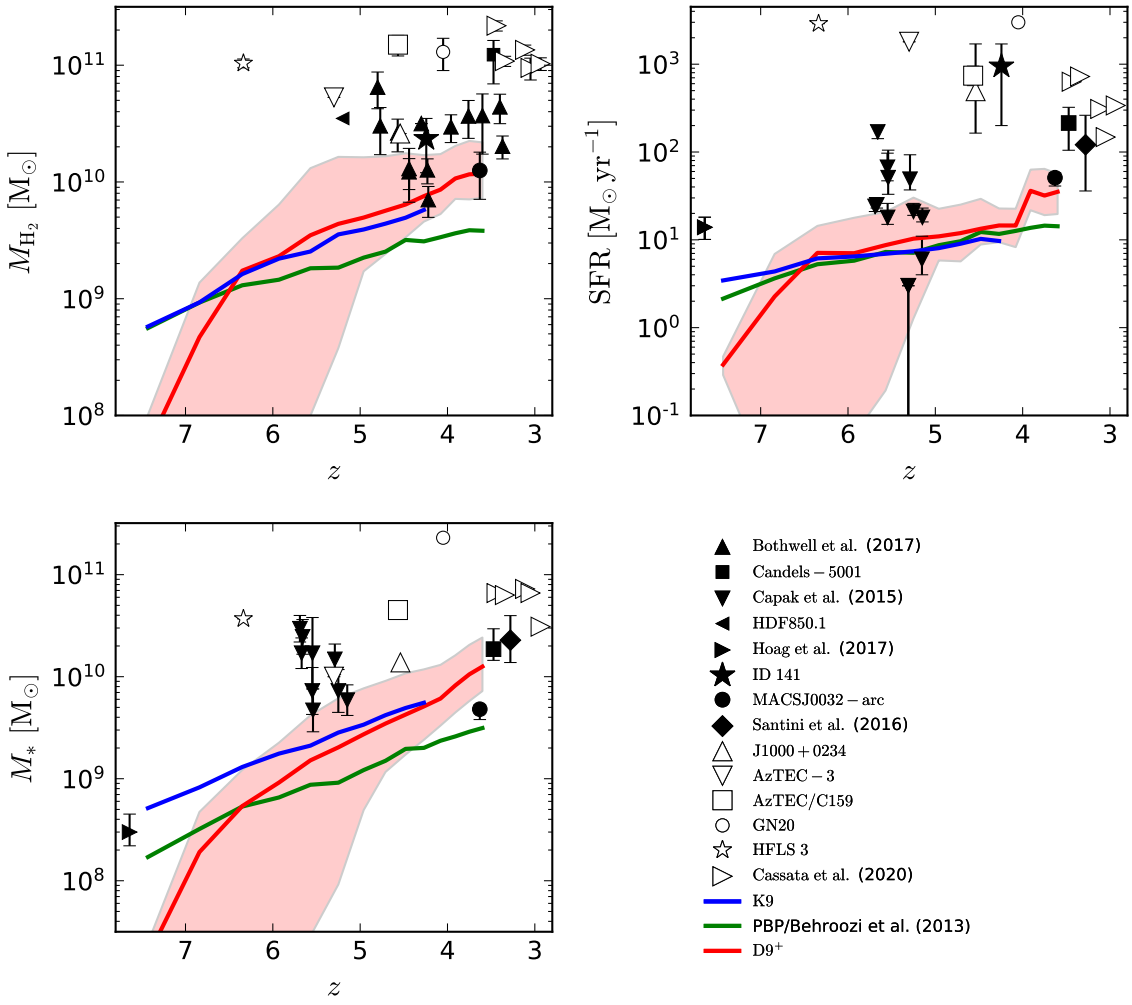

Figure 2.8: Properties of simulated and real galaxies as a function of redshift. The solid lines represent the mean trend of the 11 galaxies with robust predictions for M_{H_2} we have selected from our simulations in section 2.4.1. Their full range of variability in the D9$^+$ run is indicated with the shaded region. Symbols with errorbars refer to observed high-z galaxies for which measurements became recently available (see section 2.5.1 for details). Note that the observational sample mainly includes extreme objects that are not representative of the overall galaxy population.

strong bar/bulge plus symmetric spiral arms in the disc are noticeable in the D8$^+$ and K9 runs. Signs of interactions (with a smaller companion) appear in the D9 simulation. Finally, a grand design spiral with a small bulge is produced in the D9$^+$ run. Note that H_2 traces the densest regions of the galaxy in all simulations.

Similar conclusions can be drawn by inspecting any of the 11 selected galaxies with stable predictions for the H_2 mass in the D runs. Quite interestingly, all these objects present a disc-like morphology and show prominent spiral arms in the higher-resolution runs.

2.5 Comparison with observations

In this section, we compare the properties of the 11 galaxies we have selected from our simulations to recent observational data. First, we look at individual galaxy properties and then we examine the H_2 mass function (MF). Finally, we investigate the time evolution of the cosmic H_2 density.

2.5.1 Galaxy properties

In Fig. 2.8, we contrast the main properties of our simulated galaxies against a compilation of observational data that includes 1) a sample of 13 dusty star-forming galaxies from Bothwell et al. (2017) for which [CI] and CO observations are available, 2) Candels - 5001, a single extended object detected in CO (Ginolfi et al., 2017), 3) 12 dust-poor galaxies detected in [CII] (Capak et al., 2015), 4) ID 141, a sub-millimeter galaxy at $z = 4.24$ detected in CO, [CI] and [CII] (Cox et al., 2011); 5) HDF 850.1, a galaxy detected at $z \approx 5.2$ (Walter, Decarli, Carilli et al., 2012), 6) a spectroscopically confirmed ultra-faint galaxy at $z \sim 7.64$ which presents strong Lyman-α emission (Hoag et al., 2017), 7) MACSJ0032-arc, a lensed galaxy at $z \sim 3.6$ with detected CO emission (Dessauges-Zavadsky et al., 2017), 8) a lensed sub-millimeter galaxy at $z \sim 3.28$ detected with ALMA (Santini, Castellano et al., 2016), 9) J1000+0234, a millimeter galaxy detected in the COSMOS field (Schinnerer et al., 2008) with SFR and masses obtained by Gómez-Guijarro et al. (2018), 10) GN20, a sub-millimeter galaxy, member of a rich proto-cluster at $z = 4.05$ in the GOODS-North field (Carilli, Daddi et al., 2010), 11) AzTEC3, a sub-millimeter galaxy at $z = 5.298$ within a massive protocluster in the COSMOS field (Riechers, Capak et al., 2010), 12) HFLS3, a massive starburst galaxy at $z = 6.34$ (Riechers, Bradford et al., 2013), 13) AzTEC/C159, a star-forming disc galaxy at $z = 4.567$ (Jiménez-Andrade et al., 2018), 14) Five star-forming galaxies detected with ALMA using multiple CO transitions and the continuum (Cassata et al., 2020). It must be acknowledged that this compilation is not representative of the overall galaxy population. Only the most luminous objects with extraordinarily high SF rates can be detected at high redshift with current telescopes. This bias becomes even more extreme for the galaxies for which we can estimate the molecular mass. Many of these measurements rely on the amplification of the sources due to gravitational lensing (e.g., Cox et al., 2011; Walter, Decarli, Carilli et al., 2012; Santini, Castellano et al., 2016; Dessauges-Zavadsky et al., 2017).

In the different panels of Fig. 2.8, the H_2 masses (upper left), stellar masses (lower left), and SFRs (upper right) of the 11 simulated galaxies that have been selected in section 2.4.1 (solid lines and shaded area) are compared with the observational data (the symbols with errorbars) when available. The solid red lines show the mean values (the actual one and not the average of the log values) for the 11 galaxies extracted from the D9$^+$run while the shaded regions extend from the minimum to the maximum value. Similarly, the blue lines indicate the averages for the 11 galaxies extracted from the K9 simulation. To improve readability, we do not show their scatter, which is comparable to the shaded region. Finally, the green lines represent the mean predictions of the PBP model applied to the parent haloes of the D9$^+$ galaxies. Once again, the corresponding scatter (not shown) is comparable to the shaded region in the plot.

Fig. 2.8 shows that all models predict fast molecular enrichment of the galaxies in the redshift range $3 < z < 8$. Basically M_{H_2} increases by a factor of ten in just a Gyr. This is associated with a mild increase in the SF rate and a rapid growth of the stellar mass. On average, the differences between the models are rather small compared with the scatter among the individual galaxies. The most noticeable differences are i) the dynamical model predicts a delayed assembly of the molecular and stellar masses that approximately match the other models only for $z < 7$; and ii) the PBP model provides lower estimates for the molecular (by a factor of 2-3) and stellar masses (by a factor of 2-4) at $z < 6$.

In all cases, the simulated galaxies match well the properties and the evolutionary trends of the less extreme observed objects. This indicates good agreement. In fact, given the relatively small size of our simulation boxes, our synthetic galaxies can only be representative of the typical galaxy population and not of the tails of the distributions sampled by current observations. Fig. 2.9 provides evidence in this direction. Here we show a scatterplot of the SFR against stellar mass for the galaxies in the $D9^+$ and K9 simulations at $z \approx 4.3$. These quantities are tightly correlated. The points in the plot align in a similar fashion to the observed main sequence of star-forming galaxies (e.g. Schreiber et al., 2015; Pearson et al., 2018). Interestingly, the sequence in the simulation nicely extends to low stellar masses that are not probed by current observations. Note that the galaxies of the PBP model are assumed to lie along the green solid curve that relates the average SFR and the average M_* for different halo masses[2].

2.5.2 H_2 mass function

Studying the evolution of the H_2 mass function (MF) provides a convenient way to express information about the molecular content of the Universe. This quantity gives the comoving number density of galaxies per unit M_{H_2} interval, i.e. $dn = \phi\, dM_{H_2}$. As commonly done in the high-redshift literature, we use here the MF per unit log-mass interval

$$\Phi(M_{H_2}, z) = \frac{dn}{d\log M_{H_2}} = \frac{M_{H_2}}{\log e}\,\phi(M_{H_2}, z)\ . \tag{2.7}$$

Observationally, the MF is estimated from the CO luminosity function adopting a CO-to-H_2 conversion factor inferred for normal galaxies in the local Universe or for near-infrared galaxies at intermediate redshifts. At high z, data are still scarce and only constrain the high-mass end of the MF. In the left-hand panel of Fig. 2.10, we report two recent estimates based on the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS)[3] for sources in the redshift range $3.01 < z < 4.48$ with average $\langle z \rangle = 3.8$ (Decarli, Walter, Aravena et al., 2016; Decarli, Walter, Gónzalez-López et al., 2019). The grey box represents the measurement of the ASPECS Pilot program at $M_{H_2} \sim 10^{10}$ $M_\odot$ while the downward arrows denote the corresponding upper limits at larger M_{H_2} (Decarli, Walter, Aravena et al., 2016). On the other hand, the gold-framed sliding boxes depict the results from the ASPECS Large Program (LP) 3mm data in the same redshift interval (Decarli, Walter, Gónzalez-López et al., 2019). Superimposed, we plot the MF obtained from our simulations at $z = 3.8$ for $D9^+$ (red histogram) and PBP (green) runs and at $z = 4.3$ for the K9 (blue) run. In this case, we consider only central galaxies[4] (no satellites) with $M_{H_2} \geq 4 \times 10^7 M_\odot$. Error bars (the shaded regions) include the contributions from Poisson noise and sample variance (which we estimate from the two-point correlation function of the simulated galaxies). The horizontal purple line indicates the upper limit corresponding to counting zero galaxies in each bin of the MF extracted from the simulations. The PBP model provides the best agreement with the ASPECS Pilot data while the dynamical and equilibrium models roughly predict two to three times higher counts. On the other hand, all simulations are compatible with ASPECS LP, although the M_{H_2} ranges probed

[2] Contrary to what happens at lower and higher redshift, data at $z \sim 4$ constrain this relation only for $M_* \gtrsim$ a few $\times\, 10^9$ $M_\odot$ or, equivalently, $M_h \gtrsim$ a few $\times\, 10^{11}$ $M_\odot$ (Behroozi, Wechsler and Conroy, 2013).

[3] All the ASPECS results assume the CO[J-($J-1$)]-to-CO(1-0) luminosity ratios derived by Daddi et al. (2015) for normal star-forming galaxies, a CO(1-0)-to-H_2 conversion factor of $\alpha_{CO} = 3.6\ M_\odot\ (\mathrm{K\ km\ s^{-1}\ pc^2})^{-1}$, as well as a cosmological model with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$.

[4] Section 2.4 shows that, for low halo masses, M_{H_2} is affected by the limited spatial resolution of the K and D simulations. Therefore, the corresponding MFs could be underestimated at the low-mass end.

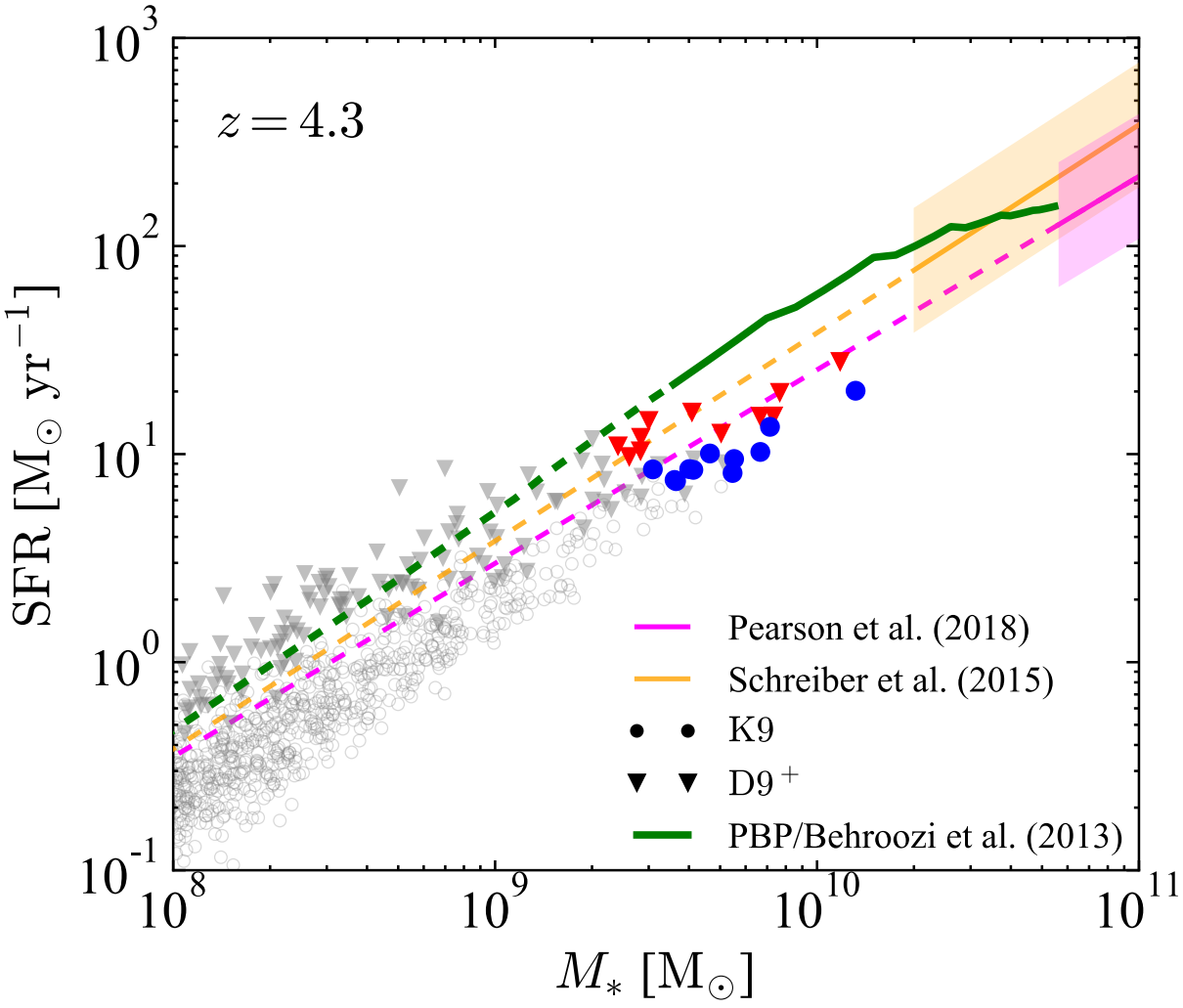

Figure 2.9: SFR versus stellar mass for the galaxies of the D9$^+$ and K9 runs at $z \approx 4.3$ (grey symbols, coloured ones highlight the 11 galaxies selected in section 2.4.1). By construction, galaxies in the PBP model lie along the green solid curve derived in Behroozi, Wechsler and Conroy (2013). For comparison, we also show fits to the main sequence of star-forming galaxies derived from observations at similar redshifts (Schreiber et al., 2015; Pearson et al., 2018). Solid lines are used in the range of M_* probed by observations. Dashed lines, instead, indicate the extrapolation of the fits to lower stellar masses. The shaded areas denote the intrinsic scatter measured about the main sequence.

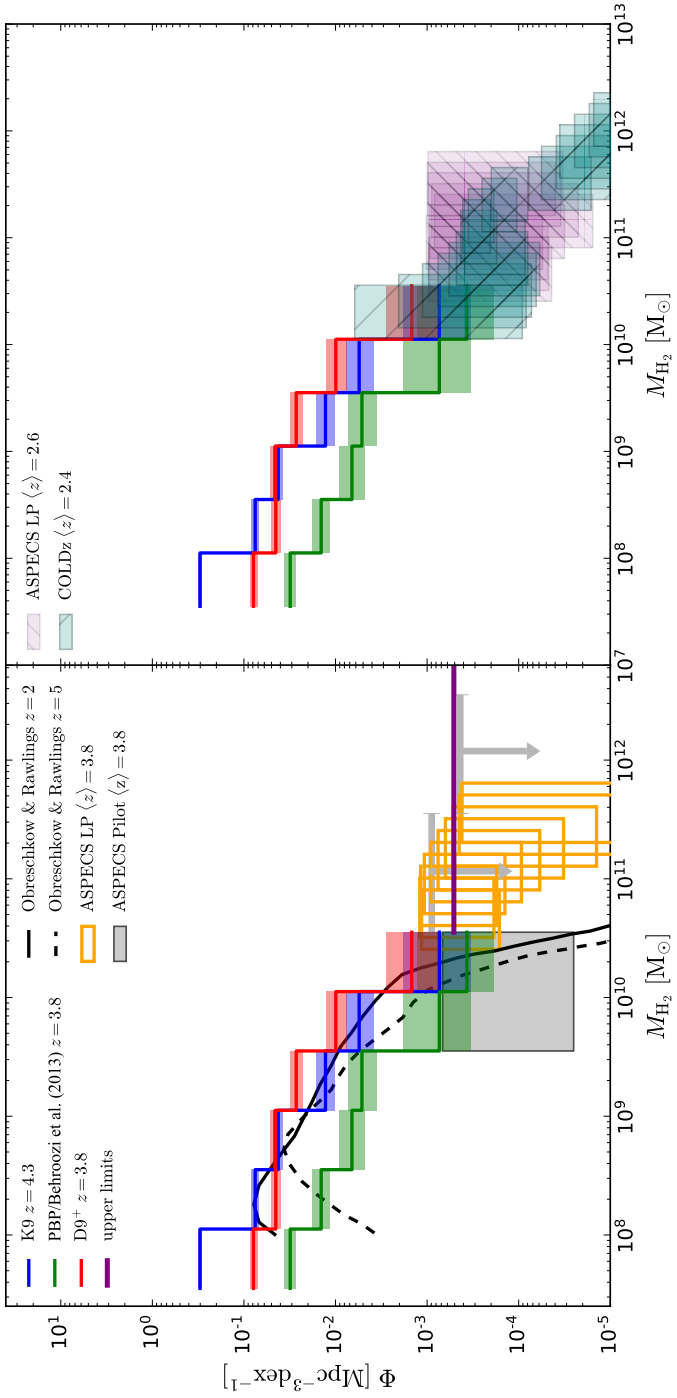

Figure 2.10: Left: the H_2 mass functions (and their 1-σ Poisson uncertainties) derived at a mean redshift of 3.8 (the full range is $3.01 < z < 4.48$) from the CO(1-0) luminosity function in the ASPECS Pilot program (gray box and downwards arrows, Decarli, Walter, Aravena et al., 2016), and in the ASPECS Large Program 3mm data (gold-framed sliding boxes, Decarli, Walter, Gónzalez-López et al., 2019) are compared with the results of our simulations (coloured histograms with errorbars that include Poisson errors and sample variance). The horizontal purple line indicates the upper limit for the MF corresponding to zero counts in the simulations in bins of 0.5 dex. The solid and dashed black curves show the models from Obreschkow and Rawlings (2009b) at $z = 2$ and $z = 5$. Right: lower-redshift data from the ASPECS Large Program (hatched purple boxes) and COLDz (hatched teal frames, Riechers, Pavesi et al., 2019) are compared with our simulation results (same as in the left-hand panel).

by the data and models do not overlap[5] given the relatively small size of our computational volume. For reference, we also show the MFs calculated at $z = 2$ and 5 by Obreschkow and Rawlings (2009b) assuming a relation between the interstellar gas pressure and the local molecular fraction (Obreschkow, D. Croton et al., 2009). These predictions were obtained by post-processing the semi-analytic galaxy catalogue of De Lucia and Blaizot (2007). They lie in the same ballpark as our simulations and are somewhat intermediate between the results of the PBP and D9$^+$ runs.

The mass function is often approximated by a Schechter function

$$\phi = \frac{n_*}{M_*} \left(\frac{M_{\mathrm{H_2}}}{M_*} \right)^{\alpha} \exp\left(-\frac{M_{\mathrm{H_2}}}{M_*} \right) , \tag{2.8}$$

or, equivalently,

$$\log \Phi = (\alpha + 1) \log \frac{M_{\mathrm{H_2}}}{M_*} - \frac{M_{\mathrm{H_2}}}{M_*} \log e + \log n_* - \log \log e \, . \tag{2.9}$$

Here, n_* is a normalisation constant, α indicates the low-mass slope, and M_* denotes the knee of the mass function, above which galaxy counts fall off exponentially. Fitting a Schechter funtion to the ASPECS LP data gives $M_* \simeq 10^{10.5}$ M$_\odot$ (Decarli, Walter, Gónzalez-López et al., 2019). On the other hand, no robust constraints can be set on α that turns out to be very sensitive to the corrections applied for fidelity and completeness (Decarli, Walter, Gónzalez-López et al., 2019). All the models displayed in Fig. 2.10 suggest that the ASPECS measurements should indeed sit around the knee of the MF. In the PBP model, the cutoff is located at $M_* \simeq 10^{9.5-10}$ M$_\odot$ while it is shifted up by approximately half a dex in the K9 and D9$^+$ runs. Another interesting aspect worth mentioning is that the faint-end slope in the D9$^+$ run ($\alpha \simeq -1.3$) is substantially shallower than in the other two simulations ($\alpha \simeq -1.9$ for K9 and $\alpha \simeq -1.7$ for PBP). Fairly flat low-mass slopes (at least down to $\simeq 10^{10}$ M$_\odot$) have been measured at $z \simeq 2 - 3$ by ASPECS LP and COLDz (Riechers, Pavesi et al., 2019) – see the right-hand panel in Fig. 2.10. The values of α found in the simulations imply that most of the H_2 in the cosmos sits within reservoirs with $M_{\mathrm{H_2}}$ slightly smaller than M_* and thus just a bit below the current detection limits. For instance, if the mass function closely follows a Schechter function with $\alpha = -1.3$ (-1.9), then only 28 (16) per cent of the total H_2 lies within objects with $M_{\mathrm{H_2}} > M_*$ while already 84 (66) per cent is found in galaxies with $M_{\mathrm{H_2}} > 0.1 M_*$. We will further discuss this in the next section.

2.5.3 Cosmic H_2 density

Determining the redshift evolution of the cosmic H_2 mass density,

$$\rho_{\mathrm{H_2}}(z) = \int M_{\mathrm{H_2}} \, \phi \, \mathrm{d}M_{\mathrm{H_2}} = \int M_{\mathrm{H_2}} \, \Phi \, \mathrm{d} \log M_{\mathrm{H_2}} \, , \tag{2.10}$$

has been the subject of continued observational effort. The current state of the art is summarized in Fig. 2.11. Although quite noisy, the data show an evolutionary trend for which $\rho_{\mathrm{H_2}}$ peaks at $z \approx 2$, in good agreement with the cosmic SF history (e.g. Madau and Dickinson, 2014). Observational constraints are looser at $z > 3$ as i) they are based on very small samples that only include the most massive galaxies; ii) the CO-to-H_2 conversion factors are uncertain, and iii) the intrinsic shape of the CO luminosity function is unknown. Given this, and the large Poisson errors, authors generally do not attempt to correct their estimates for the contribution of faint undetected objects.

[5] Decarli, Walter, Gónzalez-López et al. (2019) locate the 5σ-detection limit for ASPECS LP at approximately 2.7×10^{10} M$_\odot$ (under a series of assumptions listed in their Table 1) which nearly coincides with the H_2 mass of the most massive object in our simulation box.

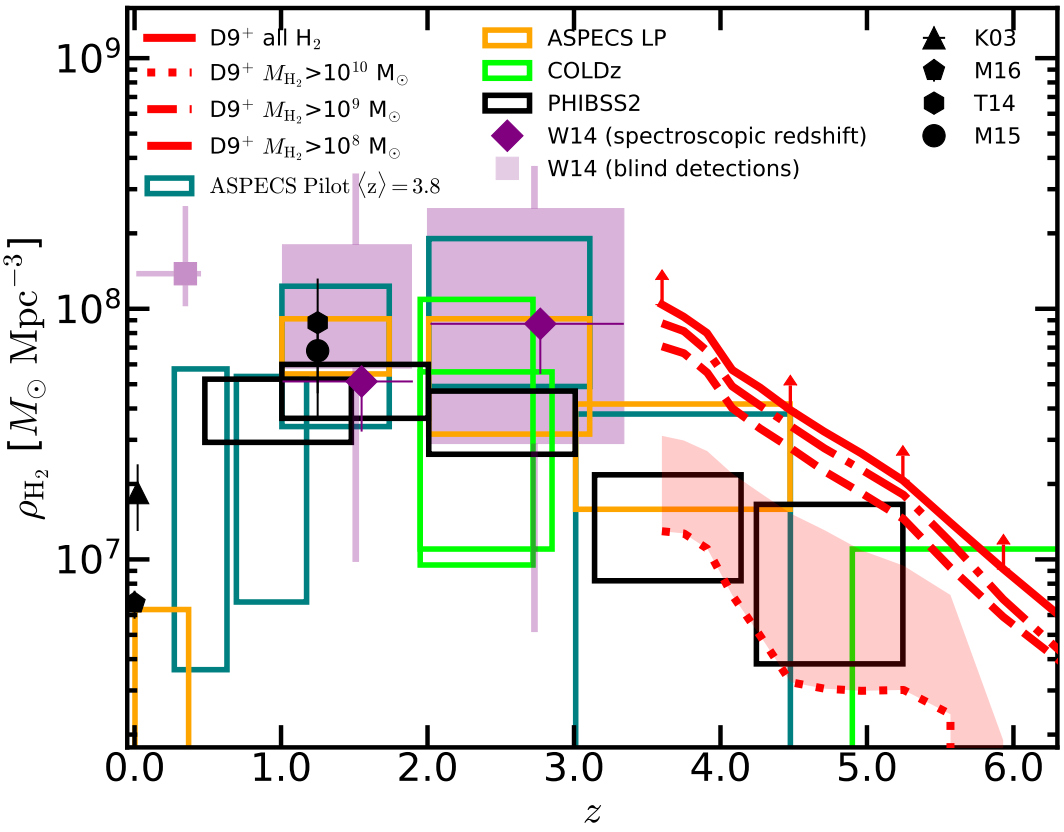

Figure 2.11: The redshift evolution of the cosmic H_2 mass density in the D9$^+$ simulation (thick red curves with different line styles corresponding to various minimum galactic H_2 masses as indicated by the labels) is compared with the observational constraints from Walter, Decarli, Sargent et al. (2014, W14), Decarli, Walter, Aravena et al. (2016, ASPECS Pilot), Decarli, Walter, Gónzalez-López et al. (2019, ASPECS LP), Lenkić et al. (2020, PHIBSS2), Riechers, Pavesi et al. (2019, COLDz), Keres, Yun and Young (2003, K03), Maeda, Ohta and Seko (2017, M17), Tomczak et al. (2014, T14) and Mortlock et al. (2015, M15). The red-shaded area indicates the expected contribution from massive haloes that are underrepresented in our simulation box.

