Leibniz-Institut für
Astrophysik Potsdam

Forschungsabteilung "Kosmologie und großräumige Strukturen"

Connecting simulations and observations in galaxy formation studies

Dissertation
zur Erlangung des akademischen Grades
"doctor rerum naturalium"
(Dr. rer. nat.)
in der Wissenschaftsdisziplin "Astrophysik"

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Giovanni Guidi

Potsdam, den 01.11.2016

Betreuer:

Prof. Dr. Matthias Steinmetz, Leibniz-Institut für Astrophysik Potsdam (AIP)
Dr. Cecilia Scannapieco, Leibniz-Institut für Astrophysik Potsdam (AIP)

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-396876
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-396876

Abstract

Observational and computational extragalactic astrophysics are two fields of research that study a similar subject from different perspectives. Observational extragalactic astrophysics aims, by recovering the spectral energy distribution of galaxies at different wavelengths, to reliably measure their properties at different cosmic times and in a large variety of environments. Analyzing the light collected by the instruments, observers try to disentangle the different processes occurring in galaxies at the scales of galactic physics, as well as the effect of larger scale processes such as mergers and accretion, in order to obtain a consistent picture of galaxy formation and evolution. On the other hand, hydrodynamical simulations of galaxy formation in cosmological context are able to follow the evolution of a galaxy along cosmic time, taking into account both external processes such as mergers, interactions and accretion, and internal mechanisms such as feedback from Supernovae and Active Galactic Nuclei. Due to the great advances in both fields of research, we have nowadays available spectral and photometric information for a large number of galaxies in the Universe at different cosmic times, which has in turn provided important knowledge about the evolution of the Universe; at the same time, we are able to realistically simulate galaxy formation and evolution in large volumes of the Universe, taking into account the most relevant physical processes occurring in galaxies.

As these two approaches are intrinsically different in their methodology and in the information they provide, the connection between simulations and observations is still not fully established, although simulations are often used in galaxies' studies to interpret observations and assess the effect of the different processes acting on galaxies on the observable properties, and simulators usually test the physical recipes implemented in their hydrodynamical codes through the comparison with observations. In this dissertation we aim to better connect the observational and computational approaches in the study of galaxy formation and evolution, using the methods and results of one field to test and validate the methods and results of the other.

In a first work we study the biases and systematics in the derivation of the galaxy properties in observations. We post-process hydrodynamical cosmological simulations of galaxy formation to calculate the galaxies' Spectral Energy Distributions (SEDs) using different approaches, including radiative transfer techniques. Comparing the direct results of the simulations with the quantities obtained applying observational techniques to these synthetic SEDs, we are able to make an analysis of the biases intrinsic in the observational algorithms, and quantify their accuracy in recovering the galaxies' properties, as well as estimating the uncertainties affecting a comparison between simulations and observations when different approaches to obtain the observables are followed. Our results show that for some quantities such as the stellar ages, metallicities and gas oxygen abundances large differences can appear, depending on the technique applied in the derivation.

In a second work we compare a set of fifteen galaxies similar in mass to the Milky Way and with a quiet merger history in the recent past (hence expected to have properties close

to spiral galaxies), simulated in a cosmological context, with data from the Sloan Digital Sky Survey (SDSS). We use techniques to obtain the observables as similar as possible to the ones applied in SDSS, with the aim of making an unbiased comparison between our set of hydrodynamical simulations and SDSS observations. We quantify the differences in the physical properties when these are obtained directly from the simulations without post-processing, or mimicking the SDSS observational techniques. We fit linear relations between the values derived directly from the simulations and following SDSS observational procedures, which in most of the cases have relatively high correlation, that can be easily used to more reliably compare simulations with SDSS data. When mimicking SDSS techniques, these simulated galaxies are photometrically similar to galaxies in the SDSS blue sequence/green valley, but have in general older ages, lower SFRs and metallicities compared to the majority of the spirals in the observational dataset.