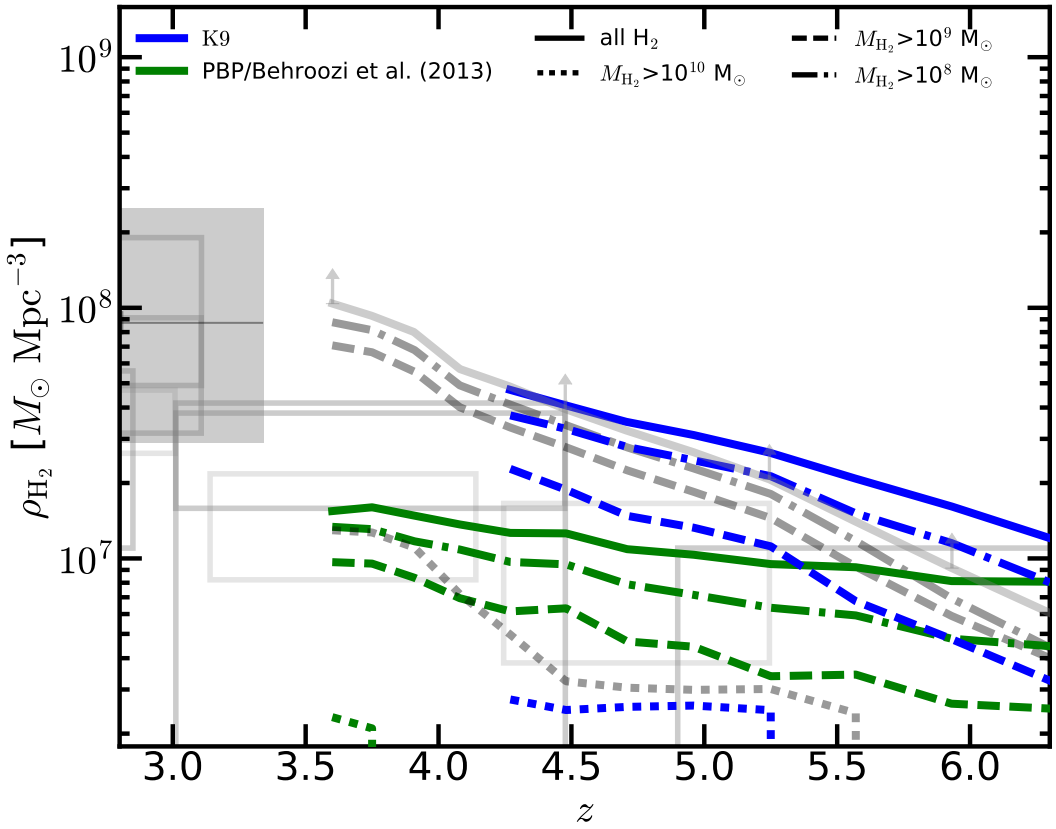

Figure 2.12: As in Fig. 2.11 but for the K9 (blue) and PBP (green) simulations and focusing on redshifts $z > 3$. To ease comparison, Fig. 2.11 is reproduced in the background with gray tones.

Overplotted is the cosmic H_2 mass density extracted from our D9$^+$ simulation (thick red lines). In particular, the dotted curve represents the contribution from the galaxies with $M_{H_2} > 10^{10}$ M$_\odot$. This threshold approximately matches the detection limit of current CO surveys at high redshift (e.g. it lies slightly below the 5σ-detection limit for ASPECS LP, Decarli, Walter, Gónzalez-López et al., 2019). For a proper comparison of the numerical results with the observations, we need to account for galaxies hosted by the rare, very massive haloes that are unlikely to form in our relatively small simulation box. To estimate their overall contribution to the cosmic H_2 density, we proceed as follows. For the most massive haloes at $4 \lesssim z \lesssim 6$, we find that $M_{H_2} \simeq M_h/30$ (see e.g. Fig. 2.3). Therefore, we first compute the total mass density contributed by haloes that are more massive than those appearing in the simulation using a fit for the halo mass function (Sheth, Mo and Tormen, 2001; Tinker et al., 2008). We then rescale the result by a factor of 30 to get a rough estimate of the corresponding H_2 density. The final contribution of the 'missing' haloes is shown as a shaded region lying above the dotted thick red line in Fig. 2.11. At $3.5 \lesssim z \lesssim 5.5$, the shaded area and the dotted line give nearly equal contributions. By considering their sum, we conclude that the agreement between the simulation and the observations is very good.

We now focus on the low-mass end. The dashed and dash-dotted curves in Fig. 2.11 represent the contribution from galaxies with $M_{H_2} > 10^9$ M$_\odot$ and $M_{H_2} > 10^8$ M$_\odot$, respectively. Finally, the solid curve

accounts for the total H_2 mass in the computational volume (without correcting for the missing haloes). Since the simulation likely underestimates the molecular mass of the galaxies residing in low-mass haloes (see section 2.4.1), we have represented this result with upward-pointing arrows to indicate that it is likely a lower limit. Our results suggest that current measurements of the H_2 mass density may be underestimated by a redshift-dependent factor that ranges between 2 and 3 at $3.5 \lesssim z \lesssim 5.5$ (after taking into account the contribution from the massive haloes that are underepresented in our box). The galaxies that host the undetected molecules should also contribute an important fraction of the cosmic SFR.

The redshift evolution of the cosmic H_2 density in the PBP and K9 simulations is presented in Fig. 2.12. The K9 predictions are largely consistent with those of the D9$^+$ run. However, they show less evolution at low M_{H_2} as low-mass molecular reservoirs are in place earlier in this model (see the top-left panel in Fig. 2.8). On the other hand, the PBP model predicts a milder redshift evolution of the cosmic H_2 density between redshift 6.5 and 3.6 with respect to the other runs. The total H_2 density at $z = 3.6$ is nearly an order of magnitude lower than in the D9$^+$ run. When one takes into account the current detection limits, it appears difficult to reconcile the PBP results with the measurements from the IRAM Plateau de Bure HIgh-z Blue Sequence Survey 2 (Lenkić et al., 2020) and ASPECS LP (Decarli, Walter, Gónzalez-López et al., 2019).

Another interesting aspect of the cosmic H_2 density, noticeable in Fig. 2.12, is that the spacing between the lines drawn with different styles varies among the different runs but, barring the PBP model, does not change much with redshift. This behaviour reflects the different shapes of the H_2 MF in the three simulations. Therefore, Fig. 2.12 extends the analysis presented in Fig. 2.10 to a broader redshift range.

2.6 Summary

In this paper, we have analyzed and compared three approximate methods for tracking the formation and evolution of H_2 in cosmological simulations of galaxy formation. The first, dubbed PBP, is a semi-empirical model that associates a H_2 mass to a galaxy based on the mass and the redshift of its host dark-matter halo (Popping, Behroozi and Peeples, 2015). The second, labelled KMT, assumes chemical equilibrium between the H_2 formation and destruction rates (Krumholz, McKee and Tumlinson, 2009). The third, called DYN, fully solves the out-of-equilibrium rate equations and accounts for the unresolved structure of molecular clouds using a sub grid model (Tomassetti et al., 2015). Both the KMT and the DYN models require as an input local estimates of the density, metallicity and the intensity of radiation in the LW band. We compute this last quantity by propagating radiation from the stellar particles in the simulations. Furthermore, in the simulations based on the KMT model, we link SF to the local density of cold gas, whereas we use the H_2 density in conjunction with the DYN model.

Each of the algorithms listed above represents a broad class of models (semi-empirical, equilibrium, non-equilibrium) that have been adopted in the literature, sometimes with different implementations. For the scientific applications, however, different authors have used wildly different spatial resolutions making it difficult to compare their results. Therefore, as a first task, we have investigated how the finite spatial resolution of the numerical simulations impacts the predictions of the H_2 content of the synthetic galaxies in the KMT and DYN models. Using the Ramses code, we have run a suite of simulations (four for DYN model and three for the KMT one) that reach different maximum levels of refinement for the dark-matter and gaseous components. All the simulations start from ICs that have the same amplitudes and phases for the mutually resolved Fourier modes so that we can easily cross match the host haloes of the galaxies in the different runs. Finally, we have compared our synthetic galaxies to a compilation of recent observational results. Our results can be summarized as follows.

(i) The conversion rate of HI into H_2 depends on the local metallicity and density of the ISM. It turns

out that the H_2 formation time can be far longer than the age of the Universe in low-mass high-z galaxies with low Z (Kuhlen, Krumholz et al., 2012; Jaacks, Thompson and Nagamine, 2013; Kuhlen, Madau and Krumholz, 2013; Thompson et al., 2014; Tomassetti et al., 2015). However, in Fig. 2.1, we have shown that the precise timing of the end of this process is resolution dependent. At higher spatial resolutions, the progenitors of a galaxy can reach higher gas densities and start forming molecules, stars and metals earlier than in lower-resolution runs. Therefore, H_2 masses that are stable with respect to resolution changes can only be obtained: a) for the objects that have reached a sufficient level of metal enrichment and where the H_2-formation timescale is shorter than the age of the Universe (see Fig. 2.1); and b) for the galaxies that are still not forming any H_2. In our simulations with resolution elements of 50-100 pc at $z = 3.6$, the first group corresponds to $M_{H_2} \gtrsim 6 \times 10^9$ $M_\odot$ and the second to $M_{H_2} \ll 10^8$ $M_\odot$ (see Fig. 2.4).

(ii) On average, in our highest-resolution runs, the KMT and the DYN models generate very similar $M_{H_2} - M_h$ relations, while the PBP one produces significantly less H_2 at fixed halo mass (by a factor of 5-6, see Fig. 2.6). On an individual basis, the KMT- and DYN-based runs yield similar H_2 masses for metal-enriched galaxies with $Z \gtrsim$ a few $\times 10^{-2}$ $Z_\odot$ (see also Krumholz and Gnedin, 2011), although the values of M_{H_2} tend to be slightly higher in the DYN model, on average (see Fig. 2.8). Note, however, that a given galaxy gets metal enriched a later time in the DYN model.

(iii) The detailed morphology of the synthetic galaxies is influenced by the maximum spatial resolution of the simulations and by the SF law (gas vs. H_2 based) that determines where and when stellar feedback injects energy into the ISM (see Fig. 2.7).

(iv) The simulated galaxies with the highest M_{H_2} (whose molecular content does not depend much on spatial resolution effects) match well the properties (star-formation rates, stellar and H_2 masses) and the evolutionary trends of the less extreme objects extracted from a compilation of recent observational data (Fig. 2.8). This is satisfactory in all respects as current observations at high-z tend to pick extreme objects that are not representative of the overall population while our relatively small simulation box targets average galaxies that populate the main sequence of star-forming objects (Fig. 2.9). Interestingly, all the synthetic galaxies that host large molecular reservoirs have disc morphologies and present spiral arms in the higher-resolution runs. Overall, the differences between the three H_2 models are smaller than the scatter among the individual galaxies (Fig. 2.8). Some systematic trends are noticeable, however. First, the dynamical model predicts a delayed assembly of the molecular and stellar masses. Second, the PBP model yields lower molecular and stellar masses at $z < 6$.

(v) The H_2 MFs extracted from the simulations at $z \approx 4$ show similar normalisations and cutoff scales ($M_{H_2} \approx 10^{10}$ $M_\odot$) as recent observational data and semi-analytic models (Fig. 2.10). The PBP model generates lower counts at fixed M_{H_2} (by a factor of $\sim 3 - 10$) with respect to the KMT and DYN ones. In all cases, the low-mass slopes of the MF are sufficiently flat that only a small fraction of molecular material is contained in galaxies with $M_{H_2} \ll 10^{10}$ $M_\odot$.

(vi) When we account for the detection limits of current CO surveys, the cosmic H_2 mass density extracted from our simulations based on the KMT or DYN models matches well with recent observations at $z > 3$ (Fig. 2.11 and 2.12). However, our results suggest that most molecular material at high-z lies yet undetected in reservoirs with $10^9 < M_{H_2} < 10^{10}$ $M_\odot$. Adding the integrated contribution of these sources to the current estimates for ρ_{H_2} increases its value by a redshift-dependent factor ranging between 2 and 3. Similarly, they should be responsible for an important fraction of the cosmic SFR at $3.5 < z < 6$. Based on our simulations,the dominant contribution to the missing H_2 comes from galaxies that lie just below the detection limits of current CO surveys and should become available with longer integration times and/or to future facilities like the next generation Very Large Array [6].

[6] https://ngvla.nrao.edu/.

CHAPTER 3

A subgrid model for molecular gas tracers[1]

3.1 Introduction

To gain a deep insight into the nature of galaxy formation and evolution, the molecular gas content of galaxies is of particular interest. Over the last years, more evidence has been found that H_2 is the fuel of star formation but at the same time it is not detectable due to its lack of a dipole moment. This makes it very difficult to observe the process of galaxy evolution directly, especially at high redshift.

Although that issue has been known for quite some time, observations still rely on a workaround by using the transitions of internal energy levels of numerous molecules, ions, and atoms as tracers for molecular gas. This was made possible by the advent of (sub-)millimetre astronomy with telescopes like ALMA and NOEMA. Converting the observed luminosity into the molecular gas mass represents the standard method for measuring the H_2 abundance of a high redshift object nowadays.

Widely used tracers for molecular gas are carbon, ionized carbon, and carbon monoxide. Their transition lines are detected to gain knowledge about the molecular gas content in galaxies. This enables telescopes to measure the luminosity of galaxies at redshifts from $z \approx 6$ towards $z \approx 3$ (e.g. Bothwell et al., 2017; Cox et al., 2011) at the (sub-)millimetre wavelengths and provides results about the molecular gas reservoir in galaxies at their different evolutionary states.

From the theoretical side, simulations provide the approach to galaxy formation. Simulations with a boxlength of a few parsecs can include many details, especially large chemical networks that calculate accurately the abundances of molecular gas tracers. In this regard, simulations covering a cosmological volume are still limited to molecular hydrogen and do not track the abundance of carbon (or molecules containing it) whose luminosity is actually measured in observations as the carbon chemistry is more complex than that of H_2.

The challenge in the use of large-scale simulations is to bring both together, capturing all effects involved in the non-linear process of galaxy formation and including the chemistry in the ISM. In general, the credibility of a simulation strongly depends on its resolution but one single resolution element in a cosmological simulation has the size of a full box of a small-scale run. Thus, large-scale simulations depend on subgrid models that describe effectively the physics on scales that are not resolved in the simulation.

The lack of a suitable subgrid model for carbon chemistry makes it problematic to compare the properties of simulated galaxies with those of observed ones. This is crucial for studying e.g. the CO luminosity and molecular gas content of galaxies in a cosmological context from a theoretical perspective. The dependencies on parameters such as redshift or metallicity of the conversion factor

[1] This chapter is in preparation for submission to MNRAS.

α_{CO} that relates $M_{\mathrm{H_2}}$ and L_{CO} (see Bolatto, Wolfire and Leroy, 2013, for a review) are essential in the connection between theory and observations.

The current situation regarding simulations is that state-of-the art simulations that include a model for the carbon chemistry are limited to boxsizes from several parsecs up to kiloparsecs (Walch et al., 2015; Girichidis et al., 2016; Pardi et al., 2017; Peters et al., 2017; Seifried et al., 2017). Here, the implementation of the different species is important for the understanding of star formation and the matter lifecycle involved as their abundances also influence the cooling of the gas. Their focus is more on the physics on a galactic scale rather than on galaxy formation.

Over the years, different methods and chemistries were used and developed to treat the chemistry in the ISM. A detailed comparison of commonly used networks is discussed in Glover and Clark (2012a). They point out that the challenge for an approximated network is not only the abundance of the different carbon species but also the timescale.

One of the therein discussed networks is the established chemistry from Nelson and Langer (1999, hereafter NL99). The network reduces its level of complexity by covering some of the molecules via the introduction of pseudo-species which enclose intermediate carbon and oxygen molecules in the CO formation channels. However, it assumes that all hydrogen is already fully molecular and the reaction channels for H_2 production is added by using the reactions presented in Glover and Mac Low (2007a) and Glover and Mac Low (2007b). Although modifications have been recently developed (Gong, E. C. Ostriker and Wolfire, 2017), the version presented in Glover and Clark (2012a) is a vastly used version of the NL99 (e.g. Seifried et al., 2017; Joshi et al., 2019; Clark et al., 2019; Mackey et al., 2019; Popping, Narayanan et al., 2019).

One of the most complex networks tracks the evolution of 32 species in 218 reactions and is introduced in Glover, Federrath et al. (2010). Its complexity comes along with high accuracy and has been applied to multiple small-scale ISM simulations (e.g. Shetty, Glover, Dullemond and Klessen, 2011; Shetty, Glover, Dullemond, E. C. Ostriker et al., 2011; Gray and E. Scannapieco, 2013).

To embed one of the mentioned chemistries in an already existing simulation code, different platforms have been developed for the calculation of the abundance of different carbon species. The most comfortable option is the post-processing that can make use of an external module. Cloudy Ferland et al. (2017) is able to calculate abundances, spectra, and emission lines, and therefore, it is used in a remarkable number of projects (e.g. Olsen, Greve et al., 2016; Pallottini, Ferrara, Decataldo et al., 2019; Lupi, Pallottini et al., 2020; Olsen, Burkhart et al., 2021). In Bisbas, Bell et al. (2012), the 3D-PDR[2] code is introduced to discuss the dependencies of different quantities – such as luminosities or abundances of different chemical species – on e.g. cosmic rays, the UV field or the metallicity (Bisbas, Schruba and van Dishoeck, 2019; Bisbas, J. C. Tan and Tanaka, 2021) after post-processing the results of a hydro simulation.

The second option is to do on-the-fly calculations with a module by including e.g. the Krome module (Grassi et al., 2014) which allows the calculation of non-equilibirum chemistries (e.g. Bovino et al., 2014; Lupi, Pallottini et al., 2020; Sillero et al., 2021).

The differences between on-the-fly and post-processing calculations is discussed in Lupi, Pallottini et al. (2020), focusing on the resulting FIR luminosity for a single high-redshift galaxy. Therein, they study in a zoom-in cosmological simulation the results obtained on the one hand by post-processing with Cloudy and on the other hand on-the-fly calculations with a non-equilibrium chemistry using Krome. Although both methods showed robust results for the [CII] luminosity, the simulation required a high resolution for a proper capturing of the chemistry.

These simulations and models are very useful for studies of the matter cycle on small scales but their

[2] PDR stands for *photodissociation region* or *photon-dominated region*.

approach is not suitable for large-scale simulations and for galaxy formation studies. Complex chemical networks require considerable computational resources and a high resolution. Thus, some studies of single galaxies assume a constant ratio of carbon species relative to molecular hydrogen (Tomassetti et al., 2014; Olsen, Greve et al., 2016).

Other approaches to track the evolution of carbon species and their luminosity in a cosmological context can also be done using post-processing (Inoue, Yoshida and Yajima, 2020) or semi-analytic models as described in Popping, Narayanan et al. (2019). These models combine simulation results with scaling relations coming from observations to infer the abundance or the luminosity.

In the following, we present a subgrid model that is designed to capture accurately and efficiently a dynamical carbon chemistry for cosmological simulations of a maximum resolution of tens of parsecs. The model is supposed to do on-the-fly calculations within an adaptive-mesh-refinement (AMR) code Ramses assuming a part of the chemistry to be in equilibrium and predict the amount of possible tracers for H_2 as CO, C^+ and C. Thus, abundance calculations are neither calculated in post-processing nor the simulations are in need to have a resolution of several parsecs. The structure of this chapter is as follows: in section 3.2, we will derive the chemistry used for our subgrid model from the NL99 network and study the impact of the approximations made by applying it to a simple GMC model in section 3.3. Section 3.4 will show how the chemistry is embedded in the subgrid model in order to calculate the abundances of C, CO, C^+ and H_2 for resolution elements with sizes of tens of parsecs. In section 3.5, we apply our subgrid model in a zoom-in simulation to a single galaxy and present the results. Then, in section 3.6 we apply our subgrid model in post-processing to a population of galaxies from a cosmological simulation. We evaluate our findings using observational results with the discussion in section 3.7 and conclude in section 3.8.

3.2 A simplified chemistry

The implementation of chemistry always represents a challenge because accuracy and computational costs have to be weighed up against each other. Mathematically, a network of chemical reactions is described by a set of time differential equations (see section 1.8). We start from the chemistry of the well-established NL99 network extended by the hydrogen reactions as presented in Glover and Clark (2012a). The NL99 network itself is already simplified by the introduction of the pseudo-species OH_x and CH_x which enclose intermediate carbon and oxygen molecules. To end up with a chemistry module which does not slow down the hydro code and does not lose too much of its accuracy at the same time, we make the following simplifications to reduce the number of differential equations even further: (i) instead of 24 reactions, we consider only 14, (ii) the number of species to consider is reduced to 5 (H, H_2, C^+, C and CO), (iii) we assume a fast consumption of the species OH_x, CH_x, and H_3^+ and (iv) we assume that the electron abundance can be approximated by $x_{C^+} \approx x_{e^-}$. Due to this, we include most of the species and reactions that are involved in the carbon chemistry, but speed up the calculations by these simplifications as the number of time differential equations that need to be solved is reduced significantly.

A full list of all reactions, species, and reaction rates for both the approximated and full network are

listed in the appendix B. Following these simplifications, one obtains the following set of equations:

$$\begin{aligned}\frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} &= \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{O},\mathrm{H}_3^+} x_{\mathrm{O}} \frac{k_{\mathrm{c.r.},\mathrm{H}_2} x_{\mathrm{H}_2}}{k_{\mathrm{C},\mathrm{H}_3^+} x_{\mathrm{C}} + k_{\mathrm{O},\mathrm{H}_3^+} x_{\mathrm{O}} + k_{\mathrm{e}^-,\mathrm{H}_3^+} x_{\mathrm{C}^+}} \alpha_{\mathrm{OH_x}} \\ &\quad + \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{C},\mathrm{H}_3^+} x_{\mathrm{C}} \frac{k_{\mathrm{c.r.},\mathrm{H}_2} x_{\mathrm{H}_2}}{k_{\mathrm{C},\mathrm{H}_3^+} x_{\mathrm{C}} + k_{\mathrm{O},\mathrm{H}_3^+} x_{\mathrm{O}} + k_{\mathrm{e}^-,\mathrm{H}_3^+} x_{\mathrm{C}^+}} \alpha_{\mathrm{CH_x}} \\ &\quad + \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{H}_2} x_{\mathrm{C}^+} x_{\mathrm{H}_2} \alpha_{\mathrm{CH_x}} - k_{\gamma,\mathrm{CO}} x_{\mathrm{CO}} \end{aligned} \tag{3.1}$$

$$\frac{\mathrm{d}x_{\mathrm{C}}}{\mathrm{d}t} = -\frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} + \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{H}_2} x_{\mathrm{C}^+} x_{\mathrm{H}_2} - k_{\gamma,\mathrm{C}} x_{\mathrm{C}} + n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{e}^-} x_{\mathrm{C}^+} x_{\mathrm{C}^+} \tag{3.2}$$

$$\frac{\mathrm{d}x_{\mathrm{C}^+}}{\mathrm{d}t} = n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{e}^-} x_{\mathrm{C}^+} x_{\mathrm{C}^+} - n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{H}_2} x_{\mathrm{C}^+} x_{\mathrm{H}_2} + k_{\gamma,\mathrm{C}} x_{\mathrm{C}} \tag{3.3}$$

$$\begin{aligned}\frac{\mathrm{d}x_{\mathrm{H}_2}}{\mathrm{d}t} &= 2 k_{\mathrm{H}_2,\mathrm{gr}} x_{\mathrm{HI}} n_{\mathrm{H}} - k_{\gamma} x_{\mathrm{H}_2} - n_{\mathrm{H}} k_{\mathrm{C}^+,\mathrm{H}_2} x_{\mathrm{C}^+} x_{\mathrm{H}_2} \\ &\quad - k_{\mathrm{c.r.},\mathrm{H}_2} x_{\mathrm{H}_2} \frac{k_{\mathrm{C},\mathrm{H}_3^+} x_{\mathrm{C}} + k_{\mathrm{O},\mathrm{H}_3^+} x_{\mathrm{O}}}{k_{\mathrm{C},\mathrm{H}_3^+} x_{\mathrm{C}} + k_{\mathrm{O},\mathrm{H}_3^+} x_{\mathrm{O}} + k_{\mathrm{e}^-,\mathrm{H}_3^+} x_{\mathrm{C}^+}} , \end{aligned} \tag{3.4}$$

where we have defined $x_{\mathrm{CO}} = n_{\mathrm{CO}}/n_{\mathrm{H}}$, $x_{\mathrm{C}} = n_{\mathrm{C}}/n_{\mathrm{H}}$, $x_{\mathrm{C}^+} = n_{\mathrm{C}^+}/n_{\mathrm{H}}$ and $x_{\mathrm{H}_2} = 2n_{\mathrm{H}_2}/n_{\mathrm{H}}$. The parameters $\alpha_{\mathrm{CH_x}}$ and $\alpha_{\mathrm{OH_x}}$ are the branching fractions and defined as $\alpha_{\mathrm{CH_x}} = \frac{k_{\mathrm{O},\mathrm{CH_x}} x_{\mathrm{O}}}{k_{\mathrm{O},\mathrm{CH_x}} x_{\mathrm{O}} + k_{\gamma,\mathrm{CH_x}}/n_{\mathrm{H}}}$ and $\alpha_{\mathrm{OH_x}} = \frac{k_{\mathrm{C},\mathrm{OH_x}} x_{\mathrm{C}}}{k_{\mathrm{C},\mathrm{OH_x}} x_{\mathrm{C}} + k_{\gamma,\mathrm{OH_x}}/n_{\mathrm{H}}}$, respectively. They account for the two different fates of the $\mathrm{CH_x}$ and $\mathrm{OH_x}$, namely to be either consumed in the reaction with an oxygen (in the case for $\mathrm{CH_x}$) or a carbon atom (in the case for $\mathrm{OH_x}$) or to be photodissociated. Note that not all differential equations have to be solved since the fraction of all carbon atoms is fixed. From that it follows e.g. $x_{\mathrm{C}} = x_{\mathrm{C,tot}} - x_{\mathrm{CO}} - x_{\mathrm{C}^+}$. The different reaction coefficients are indicated as $k_{\ldots}$ and describe at which rate two reactants form a product. They are listed in appendix A but due to the complexity and their significant role in the rest of the chapter, we will discuss the dependencies of photodissociation coefficients for H_2, C and CO in detail in the following.

In general, the photodissociation rate of a species can be reduced by effects like dust shielding, self-shielding and – if necessary – shielding from other species. For instance, the dissociation rate of H_2 depends on dust- and self-shielding and fully reads

$$k_{\gamma,\mathrm{H}_2} = G\, \kappa_{\mathrm{H}_2}\, \Phi(N_{\mathrm{H}_2})\, e^{-\tau}, \tag{3.5}$$

where G is the unshielded UV flux in Habing units and κ_{H_2} the unshielded photodissociation rate with a value of $4.2 \times 10^{-11}\ \mathrm{s}^{-1}$. $e^{-\tau}$ accounts for the dust-shielding, where $\tau = \sigma_{\mathrm{d}} N_{\mathrm{H}}$ with $\sigma_{\mathrm{d}} = 2 \times 10^{-21} (Z/\mathrm{Z}_\odot)$ cm^2 being the dust cross section and $N_{\mathrm{H}} = N_{\mathrm{HI}} + 2N_{\mathrm{H}_2}$ the total hydrogen column density. The self-shielding function Φ is defined according to Draine and Bertoldi (1996)

$$\Phi(N_{\mathrm{H}_2}) = \frac{1-\omega}{(1+x/b_5)^2} + \frac{\omega}{(1+x)^{1/2}} \exp\left[-8.5 \times 10^{-4} (1-x)^{1/2}\right] \tag{3.6}$$

with $x = N_{\mathrm{H}_2}/(5 \times 10^{14} \mathrm{cm}^{-2})$ where N_{H_2} is the column density of molecular hydrogen. The remaining parameters are set to $\omega = 0.035$ and $b_5 = 2$.

For CO, the dissociation coefficient

$$k_{\gamma,\mathrm{CO}} = G\,\kappa_{\mathrm{CO}}\,e^{-\tau}\,\Theta(N_{\mathrm{H}_2}, N_{\mathrm{CO}}) \tag{3.7}$$

consists of the dust-shielding and the UV flux, but also on the shielding function $\Theta(N_{\mathrm{H}_2}, N_{\mathrm{CO}})$ from Visser, van Dishoeck and Black (2009) and the photodissociation rate $\kappa_{\mathrm{CO}} = 2 \times 10^{-10}\ \mathrm{s}^{-1}$. At this point, we want to highlight that the shielding function depends on both, the column density of H_2 and CO.

The dissociation rate for C has the following form (Tielens and Hollenbach, 1985):

$$k_{\gamma,\mathrm{C}} = G\,\kappa_{\mathrm{C}}\,e^{-\tau}\,\Theta(N_{\mathrm{H}_2}, N_{\mathrm{C}}) \tag{3.8}$$

with $\kappa_{\mathrm{C}} = 3.1 \times 10^{-10}\ \mathrm{s}^{-1}$ and $\Theta(N_{\mathrm{H}_2}, N_{\mathrm{C}}) = \exp(-\tau_{\mathrm{C}})\Phi_{\mathrm{C}}(N_{\mathrm{H}_2})$, where $\tau_{\mathrm{C}} = 1.6 \times 10^{-17} N_{\mathrm{C}}/\mathrm{cm}^{-2}$, $\Phi_{\mathrm{C}}(N_{\mathrm{H}_2}) = \exp(-r_{\mathrm{H}_2})/(1 + r_{\mathrm{H}_2})$, and $r_{\mathrm{H}_2} = 2.8 \times 10^{-22} N_{\mathrm{H}_2}/\mathrm{cm}^{-2}$. As discussed above, G and $e^{-\tau}$ again account for the UV flux and the dust shielding.

3.3 Accuracy of approximations

3.3.1 Setup

More importantly, we need to validate the accuracy of the approximated network and that it does not cause any significant changes in the outcome. For this reason, we use a spherical cloud embedded in a uniform UV field as a test setup. The density profile $\rho(r)$ with r being the radial coordinate of the spherical cloud follows a power law of the form

$$\rho(r) \propto r^{-\alpha} \tag{3.9}$$

where $\alpha = 1.6$ (Pirogov, 2009). The radius of the cloud and its mass are fixed parameters leading to a normalization constant of the density profile. Further, we use the Larson relations (Larson, 1981) to relate the cloud size with the mass. By combining the relations for the velocity dispersion σ, size L and the mass M of the cloud,

$$\frac{\sigma}{\mathrm{km\ s}^{-1}} = 1.10 \frac{L}{\mathrm{pc}}^{0.38} \tag{3.10}$$

$$\frac{\sigma}{\mathrm{km\ s}^{-1}} = 0.42 \frac{M}{\mathrm{M}_\odot}^{0.2}, \tag{3.11}$$

we find for the maximum radius $2R_{\mathrm{max}} = L$ of a cloud

$$2\frac{R_{\mathrm{max}}}{\mathrm{pc}} = \frac{L}{\mathrm{pc}} \approx 0.08 \left(\frac{M}{\mathrm{M}_\odot}\right)^{0.5}. \tag{3.12}$$

To include the chemistry, the cloud is then subdivided into different shells and for each of these shells – going from the outer shells towards the center –, the equations (3.1), (3.2), (3.3) and (3.4) are solved. This procedure is necessary for the following reason: the column density for each species i in the j-th shell is calculated by the scheme of $N_{i,j} = N_{i,j-1} + n_j \mathrm{d}r$, where $N_{i,j-1}$ is the column density of the outer shell, n_j the density of current shell, and dr the thickness of the shell. Thus, the calculations of the abundances of H_2, CO and C can be then used to update the column density as this is an essential part for (self-)shielding and important for the chemistry of the next inner shell. As many rate coefficients are

temperature dependent, we include the temperature-density relation from Glover and Mac Low (2007a) which sets the temperature for each shell of the cloud. For the initial conditions, we assume a cloud that consists only of atomic hydrogen and atomic carbon. We decide to calculate the abundances of H_2, CO, C^+ and C for a low and a high UV field (0.1 and 100 in Habing units). The integration time is set by the assumed lifetime of a typical cloud. According to Jeffreson and Kruijssen (2018) and the references therein, the lifetimes for GMCs cover a huge range. We will use for our calculations 25 Myr, as this is an intermediate value.

We point out that this is for testing the chemistry only. In this rather simple configuration, no cloud dynamics is involved, e.g. diffusion, turbulence, collapse, heating or cooling. Thus, the complicated processes or dynamics of a giant molecular cloud are not captured accurately.

3.3.2 Results and Discussion

In Figs. 3.1 and 3.2, we present the fractions of H_2 as CO, C^+ and C in the simulated GMC as a function of the average cloud density $\langle n_H \rangle$ and its metallicity Z for UV fields of $G = 0.1$ and $G = 100$ in Habing units. The fractions are obtained via $f_{C^+,C,CO} = M_{C^+,C,CO}/M_{C,tot}$ and $f_{H_2} = M_{H_2}/M_{H,tot}$, where $M_{C,tot}$ and M_{H_2} denote the total carbon and hydrogen mass, respectively. M_C, M_{C^+}, M_{CO} are the masses of the different carbon species in the cloud and M_{H_2} is the mass of molecular hydrogen.

Turning to the differences between the networks, we find the same trends for both low and high UV radiation. The fractions for H_2 and C^+ are in excellent agreement. For C and CO however, we find deviations of $\approx 70\%$. The differences are located in the transition regions in the $\langle n_H \rangle$–Z plane. The approximated network has a slight tendency to form CO at smaller metallicities (for a fixed density). This effect is mostly relevant at low densities and low metallicities but structures with these properties do not contribute to the molecular mass budget in a resolution element in a simulation (see chapter 2). Hence, only high-mass molecular clouds are properly represented by a resolution element of the simulation and it is important that the chemistry gives accurate results for these specific elements.

Comparing both cases of high and low UV radiation, we see that the case of high UV radiation requires higher average densities to form more molecules (H_2 and CO) at fixed metallicity which is expected. Also, at fixed density an increasing metallicity promotes the formation of molecules.

In Glover and Clark (2012a), a comparison between six different chemical networks has been made and we will use their findings to compare our chemistry with other approaches. We will discuss briefly in the following their main findings to elaborate how our approximated chemistry performs in comparison to other approaches.

Two of them (denoted by G10g and G10ng therein) follow the chemistry presented in Glover, Federrath et al. (2010) which contains 218 reactions between 32 species. Compared to the G10g, G10ng was extended by the authors by recombination reactions of H^+, He^+, C^+ and O^+ on dust grains, additionally to the G10g chemistry.

Further, the therein as NL97 (Nelson and Langer, 1997) model designated chemistry is considered as well. It follows dynamically the CO and C^+ abundances in a simplified way. The assumption here is that the formation of CO depends only on the $C^+ + H_2 \rightarrow CH_2^+ + \gamma$ reaction channel. Thus, the formation part of CO in the differential equation is given by the product of the C^+ and H_2 abundance and the corresponding reaction rate. The destruction of CO happens via photodissocation releasing neutral carbon. Further, CO is destroyed by photodissociation. For the study in Glover, Federrath et al. (2010), hydrogen is not assumed to be fully molecular and thus, the NL97 chemistry has been extended by the hydrogen reaction rates according to Glover and Mac Low (2007a) and Glover and Mac Low (2007b).

The fourth model is the previously described NL99 model, which has also been extended by the hydrogen chemistry from Glover and Mac Low (2007a) and Glover and Mac Low (2007b). The last

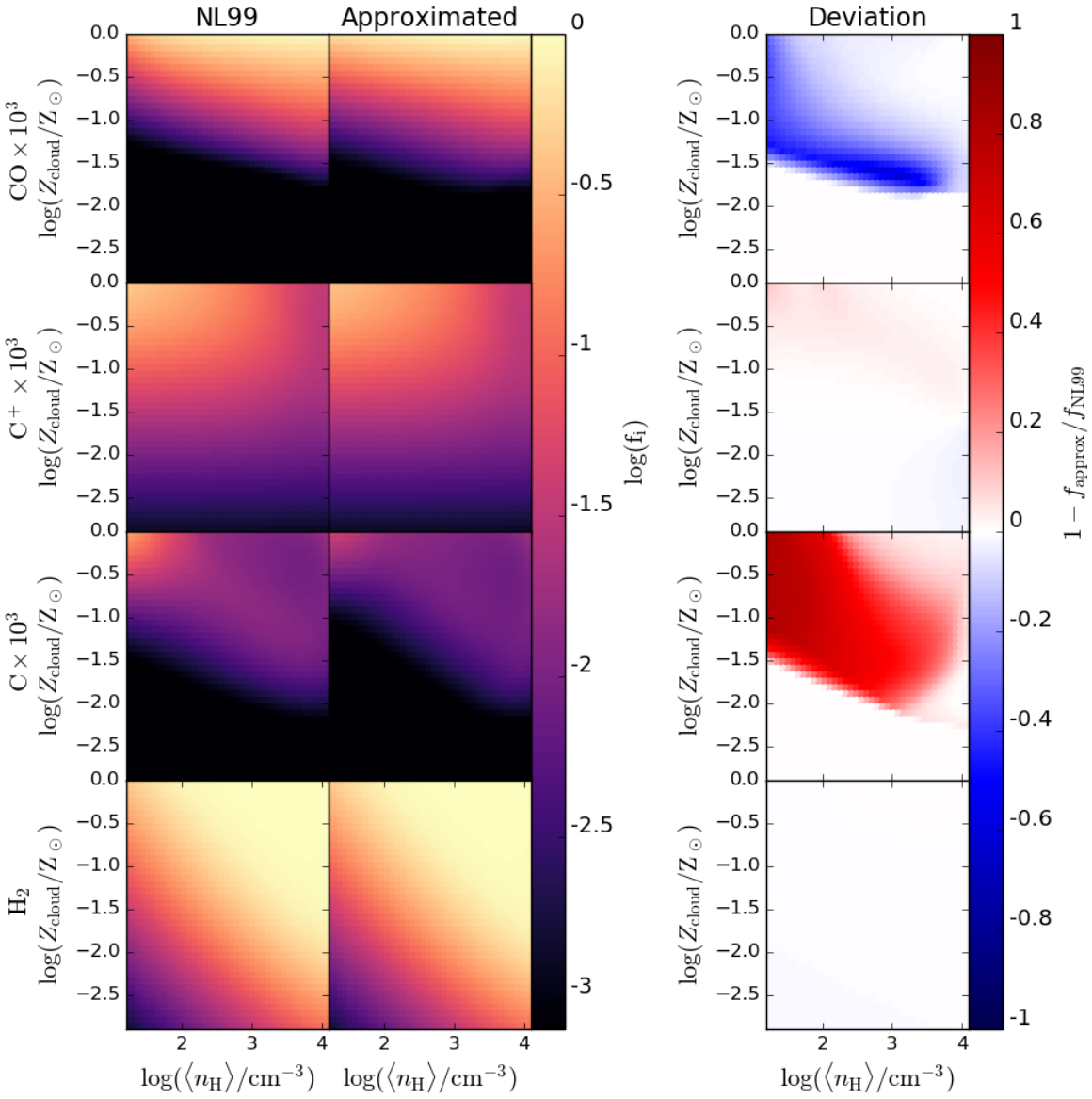

Figure 3.1: Fractions for H_2, CO, C^+ and C as indicated by the color in the GMC for a UV field of $G = 0.1$ in Habing units for different average cloud densities $\langle n_H \rangle$ and metallicities Z in units of solar metallicity. We show in the left column the results from the NL99 chemistry, the middle column presents the results from the approximated network. We also show the deviation between both chemistries in the right column.

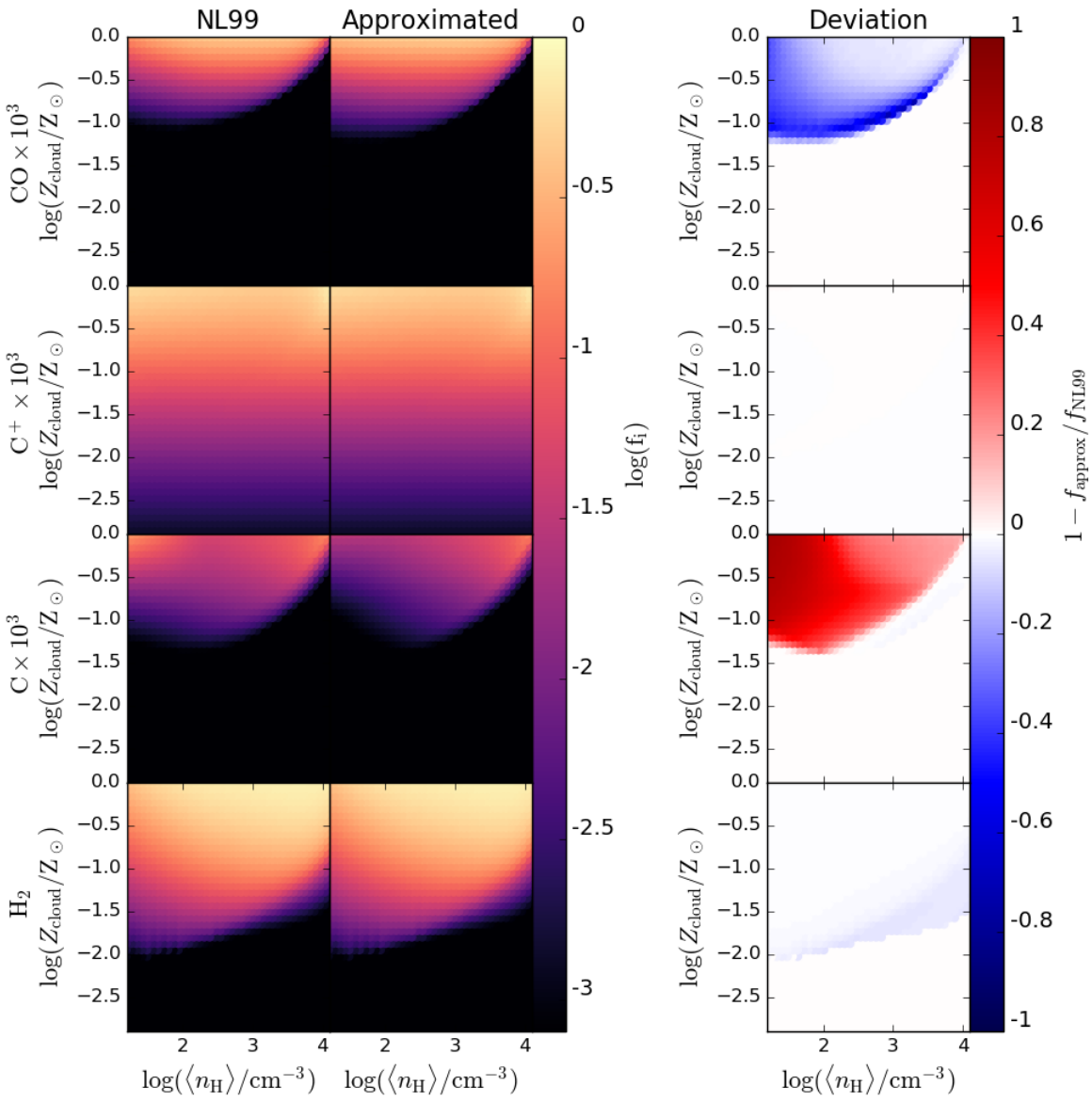

Figure 3.2: Same as Fig. 3.1 but for a UV field of $G = 100$ in Habing units.

two approximations discussed in Glover and Clark (2012a) are based on the paper by Keto and Caselli (2008). Similar to the NL97, the assumption is that the formation of CO happens via a term $k_{H_2,C^+} x_{H_2} x_{C^+}$ with k_{H_2,C^+} being the reaction rate between molecular hydrogen and ionized carbon. The reaction for the photoionization of carbon and photodissociation of carbon monoxide are included as well. The assumption of chemical equilibrium allows an analytic solution for the abundances of C^+, C and CO and the authors refer to that as KC08e. They also consider the case of the non-equilibrium solution which is denoted as KC08n.

The setup for testing the different chemistries is a cubic box with a length of 20 pc, filled with turbulent gas.

One main finding of the study presented in Glover and Clark (2012a) is that the NL99 and G10ng chemistries give very similar results with respect to the carbon and carbon monoxide abundances, but the former network takes only one third of the computational time. Models like NL97, KC08e and KC08n are even faster, but their precision suffers from their simplifications since the simpler a chemistry gets, the faster the formation timescales for different species are. We also find this effect in the comparison between our approximation and the NL99 network for CO and C. On the time scale of a few Myr, the maximum deviation is ≈45 % for an average cloud density of 100 cm^{-3} for our cloud model. Going to an integration time of tens of Myrs and higher average densities, the differences do decrease again. As our results are still very close to those of the NL99, which has according to Glover and Clark (2012a) a similar level of accuracy as a large network like G10g or G10ng, our approximated network gives reasonable results that should be very close to large networks. However, we want to point out that we use a simplified setup compared to that from Glover and Clark (2012a).

3.4 Scaling up the chemistry – the carbon subgrid model

The aim of the previous considerations is the implementation of a subgrid model in the code Ramses in order to simulate the evolution of chemistry within a galaxy. For this purpose, we need to scale up the chemistry – a process of a microscopic scale – to the scale of a resolution element which has a typical size of tens of parsecs. For the carbon chemistry, we adapt the approach to the H_2 chemistry as discussed in Tomassetti et al. (2015). There, the ISM was divided in two phases, a fully molecular and a fully atomic phase which separate at a specific density. To account for unresolved structures, we assume a mass-weighted log-normal probability function for the density n_H in a turbulent ISM (Kainulainen et al., 2009; Schneider et al., 2013):

$$\mathcal{P}_M = \frac{1}{\sqrt{2\pi} n_H \sigma} \exp\left(\frac{(\ln(n_H) - \mu)^2}{2\sigma^2} \right) . \tag{3.13}$$

σ can be obtained once a clumping factor is chosen, which is defined as

$$C_\rho = \frac{\langle n_H^2 \rangle}{\langle n_H \rangle^2} = \exp\left(\sigma^2\right) . \tag{3.14}$$

Note, that $\langle n_H \rangle$ represents the average gas density in a resolution element. The value for μ in equation (3.13) is calculated from the volume-weighted log-normal probability function $\mathcal{P}_V = \mathcal{P}_M \frac{\langle n_H \rangle}{n_H}$

with

$$\langle n_{\rm H} \rangle = \frac{\int {\rm d}n_{\rm H}\, n_{\rm H}\, \mathcal{P}_{\rm V}}{\int {\rm d}n_{\rm H}\, \mathcal{P}_{\rm V}} = \exp\left(\mu - \frac{\sigma^2}{2}\right), \tag{3.15}$$

in combination with equation (3.14). Assuming a sharp separation of the atomic and molecular phase of hydrogen, one can solve self-consistently for the density cutoff $n_{\rm c}$ using equation (10) in Tomassetti et al. (2015) for a given $\langle n_{\rm H_2} \rangle$ where this transition of the phases occurs[3]:

$$\langle n_{\rm H_2} \rangle = \frac{\langle n_{\rm H} \rangle}{4} \left[1 + {\rm erf}\left(\frac{\mu - \log(n_{\rm c})}{\sqrt{2}\sigma}\right)\right]. \tag{3.16}$$

For the hydrogen chemistry, we keep this idea and extend this ansatz for carbon to a three-phase model. The three phases are ionized (C^+ only), atomic (C only), and molecular (CO only). Similar to Tomassetti et al. (2015), we can define self-consistently the following density cutoffs for carbon:

$$\langle n_{\rm C^+} \rangle = \frac{\langle n_{\rm C,tot} \rangle}{2} \left[1 - {\rm erf}\left(\frac{\mu - \log(n_{\rm c,ion})}{\sqrt{2}\sigma}\right)\right] \tag{3.17}$$

$$\langle n_{\rm CO} \rangle = \frac{\langle n_{\rm C,tot} \rangle}{2} \left[1 + {\rm erf}\left(\frac{\mu - \log(n_{\rm c,mol})}{\sqrt{2}\sigma}\right)\right]. \tag{3.18}$$

From these equations, it follows that $n_{\rm c,ion} < n_{\rm c,mol}$, since $\langle n_{\rm C,tot} \rangle = \langle n_{\rm CO} \rangle + \langle n_{\rm C^+} \rangle + \langle n_{\rm C} \rangle$ where the number densities of all species are required to be > 0. The physical meaning is that for low densities ($n_{\rm H} < n_{\rm c,ion}$), only ionized carbon is abundant whereas carbon monoxide dominates the carbon budget at high density ($n_{\rm c,mol} > n_{\rm H}$). The neutral carbon sits in a layer between these two cutoffs. We now have to average each part of the equations (3.1), (3.2), (3.3) and (3.4) by integrating over the density continuum where two reactants exist, since these are the densities where these reactions can occur. In this way we average each part of a reaction rate over the density continuum in a resolution element and obtain reaction rates that represent the chemistry on parsec scale. Considering as an example a reaction between H_2 and C^+, we will find an integral of the form $\int_{n_{\rm c}}^{n_{\rm c,ion}} x_{\rm H_2} x_{\rm C^+} n_{\rm H} \mathcal{P}(n_{\rm H})\, {\rm d}n_{\rm H}$. Based on this method, all remaining coefficients in the rate equations for the subgrid model are calculated. They are listed in detail in appendix C. Following this method, it gives in the end the time differential equations for each species within each resolution element, specifically $\langle \frac{{\rm d}n_{\rm H_2}}{{\rm d}t} \rangle$, $\langle \frac{{\rm d}n_{\rm C^+}}{{\rm d}t} \rangle$, $\langle \frac{{\rm d}n_{\rm CO}}{{\rm d}t} \rangle$ and $\langle \frac{{\rm d}n_{\rm C}}{{\rm d}t} \rangle$, which need to be solved.

3.5 On-the-fly application of the subgrid model – simulating a high-redshift dwarf

3.5.1 Numerical setup

In the following, we apply the subgrid model to a dwarf galaxy on-the-fly in a zoom-in simulation up to $z \approx 4$. The zoom-in region represents a smaller volume of high detail and resolution embedded in a larger box of lower resolution. The setup is very similar to that in chapter 2 but since there are some differences,

[3] Note that one has to take into account for a number density that one hydrogen molecule consists of two hydrogen atoms. Thus, we have to multiply by an additional factor of 1/2, whereas the other 1/2 comes from the integration of the log-normal distribution.

we will list all simulation parameters in detail for the sake of completeness. We use the AMR code Ramses. The numerical treatment of the gas dynamics is based on a second-order Godunov scheme; stars and DM are represented by particles and their trajectories are calculated via a particle-mesh method. The ICs are generated with Music (Hahn and Abel, 2011) based on a ΛCDM cosmology. The cosmological parameters follow Planck Collaboration (2018) where $\Omega_\Lambda = 0.685$, $\Omega_0 = 0.315$, $h = 0.674$, $\sigma_8 = 0.811$, $\tau = 0.054$ and $n_s = 0.954$. The comoving box size of the parent simulation is $(6\,h^{-1}\,\mathrm{Mpc})^3$, whereas the zoom-in region covers a volume of roughly $(2\,h^{-1}\,\mathrm{Mpc})^3$. For the zoom-in region, the ICs are created from level 7 across the entire box up to level 10 in the region around the dwarf galaxy. This corresponds to a comoving resolution of 69.5 kpc and 8.7 kpc for the grid at $z = 99$. As the simulation evolves, the refinement goes up to level 15 which corresponds to a physical resolution of ≈ 55 pc at redshift $z = 4$ and is comparable to the simulation run in chapter 2. The refinements of the resolution elements are triggered if either the baryonic mass exceeds the threshold of $8 m_{\mathrm{DM}} \frac{\Omega_\mathrm{b}}{(\Omega_\mathrm{m} - \Omega_\mathrm{b})}$ or if a cell contains more than 8 dark matter particles.
The star formation is directly linked to the H_2 abundance in the following way (Tomassetti et al., 2015):

$$\dot{\rho}_{\mathrm{SF}} = \epsilon \frac{\rho_{\mathrm{H_2}}}{t_*} \,, \tag{3.19}$$

where ϵ is the star formation efficiency and $\rho_{\mathrm{H_2}}$ the local H_2 density. Here, t_* represents the minimum of the free-fall time of either the resolution element or of a cloud with a uniform density of $100\,\mathrm{cm}^{-3}$. A metallicity floor of $10^{-3}\,\mathrm{Z}_\odot$ is set at $z = 9$ which enables the H_2 production in the first place and is then followed by star formation. To avoid artificial fragmentation, we repressurize the gas using a polytropic index of 5/3. This assures that the Jeans length is always resolved by at least four grid cells (Truelove et al., 1997; Teyssier, Chapon and Bournaud, 2010).

The local UV radiation field is calculated following Tomassetti et al. (2015). Each star represents a stellar population with a mass distribution following an IMF from Kroupa (2001). Luminosities in the UV band are calculated from tabulated values from the Starburst99 (Leitherer et al., 1999) library based on the age of the considered stellar particle. We allow the UV radiation to propagate to the neighboring cells following an $1/r^2$ law and assume the gas to be optically thick after a characteristic length of r_{LW}. The UV background follows the one presented in Haardt and Madau (2012). The cooling function is calculated by the internal module of Ramses which includes the contributions from H, He, and metals.

Haloes are identified using the Amiga Halo Finder code (AHF; Gill, Knebe and Gibson, 2004; Knollmann and Knebe, 2009). Knowing the size of the halo, we define the baryonic part of a galaxy as the matter enclosed within a radius of $0.1 r_{\mathrm{h}}$.

Finally, to include the subgrid model, a separate module to Ramses has been added following the formalism presented in section 3.4.

3.5.2 Results

Properties of the simulated galaxy over time

We list the properties of our simulated object by the end of the simulation at redshift $z = 4$ in table 3.1. This object has a star formation rate of 32.6 $\mathrm{M}_\odot\ \mathrm{yr}^{-1}$ which is a quite high value compared to more massive objects at this redshift (see section 2.5.1). In contrast, the H_2 and stellar content are smaller by approximately an order of magnitude. Also the high C^+ content with respect to the other species is worth mentioning which is caused by the high star formation rate. The carbon is exposed to the UV radiation coming from the young stellar population and ionized.

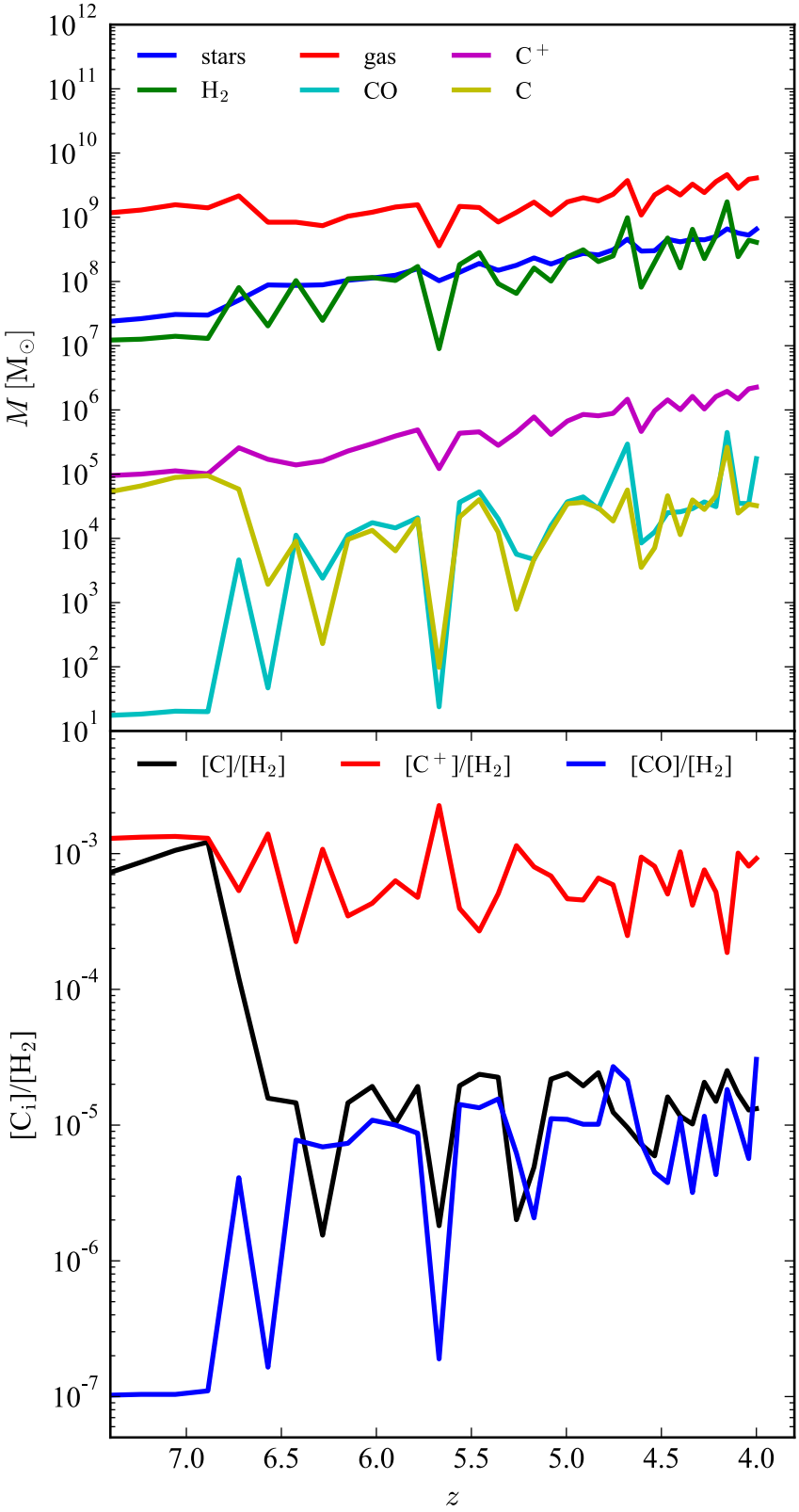

Figure 3.3: Top: The evolution of the different mass components of the simulated galaxy at different redshifts. The gas mass is fluctuating around 10^9 $M_\odot$ whereas the stellar and H_2 mass increase continuously. Their evolution is accompanied by fluctuations as well due to several merging events. The increase of H_2 mass coincides with an increase of C and CO and a decrease of C^+. Bottom: Carbon species abundance relative to the molecular hydrogen mass for different redshifts.