In a third work, we post-process hydrodynamical simulations of galaxies with radiative transfer techniques, to generate synthetic data that mimic the properties of the CALIFA Integral Field Spectroscopy (IFS) survey. We reproduce the main characteristics of the CALIFA observations in terms of field of view and spaxel physical size, data format, point spread functions and detector noise. This 3-dimensional dataset is suited to be analyzed by the same algorithms applied to the CALIFA dataset, and can be used as a tool to test the ability of the observational algorithms in recovering the properties of the CALIFA galaxies. To this purpose, we also generate the resolved maps of the simulations' properties, calculated directly from the hydrodynamical snapshots, or from the simulated spectra prior to the addition of the noise.

Our work shows that a reliable connection between the models and the data is of crucial importance both to judge the output of galaxy formation codes and to accurately test the observational algorithms used in the analysis of galaxy surveys' data. A correct interpretation of observations will be particularly important in the future, in light of the several ongoing and planned large galaxy surveys that will provide the community with large datasets of properties of galaxies (often spatially-resolved) at different cosmic times, allowing to study galaxy formation physics at a higher level of detail than ever before. We have shown that neglecting the observational biases in the comparison between simulations and an observational dataset may move the simulations to different regions in the planes of the observables, strongly affecting the assessment of the correctness of the sub-resolution physical models implemented in galaxy formation codes, as well as the interpretation of given observational results using simulations.

Contents

Chapter 1

Introduction

Physical cosmology is now in a revolutionary epoch, often referred to as the era of "precision cosmology" to underline that in the last years, thanks to the improvements in the observational techniques and instruments, several high-precision measurements have set strong constraints in the theoretical speculations about the origin and evolution of the Universe, and have established Λ-Cold Dark Matter (also called the "standard cosmological paradigm") as the most successful cosmological model, determining its free parameters at a high level of accuracy.

Modern cosmological science was born after the publication of Einstein's General Relativity in 1915, that soon became the framework for all considerations about the physical nature of the Universe. Since then the field has progressed quickly; Friedmann, Lemaître, Robertson and Walker independently determined in the 1920's the metric of a homogeneous and isotropic Universe, and the expansion of the Universe predicted by Friedmann and Lemaître was subsequently confirmed by Hubble in 1929. In the early 1930's Oort and Zwicky speculated for the first time about the existence of a form of non-luminous (i.e. dark) matter, confirmed in the 1970's by Ford and Rubin. In 1948, Gamow, Alpher, and Hermann suggested that at early times the Universe had very high densities and temperatures, and named this scenario "Big Bang" (although a similar idea was conjectured by Lemaître earlier in 1931); they also predicted that this initial state may have been imprinted into an hypothetical background radiation, at a temperature of only few Kelvin now. This radiation was later found by Penzias and Wilson in 1964 in the microwave region of the radio spectrum, and has been called the "Cosmic Microwave Background" (CMB). The COBE satellite started in 1989 the first high-precision measurement of the temperature of the CMB, beginning a new era in Cosmology, in which the predictions of theoretical and numerical calculations can be compared against observations at a great level of accuracy; it established that the CMB follows closely the spectral distribution of a black-body with temperature $T \sim 2.7$ K and tiny anisotropies ($\Delta T/T \sim 10^{-5}$) due to small inhomogeneities in the matter distribution in the Early Universe. After the discovery of the accelerated expansion of the Universe by Riess et al. in 1998 using high-redshift supernovae, which gave strong evidence for the existence of a cosmological constant Λ, later missions such as BOOMERANG (Crill et al., 2003), MAXIMA (Rabii et al., 2006), DASI (Leitch et al., 2002), WMAP (Bennett et al., 2013) and PLANCK (Planck Collaboration et al., 2015a) have confirmed that the observed CMB power spectrum matches well the predictions of the ΛCDM model, namely a Universe where the energy content is dominated by Cold Dark Matter (CDM, Peebles 1982; Bond, Szalay & Turner 1982; Blumenthal, Pagels & Primack 1982) and a cosmological constant Λ. At the same time, spectroscopic galaxy surveys such as the Two degree Field (2dF, Colless 1999) or the Sloan Digital Sky Survey (SDSS, Abazajian et al. 2003) determined with great

accuracy the spatial distribution of a large number of galaxies, allowing to infer the properties of the Large Scale Structure and to measure the galaxy power spectrum (Coil, 2013; Weinberg et al., 2004; Ostriker & Steinhardt, 1995; Kofman, Gnedin & Bahcall, 1993). In light of these recent observations, the cosmological science, that has been for a long time a purely theoretical and often speculative field of investigation, is becoming a well-established physical science described by a consistent theory subject to high-precision tests.

The recent progress in the observational techniques is also having a strong influence on the study of galaxy formation and evolution in cosmological context. Galaxies are an important aspect of Cosmology, not only because they trace the underlying dark matter distribution and can therefore be used to investigate the properties of the Universe on large scales, but also because their evolution is closely linked to the evolution of the Universe as a whole; studying the statistical properties of galaxies at different cosmic times may in principle help to constrain several important global quantities related to the processes that drive the evolution of the Universe, such as the Cosmic Star Formation History (Madau & Dickinson, 2014), the Stellar Mass function (Baldry et al., 2012; Song et al., 2015) and the Mass-Metallicity relation (Tremonti et al., 2004; Erb et al., 2006). Nowadays, large databases of galaxy photometric and spectral information at different wavelengths and redshifts are available, for instance 2dFGRS, SDSS, 2MASS, ALMA, HUDF, DEEP2, SPITZER, HERSCHEL (Colless, 1999; Abazajian et al., 2003; Werner et al., 2004; Beckwith et al., 2006; Skrutskie et al., 2006; Pilbratt et al., 2010; Hodge et al., 2013; Newman et al., 2013). Moreover, Integral Field Unit (IFU) spectrographs (e.g. MUSE, Bacon et al. 2004; WEAVE, Dalton et al. 2014; PMAS, Roth et al. 2005) are opening the possibility to study spatially-resolved properties of nearby galaxies, and thanks to Integral Field Spectroscopy (IFS) surveys such as CALIFA (Sánchez et al., 2012; García-Benito et al., 2015a), MaNGA (Bundy et al., 2015) and SAMI (Allen et al., 2015), two-dimensional spectral maps of galaxies are now available, providing in turn a more comprehensive view of a galaxy's formation history. This wealth of data may give important insight into the process of galaxy formation and evolution in cosmological context, helping to reveal the action of physical mechanisms occurring in galaxies, both internal – e.g. feedback (Fabian, 2012), cooling (Thoul & Weinberg, 1995) – and in relation to larger-scale mechanisms – mergers (Naab et al., 2007; Naab, Johansson & Ostriker, 2009), interactions (Scudder et al., 2012; Stierwalt et al., 2015), accretion (Putman, Peek & Joung, 2012). All these processes leave imprints on the shape of the Spectral Energy Distributions (SEDs) which constitute the primary source of information from large galaxy surveys. However, recovering the physical properties of galaxies from observations is often plagued by uncertainties and biases due both to instrumental effects and to the assumptions and simplifications intrinsic in the techniques applied to obtain the estimations; evaluating the uncertainties in the measurements is usually difficult, since reliable and independent tests to asses the accuracy of the observational algorithms are rarely available.