M_{gas}	4.09×10^9 M$_\odot$
M_*	6.52×10^8 M$_\odot$
$M_{\mathrm{H_2}}$	4.08×10^8 M$_\odot$
M_{CO}	1.74×10^5 M$_\odot$
M_{C}	3.23×10^4 M$_\odot$
$M_{\mathrm{C^+}}$	2.23×10^6 M$_\odot$
SFR	32.6 M$_\odot$ yr^{-1}
Z	0.16 Z$_\odot$

Table 3.1: Abundances of the different components of the object and star formation rate by $z = 4$.

We show the evolution of the different components in Fig. 3.3 starting from $z = 7.5$ until $z = 4$. One can see an evolution in all quantities as the halo the galaxy resides in is accumulating gas.

The evolution in H_2 starts at $z \approx 7.5$ around a value of $\approx 1.2 \times 10^7 M_\odot$ and due to a constant accretion of gas, the molecular hydrogen content rises as well until $z \approx 6.72$. The merging with an object, which has roughly 27% of the molecular gas mass and 37% of stellar mass with respect of the object we are tracking, leads to compression of the gas which stimulates the H_2 production and star formation.

This has also consequences for the distribution of the carbon species: the increase in H_2 and density provide the required shielding against the UV radiation to switch on the CO production. M_{CO} increases by nearly three orders of magnitude and M_{C} reduces at the same time. One output later, at $z = 6.57$, the collision causes a disruption of the dense structures and gas loss. The decrease of H_2 leads to a CO and C reduction as well due to the missing shielding coming from H_2, dust, and density.

Until $z \approx 5.78$, oscillations in $M_{\mathrm{H_2}}$, M_{C} and M_{CO} are noticeable which is related to the process of relaxation in the system. In the merging event, the dense center of the galaxy has been split up in two smaller structures that orbit for some time around each other before completing the merging process. At $z \approx 5.67$, all quantities exhibit a drop. The stellar component is less affected, but e.g. the molecular gas content drops by more than an order of magnitude and M_{C} and M_{CO} even by a factor of 10^3. Again, this is caused by a merger that destroys the densest gas structures of that galaxy. However, the gas distribution of that galaxy has already contracted again in the next time step at $z = 5.56$.

During the remaining evolution towards $z = 4$, stellar mass, H_2 mass and the C^+ masses are constantly increasing but they still exhibit fluctuations due to the constant bombardment with smaller objects that destroy the nearly spherical shape which is restored afterwards.

During the whole process between $z = 6.7$ and $z = 4$, the masses of all carbon species are increasing on average. But especially C and CO are very sensitive to the disruption of the gaseous part and the reduction of the H_2 abundance. The fractional abundances relative to H_2 are calculated via $[\mathrm{C}]/[\mathrm{H_2}]=\frac{M_{\mathrm{C}}}{6M_{\mathrm{H_2}}}$, $[\mathrm{C^+}]/[\mathrm{H_2}]=\frac{M_{\mathrm{C^+}}}{6M_{\mathrm{H_2}}}$ and $[\mathrm{CO}]/[\mathrm{H_2}]=\frac{M_{\mathrm{C^+}}}{14M_{\mathrm{H_2}}}$ and are shown in the lower panel of Fig. 3.3. From there, we find that these values stay nearly constant on average. Leaving aside the dips caused by merging events, $[\mathrm{C}]/[\mathrm{H_2}]$ reaches values between $1-2\times10^{-5}$, whereas $[\mathrm{CO}]/[\mathrm{H_2}]$ levels off between 8×10^{-6} and 3×10^{-5}. $[\mathrm{C^+}]/[\mathrm{H_2}]$ fluctuates between 2×10^{-4} and 3×10^{-3}.

Distribution of the tracked species

We present the surface densities for the different species in Fig. 3.4. For this, the content of cells within a predefined volume is integrated along the line of sight.

The middle left panel shows the molecular hydrogen distribution. Compared to the gas distribution,

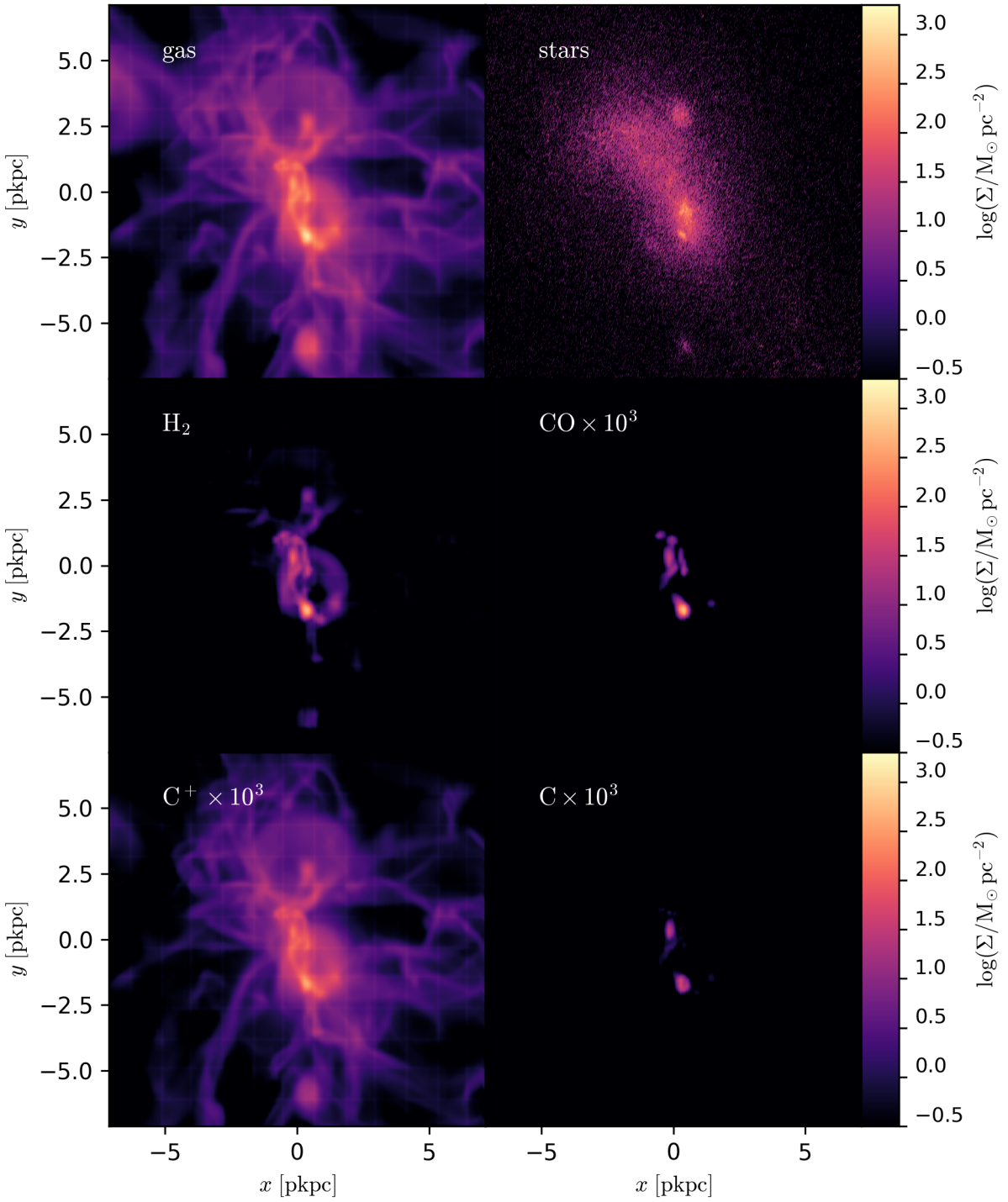

Figure 3.4: Surface densities of the simulated galaxy for gas, stars, H_2 and all tracked carbon species. The spatial scale is shown in units of pkpc (physical coordinates). The content in all cells along the line sight has been ray-traced and summed up. For better visibility, Σ_{C^+}, Σ_{CO} and Σ_C are multiplied by a factor of 10^3. Molecular hydrogen is found in the dense parts of the gas distribution. CO and C are present at high densities of H_2, whereas C^+ is found in the diffuse gas distribution of the galaxy.

one sees that most of the H_2 content sits in the dense gas regions. This is not surprising as in these regions, the shielding against UV radiation is expected to be stronger, and prevents the photodissociation. More importantly, we find that the different carbon species trace different density ranges and consequently, regions of high H_2 abundance and star formation as seen from the middle right and both lower panels.

The relatively low gas surface densities within and around the galaxy are mostly dominated by C^+. These are regions where the lack of shielding leads to quick ionization of atomic carbon. Due to permanent collisions with other objects, the ionized gas is expelled and stripped away. The very dense structures are traced almost equally well by C and CO. These regions coincide with those where H_2 is strongly abundant. Nevertheless, a higher contrast in Σ_{CO} is noticeable compared to Σ_C. Thus, it is also capturing the increasing H_2 gradient towards the center of the galaxy.

In Fig. 3.5, we present the gas-mass weighted carbon and molecular hydrogen content for different radii and densities in the simulated object at $z = 4$. In the upper left panel, we see that the carbon content is dominated by C^+, but towards higher density, the content of ionized carbon shows a decreasing trend. The abundance of molecular hydrogen is less than 10% for $\log(n_H/\mathrm{cm}^{-3}) < 2$, but exceeds this value in the next density bin. At the same time, neutral carbon and also carbon monoxide start to increase due to the present shielding effect from H_2 and the reaction channels for the CO formation that require its presence. In the highest density bin, the H_2 abundance is high enough that CO starts to form. Ionized carbon however, drops down dramatically.

In the upper right panel, we show x_i/x_{tot} versus the distance from the center of the object. Ionized carbon dominates the entire carbon abundance, starts to drop slightly for 0.5 pkpc $< r <$ 1 pkpc since this region starts to be enriched with H_2. In this region, CO and C increase their abundance as well. As one sees from Fig. 3.4, the central region contains more diffuse gas rather than a dense region, which is the explanation why C^+ increases again at larger distances.

From the lower panels, we find that $[C^+]/[H_2]$ ranges from 5×10^{-3} down to 10^{-4} across different densities and radii, whereas $[C]/[H_2]$ increases from 10^{-6} to 3×10^{-5}. In the bin with the highest density, it suddenly drops to 5×10^{-7}. For $[CO]/[H_2]$, it starts at a lower value of 5×10^{-7} at low densities, and after fluctuating around this value it reaches 10^{-4}. For the radial distribution, we find an increase for $[CO]/[H_2]$ and $[C]/[H_2]$ to 5×10^{-5} and 2×10^{-5} which is followed by a steady decrease down to 2×10^{-7} and 10^{-6} respectively. Also here, ionized carbon dominates around a value of 10^{-3}.

3.6 Application of the subgrid model – post-processing of galaxies

To get an estimate for the abundances of the different carbon species for a larger galaxy population, we can post-process one of the outputs of the $D9^+$ run from the previous chapter (see section 2.3.5). There, we run a cosmological simulation with a dynamic H_2 subgrid model in a $(12\,h^{-1}\,\mathrm{Mpc})^3$ box with a maximum physical resolution of 55 pc at $z = 4$. We will not further list any details of the simulation itself, as they are given in full detail in chapter 2.

The carbon subgrid model requires the metallicity, the cell size, the UV field and the density within a cell as an input. As integration time, we choose the age of the universe at $z = 3.6$ which is $\approx$1.7 Gyr. Since the conditions of the ISM (UV, gas density and metallicity) remain unchanged, all reactions reach equilibrium within this integration time. Making use of the galaxy catalogues, we identify all cells belonging to a galaxy and their corresponding values that describe the ISM. The returned abundances for the different species in each cell are summed up and give the masses for carbon (C, C^+ and CO) and molecular hydrogen. These masses of the different galaxies versus the halo mass can be found in appendix D.1.

In Fig. 3.6, we show the abundance ratios for the different carbon species relative to H_2 for halo masses

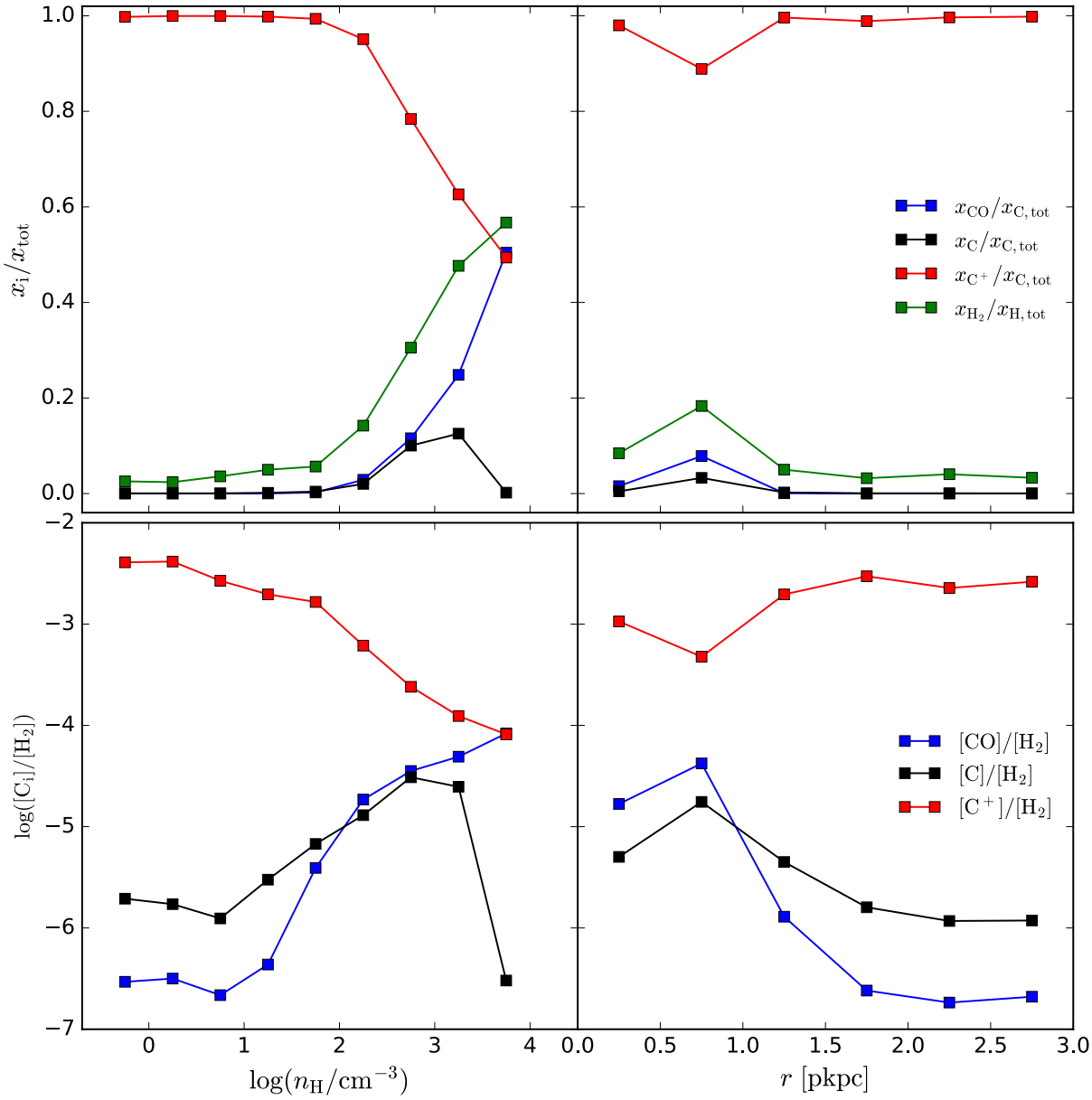

Figure 3.5: Distribution of different species in different density bins and distances from the galaxy center. In the upper two panels, we show the fractional abundance normalized to the carbon atoms for ionized, atomic and molecular carbon in the form of CO, whereas the two lower panels present the abundances relative to molecular hydrogen.

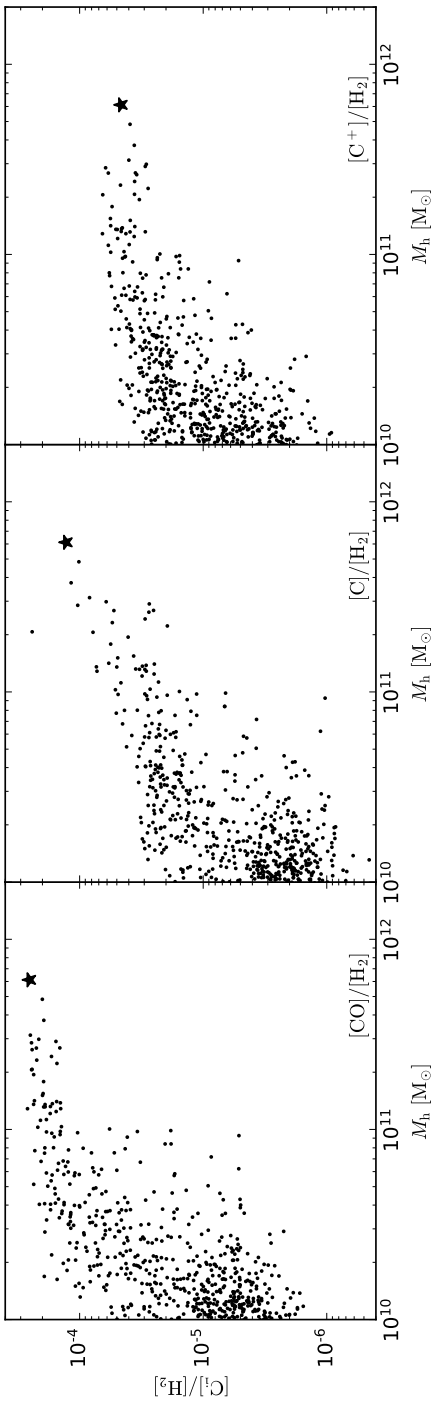

Figure 3.6: Fractional abundances for ionized and neutral carbon and carbon monoxide versus the halo masses. The star highlights the object that will be studied in detail in Fig. 3.7. The different masses for the single species of these objects can be found in Fig. D.1.

of $M_h \geq 10^{10}$ $M_\odot$. For CO, the entire sample exhibits a scatter ranging from 2×10^{-6} to 2×10^{-4}. We find an increase of M_{CO} with growing halo mass and a reduction of the scatter towards $M_h \leq 10^{11}$ $M_\odot$. Beyond this value, [CO]/[H_2] saturates around a value of 2×10^{-4}. The fractional abundances for neutral carbon of the galaxy population start in a range from 6×10^{-7} up to 3×10^{-5}. Towards higher halo masses, the scatter reduces and [C]/[H_2] increases to $1 - 2 \times 10^{-4}$. We find for [C^+]/[H_2] a scatter of more than an order of magnitude for halo masses around 10^{10} $M_\odot$ and a plateau at higher masses in a range of $3 - 7 \times 10^{-5}$. Hence, the galaxy population tends to contain more carbon monoxide than neutral and ionized carbon.

We also show the distributions of CO, C^+ and C for different density bins in the upper left panel of Fig. 3.7. Going from lower densities to higher ones, the amount of H_2 increases due to improved dust and self-shielding. Since the shielding functions for both CO and C depend on N_{H_2} and the formation channels for CO depend on the ionization by cosmic rays reaction with H_2, the drop of C^+ and the increase of CO is consistent. The C fraction however stays nearly constant around a value of 0.15 up to densities of $n_H \sim 10^3$ cm^{-3}. From that density on, neutral carbon is not ionized anymore, but it is directly converted into CO causing the continuous decrease of neutral carbon.

Looking at the radial distribution in the upper right panel, we find a steady decrease of C and C^+ towards the center of the galaxy. In the outmost region, we find an equal fraction of C^+ and CO, but for $r < 4.5$ pkpc, the CO and H_2 exhibit a sudden increase.

For the fraction [CO]/[H_2], we find values of $\sim 3 \times 10^{-5}$ for densities of $n_H \approx 1$ cm^{-3} up to $\sim 3 \times 10^{-4}$ at $n_H \approx 10^4$ cm^{-3}. The latter is also found for the central part of the galaxy (see lower right panel of Fig. 3.7). At the same time [C]/[H_2] drops from 10^{-4} by more than an order of magnitude between $n \approx 1$ cm^{-3} and $n_H \approx 10^4$ cm^{-3}. The fraction [C^+]/[H_2] makes an even more dramatic evolution in this density range starting at 6×10^{-3} and falling down to 3×10^{-6}. For $r < 6$ pkpc, [C^+]/[H_2] and [C]/[H_2] fluctuate between values of $1 - 3 \times 10^{-5}$ and $10^{-5} - 10^{-4}$ respectively.[C^+]/[H_2] is in the outer parts (4.5 pkpc $< r <$ 6 pkpc) more abundant than [C]/[H_2] by ≈ 0.5 dex, but after a sudden drop, both values increase towards 1.3×10^{-5}. Across the entire galaxy,[CO]/[H_2] stays nearly constant around a value of 3×10^{-4}.

In the following, we will compare our obtained values with observational data.

3.7 Discussion

The available amount of observational data for the abundances of C, CO and C^+ is limited, especially for high-redshift objects. Thus, many observational estimates are based on local observations.

The Galactic value found by Frerking et al. (1989) is [C]/[H_2]= 2.2×10^{-5}. Nevertheless, Danielson et al. (2011) present a detailed study of the ISM properties in the lensed galaxy SMM J2135-0102 at $z = 2.3$ and determine a rather high [C]/[H_2] fraction of $3.8 \pm 0.1 \times 10^{-5}$. Weiß et al. (2005) determine for a high-redshift galaxy at $z \sim 2.5$ a value of 5×10^{-5} for [C]/[H_2]. In Harrington et al. (2021), they obtain after a discussion of different models a ratio of [C]/[H_2] of $6.82 \pm 3.04 \times 10^{-5}$. However, the calculation for [C]/[H_2] requires to keep the value of [CO]/[H_2] fixed to $0.5 - 1 \times 10^{-4}$. Jiao et al. (2019) obtain an average value of [C]/[H_2]= $2.5 \pm 1.0 \times 10^{-5}$ from 15 nearby galaxies.

Observations of Galactic clouds by Blake et al. (1987) obtain a value of $5 - 6 \times 10^{-5}$ for [CO]/[H_2]. Israel and Baas (2002) find for the ratio of [C]/[CO] a range between 1 and 3, but in quiescent galaxies, values from 0.3 to 1 for star-forming regions. In Israel, Rosenberg and van der Werf (2015), with far-infrared observations of 76 local starburst galaxies an average value of [CO]/[C]= 8 ± 3 is obtained.

A rare theoretical estimate for ionized carbon is presented in Nordon and Sternberg (2016). The [C^+]/[H_2] fraction is estimated to ≈ 0.4.

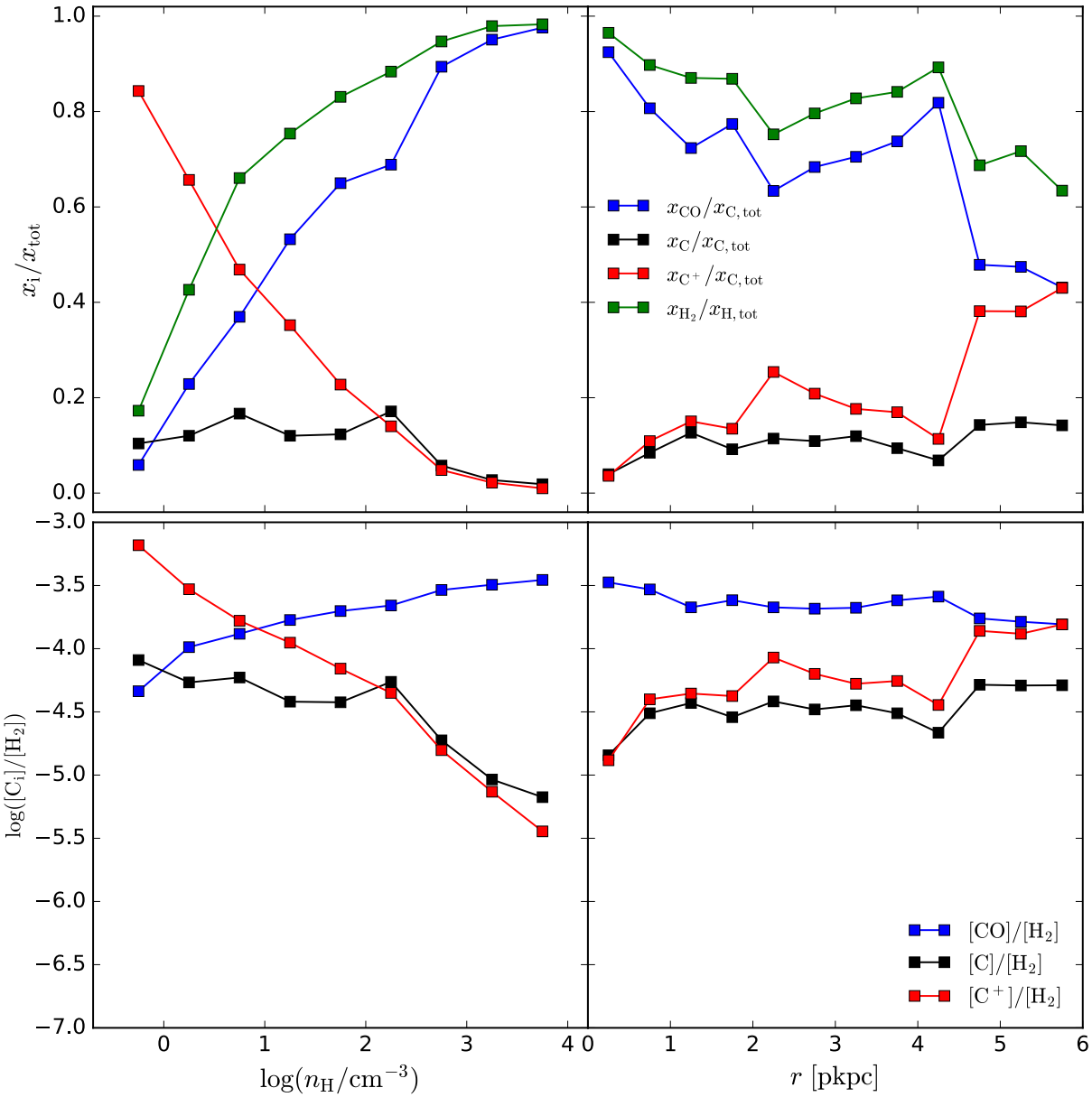

Figure 3.7: Gas mass-weighted fractions of H_2, CO, C and C^+ in the most massive halo of the post-processed galaxy distribution. This object is highlighted by a star symbol in Fig. 3.6. In the upper two panels, we show the fractions for the different carbon species normalized on the total number of carbon atoms and the molecular hydrogen fraction for different densities and across the galaxy. The lower two panels show the relative abundances of CO, C and C^+ with respect to H_2 for different density and radial bins.