A complementary way to study galaxy formation and evolution in cosmological context are numerical simulations, usually run in big computing facilities with a large number of CPUs. With the improvement of the numerical techniques in recent years, it is now possible with hydrodynamical codes to self-consistently simulate the formation of galaxies in a cosmological environment, following the linked evolution of gas and dark matter from the early stages of collapse. Since the work of Navarro, Frenk & White (1997); Navarro & Steinmetz (2000), it has been evident that, in order to simulate the formation of disk galaxies, it is required to additionally model at least some of the most relevant physical processes at the galactic scale, such as supernovae explosions, gas cooling, chemical enrichment, stellar winds, and feedback from central supermassive black holes, as these processes may have a great impact on the final properties of galaxies, in particular in relation to the formation

of the disk (Governato et al., 2004; Scannapieco et al., 2008; Piontek & Steinmetz, 2011; Springel, 2012). The amount of data available about the properties of galaxies at different cosmic times provides a direct way to test the reliability of the description of the various galactic-scale physical processes in the hydrodynamical codes, as comparing the results of simulations with observations is often the only way to evaluate whether the physical recipes embedded in the codes are somehow realistic. In fact, most of the relevant processes - e.g. star formation, SN and BH feedback - are implemented in the codes with a certain degree of approximation, because the physical scales at which they act are unresolved in the simulations, and a detailed understanding of these processes is still lacking. Despite the progress in the development of simulations codes in recent years, it still remains difficult to form fully realistic spirals from cosmological initial conditions, although several works have been quite successful in producing galaxies with larger disks and properties that resemble observational results (e.g. Governato et al. 2010; Aumer et al. 2013; Vogelsberger et al. 2014; Wang et al. 2015; Governato et al. 2007; Scannapieco et al. 2008; Nelson et al. 2015; Schaye et al. 2015).

Albeit in the last years observations and simulations have strongly progressed individually, much uncertainty remains in the comparison between them, even when observations are used to decide on the successes and failures of the models, and comparison with simulations is an important aspect of the interpretation of observational results. In fact, the information given by the simulations is intrinsically different from what is recovered from observations. While the former are based on the mass distribution, the latter are recovered from the analysis of the light collected by the photometric cameras or spectrographs. In addition, each galaxy survey has its own observational limitation and biases, which must be taken into account when simulations are compared with observations in order to validate a physical model, or if possible improvements to the model are examined. Previous works using simulations have shown that, as the methods usually applied to derive the properties of simulated galaxies are very different from observational techniques, several biases might be introduced making the comparisons unreliable (e.g. Abadi et al., 2003; Governato et al., 2009; Scannapieco et al., 2010; Snyder et al., 2011; Munshi et al., 2013; Christensen et al., 2014). A possible solution, already employed in some works (e.g. Governato et al. 2009; Scannapieco et al. 2010; Snyder et al. 2015), is to use techniques similar to the observational ones to recover the properties of the simulated galaxies. This requires a conversion of the simulations into synthetic observations, to which the observational algorithms that estimate the properties of the galaxies can be applied (see e.g. Bellovary et al. 2014; Michałowski et al. 2014; Hayward et al. 2014; Smith & Hayward 2015; Hayward & Smith 2015).

This thesis

This thesis addresses the problem of improving the techniques to compare the simulations and observations in a more meaningful and consistent way, in order to better establish the connection among them. On one side, simulators can take advantage of more consistently comparing the results of simulations with observations to better assess the validity of the physical recipes and numerical schemes implemented in their hydrodynamical codes; on the other side, observers can use simulated data to test their observational algorithms, with the advantage that the quantities to be recovered are known by construction.

In this dissertation we discuss results obtained converting the outputs of recent hydrodynamical simulations of galaxy formation in cosmological context into synthetic observations using different techniques, including state-of-the-art radiative transfer algorithms. We estimate the differences in the galaxies' SEDs arising by the different assumptions in the radiative transfer calculations, such as the amount of dust extinction, the choice of the Initial Mass Function, the Stellar Population Synthesis model assumed, and the orientation.

We apply to the synthetic observations several algorithms commonly used by observers to recover the physical properties of the galaxies, testing their accuracy and the effects of different observational biases. We also use these synthetic observations to consistently compare the simulations with data from the Sloan Digital Sky Survey (SDSS, Abazajian et al. 2003), applying the SDSS techniques to the synthetic spectra to recover the galaxies' physical properties, in order to carefully test the sub-resolution physics implemented in the hydrodynamical codes. The hydrodynamical simulations have been also post-processed to generate, with a radiative transfer code, a synthetic dataset mimicking the Integral Field Spectroscopy (IFS) CALIFA survey (Sánchez et al., 2012), that can be used to test the algorithms that recover the physical properties of galaxies in IFS survey such as CALIFA.

Structure of the thesis

After the general introduction given in this Chapter, in Chapter 2 we present an overview to the theory of galaxy formation in a ΛCDM Universe, describing also the most common numerical methods used to simulate galaxy formation in cosmological context.

In Chapter 3 we discuss the conversion of simulations into mock observations, and we study the systematic effects arising when various techniques commonly applied in galaxy surveys are used to recover the physical properties of simulated galaxies. We estimate the accuracy of each method and the effects of the most relevant biases, providing also a comparison among the methods.

Next, in Chapter 4, we analyze mock data generated from hydrodynamical simulations mimicking the techniques and biases of the Sloan Digital Sky Survey (SDSS), and we consistently compare the properties of simulated galaxies with the SDSS dataset, in order to estimate how close the simulations are to the galaxies in SDSS. We also study how the properties change when simulations are compared with observations without post-processing, as is usually done in simulation studies.

In Chapter 5 we describe the generation of mock datacubes from hydrodynamical simulations mimicking the properties of the CALIFA observations. We also describe how we create the maps of the resolved properties of the simulated galaxies, that provide a benchmark to test the results recovered by the observational algorithms applied to our resolved synthetic spectra.

Finally, we conclude and summarize in Chapter 6.

Remarks on the authorship

Chapter 3 has been published in the journal MNRAS (Guidi, Scannapieco & Walcher, 2015). The cosmological simulations used in this work have been provided by Cecilia Scannapieco (using for five of them an update to the chemical enrichment model done by Pierre-Antoine Poulhazan) and Michael Aumer, while I developed the techniques used to convert simulations into mock observations, and ran the radiative transfer code SUNRISE. Most of the analysis has been carried out by me, while Anna Gallazzi determined the mean stellar ages and metallicities from the mock data using the SDSS pipeline and Jakob Walcher provided the code for estimating the photometric stellar mass. I wrote the paper with the help of Cecilia Scannapieco and comments from my co-authors.

Chapter 4 has been also published in MNRAS (Guidi et al., 2016). The simulations have been provided by Cecilia Scannapieco and Michael Aumer. Anna Gallazzi gave several suggestions to improve the Lick indices analysis, and used the SDSS pipeline to estimate the mean age and metallicity of the simulations. The paper was written by me, while all co-authors gave various advices and minor comments to the manuscript.

Chapter 5 has been submitted to MNRAS. The cosmological simulations have been run

by Cecilia Scannapieco and Michael Aumer, while Javier Casado made the analysis and the modelling of the noise in CALIFA, and added the point spread functions to the synthetic datacubes. I ran the radiative transfer simulations with SUNRISE and converted them into the CALIFA dataformat. Sec. 5.3 of this chapter has been written by Javier Casado; his contribution is marked in the text. The other parts of the text have been written by me, while all co-authors helped with discussions and comments.

Chapter 2

Galaxy formation in cosmological context

The formation of galaxies in a cosmological context is a complex process that involves highly non-linear physics on scales over several orders of magnitude, ranging from the microscopic scales of atomic physics up to the cosmological distances of the Large Scale Structure (LSS). Therefore, a fully-descriptive galaxy formation theory must consistently link the physics on the smaller scales with the one on the larger.