From the simulation of the dwarf presented in section 3.5, we find maximum values for [C]/[H_2] and [CO]/[H_2] of $\approx 2 \pm 1.0 \times 10^{-5}$ (see lower panel in Fig. 3.3) which is – given the large range found by observations – in good agreement for neutral carbon. Considering the carbon monoxide abundance, we find that our object exhibits a relatively low abundance in comparison with observational data. Due to the strong interaction with smaller objects that influence the gas distribution significantly by the constant destruction and disruption of very dense regions, the missing shielding leads to a very high abundance fraction of C^+ relative to H_2 between 2×10^{-3} and 1.5×10^{-4}. This value is difficult to evaluate given the sparse observational data. A more massive object with a pronounced and stable bulge might not be that strongly affected by collisions with other objects as our simulated galaxy which is always forced to recover its CO and C abundance afterwards.

Considering the results from the post-processing of galaxies in a cosmological volume presented in section 3.6, we fully ignore the dynamics in the ISM since the assumption is e.g. a static density, constant metallicity or UV field for 1.7 Gyr which does not represent a realistic galaxy. Comparing the H_2 masses in D.1 with those from Fig. 2.3, we see the consequences of this assumption: the scatter in the masses in the latter one is much larger, ranging from 10^7 $M_\odot$ to $M_{H_2} = 3 \times 10^9$ $M_\odot$ for $M_h = 10^{10}$ $M_\odot$. However, for haloes with masses of $M_h > 2 \times 10^{11}$ $M_\odot$, the H_2 masses are very similar between the post-processed and on-the-fly calculated objects and show a deviation of $< 40\%$. We have shown in chapter 2 that these masses are in agreement with observations. The impact of post-processing on the carbon species is difficult to quantify, however we have shown in section 3.5 that the merging of two galaxies and the disrupting of high-density structures leads to a drop in CO and C. H_2 is affected as well, but the influence on carbon species like CO is larger. Thus, we can assume that the CO and C masses of the post-processed objects represent an upper limit since their budget would have been reduced due to the collisions with smaller objects during their formation young, low mass phase. As upper limit for the abundance ratios of CO and C relative to H_2, we find [CO]/[H_2]=10^{-4} and [C]/[H_2]=6×10^{-5} which is consistent with the variety of values found by observations. As an upper limit for the C^+ abundance fraction, we find for these haloes [C^+]/[H_2]=6×10^{-5} but due to the lack of observational data, we can not evaluate how realistic this value is.

3.8 Summary

We presented a new subgrid model for the non-equilibrium chemistry for C, CO and C^+. Our main findings are the following:

(i) Our approximated chemistry gives reasonable results with respect to the abundances of the considered carbon species and molecular hydrogen. For that, we modelled a spherical GMC with a power-law density profile and compared the masses of H_2, C, C^+ and CO after an integration time that corresponds to the typical lifetime of a molecular cloud. Although some deviations at low average densities and low metallicities occur, we find a decent agreement between the network from Glover and Clark (2012a) and our approximated one in the high density region. This is especially important for further considerations since the simulation which uses the subgrid model will resolve large and massive clouds with high densities.

(ii) In section 3.5, we embedded this chemistry in a subgrid model by assuming a log-normal PDF to describe the density distribution of the turbulent ISM within a cell. The subgrid model uses the metal abundance, the strength of the UV field and the density to calculate dynamically the abundances of H_2, C, C^+ and CO. As an application, we simulated a dwarf galaxy until redshift $z = 4$ with a physical resolution of $\Delta x \approx 55$ pc at the end of the run. The on-the-fly use of the subgrid model gives plausible results compared to the available observational data today. We find maximum values for [C]/[H_2] and [CO]/[H_2]

of $\approx 2 \times 10^{-5}$ and 3×10^{-5} at $z = 4$ for our simulated galaxy respectively. Due to several merging events of the galaxy with smaller objects from its environment, we find fluctuations in the abundances of the different carbon species during its evolution.

(iii) In section 3.6, we post-processed the objects from an output of the $D9^+$ simulation at $z = 3.6$ presented in chapter 2 using the subgrid model. For high-mass haloes, the H_2 content of post-processed objects is very similar to the on-the-fly calculations. As we have observed a high sensitivity of the CO and C abundance in a low mass object due to collisions, we conclude that the calculated masses for atomic carbon and carbon monoxide represent upper limits. Thus, we find typical values of $[C]/[H_2]=10^{-4}$ and $[CO]/[H_2]=6 \times 10^{-5}$.

State-of-the art cosmological simulations typically have a spatial resolution of tens of parsecs and at the same time, simulations that include the carbon chemistry require an even higher resolution for an accurate calculation of abundances. The previously presented subgrid model offers the possibility to run cosmological simulations and allows us to study the relation between the molecular gas content, star formation and the luminosities of different tracers on large scales.

CHAPTER 4

Summary & Outlook

In chapter 2, we analyze the outcome of a cosmological simulation with a H_2 subgrid model from Tomassetti et al. (2015) for Ramses. This allows us to investigate the significance of molecular gas in a cosmological context and establishing the connection with observational quantities such as the moelcular gas density and properties of high-redshift galaxies. Additionally to that, the impact of resolution and of different H_2 models on the formation and evolution of galaxies are studied. Our main findings are:

(i) The driving factors for the H_2 formation are density and metallicity. In a resolution study, we find a very similar amount of H_2 in galaxies in high-mass haloes for simulations with a maximum spatial resolution of 55 pc and 110 pc. The requirement for this is both a sufficiently high gas density and metal enrichment at high redshift.

(ii) We compare the outcome of three different H_2 models by applying them in the simulation or on specific haloes. We use the dynamic subgrid model from Tomassetti et al. (2015), where H_2 is actively destroyed, and Krumholz, McKee and Tumlinson (2009), which assumes equilibrium between dust grain formation and UV destruction of H_2, for on-the-fly calculations. The third model is the semi-analytical model from Popping, Behroozi and Peeples (2015) which calculates the H_2 mass directly for a given redshift and halo mass. All models give very similar results for the whole H_2 content for galaxies. For a detailed morphology study for the molecular gas content, the dynamical subgrid model is recommended, since more features are visible in the H_2 maps.

(iii) Finally, we calculate the cosmological H_2 content. Applying the mass limit of state-of-the-art surveys of $M_{H_2} \geq 1 \times 10^{10}\,M_\odot$ to our simulated galaxies exhibits a good agreement of theory and observations. Including objects with $< 1 \times 10^{10}\,M_\odot$, we find an even higher H_2 content that current surveys still miss. Since these objects and the high-gas-density regions are poorly resolved, their actual molecular gas mass is even higher. Therefore, we can give a lower limit for the cosmological H_2 content.

In chapter 3, we develop an extension of the subgrid model for Ramses to include the carbon species C, C^+ and CO as the luminosity coming from these are used in high-z observations to calculate the molecular gas mass. For that, we start from a larger and widely used chemical network by Nelson and Langer (1999), reduce the number of reaction channels and obtain the time differential equations for the carbon species C, CO and C^+. By assuming a lognormal distribution for the density in each resolution element, we are able to scale the chemistry up to scales of tens of parsecs making it applicable to on-the-fly and post-processing calculations for our simulation code. Our main findings are:

(i) We apply the approximated chemistry and the original chemical network from Nelson and Langer (1999) to a simplified spherical cloud. Both show an overall good agreement for metallicities of $>0.1\,Z_\odot$ and average densities of $>10^3$ cm^{-3}. These ranges are of particular interest since most of the molecular gas content is abundant in regions of high density and metal enrichment.

(ii) We run a zoom-in simulation of a galaxy until $z = 4$ in Ramses and apply the subgrid model on-the-fly that is based on the simplified chemistry. During its evolution, the galaxy merges with several objects from its environment. The abundances of C and CO are severely influenced by this as these collisions disrupted high-density regions enabling the UV radiation to photodissociate CO and ionize C. For the abundance fractions, we find [C]/[H_2] and [CO]/[H_2] of $\approx 2 \times 10^{-5}$ at $z = 4$ which seems plausible given the situation of observational data.

(iii) The subgrid model is also used to post-process the galaxy population from chapter 2 at $z = 3.6$. We find that for haloes with $M_h > 2 \times 10^{11}$ $M_\odot$ the H_2 mass is very similar to the one calculated on the fly. Since the merging history of all objects is ignored for the post-processing, we obtain an upper limit for [C]/[H_2] and [CO]/[H_2] of 10^{-4} and 6×10^{-5} respectively which is in agreement with current observational data.

Knowing the abundances of H_2 and its tracers is the first step for luminosity calculations. This requires the solution of the radiative transfer equation. However, there are a lot of subtleties to it making this not a trivial task. For a reliable result, the lack of resolution of our simulation has to be compensated. A good model to describe the temperature and velocity of the different ISM components is necessary. Also, a better dust model is required as its absorption influences significantly the luminosity. On top of that, a detailed resolution study for carbon abundances would be of interest, especially with the focus on luminosity and the inferred molecular gas content. Since nowadays the first choice for the luminosity calculations in cosmological simulations is to post-process their outputs, a test of how accurate this method is would be interesting as well.

If these things from the theoretical and numerical side are clarified, a detailed study of the different conversion factors from a theoretical point of view is possible. This includes metallicity and redshift dependence, and even an evaluation of the reliability of the different tracers that have been discussed here.

Bibliography

Aarseth, S. J. (1963), *Dynamical evolution of clusters of galaxies, I*, Mon. Not. R. Astron. Soc. **126** 223 (cit. on p. 23).

Abbott, B. P. et al. (2016), *Observation of Gravitational Waves from a Binary Black Hole Merger*, **116**, 061102 061102, arXiv: 1602.03837 [gr-qc] (cit. on p. 3).

Abdo, A. A. et al. (2010), *Fermi Observations of Cassiopeia and Cepheus: Diffuse Gamma-ray Emission in the Outer Galaxy*, Astrophys. J. **710** 133, arXiv: 0912.3618 [astro-ph.HE] (cit. on p. 13).

Abraham, R. G. (1998), *Perspectives in Physical Morphology*, arXiv e-prints, astro-ph/9809131 astro, arXiv: astro-ph/9809131 [astro-ph] (cit. on p. 7).

Agertz, O. and A. V. Kravtsov (2016), *The Impact of Stellar Feedback on the Structure, Size, and Morphology of Galaxies in Milky-Way-sized Dark Matter Halos*, Astrophys. J. **824**, 79 79, arXiv: 1509.00853 [astro-ph.GA] (cit. on p. 37).

Allen, R. J., D. E. Hogg and P. D. Engelke (2015), *The Structure of Dark Molecular Gas in the Galaxy. I. A Pilot Survey for 18 cm OH Emission Toward l≈ 105°, b≈ +1°*, **149**, 123 123, arXiv: 1502.00657 [astro-ph.GA] (cit. on p. 14).

Alpher, R. A., H. Bethe and G. Gamow (1948), *The Origin of Chemical Elements*, Physical Review **73** 803 (cit. on p. 2).

Alpher, R. A. and R. Herman (1948), *Evolution of the Universe*, Nature **162** (4124) 774 (cit. on p. 2).

Angulo, R. E. et al. (2017), *Earth-mass haloes and the emergence of NFW density profiles*, Mon. Not. R. Astron. Soc. **471** 4687, arXiv: 1604.03131 [astro-ph.CO] (cit. on p. 38).

Baczynski, C., S. C. O. Glover and R. S. Klessen (2015), *Fervent: chemistry-coupled, ionizing and non-ionizing radiative feedback in hydrodynamical simulations*, Mon. Not. R. Astron. Soc. **454** 380, arXiv: 1503.08987 [astro-ph.IM] (cit. on p. 34).

Barnard, E. E. (1892), *Nova Aurigæ*, Mon. Not. R. Astron. Soc. **53** 35 (cit. on p. 19).

Barnes, J. and P. Hut (1986), *A hierarchical O(N log N) force-calculation algorithm*, **324** 446 (cit. on p. 23).

Behroozi, P. S., R. H. Wechsler and C. Conroy (2013), *The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8*, Astrophys. J. **770**, 57 57, arXiv: 1207.6105 [astro-ph.CO] (cit. on pp. 34, 50, 51).

Bigiel, F., A. K. Leroy, F. Walter, E. Brinks, W. J. G. de Blok, C. Kramer et al. (2011), *A Constant Molecular Gas Depletion Time in Nearby Disk Galaxies*, Astrophys. J. Lett **730**, L13 L13, arXiv: 1102.1720 [astro-ph.CO] (cit. on p. 20).

Bigiel, F., A. K. Leroy, F. Walter, E. Brinks, W. J. G. de Blok, B. Madore et al. (2008), *The Star Formation Law in Nearby Galaxies on Sub-Kpc Scales*, **136** 2846, arXiv: 0810.2541 [astro-ph] (cit. on pp. 20, 31, 34).

Bisbas, T. G., T. A. Bell et al. (2012), *3D-PDR: a new three-dimensional astrochemistry code for treating photodissociation regions*, Mon. Not. R. Astron. Soc. **427** 2100, arXiv: 1209.1091 [astro-ph.SR] (cit. on p. 60).

Bisbas, T. G., A. Schruba and E. F. van Dishoeck (2019), *Simulating the atomic and molecular content of molecular clouds using probability distributions of physical parameters*, Mon. Not. R. Astron. Soc. **485** 3097, arXiv: 1901.11306 [astro-ph.GA] (cit. on p. 60).

Bisbas, T. G., J. C. Tan and K. E. I. Tanaka (2021), *Photodissociation region diagnostics across galactic environments*, Mon. Not. R. Astron. Soc. **502** 2701, arXiv: 2012.06773 [astro-ph.GA] (cit. on p. 60).

Black, J. H. and E. F. van Dishoeck (1987), *Fluorescent Excitation of Interstellar H 2*, Astrophys. J. **322** 412 (cit. on p. 32).

Blake, G. A. et al. (1987), *Molecular Abundances in OMC-1: The Chemical Composition of Interstellar Molecular Clouds and the Influence of Massive Star Formation*, Astrophys. J. **315** 621 (cit. on p. 76).

Blitz, L. and E. Rosolowsky (2004), *The Role of Pressure in Giant Molecular Cloud Formation*, Astrophys. J. Lett **612** L29, arXiv: astro-ph/0407492 [astro-ph] (cit. on p. 32).

– (2006), *The Role of Pressure in GMC Formation II: The H_2-Pressure Relation*, Astrophys. J. **650** 933, arXiv: astro-ph/0605035 [astro-ph] (cit. on pp. 32, 34).

Blumenthal, G. R. et al. (1984), *Formation of galaxies and large-scale structure with cold dark matter.*, **311** 517 (cit. on p. 7).

Bok, B. J. and E. F. Reilly (1947), *Small Dark Nebulae.*, Astrophys. J. **105** 255 (cit. on p. 19).

Bolatto, A. D., M. Wolfire and A. K. Leroy (2013), *The CO-to-H_2 Conversion Factor*, Annual Review of Astronomy and Astrophysics **51** 207 (cit. on pp. 12, 13, 60).

Bolatto, A. D., A. K. Leroy et al. (2011), *The State of the Gas and the Relation between Gas and Star Formation at Low Metallicity: The Small Magellanic Cloud*, Astrophys. J. **741**, 12 12, arXiv: 1107.1717 [astro-ph.CO] (cit. on p. 32).

Bothwell, M. S. et al. (2017), *ALMA observations of atomic carbon in $z \sim 4$ dusty star-forming galaxies*, Mon. Not. R. Astron. Soc. **466** 2825, arXiv: 1612.04380 [astro-ph.GA] (cit. on pp. 31, 49, 59).

Bovino, S. et al. (2014), *Formation of Carbon-enhanced Metal-poor Stars in the Presence of Far-ultraviolet Radiation*, Astrophys. J. Lett **790**, L35 L35, arXiv: 1406.4450 [astro-ph.SR] (cit. on p. 60).

Brown, R. L. and P. A. Vanden Bout (1991), *CO Emission at z=2.2867 in the Galaxy IRAS F10214+4724*, **102** 1956 (cit. on p. 12).

Bryan, G. L. and M. L. Norman (1998), *Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons*, Astrophys. J. **495** 80 (cit. on p. 24).

Capak, P. L. et al. (2015), *Galaxies at redshifts 5 to 6 with systematically low dust content and high [C II] emission*, **522** 455, arXiv: 1503.07596 [astro-ph.GA] (cit. on pp. 31, 49).

Carilli, C. L., E. Daddi et al. (2010), *Imaging the Molecular Gas in a Submillimeter Galaxy at z = 4.05: Cold Mode Accretion or a Major Merger?*, Astrophys. J. **714** 1407, arXiv: 1002.3838 [astro-ph.CO] (cit. on p. 49).

Carilli, C. L. and F. Walter (2013), *Cool Gas in High-Redshift Galaxies*, **51** 105, arXiv: 1301.0371 [astro-ph.CO] (cit. on pp. 13, 31).

Cassata, P. et al. (2020), *ALMA Reveals the Molecular Gas Properties of Five Star-forming Galaxies across the Main Sequence at 3*, Astrophys. J. **891**, 83 83, arXiv: 2002.04040 [astro-ph.GA] (cit. on p. 49).

Cazaux, S. and M. Spaans (2004), *Molecular Hydrogen Formation on Dust Grains in the High-Redshift Universe*, Astrophys. J. **611** 40, arXiv: astro-ph/0404340 [astro-ph] (cit. on p. 101).

Cen, R. and J. P. Ostriker (1993), *Cold Dark Matter Cosmogony with Hydrodynamics and Galaxy Formation: Galaxy Properties at Redshift Zero*, Astrophys. J. **417** 415 (cit. on p. 24).

Chen, L.-H. et al. (2018), *Populating H_2 and CO in galaxy simulation with dust evolution*, Mon. Not. R. Astron. Soc. **474** 1545, arXiv: 1711.00600 [astro-ph.GA] (cit. on p. 32).

Christensen, C. et al. (2012), *Implementing molecular hydrogen in hydrodynamic simulations of galaxy formation*, Mon. Not. R. Astron. Soc. **425** 3058, arXiv: 1205.5567 [astro-ph.CO] (cit. on pp. 33, 35).

Clark, P. C. et al. (2019), *Tracing the formation of molecular clouds via [C II], [C I], and CO emission*, Mon. Not. R. Astron. Soc. **486** 4622, arXiv: 1809.00489 [astro-ph.GA] (cit. on p. 60).

Cortzen, I. et al. (2019), *PAHs as tracers of the molecular gas in star-forming galaxies*, Mon. Not. R. Astron. Soc. **482** 1618, arXiv: 1810.05178 [astro-ph.GA] (cit. on p. 14).

Cox, P. et al. (2011), *Gas and Dust in a Submillimeter Galaxy at z = 4.24 from the Herschel Atlas*, Astrophys. J. **740**, 63 63, arXiv: 1107.2924 [astro-ph.CO] (cit. on pp. 49, 59).

Daddi, E. et al. (2015), *CO excitation of normal star-forming galaxies out to z = 1.5 as regulated by the properties of their interstellar medium*, Astron. Astrophys. **577**, A46 A46, arXiv: 1409.8158 [astro-ph.GA] (cit. on p. 50).

Dame, T. M., D. Hartmann and P. Thaddeus (2001), *The Milky Way in Molecular Clouds: A New Complete CO Survey*, Astrophys. J. **547** 792, arXiv: astro-ph/0009217 [astro-ph] (cit. on p. 13).

Danielson, A. L. R. et al. (2011), *The properties of the interstellar medium within a star-forming galaxy at z= 2.3*, Mon. Not. R. Astron. Soc. **410** 1687, arXiv: 1008.3183 [astro-ph.CO] (cit. on p. 76).

Davé, R., R. Thompson and P. F. Hopkins (2016), *MUFASA: galaxy formation simulations with meshless hydrodynamics*, Mon. Not. R. Astron. Soc. **462** 3265, arXiv: 1604.01418 [astro-ph.GA] (cit. on pp. 32, 34, 35).

de los Reyes, M. A. C. and J. Kennicutt Robert C. (2019), *Revisiting the Integrated Star Formation Law. I. Non-starbursting Galaxies*, Astrophys. J. **872**, 16 16, arXiv: 1901.01283 [astro-ph.GA] (cit. on p. 20).

De Lucia, G. and J. Blaizot (2007), *The hierarchical formation of the brightest cluster galaxies*, Mon. Not. R. Astron. Soc. **375** 2, arXiv: astro-ph/0606519 [astro-ph] (cit. on p. 53).

Decarli, R., F. Walter, M. Aravena et al. (2016), *ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO Luminosity Functions and the Evolution of the Cosmic Density of Molecular Gas*, Astrophys. J. **833**, 69 69, arXiv: 1607.06770 [astro-ph.GA] (cit. on pp. 31, 50, 52, 54).

Decarli, R., F. Walter, J. Gónzalez-López et al. (2019), *The ALMA Spectroscopic Survey in the HUDF: CO Luminosity Functions and the Molecular Gas Content of Galaxies through Cosmic History*, Astrophys. J. **882**, 138 138, arXiv: 1903.09164 [astro-ph.GA] (cit. on pp. 50, 52–56).

Dessauges-Zavadsky, M. et al. (2017), *Molecular gas properties of a lensed star-forming galaxy at z 3.6: a case study*, Astron. Astrophys. **605**, A81 A81, arXiv: 1610.08065 [astro-ph.GA] (cit. on p. 49).

Dicke, R. H. et al. (1965), *Cosmic Black-Body Radiation.*, Astrophys. J. **142** 414 (cit. on p. 2).

Diemer, B. et al. (2018), *Modeling the Atomic-to-molecular Transition in Cosmological Simulations of Galaxy Formation*, Astrophys. J. Suppl. Ser. **238**, 33 33, arXiv: 1806.02341 [astro-ph.GA] (cit. on p. 32).

Downes, D. and P. M. Solomon (1998), *Rotating Nuclear Rings and Extreme Starbursts in Ultraluminous Galaxies*, Astrophys. J. **507** 615, arXiv: astro-ph/9806377 [astro-ph] (cit. on p. 13).

Draine, B. T. (1978), *Photoelectric heating of interstellar gas.*, Astrophys. J. Suppl. Ser. **36** 595 (cit. on p. 32).

– (2003), *Scattering by Interstellar Dust Grains. I. Optical and Ultraviolet*, Astrophys. J. **598** 1017, arXiv: astro-ph/0304060 [astro-ph] (cit. on p. 101).

– (2011), *Physics of the Interstellar and Intergalactic Medium* (cit. on p. 10).

Draine, B. T. and F. Bertoldi (1996), *Structure of Stationary Photodissociation Fronts*, Astrophys. J. **468** 269, arXiv: astro-ph/9603032 [astro-ph] (cit. on pp. 32, 62, 101).

Draine, B. T., D. A. Dale et al. (2007), *Dust Masses, PAH Abundances, and Starlight Intensities in the SINGS Galaxy Sample*, Astrophys. J. **663** 866, arXiv: astro-ph/0703213 [astro-ph] (cit. on p. 36).

Dutton, A. A. and F. C. van den Bosch (2009), *The impact of feedback on disc galaxy scaling relations*, Mon. Not. R. Astron. Soc. **396** 141, arXiv: 0810.4963 [astro-ph] (cit. on p. 32).

Efstathiou, G. and J. W. Eastwood (1981), *On the clustering of particles in an expanding universe*, Mon. Not. R. Astron. Soc. **194** 503 (cit. on p. 23).

Einstein, A. (1905a), *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*, Annalen der Physik **323** 639 (cit. on p. 1).

– (1905b), *Zur Elektrodynamik bewegter Körper*, Annalen der Physik **322** 891 (cit. on p. 1).

Einstein, A. (1915), *Die Feldgleichungen der Gravitation*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin 844 (cit. on p. 1).

Elmegreen, B. G. (2007), *On the Rapid Collapse and Evolution of Molecular Clouds*, Astrophys. J. **668** 1064, arXiv: 0707.2252 [astro-ph] (cit. on p. 31).

Engelke, P. D. and R. J. Allen (2019), *OH as an Alternate Tracer for Molecular Gas: Quantity and Structure of Molecular Gas in W5*, Astrophys. J. **874**, 49 49, arXiv: 1901.02515 [astro-ph.GA] (cit. on p. 14).

Feldmann, R., N. Y. Gnedin and A. V. Kravtsov (2011), *How Universal is the Σ_{SFR}–Σ_{H_2} Relation?*, Astrophys. J. **732**, 115 115, arXiv: 1010.1539 [astro-ph.CO] (cit. on p. 33).

Ferland, G. J. et al. (2017), *The 2017 Release Cloudy*, **53** 385, arXiv: 1705.10877 [astro-ph.GA] (cit. on p. 60).

Ferrière, K. M. (2001), *The interstellar environment of our galaxy*, Rev. Mod. Phys. **73** (4) 1031, URL: https://link.aps.org/doi/10.1103/RevModPhys.73.1031 (cit. on p. 9).

Förster Schreiber, N. M. and S. Wuyts (2020), *Star-Forming Galaxies at Cosmic Noon*, **58** 661, arXiv: 2010.10171 [astro-ph.GA] (cit. on p. 8).

Frerking, M. A. et al. (1989), *The Abundances of Atomic Carbon and Carbon Monoxide Compared with Visual Extinction in the Ophiuchus Molecular Cloud Complex*, Astrophys. J. **344** 311 (cit. on p. 76).

Friedman, A. (1922), *Über die Krümmung des Raumes*, Zeitschrift für Physik **10** 377 (cit. on p. 2).

Fu, J., Q. Guo et al. (2010), *The atomic-to-molecular transition and its relation to the scaling properties of galaxy discs in the local Universe*, Mon. Not. R. Astron. Soc. **409** 515, arXiv: 1004.2325 [astro-ph.CO] (cit. on pp. 32, 34).

Fu, J., G. Kauffmann et al. (2012), *The effect of star formation on the redshift evolution of interstellar metals, atomic and molecular gas in galaxies*, Mon. Not. R. Astron. Soc. **424** 2701, arXiv: 1203.5280 [astro-ph.GA] (cit. on pp. 32, 34).

Fumagalli, M., M. R. Krumholz and L. K. Hunt (2010), *Testing Models for Molecular Gas Formation in Galaxies: Hydrostatic Pressure or Gas and Dust Shielding?*, Astrophys. J. **722** 919, arXiv: 1008.4137 [astro-ph.CO] (cit. on p. 32).

Gao, Y. and P. M. Solomon (2004), *The Star Formation Rate and Dense Molecular Gas in Galaxies*, Astrophys. J. **606** 271, arXiv: astro-ph/0310339 [astro-ph] (cit. on p. 14).

Gelb, J. M. and E. Bertschinger (1994), *Cold Dark Matter. II. Spatial and Velocity Statistics*, Astrophys. J. **436** 491 (cit. on p. 23).

Gill, S. P. D., A. Knebe and B. K. Gibson (2004), *The evolution of substructure - I. A new identification method*, Mon. Not. R. Astron. Soc. **351** 399, arXiv: astro-ph/0404258 [astro-ph] (cit. on pp. 38, 69).

Gingold, R. A. and J. J. Monaghan (1977), *Smoothed particle hydrodynamics: theory and application to non-spherical stars*, Mon. Not. R. Astron. Soc. **181** 375 (cit. on p. 24).