The modern theory of galaxy formation developed together with studies of cosmology and LSS formation, and is now well-established within the so-called ΛCDM (Lambda-Cold Dark Matter) cosmological paradigm, also known as 'concordance model' (Weinberg, 1972; Peebles, 1993; Peacock, 1999). According to the ΛCDM model the total energy content of the Universe is divided into a cosmological constant or Dark Energy Λ ($\sim 70\%$) responsible for the observed accelerated expansion of the Universe (Perlmutter et al., 1999), an hypothetical form of matter which moves slowly compared to the speed of light called Cold Dark Matter (CDM, $\sim 25\%$), and the ordinary baryonic matter ($\sim 5\%$). Although the nature of Cold Dark Matter (which is the dominant component of the matter content of the Universe) is still under debate, physicists believe it could consist of new and yet undiscovered weakly-interacting massive particles (WIMPs); these particles appear for instance in super-symmetric extensions of the Standard Model of particle physics (Jungman, Kamionkowski & Griest, 1996).

In the ΛCDM model, all the cosmic structures that we observe now grew from the quantum fluctuations of a primordial field in the early Universe, transformed into macroscopic density fluctuations right after the Big Bang during a period of Inflation, when the scale factor grew exponentially (Guth, 1981; Liddle & Lyth, 2000). During Inflation, the fluctuations in the primordial quantum field have been stretched outside the cosmological horizon where no causal physics can affect them, and then remained with constant amplitude until they re-entered the horizon at later times, in the period of the standard expansion of the Universe (Baumann, 2009). Once they re-enter the horizon, these density fluctuations grew under the effect of gravity, collapsing hierarchically forming first small dark matter halos and then aggregating into progressively larger systems (hence ΛCDM is also known as a 'bottom-up' formation scenario). In the first stages of collapse the baryons trace the dark matter distribution (which is the main component of the density fluctuations) and later move to the center of the potential well generated by the dark matter halo, forming a protogalaxy (White & Rees, 1978; Gnedin & Hui, 1998). Star formation transforms the high-density gas of the protogalaxy into stars, that then enrich the InterStellar Medium (ISM) with metals produced during stellar nucleosynthesis. Infalling gas from the surroundings may replenishing the galaxy's ISM with hydrogen and helium and further fuel star formation, while

merging processes make a galaxy incorporate the stellar and gas content of other objects.

Although the ΛCDM model is largely supported by the results of different experimental measurements of the Cosmic Microwave Background (CMB) acoustic peaks (e.g. Planck Collaboration et al. 2015a) and the LSS (e.g. Doroshkevich et al. 2004), alternative cosmological paradigms are still under debate, in particular regarding the nature of dark matter (Bertone, 2010) and dark energy (Amendola et al., 2007; Frieman, Turner & Huterer, 2008). Famous alternatives to CDM are the Hot Dark Matter (HDM) scenarios, according to which larger structures form first and then fragment into smaller objects (top-down structure formation), albeit these scenarios have been ruled out by observations already in the 80's (see Frenk & White 2012), or the Warm Dark Matter (WDM) paradigm, in which the dark matter particle is a sterile neutrino or gravitino (Bode, Ostriker & Turok, 2001) forming structures hierarchically at larger scales, while suppressing the collapse below a characteristic mass scale by free streaming of the particles (Schneider, Smith & Reed, 2013). Alternatives to the cosmological constant Λ are for instance quintessence theories, which describe dark energy as a minimally-coupled scalar field ϕ (Ratra & Peebles, 1988), or modifications to General Relativity which are able to explain the late-time accelerated expansion of the Universe without introducing dark energy (for instance the $f(r)$ theories, see e.g. de Felice & Tsujikawa 2010).

From this general picture it emerges that galaxy formation is a process driven at large scales by the formation and evolution of dark matter halos, and later, when baryonic processes start to be relevant, by the physics at smaller scales. The goal of galaxy formation theories is then to connect the properties of the galaxies that we observe in the sky, such as their morphologies, spatial distributions, magnitudes, colours, stellar and gas content, amount of metals, with the various physical processes acting at the different scales during their evolution.

In the following paragraphs we will review the most relevant aspects of galaxy formation theory in a ΛCDM cosmological context.

2.1 ΛCDM cosmological background

Modern cosmology is based on the assumption that the Universe is, on large scales, homogeneous and isotropic, the so-called "cosmological principle". In the framework of General Relativity, this assumption determines the metric of the Universe, that must be of the Robertson-Walker form, where the line element (in spherical comoving coordinates r, θ, ϕ) is

$$ds^2 = c^2 dt^2 - a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta \, d\phi^2) \right). \tag{2.1}$$

The integer parameter k can take the values {-1, 0, 1}, respectively, for a Universe with negative, flat, and positive curvature. The function $a(t)$ is called the *scale factor*, and is related to the physical (or proper) distance $d(t)$ in the following way:

$$d(t) = a(t) \, d_0 \tag{2.2}$$

where d_0 is the distance at the reference time t_0; usually, $a(t_0)$ is assumed to be 1 at the present time. The *Hubble parameter* $H(t)$ is defined from the scale factor as

$$H = \frac{\dot{a}}{a} \tag{2.3}$$

often expressed in terms of the dimensionless factor h as $H = 100 \cdot h$ km s^{-1}Mpc^{-1}.

Having defined the Hubble parameter, taking the derivative of the proper distance with respect to time we obtain the Hubble law, determined observationally by Hubble combining measurements of recessional velocity v and distances of galaxies (Hubble, 1929):

$$v = H_0\, d \tag{2.4}$$

where the subscript '0' means the value of the Hubble parameter at the present time.

Another important definition in an expanding Universe described by the Robertson-Walker metric is the redshift z, which measures the shift in wavelength (frequency) of a photon travelling in the Universe, and is easily measured in observations from the shift of the galaxies' spectral features; it is related to $a(t)$ as:

$$z \equiv \frac{\lambda_o}{\lambda_e} - 1 = \frac{a_o}{a_e} - 1 \tag{2.5}$$

where λ_o and λ_e are the wavelengths of the photon respectively at the time of observation and emission, and a_o and a_e the corresponding scale factors.

In General Relativity, the space-time and the energy-momentum density satisfy locally the Einstein equation (with $c \equiv 1$):

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} + \Lambda\, g_{\mu\nu} = 8\pi G\, T_{\mu\nu} \tag{2.6}$$

where $g_{\mu\nu}$ is the metric, $R_{\mu\nu}$ and R are respectively the Ricci curvature tensor and Ricci scalar, $T_{\mu\nu}$ is the energy-momentum tensor, and Λ the cosmological constant.

If we assume that the energy-momentum tensor $T_{\mu\nu}$ is in the form of an ideal fluid with four-velocity u_μ, matter density ρ and pressure p, i.e.

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + p\, g_{\mu\nu} \tag{2.7}$$

and we insert the Robertson-Walker metric (2.1) into the Einstein equation (2.6), we obtain the Friedmann equations

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \\ \frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \end{aligned} \tag{2.8}$$

that determine the evolution of a given ρ, p, Λ and k.

For a perfect fluid, the equation of state $p(\rho)$ follows the general relation

$$p = w\, \rho \tag{2.9}$$

with w constant. A perfect fluid whose pressure is identically zero ($w = 0$) is called dust. An example of such a fluid is the ordinary baryonic matter, for which the pressure is negligible compared to its energy density; according to the Friedmann equations the matter density evolves as the Universe expands as $\rho_m \propto a^{-3}$. On the other hand, radiation (and neutrinos) satisfy the equation of state (2.9) with $w = 1/3$, therefore the energy density decays as $\rho_r \propto a^{-4}$ (faster), with the additional factor $1/a$ due to cosmological redshift. From the cosmological constant term in the Einstein equation (2.6), we can define the energy-momentum tensor of the vacuum as

$$T_{\mu\nu}^{(\mathrm{vac})} = -\frac{\Lambda}{8\pi G} g_{\mu\nu} \tag{2.10}$$

which has the form of a perfect fluid with

$$p = -\rho = -\frac{\Lambda}{8\pi G} \tag{2.11}$$

therefore the equation of state for the cosmological constant has $w = -1$, and the energy density is independent of the scale factor a.