Ginolfi, M. et al. (2017), *Molecular gas on large circumgalactic scales at z = 3.47*, Mon. Not. R. Astron. Soc. **468** 3468, arXiv: 1611.07026 [astro-ph.GA] (cit. on p. 49).
Girichidis, P. et al. (2016), *The SILCC (SImulating the LifeCycle of molecular Clouds) project - II. Dynamical evolution of the supernova-driven ISM and the launching of outflows*, Mon. Not. R. Astron. Soc. **456** 3432, arXiv: 1508.06646 [astro-ph.GA] (cit. on p. 60).
Glover, S. C. O. and P. C. Clark (2012a), *Approximations for modelling CO chemistry in giant molecular clouds: a comparison of approaches*, Mon. Not. R. Astron. Soc. **421** 116, arXiv: 1102.0670 [astro-ph.GA] (cit. on pp. 60, 61, 64, 67, 78, 103).
– (2012b), *Is molecular gas necessary for star formation?*, Mon. Not. R. Astron. Soc. **421** 9, arXiv: 1105.3073 [astro-ph.GA] (cit. on pp. 20, 31).
– (2012c), *Star formation in metal-poor gas clouds*, Mon. Not. R. Astron. Soc. **426** 377, arXiv: 1203.4251 [astro-ph.GA] (cit. on p. 20).
Glover, S. C. O., C. Federrath et al. (2010), *Modelling CO formation in the turbulent interstellar medium*, Mon. Not. R. Astron. Soc. **404** 2, arXiv: 0907.4081 [astro-ph.SR] (cit. on pp. 19, 60, 64).
Glover, S. C. O. and M.-M. Mac Low (2007a), *Simulating the Formation of Molecular Clouds. I. Slow Formation by Gravitational Collapse from Static Initial Conditions*, Astrophys. J. Suppl. Ser. **169** 239, arXiv: astro-ph/0605120 [astro-ph] (cit. on pp. 35, 60, 64, 101).
– (2007b), *Simulating the Formation of Molecular Clouds. II. Rapid Formation from Turbulent Initial Conditions*, Astrophys. J. **659** 1317, arXiv: astro-ph/0605121 [astro-ph] (cit. on pp. 60, 64).
Gnedin, N. Y. and L. Hui (1998), *Probing the Universe with the Lyalpha forest - I. Hydrodynamics of the low-density intergalactic medium*, Mon. Not. R. Astron. Soc. **296** 44 (cit. on p. 24).
Gnedin, N. Y. and A. V. Kravtsov (2011), *Environmental Dependence of the Kennicutt-Schmidt Relation in Galaxies*, Astrophys. J. **728**, 88 88, arXiv: 1004.0003 [astro-ph.CO] (cit. on pp. 34, 37).
Gnedin, N. Y., K. Tassis and A. V. Kravtsov (2009), *Modeling Molecular Hydrogen and Star Formation in Cosmological Simulations*, Astrophys. J. **697** 55, arXiv: 0810.4148 [astro-ph] (cit. on pp. 33, 35, 37).
Goicoechea, J. R. et al. (2019), *Molecular tracers of radiative feedback in Orion (OMC-1). Widespread CH^+ (J = 1-0), CO (10-9), HCN (6-5), and HCO^+ (6-5) emission*, Astron. Astrophys. **622**, A91 A91, arXiv: 1812.00821 [astro-ph.GA] (cit. on p. 14).
Gómez-Guijarro, C. et al. (2018), *Starburst to Quiescent from HST/ALMA: Stars and Dust Unveil Minor Mergers in Submillimeter Galaxies at z 4.5*, Astrophys. J. **856**, 121 121, arXiv: 1802.07751 [astro-ph.GA] (cit. on p. 49).
Gong, M., E. C. Ostriker and M. Wolfire (2017), *A Simple and Accurate Network for Hydrogen and Carbon Chemistry in the Interstellar Medium*, Astrophys. J. **843**, 38 38, arXiv: 1610.09023 [astro-ph.GA] (cit. on pp. 19, 60).
Gott, J. R., E. L. Turner and S. J. Aarseth (1979), *N-body simulations of galaxy clustering. III - The covariance function*, Astrophys. J. **234** 13 (cit. on p. 23).
Governato, F. et al. (1999), *Properties of galaxy clusters: mass and correlation functions*, Mon. Not. R. Astron. Soc. **307** 949 (cit. on p. 23).
Graciá-Carpio, J. et al. (2006), *Is HCN a True Tracer of Dense Molecular Gas in Luminous and Ultraluminous Infrared Galaxies?*, Astrophys. J. Lett **640** L135, arXiv: astro-ph/0602339 [astro-ph] (cit. on p. 14).
Grassi, T. et al. (2014), *KROME - a package to embed chemistry in astrophysical simulations*, Mon. Not. R. Astron. Soc. **439** 2386, arXiv: 1311.1070 [astro-ph.GA] (cit. on p. 60).
Gray, W. J. and E. Scannapieco (2013), *Thermal and Chemical Evolution of Collapsing Filaments*, Astrophys. J. **768**, 174 174, arXiv: 1304.0771 [astro-ph.CO] (cit. on p. 60).

Greengard, L. and V. Rokhlin (1986), *A fast algorithm for particle simulations*, Computational Physics **73** 325 (cit. on p. 23).

Haardt, F. and P. Madau (2012), *Radiative Transfer in a Clumpy Universe. IV. New Synthesis Models of the Cosmic UV/X-Ray Background*, Astrophys. J. **746**, 125 125, arXiv: 1105.2039 [astro-ph.CO] (cit. on pp. 37, 69).

Hahn, O. and T. Abel (2011), *Multi-scale initial conditions for cosmological simulations*, Mon. Not. R. Astron. Soc. **415** 2101, arXiv: 1103.6031 [astro-ph.CO] (cit. on pp. 29, 36, 69).

Harrington, K. C. et al. (2021), *Turbulent Gas in Lensed Planck-selected Starbursts at $z \sim 1$-3.5*, Astrophys. J. **908**, 95 95, arXiv: 2010.16231 [astro-ph.GA] (cit. on p. 76).

Heays, A. N., A. D. Bosman and E. F. van Dishoeck (2017), *Photodissociation and photoionisation of atoms and molecules of astrophysical interest*, Astron. Astrophys. **602**, A105 A105, arXiv: 1701.04459 [astro-ph.SR] (cit. on p. 100).

Herschel, W. (1811), *Astronomical Observations Relating to the Construction of the Heavens, Arranged for the Purpose of a Critical Examination, the Result of Which Appears to Throw Some New Light upon the Organization of the Celestial Bodie*, Philosophical Transactions of the Royal Society of London **101** 269 (cit. on p. 19).

Hoag, A. et al. (2017), *Spectroscopic confirmation of an ultra-faint galaxy at the epoch of reionization*, Nature Astronomy **1**, 0091 0091, arXiv: 1704.02970 [astro-ph.GA] (cit. on p. 49).

Hockney, R. W. and J. W. Eastwood (1981), *Computer Simulation Using Particles* (cit. on p. 23).

Hollenbach, D. and C. F. McKee (1979), *Molecule formation and infrared emission in fast interstellar shocks. I. Physical processes.*, Astrophys. J. Suppl. Ser. **41** 555 (cit. on p. 32).

Holmberg, E. (1941), *On the Clustering Tendencies among the Nebulae. II. a Study of Encounters Between Laboratory Models of Stellar Systems by a New Integration Procedure*, Astrophys. J. **1** 385 (cit. on p. 23).

Hopkins, P. F. et al. (2014), *Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation*, Mon. Not. R. Astron. Soc. **445** 581, arXiv: 1311.2073 [astro-ph.CO] (cit. on pp. 32, 34).

Hu, C.-Y. et al. (2016), *Star formation and molecular hydrogen in dwarf galaxies: a non-equilibrium view*, Mon. Not. R. Astron. Soc. **458** 3528, arXiv: 1510.05644 [astro-ph.GA] (cit. on p. 33).

Hubble, E. (1927), *The classification of spiral nebulae*, The Observatory **50** 276 (cit. on p. 7).

– (1929), *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, Proceedings of the National Academy of Science **15** 168 (cit. on p. 2).

Hubble, E. and M. L. Humason (1931), *The Velocity-Distance Relation among Extra-Galactic Nebulae*, Philosophical Transactions of the Royal Society of London **74** 43 (cit. on p. 2).

Hultman, J. and A. Pharasyn (1999), *Hierarchical, dissipative formation of elliptical galaxies: is thermal instability the key mechanism?. Hydrodynamical simulations including supernova feedback, multi-phase gas and metal enrichment in CDM: structure and dynamics of elliptical galaxies*, Astron. Astrophys. **347** 769 (cit. on p. 24).

Inoue, S., N. Yoshida and H. Yajima (2020), *The CO universe: modelling CO emission and H_2 abundance in cosmological galaxy formation simulations*, Mon. Not. R. Astron. Soc. **498** 5960, arXiv: 2008.12484 [astro-ph.GA] (cit. on p. 61).

Israel, F. P. (2020), *Central molecular zones in galaxies: ^{12}CO-to-^{13}CO ratios, carbon budget, and X factors*, Astron. Astrophys. **635**, A131 A131, arXiv: 2001.04975 [astro-ph.GA] (cit. on p. 14).

Israel, F. P. and F. Baas (2002), *Neutral atomic carbon in centers of galaxies*, Astron. Astrophys. **383** 82, arXiv: astro-ph/0112187 [astro-ph] (cit. on p. 76).

Israel, F. P., M. J. F. Rosenberg and P. van der Werf (2015), *Neutral carbon and CO in 76 (U)LIRGs and starburst galaxy centers. A method to determine molecular gas properties in luminous galaxies*, Astron. Astrophys. **578**, A95 A95, arXiv: 1504.08005 [astro-ph.GA] (cit. on p. 76).

Jaacks, J., R. Thompson and K. Nagamine (2013), *Impact of H_2-based Star Formation Model on the z >= 6 Luminosity Function and the Ionizing Photon Budget for Reionization*, Astrophys. J. **766**, 94 94, arXiv: 1301.5270 [astro-ph.CO] (cit. on pp. 32, 34, 57).

Jeans, J. H. (1902), *The Stability of a Spherical Nebula*, Philosophical Transactions of the Royal Society A. **199** 312 (cit. on p. 22).

Jeffreson, S. M. R. and J. M. D. Kruijssen (2018), *A general theory for the lifetimes of giant molecular clouds under the influence of galactic dynamics*, Mon. Not. R. Astron. Soc. **476** 3688, arXiv: 1803.01850 [astro-ph.GA] (cit. on p. 64).

Jenkins, A. et al. (1998), *Evolution of Structure in Cold Dark Matter Universes*, Astrophys. J. **499** 20 (cit. on p. 23).

Jiao, Q. et al. (2019), *Resolved Neutral Carbon Emission in Nearby Galaxies: [C I] Lines as Total Molecular Gas Tracers*, Astrophys. J. **880**, 133 133, arXiv: 1906.05671 [astro-ph.GA] (cit. on p. 76).

Jiménez-Andrade, E. F. et al. (2018), *Molecular gas in AzTEC/C159: a star-forming disk galaxy 1.3 Gyr after the Big Bang*, Astron. Astrophys. **615**, A25 A25, arXiv: 1710.10181 [astro-ph.GA] (cit. on p. 49).

Joshi, P. R. et al. (2019), *On the resolution requirements for modelling molecular gas formation in solar neighbourhood conditions*, Mon. Not. R. Astron. Soc. **484** 1735, arXiv: 1812.03725 [astro-ph.GA] (cit. on p. 60).

Kainulainen, J. et al. (2009), *Probing the evolution of molecular cloud structure. From quiescence to birth*, Astron. Astrophys. **508** L35, arXiv: 0911.5648 [astro-ph.GA] (cit. on pp. 35, 67).

Katz, H. et al. (2017), *Interpreting ALMA observations of the ISM during the epoch of reionization*, Mon. Not. R. Astron. Soc. **468** 4831, arXiv: 1612.01786 [astro-ph.GA] (cit. on p. 33).

Kauffmann, J. et al. (2017), *Molecular Line Emission as a Tool for Galaxy Observations (LEGO). I. HCN as a tracer of moderate gas densities in molecular clouds and galaxies*, Astron. Astrophys. **605**, L5 L5, arXiv: 1707.05352 [astro-ph.GA] (cit. on p. 14).

Kennicutt Robert C., J. (1989), *The Star Formation Law in Galactic Disks*, Astrophys. J. **344** 685 (cit. on p. 37).

– (1998), *The Global Schmidt Law in Star-forming Galaxies*, Astrophys. J. **498** 541, arXiv: astro-ph/9712213 [astro-ph] (cit. on pp. 20, 21, 37).

Kennicutt Robert C., J., D. Calzetti et al. (2007), *Star Formation in NGC 5194 (M51a). II. The Spatially Resolved Star Formation Law*, Astrophys. J. **671** 333, arXiv: 0708.0922 [astro-ph] (cit. on pp. 20, 31).

Kennicutt Robert C., J. and M. A. C. de los Reyes (2021), *Revisiting the Integrated Star Formation Law. II. Starbursts and the Combined Global Schmidt Law*, Astrophys. J. **908**, 61 61, arXiv: 2012.05363 [astro-ph.GA] (cit. on p. 20).

Keres, D., M. S. Yun and J. S. Young (2003), *CO Luminosity Functions for Far-Infrared- and B-Band-selected Galaxies and the First Estimate for Ω_{HI}+H_2*, Astrophys. J. **582** 659, arXiv: astro-ph/0209413 [astro-ph] (cit. on p. 54).

Keto, E. and P. Caselli (2008), *The Different Structures of the Two Classes of Starless Cores*, Astrophys. J. **683** 238, arXiv: 0804.0822 [astro-ph] (cit. on p. 67).

Knollmann, S. R. and A. Knebe (2009), *AHF: Amiga's Halo Finder*, Astrophys. J. Suppl. Ser. **182** 608, arXiv: 0904.3662 [astro-ph.CO] (cit. on pp. 38, 69).

Krips, M. et al. (2016), *ACA [CI] observations of the starburst galaxy NGC 253*, Astron. Astrophys. **592**, L3 L3 (cit. on p. 14).

Kroupa, P. (2001), *On the variation of the initial mass function*, Mon. Not. R. Astron. Soc. **322** 231, arXiv: astro-ph/0009005 [astro-ph] (cit. on pp. 37, 69).

Krumholz, M. R. (2015), *Notes on Star Formation*, arXiv e-prints, arXiv:1511.03457 arXiv:1511.03457, arXiv: 1511.03457 [astro-ph.GA] (cit. on p. 19).

Krumholz, M. R. and A. Dekel (2012), *Metallicity-dependent Quenching of Star Formation at High Redshift in Small Galaxies*, Astrophys. J. **753**, 16 16, arXiv: 1106.0301 [astro-ph.CO] (cit. on pp. 32, 34).

Krumholz, M. R. and N. Y. Gnedin (2011), *A Comparison of Methods for Determining the Molecular Content of Model Galaxies*, Astrophys. J. **729**, 36 36, arXiv: 1011.4065 [astro-ph.CO] (cit. on pp. 33, 57).

Krumholz, M. R., A. K. Leroy and C. F. McKee (2011), *Which Phase of the Interstellar Medium Correlates with the Star Formation Rate?*, Astrophys. J. **731**, 25 25, arXiv: 1101.1296 [astro-ph.CO] (cit. on pp. 20, 31).

Krumholz, M. R. and C. F. McKee (2005), *A General Theory of Turbulence-regulated Star Formation, from Spirals to Ultraluminous Infrared Galaxies*, Astrophys. J. **630** 250, arXiv: astro-ph/0505177 [astro-ph] (cit. on pp. 20, 31).

Krumholz, M. R., C. F. McKee and J. Tumlinson (2008), *The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Photodissociation Fronts in Finite Clouds*, Astrophys. J. **689** 865, arXiv: 0805.2947 [astro-ph] (cit. on pp. 32, 34).

– (2009), *The Star Formation Law in Atomic and Molecular Gas*, Astrophys. J. **699** 850, arXiv: 0904.0009 [astro-ph.GA] (cit. on pp. 31–34, 56, 81).

Kuhlen, M., M. R. Krumholz et al. (2012), *Dwarf Galaxy Formation with H_2-regulated Star Formation*, Astrophys. J. **749**, 36 36, arXiv: 1105.2376 [astro-ph.CO] (cit. on pp. 31, 32, 34, 37, 57).

Kuhlen, M., P. Madau and M. R. Krumholz (2013), *Dwarf Galaxy Formation with H_2-regulated Star Formation. II. Gas-rich Dark Galaxies at Redshift 2.5*, Astrophys. J. **776**, 34 34, arXiv: 1305.5538 [astro-ph.CO] (cit. on pp. 31, 32, 34, 37, 43, 57).

Lacey, C. G. et al. (2016), *A unified multiwavelength model of galaxy formation*, Mon. Not. R. Astron. Soc. **462** 3854, arXiv: 1509.08473 [astro-ph.GA] (cit. on p. 32).

Lagos, C. d. P., R. A. Crain et al. (2015), *Molecular hydrogen abundances of galaxies in the EAGLE simulations*, Mon. Not. R. Astron. Soc. **452** 3815, arXiv: 1503.04807 [astro-ph.GA] (cit. on pp. 32, 34).

Lagos, C. d. P., R. J. Tobar et al. (2018), *Shark: introducing an open source, free, and flexible semi-analytic model of galaxy formation*, Mon. Not. R. Astron. Soc. **481** 3573, arXiv: 1807.11180 [astro-ph.GA] (cit. on p. 32).

Lagos, C. D. P. et al. (2011), *Cosmic evolution of the atomic and molecular gas contents of galaxies*, Mon. Not. R. Astron. Soc. **418** 1649, arXiv: 1105.2294 [astro-ph.CO] (cit. on pp. 32, 34).

Larson, R. B. (1981), *Turbulence and star formation in molecular clouds.*, Mon. Not. R. Astron. Soc. **194** 809 (cit. on pp. 21, 63).

Leitherer, C. et al. (1999), *Starburst99: Synthesis Models for Galaxies with Active Star Formation*, Astrophys. J. Suppl. Ser. **123** 3, arXiv: astro-ph/9902334 [astro-ph] (cit. on pp. 37, 69).

Lemaître, G. (1927), *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, Annales de la Société; Scientifique de Bruxelles **47** 49 (cit. on p. 2).

Lenkić, L. et al. (2020), *Plateau de Bure High-z Blue Sequence Survey 2 (PHIBSS2): Search for Secondary Sources, CO Luminosity Functions in the Field, and the Evolution of Molecular Gas Density through Cosmic Time*, Astrophys. J. **159** 190 (cit. on pp. 8, 9, 54, 56).

Leroy, A. K. et al. (2008), *The Star Formation Efficiency in Nearby Galaxies: Measuring Where Gas Forms Stars Effectively*, **136** 2782, arXiv: 0810.2556 [astro-ph] (cit. on pp. 31, 32).

Lupi, A., S. Bovino et al. (2018), *The natural emergence of the correlation between H_2 and star formation rate surface densities in galaxy simulations*, Mon. Not. R. Astron. Soc. **474** 2884, arXiv: 1710.01315 [astro-ph.GA] (cit. on pp. 33, 35, 37).

Lupi, A., A. Pallottini et al. (2020), *Predicting FIR lines from simulated galaxies*, Mon. Not. R. Astron. Soc. **496** 5160, arXiv: 2004.06118 [astro-ph.GA] (cit. on p. 60).

Lupi, A., M. Volonteri et al. (2019), *High-redshift quasars and their host galaxies - I. Kinematical and dynamical properties and their tracers*, Mon. Not. R. Astron. Soc. **488** 4004, arXiv: 1901.02464 [astro-ph.GA] (cit. on p. 33).

Mac Low, M.-M. and S. C. O. Glover (2012), *The Abundance of Molecular Hydrogen and Its Correlation with Midplane Pressure in Galaxies: Non-equilibrium, Turbulent, Chemical Models*, Astrophys. J. **746**, 135 135, arXiv: 1011.3054 [astro-ph.GA] (cit. on p. 33).

Mackey, J. et al. (2019), *Non-equilibrium chemistry and destruction of CO by X-ray flares*, Mon. Not. R. Astron. Soc. **486** 1094, arXiv: 1803.10367 [astro-ph.GA] (cit. on p. 60).

Madau, P., L. Pozzetti and M. Dickinson (1998), *The Star Formation History of Field Galaxies*, **498** 106 (cit. on pp. 3, 8).

Madau, P. and M. Dickinson (2014), *Cosmic Star-Formation History*, **52** 415, arXiv: 1403.0007 [astro-ph.CO] (cit. on pp. 8, 9, 53).

Maeda, F., K. Ohta and A. Seko (2017), *Evolution of Cosmic Molecular Gas Mass Density from z 0 to z = 1-1.5*, Astrophys. J. **835**, 120 120, arXiv: 1612.04816 [astro-ph.GA] (cit. on p. 54).

Maiolino, R. et al. (2015), *The assembly of 'normal' galaxies at z ~ 7 probed by ALMA*, Mon. Not. R. Astron. Soc. **452** 54, arXiv: 1502.06634 [astro-ph.GA] (cit. on p. 31).

McKee, C. F. and M. R. Krumholz (2010), *The Atomic-to-Molecular Transition in Galaxies. III. A New Method for Determining the Molecular Content of Primordial and Dusty Clouds*, Astrophys. J. **709** 308, arXiv: 0908.0330 [astro-ph.GA] (cit. on pp. 32, 34).

McKee, C. F. and E. C. Ostriker (2007), *Theory of Star Formation*, **45** 565, arXiv: 0707.3514 [astro-ph] (cit. on p. 19).

McMillan, S. L. W. (1993), *An O(N N) Integration Scheme for Collisional Stellar Systems*, Astrophys. J. **414** 200 (cit. on p. 23).

Micic, M. et al. (2012), *Modelling H_2 formation in the turbulent interstellar medium: solenoidal versus compressive turbulent forcing*, Mon. Not. R. Astron. Soc. **421** 2531, arXiv: 1103.3056 [astro-ph.SR] (cit. on p. 35).

Mortlock, A. et al. (2015), *Deconstructing the galaxy stellar mass function with UKIDSS and CANDELS: the impact of colour, structure and environment*, Mon. Not. R. Astron. Soc. **447** 2, arXiv: 1411.3339 [astro-ph.GA] (cit. on p. 54).

Murante, G., P. Monaco, S. Borgani et al. (2015), *Simulating realistic disc galaxies with a novel sub-resolution ISM model*, Mon. Not. R. Astron. Soc. **447** 178 (cit. on p. 32).

Murante, G., P. Monaco, M. Giovalli et al. (2010), *A subresolution multiphase interstellar medium model of star formation and supernova energy feedback*, Mon. Not. R. Astron. Soc. **405** 1491, arXiv: 1002.4122 [astro-ph.CO] (cit. on p. 32).

Navarro, J. F., C. S. Frenk and S. D. M. White (1996), *The Structure of Cold Dark Matter Halos*, Astrophys. J. **462** 563 (cit. on p. 3).

Nelson, R. P. and W. D. Langer (1997), *The Dynamics of Low-Mass Molecular Clouds in External Radiation Fields*, Astrophys. J. **482** 796 (cit. on p. 64).

– (1999), *On the Stability and Evolution of Isolated BOK Globules*, Astrophys. J. **524** 923 (cit. on pp. 60, 81, 100).

Nguyen, H. et al. (2018), *Dust-Gas Scaling Relations and OH Abundance in the Galactic ISM*, Astrophys. J. **862**, 49 49, arXiv: 1805.11787 [astro-ph.GA] (cit. on p. 14).

Nickerson, S., R. Teyssier and J. Rosdahl (2018), *A simple model for molecular hydrogen chemistry coupled to radiation hydrodynamics*, Mon. Not. R. Astron. Soc. **479** 3206, arXiv: 1802.00445 [astro-ph.GA] (cit. on p. 33).

Nishimura, A. et al. (2015), *Revealing the Physical Properties of Molecular Gas in Orion with a Large-scale Survey in J = 2-1 Lines of ^{12}CO, ^{13}CO, and $C^{18}O$*, Astrophys. J. Suppl. Ser. **216**, 18 18, arXiv: 1412.0790 [astro-ph.GA] (cit. on p. 14).

Nordon, R. and A. Sternberg (2016), *C^+/H_2 gas in star-forming clouds and galaxies*, Mon. Not. R. Astron. Soc. **462** 2804, arXiv: 1603.02300 [astro-ph.GA] (cit. on p. 76).

Obreschkow, D., D. Croton et al. (2009), *Simulation of the Cosmic Evolution of Atomic and Molecular Hydrogen in Galaxies*, Astrophys. J. **698** 1467, arXiv: 0904.2221 [astro-ph.CO] (cit. on pp. 32, 53).

Obreschkow, D. and S. Rawlings (2009a), *Compactness of cold gas in high-redshift galaxies*, Mon. Not. R. Astron. Soc. **400** 665, arXiv: 0908.0724 [astro-ph.CO] (cit. on p. 32).

– (2009b), *The Cosmic Decline in the H_2/H I Ratio in Galaxies*, Astrophys. J. Lett **696** L129, arXiv: 0904.0213 [astro-ph.GA] (cit. on pp. 52, 53).

Olsen, K. P., B. Burkhart et al. (2021), *SIGAME v3: Gas Fragmentation in Post-processing of Cosmological Simulations for More Accurate Infrared Line Emission Modeling*, arXiv e-prints, arXiv:2102.02868 arXiv:2102.02868, arXiv: 2102.02868 [astro-ph.GA] (cit. on p. 60).

Olsen, K. P., T. R. Greve et al. (2016), *SImulator of GAlaxy Millimetre/submillimetre Emission (SÍGAME): CO emission from massive z = 2 main-sequence galaxies*, Mon. Not. R. Astron. Soc. **457** 3306, arXiv: 1507.00012 [astro-ph.GA] (cit. on pp. 60, 61).

Orr, M. E. et al. (2017), *Stacked Star Formation Rate Profiles of Bursty Galaxies Exhibit "Coherent" Star Formation*, Astrophys. J. Lett **849**, L2 L2, arXiv: 1709.10099 [astro-ph.GA] (cit. on p. 37).

Pallottini, A., A. Ferrara, D. Decataldo et al. (2019), *Deep into the structure of the first galaxies: SERRA views*, Mon. Not. R. Astron. Soc. **487** 1689, arXiv: 1905.08254 [astro-ph.GA] (cit. on pp. 33, 60).

Pallottini, A., A. Ferrara, S. Gallerani et al. (2017), *Zooming on the internal structure of z6 galaxies*, Mon. Not. R. Astron. Soc. **465** 2540, arXiv: 1609.01719 [astro-ph.GA] (cit. on pp. 32, 33).

Papadopoulos, P. P. (2007), *HCN versus HCO^+ as Dense Molecular Gas Mass Tracers in Luminous Infrared Galaxies*, Astrophys. J. **656** 792, arXiv: astro-ph/0610477 [astro-ph] (cit. on p. 14).

Papadopoulos, P. P. and T. R. Greve (2004), *C I Emission in Ultraluminous Infrared Galaxies as a Molecular Gas Mass Tracer*, Astrophys. J. Lett **615** L29, arXiv: astro-ph/0409559 [astro-ph] (cit. on p. 14).

Pardi, A. et al. (2017), *The impact of magnetic fields on the chemical evolution of the supernova-driven ISM*, Mon. Not. R. Astron. Soc. **465** 4611, arXiv: 1611.00585 [astro-ph.GA] (cit. on p. 60).

Pearson, W. J. et al. (2018), *Main sequence of star forming galaxies beyond the Herschel confusion limit*, Astron. Astrophys. **615**, A146 A146, arXiv: 1804.03482 [astro-ph.GA] (cit. on pp. 50, 51).

Pelupessy, F. I., P. P. Papadopoulos and P. van der Werf (2006), *Incorporating the Molecular Gas Phase in Galaxy-sized Numerical Simulations: First Applications in Dwarf Galaxies*, Astrophys. J. **645** 1024, arXiv: astro-ph/0603581 [astro-ph] (cit. on p. 33).

Pelupessy, F. I. and P. P. Papadopoulos (2009), *Molecular Gas, CO, and Star Formation in Galaxies: Emergent Empirical Relations, Feedback, and the Evolution of Very Gas-Rich Systems*, Astrophys. J. **707** 954, arXiv: 0910.4905 [astro-ph.CO] (cit. on p. 33).

Penzias, A. A. and R. W. Wilson (1965), *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, Astrophys. J. **142** 419 (cit. on p. 2).

Perlmutter, S. et al. (1999), *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, Astrophys. J. **517** 565 (cit. on p. 3).

Peters, T. et al. (2017), *The SILCC project - IV. Impact of dissociating and ionizing radiation on the interstellar medium and Hα emission as a tracer of the star formation rate*, Mon. Not. R. Astron. Soc. **466** 3293, arXiv: 1610.06569 [astro-ph.GA] (cit. on p. 60).

Pillepich, A. et al. (2018), *Simulating galaxy formation with the IllustrisTNG model*, Mon. Not. R. Astron. Soc. **473** 4077, arXiv: 1703.02970 [astro-ph.GA] (cit. on p. 24).

Pirogov, L. E. (2009), *Density profiles in molecular cloud cores associated with high-mass star-forming regions*, Astronomy Reports **53** 1127, arXiv: 0911.4421 [astro-ph.GA] (cit. on p. 63).

Planck Collaboration (2014), *Planck 2013 results. XVI. Cosmological parameters*, Astron. Astrophys. **571**, A16 A16, arXiv: 1303.5076 [astro-ph.CO] (cit. on pp. 6, 36).

– (2018), *Planck 2018 results. VI. Cosmological parameters*, arXiv e-prints, arXiv: 1807.06209 (cit. on pp. 6, 69).

Popping, G., P. S. Behroozi and M. S. Peeples (2015), *Evolution of the atomic and molecular gas content of galaxies in dark matter haloes*, Mon. Not. R. Astron. Soc. **449** 477, arXiv: 1409.1574 [astro-ph.GA] (cit. on pp. 32, 34, 56, 81).

Popping, G., D. Narayanan et al. (2019), *The art of modelling CO, [C I], and [C II] in cosmological galaxy formation models*, Mon. Not. R. Astron. Soc. **482** 4906, arXiv: 1805.11093 [astro-ph.GA] (cit. on pp. 60, 61).

Popping, G., R. S. Somerville and S. C. Trager (2014), *Evolution of the atomic and molecular gas content of galaxies*, Mon. Not. R. Astron. Soc. **442** 2398, arXiv: 1308.6764 [astro-ph.CO] (cit. on p. 32).

Rasera, Y. and R. Teyssier (2006), *The history of the baryon budget. Cosmic logistics in a hierarchical universe*, Astron. Astrophys. **445** 1, arXiv: astro-ph/0505473 [astro-ph] (cit. on p. 37).

Rees, M. J. and J. P. Ostriker (1977), *Cooling, dynamics and fragmentation of massive gas clouds: clues to the masses and radii of galaxies and clusters.*, Mon. Not. R. Astron. Soc. **179** 541 (cit. on p. 7).

Richings, A. J. and J. Schaye (2016), *Chemical evolution of giant molecular clouds in simulations of galaxies*, Mon. Not. R. Astron. Soc. **460** 2297, arXiv: 1604.01035 [astro-ph.GA] (cit. on p. 33).

Riechers, D. A., C. M. Bradford et al. (2013), *A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34*, **496** 329, arXiv: 1304.4256 [astro-ph.CO] (cit. on p. 49).

Riechers, D. A., P. L. Capak et al. (2010), *A Massive Molecular Gas Reservoir in the z = 5.3 Submillimeter Galaxy AzTEC-3*, Astrophys. J. Lett **720** L131, arXiv: 1008.0389 [astro-ph.CO] (cit. on pp. 31, 49).

Riechers, D. A., R. Pavesi et al. (2019), *COLDz: Shape of the CO Luminosity Function at High Redshift and the Cold Gas History of the Universe*, Astrophys. J. **872**, 7 7, arXiv: 1808.04371 [astro-ph.GA] (cit. on pp. 52–54).

Riess, A. G., S. Casertano et al. (2019), *Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond* Λ*CDM*, Astrophys. J. **876**, 85 85, arXiv: 1903.07603 [astro-ph.CO].

Riess, A. G., A. V. Filippenko et al. (1998), *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, **116** 1009, arXiv: astro-ph/9805201 [astro-ph] (cit. on p. 3).

Robertson, B. E. and A. V. Kravtsov (2008), *Molecular Hydrogen and Global Star Formation Relations in Galaxies*, Astrophys. J. **680** 1083, arXiv: 0710.2102 [astro-ph] (cit. on p. 33).

Roueff, A. et al. (2021), *$C^{18}O$, ^{13}CO, and ^{12}CO abundances and excitation temperatures in the Orion B molecular cloud. Analysis of the achievable precision in modeling spectral lines within the approximation of the local thermodynamic equilibrium*, Astron. Astrophys. **645**, A26 A26, arXiv: 2005.08317 [astro-ph.GA] (cit. on p. 14).

Rubin, V. C., J. Ford W. K. and N. Thonnard (1978), *Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa -> Sc.*, Astrophys. J. Lett **225** L107 (cit. on p. 2).

Rubin, V. C. and J. Ford W. Kent (1970), *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** 379 (cit. on p. 2).

Saito, T. et al. (2020), *The 300-pc Scale ALMA View of [C I] J = 1-0, CO J = 1-0, and 609 μm Dust Continuum in A Luminous Infrared Galaxy*, Mon. Not. R. Astron. Soc. (cit. on p. 14).

Salak, D. et al. (2019), *ALMA Observations of Atomic Carbon [C I] ($^3P_1 \rightarrow {}^3P_0$) and Low-J CO Lines in the Starburst Galaxy NGC 1808*, Astrophys. J. **887**, 143 143, arXiv: 1911.04077 [astro-ph.GA] (cit. on p. 14).

Santini, P., M. Castellano et al. (2016), *Characterizing elusive, faint dusty star-forming galaxies: a lensed, optically undetected ALMA galaxy at z 3.3*, Astron. Astrophys. **596**, A75 A75, arXiv: 1610.08515 [astro-ph.GA] (cit. on p. 49).

Santini, P., E. Merlin et al. (2019), *Passive galaxies in the early Universe: ALMA confirmation of $z \sim$ 3-5 candidates in the CANDELS GOODS-South field*, Mon. Not. R. Astron. Soc. **486** 560, arXiv: 1902.09548 [astro-ph.GA] (cit. on p. 31).

Scannapieco, C. et al. (2012), *The Aquila comparison project: the effects of feedback and numerical methods on simulations of galaxy formation*, Mon. Not. R. Astron. Soc. **423** 1726, arXiv: 1112.0315 [astro-ph.GA] (cit. on pp. 36, 38).

Schäbe, A. et al. (2020), *A comparison of H_2 formation models at high redshift*, Mon. Not. R. Astron. Soc. **497** 5008, arXiv: 2003.04329 [astro-ph.GA] (cit. on p. 31).

Schaye, J. et al. (2015), *The EAGLE project: simulating the evolution and assembly of galaxies and their environments*, Mon. Not. R. Astron. Soc. **446** 521, arXiv: 1407.7040 [astro-ph.GA] (cit. on p. 24).

Schinnerer, E. et al. (2008), *Molecular Gas in a Submillimeter Galaxy at z = 4.5: Evidence for a Major Merger at 1 Billion Years after the Big Bang*, Astrophys. J. Lett **689** L5, arXiv: 0810.3405 [astro-ph] (cit. on p. 49).

Schmidt, M. (1959), *The rate of star formation*, Astrophys. J. **129** 243 (cit. on pp. 19, 37).

Schneider, N. et al. (2013), *What Determines the Density Structure of Molecular Clouds? A Case Study of Orion B with Herschel*, Astrophys. J. Lett **766**, L17 L17, arXiv: 1304.0327 [astro-ph.GA] (cit. on pp. 35, 67).

Schreiber, C. et al. (2015), *The Herschel view of the dominant mode of galaxy growth from z = 4 to the present day*, Astron. Astrophys. **575**, A74 A74, arXiv: 1409.5433 [astro-ph.GA] (cit. on pp. 50, 51).

Seifried, D. et al. (2017), *SILCC-Zoom: the dynamic and chemical evolution of molecular clouds*, Mon. Not. R. Astron. Soc. **472** 4797, arXiv: 1704.06487 [astro-ph.GA] (cit. on pp. 32, 60).

Sellwood, J. A. (1987), *The art of N-body building.*, **25** 151 (cit. on p. 23).

Semenov, V. A., A. V. Kravtsov and N. Y. Gnedin (2018), *How Galaxies Form Stars: The Connection between Local and Global Star Formation in Galaxy Simulations*, Astrophys. J. **861**, 4 4, arXiv: 1803.00007 [astro-ph.GA] (cit. on p. 37).

Shapley, H. and H. D. Curtis (1921), *The Scale of the Universe*, Bulletin of the National Research Council **2** 171 (cit. on p. 1).

Sheth, R. K., H. J. Mo and G. Tormen (2001), *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, Mon. Not. R. Astron. Soc. **323** 1, arXiv: astro-ph/9907024 [astro-ph] (cit. on p. 55).

Shetty, R., S. C. O. Glover, C. P. Dullemond and R. S. Klessen (2011), *Modelling CO emission - I. CO as a column density tracer and the X factor in molecular clouds*, Mon. Not. R. Astron. Soc. **412** 1686, arXiv: 1011.2019 [astro-ph.GA] (cit. on p. 60).

Shetty, R., S. C. O. Glover, C. P. Dullemond, E. C. Ostriker et al. (2011), *Modelling CO emission - II. The physical characteristics that determine the X factor in Galactic molecular clouds*, Mon. Not. R. Astron. Soc. **415** 3253, arXiv: 1104.3695 [astro-ph.GA] (cit. on p. 60).

Silk, J. (1997), *Feedback, Disk Self-Regulation, and Galaxy Formation*, Astrophys. J. **481** 703, arXiv: astro-ph/9612117 [astro-ph] (cit. on p. 20).

Sillero, E. et al. (2021), *Modelling H_2 and its effects on star formation using a joint implementation of GADGET-3 and KROME*, Mon. Not. R. Astron. Soc. **504** 2325, arXiv: 2104.04357 [astro-ph.GA] (cit. on p. 60).

Solomon, P. M., D. Downes and S. J. E. Radford (1992), *Warm Molecular Gas in the Primeval Galaxy IRAS 10214+4724*, Astrophys. J. Lett **398** L29 (cit. on p. 13).

Solomon, P. M., D. Downes, S. J. E. Radford and J. W. Barrett (1997), *The Molecular Interstellar Medium in Ultraluminous Infrared Galaxies*, Astrophys. J. **478** 144, arXiv: astro-ph/9610166 [astro-ph] (cit. on p. 13).

Somerville, R. S., G. Popping and S. C. Trager (2015), *Star formation in semi-analytic galaxy formation models with multiphase gas*, Mon. Not. R. Astron. Soc. **453** 4337, arXiv: 1503.00755 [astro-ph.GA] (cit. on pp. 32, 34).

Spitzer Lyman, J. (1941), *The Dynamics of the Interstellar Medium. II. Radiation Pressure.*, Astrophys. J. **94** 232 (cit. on p. 19).

Springel, V. (2010), *E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh*, Mon. Not. R. Astron. Soc. **401** (2) 791 (cit. on pp. 24, 26).

Springel, V., S. D. M. White et al. (2005), *Simulations of the formation, evolution and clustering of galaxies and quasars*, **435** (7042) 629 (cit. on pp. 3, 24).

Springel, V., N. Yoshida and S. D. M. White (2001), *GADGET: a code for collisionless and gasdynamical cosmological simulations*, New Astronomy **6** (2) 79 (cit. on p. 24).

Sternberg, A. (1988), *The Infrared Response of Molecular Hydrogen Gas to Ultraviolet Radiation: A Scaling Law*, Astrophys. J. **332** 400 (cit. on p. 32).

Sternberg, A. et al. (2014), *H I-to-H_2 Transitions and H I Column Densities in Galaxy Star-forming Regions*, Astrophys. J. **790**, 10 10, arXiv: 1404.5042 [astro-ph.GA] (cit. on p. 101).

Stevens, A. R. H., D. J. Croton and S. J. Mutch (2016), *Building disc structure and galaxy properties through angular momentum: the DARK SAGE semi-analytic model*, Mon. Not. R. Astron. Soc. **461** 859, arXiv: 1605.00647 [astro-ph.GA] (cit. on p. 32).

Stevens, A. R. H., M. Martig et al. (2014), *Where do galaxies end? Comparing measurement techniques of hydrodynamic-simulation galaxies' integrated properties*, Mon. Not. R. Astron. Soc. **445** 239, arXiv: 1404.4053 [astro-ph.GA] (cit. on p. 38).

Strateva, I. et al. (2001), *Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data*, **122** 1861, arXiv: astro-ph/0107201 [astro-ph] (cit. on p. 8).

Strömgren, B. (1939), *The Physical State of Interstellar Hydrogen.*, Astrophys. J. **89** 526 (cit. on p. 11).

Strong, A. W. and J. R. Mattox (1996), *Gradient model analysis of EGRET diffuse Galactic γ-ray emission.*, Astron. Astrophys. **308** L21 (cit. on p. 13).

Tacconi, L. J. et al. (2008), *Submillimeter Galaxies at z 2: Evidence for Major Mergers and Constraints on Lifetimes, IMF, and CO-H_2 Conversion Factor*, Astrophys. J. **680** 246, arXiv: 0801.3650 [astro-ph] (cit. on p. 13).

Tajiri, Y. and M. Umemura (1998), *A Criterion for Photoionization of Pregalactic Clouds Exposed to Diffuse Ultraviolet Background Radiation*, Astrophys. J. **502** 59, arXiv: astro-ph/9806046 [astro-ph] (cit. on p. 37).

Tan, Q.-H. et al. (2011), *^{12}CO, ^{13}CO and $C^{18}O$ observations along the major axes of nearby bright infrared galaxies*, Research in Astronomy and Astrophysics **11** 787, arXiv: 1103.5540 [astro-ph.CO] (cit. on p. 14).

Teyssier, R. (2002), *Cosmological Hydrodynamics with Adaptive Mesh Refinement: a new high resolution code called RAMSES*, Astron. Astrophys. **385** 337 (cit. on pp. 24, 36).

Teyssier, R. (2015), *Grid-Based Hydrodynamics in Astrophysical Fluid Flows*, **53** 325 (cit. on p. 26).

Teyssier, R., D. Chapon and F. Bournaud (2010), *The Driving Mechanism of Starbursts in Galaxy Mergers*, Astrophys. J. Lett **720** L149, arXiv: 1006.4757 [astro-ph.CO] (cit. on pp. 36, 69).

Thompson, R. et al. (2014), *Molecular Hydrogen Regulated Star Formation in Cosmological Smoothed Particle Hydrodynamics Simulations*, Astrophys. J. **780**, 145 145, arXiv: 1301.0063 [astro-ph.CO] (cit. on pp. 32, 34, 57).

Tielens, A. G. G. M. and D. Hollenbach (1985), *Photodissociation regions. I. Basic model.*, Astrophys. J. **291** 722 (cit. on pp. 63, 100, 101).

Tinker, J. et al. (2008), *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*, Astrophys. J. **688** 709, arXiv: 0803.2706 [astro-ph] (cit. on p. 55).

Tomassetti, M. et al. (2014), *Atomic carbon as a powerful tracer of molecular gas in the high-redshift Universe: perspectives for ALMA.*, Mon. Not. R. Astron. Soc. **445** L124, arXiv: 1405.0411 [astro-ph.GA] (cit. on p. 61).

– (2015), *Simulating the H_2 content of high-redshift galaxies*, Mon. Not. R. Astron. Soc. **446** 3330, arXiv: 1403.7132 [astro-ph.GA] (cit. on pp. 29, 32–35, 37, 43, 56, 57, 67–69, 81).

Tomczak, A. R. et al. (2014), *Galaxy Stellar Mass Functions from ZFOURGE/CANDELS: An Excess of Low-mass Galaxies since $z = 2$ and the Rapid Buildup of Quiescent Galaxies*, Astrophys. J. **783**, 85 85, arXiv: 1309.5972 [astro-ph.CO] (cit. on p. 54).

Truelove, J. K. et al. (1997), *The Jeans Condition: A New Constraint on Spatial Resolution in Simulations of Isothermal Self-gravitational Hydrodynamics*, Astrophys. J. Lett **489** L179 (cit. on pp. 28, 36, 69).

van der Wel, A. et al. (2014), *3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since $z = 3$*, Astrophys. J. **788**, 28 28, arXiv: 1404.2844 [astro-ph.GA] (cit. on p. 34).

van Dishoeck, E. F. and J. H. Black (1986), *Comprehensive Models of Diffuse Interstellar Clouds: Physical Conditions and Molecular Abundances*, Astrophys. J. Suppl. Ser. **62** 109 (cit. on p. 32).

Visser, R., E. F. van Dishoeck and J. H. Black (2009), *The photodissociation and chemistry of CO isotopologues: applications to interstellar clouds and circumstellar disks*, Astron. Astrophys. **503** 323, arXiv: 0906.3699 [astro-ph.GA] (cit. on pp. 63, 100).

Vogelsberger, M. et al. (2014), *Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe*, Mon. Not. R. Astron. Soc. **444** 1518, arXiv: 1405.2921 [astro-ph.CO] (cit. on p. 24).

von Hoerner, S. (1960), *Die numerische Integration des n-Körper-Problemes für Sternhaufen. I*, Zeitschrift für Astrophysik **50** 184 (cit. on p. 23).

– (1963), *Die numerische Integration des n-Körper-Problemes für Sternhaufen. II*, Zeitschrift für Astrophysik **57** 47 (cit. on p. 23).

Wakelam, V. et al. (2010), *Reaction Networks for Interstellar Chemical Modelling: Improvements and Challenges*, **156** 13, arXiv: 1011.1184 [astro-ph.GA] (cit. on p. 100).

Walch, S. et al. (2015), *The SILCC (SImulating the LifeCycle of molecular Clouds) project - I. Chemical evolution of the supernova-driven ISM*, Mon. Not. R. Astron. Soc. **454** 238, arXiv: 1412.2749 [astro-ph.GA] (cit. on p. 60).

Walter, F., R. Decarli, M. Sargent et al. (2014), *A Molecular Line Scan in the Hubble Deep Field North: Constraints on the CO Luminosity Function and the Cosmic H_2 Density*, Astrophys. J. **782**, 79 79, arXiv: 1312.6365 [astro-ph.CO] (cit. on p. 54).

Walter, F., R. Decarli, C. L. Carilli et al. (2012), *The intense starburst HDF 850.1 in a galaxy overdensity at $z \approx 5.2$ in the Hubble Deep Field*, **486** 233, arXiv: 1206.2641 [astro-ph.CO] (cit. on p. 49).

Warren, M. S. et al. (1992), *Dark Halos Formed via Dissipationless Collapse. I. Shapes and Alignment of Angular Momentum*, Astrophys. J. **399** 405 (cit. on p. 23).

Weingartner, J. C. and B. T. Draine (2001), *Dust Grain-Size Distributions and Extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud*, Astrophys. J. **548** 296, arXiv: astro-ph/0008146 [astro-ph] (cit. on p. 101).

Weiß, A. et al. (2005), *Atomic carbon at redshift ~2.5*, Astron. Astrophys. **429** L25, arXiv: astro-ph/0409711 [astro-ph] (cit. on p. 76).

Whipple, F. L. (1946), *Concentrations of the Interstellar Medium.*, Astrophys. J. **104** 1 (cit. on p. 19).

White, S. D. M. and M. J. Rees (1978), *Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering.*, Mon. Not. R. Astron. Soc. **183** 341 (cit. on p. 7).

Wise, J. H. et al. (2012), *The Birth of a Galaxy: Primordial Metal Enrichment and Stellar Populations*, Astrophys. J. **745**, 50 50, arXiv: 1011.2632 [astro-ph.CO] (cit. on p. 37).

Wong, T. and L. Blitz (2002), *The Relationship between Gas Content and Star Formation in Molecule-rich Spiral Galaxies*, Astrophys. J. **569** 157, arXiv: astro-ph/0112204 [astro-ph] (cit. on pp. 20, 31, 32).

Wong, T., R. Xue et al. (2013), *CARMA Survey toward Infrared-bright Nearby Galaxies (STING). III. The Dependence of Atomic and Molecular Gas Surface Densities on Galaxy Properties*, Astrophys. J. Lett **777**, L4 L4, arXiv: 1309.6513 [astro-ph.GA] (cit. on p. 32).

Wu, J. et al. (2010), *The Properties of Massive, Dense Clumps: Mapping Surveys of HCN and CS*, Astrophys. J. Suppl. Ser. **188** 313, arXiv: 1004.0398 [astro-ph.GA] (cit. on p. 14).

Zanella, A. et al. (2018), *The [C II] emission as a molecular gas mass tracer in galaxies at low and high redshifts*, Mon. Not. R. Astron. Soc. **481** 1976, arXiv: 1808.10331 [astro-ph.GA] (cit. on p. 14).

APPENDIX A

Reaction rates

reaction	coefficent	reference
$\mathrm{cr} + \mathrm{H_2} \rightarrow \mathrm{H_3^+} + \mathrm{e^-} + \mathrm{H}$	1.2×10^{-17}	UMIST[a]
$\mathrm{cr} + \mathrm{He} \rightarrow \mathrm{He^+} + \mathrm{e^-}$	6.5×10^{-18}	UMIST
$\mathrm{C} + \mathrm{H_3^+} \rightarrow \mathrm{CH_x} + \mathrm{H_2}$	2.0×10^{-9}	Nelson and Langer (1999)
$\mathrm{O} + \mathrm{H_3^+} \rightarrow \mathrm{OH_x} + \mathrm{H_2}$	8.0×10^{-10}	Nelson and Langer (1999)
$\mathrm{CO} + \mathrm{H_3^+} \rightarrow \mathrm{HCO^+} + \mathrm{H_2}$	$1.3 \times 10^{-9} \left(\frac{T}{300\mathrm{K}}\right)^{-0.14} \exp\left(\frac{3.40\mathrm{K}}{T}\right)$	UMIST
$\mathrm{He^+} + \mathrm{H_2} \rightarrow \mathrm{He} + \mathrm{H} + \mathrm{H^+}$	$3.7 \times 10^{-14} \exp\left(-\frac{35\mathrm{K}}{T}\right)$	UMIST
$\mathrm{He^+} + \mathrm{CO} \rightarrow \mathrm{C^+} + \mathrm{O} + \mathrm{He}$	1.6×10^{-9}	UMIST
$\mathrm{C^+} + \mathrm{H_2} \rightarrow \mathrm{CH_x} + \mathrm{H}$	$2.31 \times 10^{-13} T^{-1.3} \exp\left(-\frac{23\mathrm{K}}{T}\right)$	Wakelam et al. (2010)
$\mathrm{C^+} + \mathrm{OH_x} \rightarrow \mathrm{HCO^+}$	$9.15 \times 10^{-10}(0.62 + 45.41\sqrt{T})$	Wakelam et al. (2010)
$\mathrm{O} + \mathrm{CH_x} \rightarrow \mathrm{CO} + \mathrm{H}$	7.7×10^{-11}	Wakelam et al. (2010)
$\mathrm{C} + \mathrm{OH_x} \rightarrow \mathrm{CO} + \mathrm{H}$	$7.95 \times 10^{-10} T^{-0.339} \exp\left(\frac{0.108\mathrm{K}}{T}\right)$	Wakelam et al. (2010)
$\mathrm{He^+} + \mathrm{e^-} \rightarrow \mathrm{He} + \gamma$	$5.36 \times 10^{-12} \left(\frac{T}{300\mathrm{K}}\right)^{-0.5}$	UMIST
$\mathrm{H_3^+} + \mathrm{e^-} \rightarrow \mathrm{H_2} + \mathrm{H}$	$2.34 \times 10^{-8} \left(\frac{T}{300\mathrm{K}}\right)^{-0.52}$	UMIST
$\mathrm{C^+} + \mathrm{e^-} \rightarrow \mathrm{C} + \gamma$	$2.36 \times 10^{-12} \left(\frac{T}{300\mathrm{K}}\right)^{-0.29} \exp\left(\frac{17.60\mathrm{K}}{T}\right)$	UMIST
$\mathrm{HCO^+} + \mathrm{e^-} \rightarrow \mathrm{CO} + \mathrm{H}$	$2.4 \times 10^{-7} \left(\frac{T}{300\mathrm{K}}\right)^{-0.69}$	UMIST
$\mathrm{M^+} + \mathrm{e^-} \rightarrow \mathrm{M} + \gamma$	$3.8 \times 10^{-10} \left(\frac{T}{\mathrm{K}}\right)^{-0.65}$	Nelson and Langer (1999)
$\mathrm{H_3^+} + \mathrm{M} \rightarrow \mathrm{M^+} + \mathrm{H_2}$	2.0×10^{-9}	Nelson and Langer (1999)
$\mathrm{C} + \gamma \rightarrow \mathrm{C^+} + \mathrm{e^-}$	$3.1 \times 10^{-10}\, G\, \exp(-3.3A_V)\, \Theta(N_{\mathrm{H_2}}, N_{\mathrm{C}})$	UMIST, Tielens and Hollenbach (1985)
$\mathrm{CH_x} + \gamma \rightarrow \mathrm{C} + \mathrm{H}$	9.1×10^{-10}	Heays, Bosman and van Dishoeck (2017)
$\mathrm{OH_x} + \gamma \rightarrow \mathrm{O} + \mathrm{H}$	3.8×10^{-10}	Heays, Bosman and van Dishoeck (2017)
$\mathrm{CO} + \gamma \rightarrow \mathrm{C} + \mathrm{O}$	$2.0 \times 10^{-10}\, G\, \exp(-3.5A_V)\, \Theta(N_{\mathrm{H_2}}, N_{\mathrm{CO}})$	UMIST,Visser, van Dishoeck and Black (2009)
$\mathrm{M} + \gamma \rightarrow \mathrm{M^+} + \mathrm{e^-}$	2.0×10^{-10}	Nelson and Langer (1999)
$\mathrm{HCO^+} + \gamma \rightarrow \mathrm{CO} + \mathrm{H^+}$	5.4×10^{-12}	UMIST

Table A.1: Reaction and their corresponding rates used in this thesis.

[a] http://www.rate99.co.uk

reaction	coefficent	reference
$H + H + grain \rightarrow H_2 + grain$	$3.025 \times 10^{-17} S_H \sqrt{0.01\,T}\,(Z/Z_\odot)$	
	$S_H = (1 + 0.4\sqrt{0.01(T + T_d)} + 0.2(0.01\,T) + 0.08(0.01\,T)^2)^{-1}$	Tielens and Hollenbach (1985),
	with $T_d = 10$ K	Cazaux and Spaans (2004)
$H_2 + H \rightarrow 3H$	$4.670 \times 10^{-7}\ \exp\left(-\frac{55000\ \mathrm{K}}{T}\right)\left(\frac{T}{300\mathrm{K}}\right)^{-1}$	UMIST
$H_2 + H_2 \rightarrow 2H + H_2$	$1.0 \times 10^{-8}\ \exp\left(-\frac{84100\mathrm{K}}{T}\right)$	UMIST
$H_2 + \gamma \rightarrow 2H$	$4.200 \times 10^{-11}\, G\, \Phi(N_{H_2})\, e^{-\tau}$	
	$\tau = \sigma_d N_H$ with $\sigma_d = 2 \times 10^{-21} (Z/Z_\odot)\ \mathrm{cm}^2$	
	$\Phi(N_{H_2}) = \frac{1-\omega}{(1+x/b_5)^2} + \frac{\omega}{(1+x)^{1/2}} \exp\left[-8.5 \times 10^{-4}(1-x)^{1/2}\right]$	Draine and Bertoldi (1996),
	$x = N_{H_2}/(5 \times 10^{14}\mathrm{cm}^{-2})$, $\omega = 0.035$, $b_5 = 2$	Sternberg et al. (2014)
$cr + H \rightarrow H^+ + e^-$	1.300×10^{-17}	UMIST
$H + e^- \rightarrow H^+ + 2e^-$	$\exp(-33.1 + 13.5\log(T_e) - 5.79\log(T_e)^2$	
	$-5.79\log(T_e)^2 + 1.56\log(T_e)^3$	
	$-0.288\log(T_e)^4 + 3.48 \times 10^{-2}\log(T_e)^5$	
	$-2.63 \times 10^{-3}\log(T_e)^6 + 1.12 \times 10^{-4}\log(T_e)^7$	
	$-2.04 \times 10^{-6}\log(T_e)^8)$	Glover and Mac Low (2007a)
$H^+ + e^- \rightarrow H + \gamma$	$3.500 \times 10^{12}\left(\frac{T}{300\mathrm{K}}\right)^{-0.75}$	UMIST
$H^+ + e^- + grain \rightarrow H + grain$	$1 \times 10^{-14} \times \frac{12.25}{1+(8.074\times10^{-6}\Psi)^{1.378}(1+5.087\times10^{2}T^{-0.01586}\Psi^{-0.4723-0.00001102\log(T)})}$	
	$\Psi = \sqrt{T}\frac{G}{x_{e^-}}$	Weingartner and Draine (2001),
		Draine (2003)

Table A.2: Continuation of table A.1.

APPENDIX B

Approximated Network

We start from the extended version of NL99 network from Glover and Clark (2012a). In order to limit the computational costs, we need to reduce the network to the most significant reaction channels. We start with the following reactions:

$$H_2 + \text{c.r.} \rightarrow H_2^+ + e^- \tag{B.1}$$
$$C + H_3^+ \rightarrow CH_x + H_2 \tag{B.2}$$
$$O + H_3^+ \rightarrow OH_x + H_2 \tag{B.3}$$
$$C^+ + H_2 \rightarrow CH_x + H \tag{B.4}$$
$$O + CH_x \rightarrow CO + H \tag{B.5}$$
$$C + OH_x \rightarrow CO + H \tag{B.6}$$
$$H_3^+ + e^- \rightarrow H_2 + H \tag{B.7}$$
$$C^+ + e^- \rightarrow C + \gamma \tag{B.8}$$
$$C + \gamma \rightarrow C^+ + e^- \tag{B.9}$$
$$CH_x + \gamma \rightarrow C + H \tag{B.10}$$
$$CO + \gamma \rightarrow C + O \tag{B.11}$$
$$OH_x + \gamma \rightarrow O + H \tag{B.12}$$
$$H + H + \text{grain} \rightarrow H_2 + \text{grain} \tag{B.13}$$
$$H_2 + \gamma \rightarrow 2H \tag{B.14}$$

The CO formation severely depends on the formation of OH_x and CH_x which is induced by the H_3^+ formation. We start by estimating the amount of OH_x and CH_x by setting up the rate equation under the

assumption that the consumption/destruction to happen with the same rate as the formation.

$$\frac{\mathrm{d}x_{\mathrm{OH_x}}}{\mathrm{d}t} = \frac{1}{3}k_{\mathrm{O,H_3^+}}x_{\mathrm{O}}x_{\mathrm{H_3^+}}n_{\mathrm{H}} - k_{\mathrm{C,OH_x}}x_{\mathrm{C}}x_{\mathrm{OH_x}}n_{\mathrm{H}} - k_{\gamma,\mathrm{OH_x}}x_{\mathrm{OH_x}} \tag{B.15}$$

$$\frac{\mathrm{d}x_{\mathrm{CH_x}}}{\mathrm{d}t} = \frac{1}{3}k_{\mathrm{C,H_3^+}}x_{\mathrm{C}}x_{\mathrm{H_3^+}}n_{\mathrm{H}} + \frac{1}{2}k_{\mathrm{C^+,H_2}}x_{\mathrm{C^+}}x_{\mathrm{H_2}}n_{\mathrm{H}} - k_{\mathrm{O,CH_x}}x_{\mathrm{O}}x_{\mathrm{CH_x}}n_{\mathrm{H}} - k_{\gamma,\mathrm{CH_x}}x_{\mathrm{CH_x}} \tag{B.16}$$

$$\frac{\mathrm{d}x_{\mathrm{OH_x}}}{\mathrm{d}t} = 0 \tag{B.17}$$

$$\frac{\mathrm{d}x_{\mathrm{CH_x}}}{\mathrm{d}t} = 0 \tag{B.18}$$

$$x_{\mathrm{OH_x}} = \frac{\frac{1}{3}k_{\mathrm{O,H_3^+}}x_{\mathrm{O}}x_{\mathrm{H_3^+}}}{k_{\mathrm{C,OH_x}}x_{\mathrm{C}} + k_{\gamma,\mathrm{OH_x}}/n_{\mathrm{H}}} \tag{B.19}$$

$$x_{\mathrm{CH_x}} = \frac{\frac{1}{3}k_{\mathrm{C,H_3^+}}x_{\mathrm{C}}x_{\mathrm{H_3^+}} + \frac{1}{2}k_{\mathrm{C^+,H_2}}x_{\mathrm{C^+}}x_{\mathrm{H_2}}}{k_{\mathrm{O,CH_x}}x_{\mathrm{O}} + k_{\gamma,\mathrm{CH_x}}/n_{\mathrm{H}}} \tag{B.20}$$

The rate equation for CO is according to the reaction rates given above

$$\frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} = k_{\mathrm{C,OH_x}}x_{\mathrm{C}}x_{\mathrm{OH_x}}n_{\mathrm{H}} + k_{\mathrm{O,CH_x}}x_{\mathrm{O}}x_{\mathrm{CH_x}}n_{\mathrm{H}} - k_{\gamma,\mathrm{CO}}x_{\mathrm{CO}} \tag{B.21}$$

$$\begin{aligned} &= k_{\mathrm{C,OH_x}}x_{\mathrm{C}}n_{\mathrm{H}}\frac{\frac{1}{3}k_{\mathrm{O,H_3^+}}x_{\mathrm{O}}x_{\mathrm{H_3^+}}}{k_{\mathrm{C,OH_x}}x_{\mathrm{C}} + k_{\gamma,\mathrm{OH_x}}/n_{\mathrm{H}}} \\ &+ k_{\mathrm{O,CH_x}}x_{\mathrm{O}}n_{\mathrm{H}}\frac{\frac{1}{3}k_{\mathrm{C,H_3^+}}x_{\mathrm{C}}x_{\mathrm{H_3^+}} + \frac{1}{2}k_{\mathrm{C^+,H_2}}x_{\mathrm{C^+}}x_{\mathrm{H_2}}}{k_{\mathrm{O,CH_x}}x_{\mathrm{O}} + k_{\gamma,\mathrm{CH_x}}/n_{\mathrm{H}}} - k_{\gamma,\mathrm{CO}}x_{\mathrm{CO}} \end{aligned} \tag{B.22}$$

and we define

$$\alpha_{\mathrm{CH_x}} = \frac{k_{\mathrm{O,CH_x}}x_{\mathrm{O}}}{k_{\mathrm{O,CH_x}}x_{\mathrm{O}} + k_{\gamma,\mathrm{CH_x}}/n_{\mathrm{H}}} \tag{B.23}$$

$$\alpha_{\mathrm{OH_x}} = \frac{k_{\mathrm{C,OH_x}}x_{\mathrm{C}}}{k_{\mathrm{C,OH_x}}x_{\mathrm{C}} + k_{\gamma,\mathrm{OH_x}}/n_{\mathrm{H}}} \tag{B.24}$$

which are also known as the *branching ratio* or *branching fraction* and takes all possible fates of a species into account. This gives for the CO rate equations

$$\begin{aligned} \frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} &= \frac{1}{3}n_{\mathrm{H}}k_{\mathrm{O,H_3^+}}x_{\mathrm{O}}x_{\mathrm{H_3^+}}\alpha_{\mathrm{OH_x}} + \frac{1}{3}n_{\mathrm{H}}k_{\mathrm{C,H_3^+}}x_{\mathrm{C}}x_{\mathrm{H_3^+}}\alpha_{\mathrm{CH_x}} \\ &+ \frac{1}{2}n_{\mathrm{H}}k_{\mathrm{C^+,H_2}}x_{\mathrm{C^+}}x_{\mathrm{H_2}}\alpha_{\mathrm{CH_x}} - k_{\gamma,\mathrm{CO}}x_{\mathrm{CO}} \ , \end{aligned} \tag{B.25}$$

where we have also included that CH_x is also formed via the C^+-H_2 channel. We still need to estimate $x_{H_3^+}$. Further, we assume charge conservation as $x_{C^+} \approx x_{e^-}$.

$$\frac{\mathrm{d}x_{H_3^+}}{\mathrm{d}t} = 0 \tag{B.26}$$

$$0 = \frac{3}{2}k_{\mathrm{c.r.},H_2}x_{H_2} - n_H k_{C,H_3^+}x_C x_{H_3^+} - n_H k_{O,H_3^+}x_O x_{H_3^+} - n_H k_{e^-,H_3^+}x_{e^-}x_{H_3^+} \tag{B.27}$$

$$0 = n\left(k_{C,H_3^+}x_C + k_{O,H_3^+}x_O + k_{e^-,H_3^+}x_{C^+}\right)x_{H_3^+} - \frac{3}{2}k_{\mathrm{c.r.},H_2}x_{H_2} \tag{B.28}$$

$$x_{H_3^+} = \frac{3}{2}\frac{k_{\mathrm{c.r.},H_2}x_{H_2}}{k_{C,H_3^+}x_C + k_{O,H_3^+}x_O + k_{e^-,H_3^+}x_{C^+}} \tag{B.29}$$

The only physical solution to that quadratic equation is the positive one. Knowing $x_{H_3^+}$, we can obtain the x_{e^-} as well. We write the rate equation for carbon as follows:

$$\begin{aligned}\frac{\mathrm{d}x_C}{\mathrm{d}t} &= -n_H k_{C,OH_x}x_C x_{OH_x} - \frac{1}{3}n_H k_{C,H_3^+}x_C x_{H_3^+} - k_{\gamma,C}x_C + k_{\gamma,CO}x_{CO} \\ &\quad + n_H k_{C^+,e^-}x_{C^+}x_{e^-} + k_{\gamma,CH_x}x_{CH_x} \end{aligned} \tag{B.30}$$

$$\begin{aligned}&= -\frac{1}{3}n_H k_{O,H_3^+}x_O x_{H_3^+}\alpha_{OH_x} - \frac{1}{3}n_H k_{C,H_3^+}x_C x_{H_3^+} - k_{\gamma,C}x_C + k_{\gamma,CO}x_{CO} \\ &\quad + n_H k_{C^+,e^-}x_{C^+}x_{e^-} + k_{\gamma,CH_x}x_{CH_x}\end{aligned} \tag{B.31}$$

Inserting the previously found expression for x_{CH_x} into the last term:

$$k_{\gamma,CH_x}x_{CH_x} = k_{\gamma,CH_x}\frac{\frac{1}{3}k_{C,H_3^+}x_C x_{H_3^+} + \frac{1}{2}k_{C^+,H_2}x_{C^+}x_{H_2}}{k_{O,CH_x}x_O + k_{\gamma,CH_x}/n_H} \tag{B.32}$$

$$= (1-\alpha_{CH_x})\left(\frac{1}{3}k_{C,H_3^+}x_C x_{H_3^+} + \frac{1}{2}k_{C^+,H_2}x_{C^+}x_{H_2}\right)n_H \tag{B.33}$$

This gives:

$$\begin{aligned}\frac{\mathrm{d}x_C}{\mathrm{d}t} &= -\frac{1}{3}n_H k_{O,H_3^+}x_O x_{H_3^+}\alpha_{OH_x} - \frac{1}{3}n_H k_{C,H_3^+}x_C x_{H_3^+}\alpha_{CH_x} - k_{\gamma,C}x_C + k_{\gamma,CO}x_{CO} \\ &\quad + n_H k_{C^+,e^-}x_{C^+}x_{e^-} + (1-\alpha_{CH_x})n_H\frac{1}{2}k_{C^+,H_2}x_{C^+}x_{H_2}\end{aligned} \tag{B.34}$$

For H_2, we find:

$$\frac{\mathrm{d}x_{H_2}}{\mathrm{d}t} = 2k_{H_2,gr}x_{HI}n_H - k_\gamma x_{H_2} - k_{\mathrm{c.r.},H_2}x_{H_2} - n_H k_{C^+,H_2}x_{C^+}x_{H_2} + \frac{1}{3}n_H k_{e^-,H_3^+}x_{e^-}x_{H_3^+} \tag{B.35}$$

And we obtain x_{C^+} by

$$x_{C^+} = x_{C,\mathrm{tot}} - x_{CO} - x_C \tag{B.36}$$

and we know further:

$$x_{\mathrm{O}} = x_{\mathrm{O,tot}} - x_{\mathrm{CO}} \tag{B.37}$$

$$x_{\mathrm{HI}} = 1 - x_{\mathrm{H_2}} \tag{B.38}$$

Summarizing the rate equations, we find:

$$\frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} = \frac{1}{3} n_{\mathrm{H}} k_{\mathrm{O,H_3^+}} x_{\mathrm{O}} x_{\mathrm{H_3^+}} \alpha_{\mathrm{OH_x}} + \frac{1}{3} n_{\mathrm{H}} k_{\mathrm{C,H_3^+}} x_{\mathrm{C}} x_{\mathrm{H_3^+}} \alpha_{\mathrm{CH_x}} + \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} \alpha_{\mathrm{CH_x}} - k_{\gamma,\mathrm{CO}} x_{\mathrm{CO}} \tag{B.39}$$

$$\begin{aligned}\frac{\mathrm{d}x_{\mathrm{C}}}{\mathrm{d}t} &= -\frac{1}{3} n_{\mathrm{H}} k_{\mathrm{O,H_3^+}} x_{\mathrm{O}} x_{\mathrm{H_3^+}} \alpha_{\mathrm{OH_x}} - \frac{1}{3} n_{\mathrm{H}} k_{\mathrm{C,H_3^+}} x_{\mathrm{C}} x_{\mathrm{H_3^+}} \alpha_{\mathrm{CH_x}} - k_{\gamma,\mathrm{C}} x_{\mathrm{C}} + k_{\gamma,\mathrm{CO}} x_{\mathrm{CO}} \\ &\quad + n_{\mathrm{H}} k_{\mathrm{C^+,e^-}} x_{\mathrm{C^+}} x_{\mathrm{e^-}} + (1 - \alpha_{\mathrm{CH_x}}) n_{\mathrm{H}} \frac{1}{2} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} \\ &= -\frac{\mathrm{d}x_{\mathrm{CO}}}{\mathrm{d}t} + \frac{1}{2} n_{\mathrm{H}} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} - k_{\gamma,\mathrm{C}} x_{\mathrm{C}} + n_{\mathrm{H}} k_{\mathrm{C^+,e^-}} x_{\mathrm{C^+}} x_{\mathrm{e^-}}\end{aligned} \tag{B.40}$$

$$\frac{\mathrm{d}x_{\mathrm{C^+}}}{\mathrm{d}t} = n_{\mathrm{H}} k_{\mathrm{C^+,e^-}} x_{\mathrm{C^+}} x_{\mathrm{C^+}} - n_{\mathrm{H}} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} + k_{\gamma,\mathrm{C}} x_{\mathrm{C}} \tag{B.41}$$

$$\frac{\mathrm{d}x_{\mathrm{H_2}}}{\mathrm{d}t} = 2 k_{\mathrm{H_2,gr}} x_{\mathrm{HI}} n_{\mathrm{H}} - k_{\gamma} x_{\mathrm{H_2}} - k_{\mathrm{c.r.,H_2}} x_{\mathrm{H_2}} - n_{\mathrm{H}} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} + \frac{2}{3} n_{\mathrm{H}} k_{\mathrm{e^-,H_3^+}} x_{\mathrm{e^-}} x_{\mathrm{H_3^+}} \tag{B.42}$$

The last equation can be rewritten using the obtained expression for $x_{\mathrm{H_3^+}}$:

$$\frac{\mathrm{d}x_{\mathrm{H_2}}}{\mathrm{d}t} = 2 k_{\mathrm{H_2,gr}} x_{\mathrm{HI}} n_{\mathrm{H}} - k_{\gamma} x_{\mathrm{H_2}} - n_{\mathrm{H}} k_{\mathrm{C^+,H_2}} x_{\mathrm{C^+}} x_{\mathrm{H_2}} \tag{B.43}$$

$$- k_{\mathrm{c.r.,H_2}} x_{\mathrm{H_2}} \frac{k_{\mathrm{C,H_3^+}} x_{\mathrm{C}} + k_{\mathrm{O,H_3^+}} x_{\mathrm{O}}}{k_{\mathrm{C,H_3^+}} x_{\mathrm{C}} + k_{\mathrm{O,H_3^+}} x_{\mathrm{O}} + k_{\mathrm{e^-,H_3^+}} x_{\mathrm{C^+}}} \tag{B.44}$$

One has to solve only 2 of the differential equations for the carbon species as the remaining one can be simply obtained by the conservation law.

APPENDIX C

Integrals for the subgrid model

We end up with 9 integrals and 3 density cutoffs n_c, $n_{c,ion}$ and $n_{c,mol}$ which are calculated according to equations (3.16), (3.17) and (3.18). $\mathcal{P}(n_H)$ and C_ρ are defined in equations (3.13) and (3.14). The integrals are found to be:

1.

$$\langle 2k_{H_2,gr} x_{HI} n_H \rangle = \int_0^{n_c} 2k_{H_2,gr} n_H \mathcal{P}(n_H) dn_H \tag{C.1}$$

The formation of H_2 takes place in the atomic hydrogen phase. In the code, this integral has the designation INT_I.

2.

$$\begin{aligned}
&\langle -k_\gamma x_{H_2} - k_{c.r.,H_2} + \frac{2}{3} n_H k_{e^-,H_3^+} x_{e^-} x_{H_3^+} \rangle \\
&= -\int_{n_c}^{\infty} \left(k_\gamma x_{H_2} + k_{c.r.,H_2} x_{H_2} - \frac{2}{3} n_H k_{e^-,H_3^+} x_{e^-} x_{H_3^+} \right) \mathcal{P}(n_H) dn_H \qquad \text{(C.2)} \\
&= \underbrace{-\int_{n_c}^{\infty} k_\gamma x_{H_2} \mathcal{P}(n_H) dn_H}_{\text{INT_IIa}} \underbrace{-\int_{n_c}^{\infty} \left(k_{c.r.,H_2} x_{H_2} - \frac{2}{3} n_H k_{e^-,H_3^+} x_{e^-} x_{H_3^+} \right) \mathcal{P}(n_H) dn_H}_{\text{INT_IIb_IIc}} \qquad \text{(C.3)}
\end{aligned}$$

The photodissociation and interaction with cosmic rays of H_2 can only happen in the phase where hydrogen is purely molecular. In the code, these integrals have the designation INT_IIa and INT_IIb_IIc.

3.

$$-\langle n_H k_{C^+,H_2} x_{C^+} x_{H_2} \rangle = -\int_{n_c}^{n_{c,ion}} n_H k_{C^+,H_2} x_{C,tot} \mathcal{P}(n_H) dn_H \tag{C.4}$$

The integration limits are set for the full range of densities, since H_2 exists in high density regions and C^+ in the ionized phase. To give a contribution, the condition $n_c < n_{c,ion}$ has to be fulfilled. Otherwise, there is not overlap of the ionized phase of carbon and the molecular phase of hydrogen and no such reaction can take place. In the code, this integral has the designation INT_III.

4.

$$-\langle k_{\gamma,C} x_C \rangle = -\int_{n_{c,ion}}^{n_{c,mol}} k_{\gamma,C} x_{C,tot} \mathcal{P}(n_H) \mathrm{d}n_H \tag{C.5}$$

The ionization of carbon happens between the density cutoffs where C is ionized and molecular in form of CO. In the code, this integral has the designation INT_IV.

5.

$$\langle n_H k_{C^+,e^-} x_{C^+} x_{e^-} \rangle \approx \langle n_H k_{C^+,e^-} x_{C^+} x_{C^+} \rangle = \int_0^{n_{c,ion}} n_H k_{C^+,e^-} x_{C,tot} x_{C,tot} \mathcal{P}(n_H) \mathrm{d}n_H \tag{C.6}$$

The recombination of C^+ happens in the ionized phase. In the code, this integral has the designation INT_V.

6. Integrals for $\langle \frac{\mathrm{d}x_{CO}}{\mathrm{d}t} \rangle$:

a)

$$-\langle k_{\gamma,CO} x_{CO} \rangle = -\int_{n_{c,mol}}^{\infty} k_{\gamma,CO} x_{C,tot} \mathcal{P}(n_H) \mathrm{d}n_H \tag{C.7}$$

The photodissociation only at those densities where CO is present. In the code, this integral has the designation INT_VIa.

b)

$$\frac{1}{2} \langle n_H k_{C^+,H_2} x_{C^+} x_{H_2} \alpha_{CH_x} \rangle = \int_{n_c}^{n_{c,ion}} \frac{1}{2} n_H k_{C^+,H_2} x_{C,tot} \alpha_{CH_x} \mathcal{P}(n_H) \mathrm{d}n_H \tag{C.8}$$

See integral INT_3 for the explanation of the integration limits. In the code, this integral has the designation INT_VIb.

c)

$$\begin{aligned} &\frac{1}{3} \langle n_H k_{O,H_3^+} x_O x_{H_3^+} \alpha_{OH_x} + n_H k_{C,H_3^+} x_C x_{H_3^+} \alpha_{CH_x} \rangle \\ &= \frac{1}{3} \int_{n_{c,ion}}^{n_{c,mol}} \left(n_H k_{O,H_3^+} x_O x_{H_3^+} \alpha_{OH_x} + n_H k_{C,H_3^+} x_C x_{H_3^+} \alpha_{CH_x} \right) \mathcal{P}(n_H) \mathrm{d}n_H \end{aligned} \tag{C.9}$$

For this reaction, the presence of neutral carbon needed, that exists between densities $n_{c,ion}$

and $n_{\mathrm{c,mol}}$. The $x_{\mathrm{H}_3^+}$ has been replaced by the expression obtained in equation B.29. In the code, this integral has the designation INT_VIc.

Since we have now the density cutoffs, we have now the integration limits as well. For the integration, we use the Gauss–Legendre quadrature method. This uses that an integral between -1 and 1 can be rewritten to

$$\int_{-1}^{1} f(x)\mathrm{d}x = \sum_{i=1}^{n} w_i f(x_i) \tag{C.10}$$

where x_i denote the position of a node within the integration interval and n represents the number of nodes. For arbitrary integration limits, we use

$$\int_{a}^{b} f(x)\,\mathrm{d}x = \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{b-a}{2}x + \frac{a+b}{2}\right)\mathrm{d}x. \tag{C.11}$$

We still need to find the values for the w_i which are simply defined by

$$w_i = \frac{2}{\left(1-x_i^2\right)\left[P_n'(x_i)\right]^2}. \tag{C.12}$$

In the case of the Gauss–Legendre quadrature method, the x_i are the roots of the n-th Legendre polynomials, $P_n'(x_i)$ are the values for the derivatives of the $n-$th polynomial at x_i.
The roots of the polynomials are not that easy to find. However, there is an expression to approximate the x_i given by

$$x_i \approx \cos\left(\pi\,\frac{4i-1}{4n+2}\right), \quad i = 1,\ldots,n. \tag{C.13}$$

The error reduces with increasing n.
Additionally to the integrals, we can now calculate the column densities as well.

$$N_{\mathrm{H}}(n_{\mathrm{H}}) = \langle n_{\mathrm{H}}\rangle \Delta x \int_{0}^{n_{\mathrm{H}}} \mathrm{d}n'_{\mathrm{H}}\mathcal{P}_{\mathrm{M}}(n'_{\mathrm{H}}) \tag{C.14}$$

$$= \frac{\langle n_{\mathrm{H}}\rangle \Delta x}{2}\left[1 - \mathrm{erf}\left(\frac{\mu - \ln(n_H)}{\sqrt{2}\sigma}\right)\right] \tag{C.15}$$

$$N_{\mathrm{H}_2}(n_{\mathrm{H}}) = \frac{1}{2}\langle n_{\mathrm{H}}\rangle \Delta x \int_{n_{\mathrm{c}}}^{n_{\mathrm{H}}} \mathrm{d}n'_{\mathrm{H}}\mathcal{P}_{\mathrm{M}}(n'_{\mathrm{H}}) \tag{C.16}$$

$$= \frac{\langle n_{\mathrm{H}}\rangle \Delta x}{4}\left[\mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{c}})}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{H}})}{\sqrt{2}\sigma}\right)\right] \tag{C.17}$$

$$N_{\mathrm{CO}}(n_{\mathrm{H}}) = \langle n_{\mathrm{H}} \rangle x_{\mathrm{C,tot}} \Delta x \int_{n_{\mathrm{c,mol}}}^{n_{\mathrm{H}}} \mathrm{d}n'_{\mathrm{H}} \mathcal{P}_{\mathrm{M}}(n'_{\mathrm{H}}) \tag{C.18}$$

$$= \frac{\langle n_{\mathrm{H}} \rangle x_{\mathrm{C,tot}} \Delta x}{2} \left[\mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{c,mol}})}{\sqrt{2}\sigma} \right) - \mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{H}})}{\sqrt{2}\sigma} \right) \right] \tag{C.19}$$

$$N_{\mathrm{C}}(n_{\mathrm{H}}) = \langle n_{\mathrm{H}} \rangle x_{\mathrm{C,tot}} \Delta x \int_{n_{\mathrm{c,ion}}}^{n_{\mathrm{H}}} \mathrm{d}n'_{\mathrm{H}} \mathcal{P}_{\mathrm{M}}(n'_{\mathrm{H}}) \tag{C.20}$$

$$= \frac{\langle n_{\mathrm{H}} \rangle x_{\mathrm{C,tot}} \Delta x}{2} \left[\mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{c,ion}})}{\sqrt{2}\sigma} \right) - \mathrm{erf}\left(\frac{\mu - \ln(n_{\mathrm{H}})}{\sqrt{2}\sigma} \right) \right] \tag{C.21}$$

APPENDIX D

Masses of H_2, CO, C^+ and C for the post-processed galaxy population

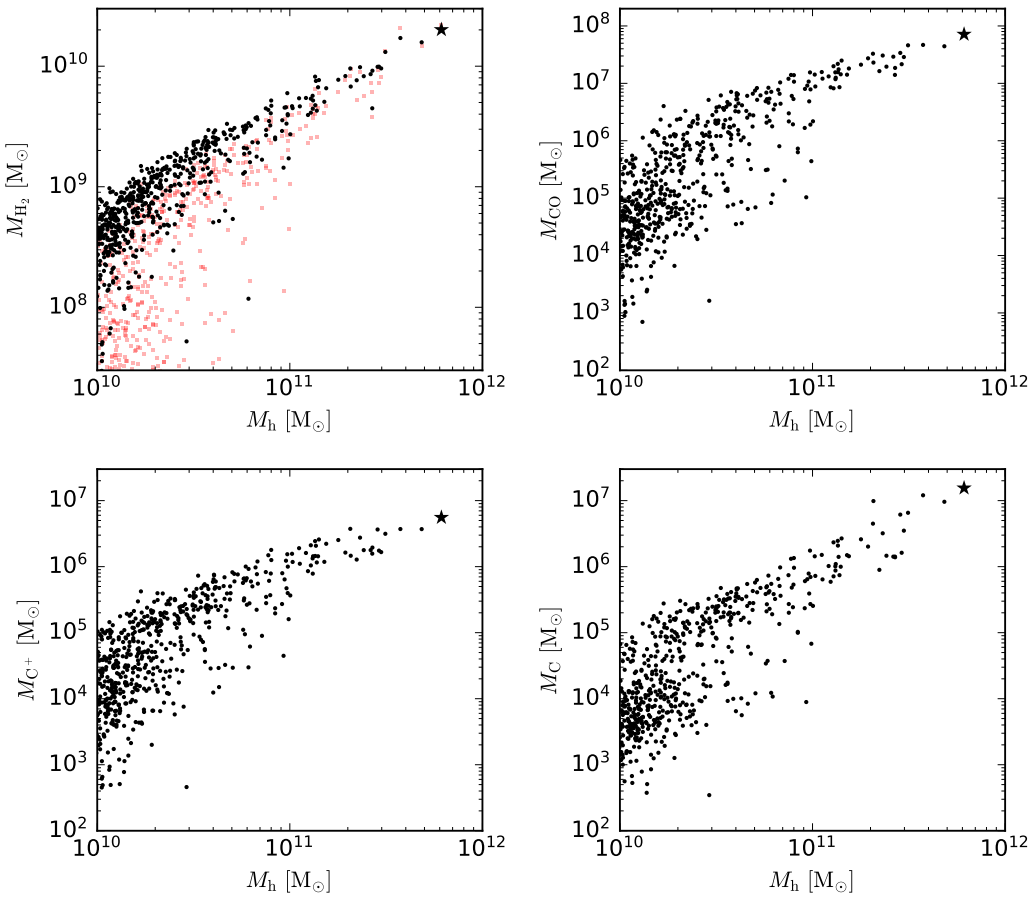

Figure D.1: Abundances of H_2, CO, C and C^+ for different halo masses in the D9$^+$ simulation (see chapter 2) for $z = 3.6$. The distribution of the different species across densities and radius in this galaxy represented by the star symbol is shown in figure 3.7.

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank my supervisor Prof. Dr. Cristiano Porciani who gave me the opportunity to do the PhD, to learn new things and broaden my knowledge.

I would like to thank Dr. Emilio Romano-Díaz who introduced me into the world of galaxy simulations.

Also, a special thanks goes to the AIfA secretaries Christina Stein-Schmitz, Sabine Derdau, Ellen Fuhrmann and Elisabeth Kramer. Their help with administrative things, documents or problems of any kind was always appreciated. I will also never forget the collaboration to maintain and clean the coffee maker.

Thanks also to Sandra, Steffi, and Helge for being good friends.

Also a very furry thanks goes to our two cats Hikari and Momo. They never left any opportunity to cheer me up. Or to annoy me when feeding time was approaching.

Particular thanks goes to my wonderful wife, Kim, who accompanied me for the entire journey. I am grateful and happy that I met you and that our journey will still continue.

Thank you, Jonathan, for being the best son I could imagine.

I also want to thank my parents and my brother for their support throughout the years.

Finally, I acknowledge the financial support I received through the SFB 956 'The Conditions and Impact of Star Formation', sub-project C4 by the Deutsche Forschungsgemeinschaft (DFG).