The Friedmann equations have analytical solution in some ideal cases; for instance, for the Einstein-de Sitter model, i.e. a flat, matter-dominated Universe without cosmological constant ($k, \Lambda = 0$, $\rho_m \gg \rho_r$) the solution is $a \propto t^{2/3}$. If the same Universe is dominated by radiation ($k, \Lambda = 0$, $\rho_r \gg \rho_m$) the scale factor is a function of time as $a \propto t^{1/2}$. For a flat Universe with only the cosmological constant term ($k = 0$, $\rho_m, \rho_r = 0$) the evolution of the scale factor is exponential $a \propto \exp(\sqrt{\Lambda/3}\, t)$.

For each matter component, we can define the dimensionless *density parameter* as

$$\Omega_{i,0} \equiv \frac{8\pi G}{3H_0^2}\rho_{i,0} \tag{2.12}$$

where the quantities with the subscript '0' are evaluated at present time, and we will denote the matter and radiation density parameters respectively Ω_m and Ω_r. Similarly, we define the cosmological constant and curvature density parameters according to

$$\begin{aligned} \Omega_\Lambda &\equiv \frac{\Lambda}{3H_0^2} \\ \Omega_k &\equiv k/H_0^2. \end{aligned} \tag{2.13}$$

If we define the 'critical density' as

$$\rho_c \equiv \frac{3H_0}{8\pi G} \tag{2.14}$$

we can rewrite the present-day density parameters as $\Omega_{i,0} = \rho_{i,0}/\rho_c$. The first Friedmann equations becomes

$$H^2 = H_0^2 \left(\Omega_{r,0}/a^4 + \Omega_{m,0}/a^3 + \Omega_{k,0}/a^2 + \Omega_{\Lambda,0}\right) \tag{2.15}$$

where the Hubble parameter is written as a function of the current values of the density parameters.

Table 2.1 shows the values of the most relevant cosmological parameters as measured by the Planck mission. Observations of the Cosmic Microwave Background (CMB) indicate that the Universe is flat ($\Omega_k \simeq 0$) at a high level of precision (Planck Collaboration et al., 2015b), consistent with the predictions of inflation theories (Baumann, 2009).

2.2 Structure formation

The Friedmann equations describe the evolution of the cosmic fluid over scales for which the assumptions of homogeneity and isotropy hold. These assumptions are satisfied in the early stages of the evolution of the Universe, as demonstrated by the high degree of isotropy of the CMB (Boggess et al., 1992), and at later times for very large scales (Hogg et al., 2005). The scale over which the Universe looks homogeneous at present time is still under debate; the upper limit of $\sim$ 370 Mpc given by Yadav, Bagla & Khandai (2010) has been later challenged by Clowes et al. (2013), who claimed that structures over scales $\sim$ 500 Mpc exist in the Universe. On smaller scales, however, the Universe is highly non-homogeneous, and

Parameter	Symbol	Value and uncertainty
Baryon density	$\Omega_b h^2$	0.02230 ± 0.00014
Cold dark matter density	$\Omega_c h^2$	0.1188 ± 0.0010
$100 \times r_*/D_A$ computed with CosmoMC chains	$100\,\theta_{\rm MC}$	1.04093 ± 0.00030
Thomson scattering optical depth	τ	0.066 ± 0.012
Primordial curvature perturbations ($k_0 = 0.05$ Mpc^{-1})	$\ln(10^{10} A_s)$	3.064 ± 0.023
Scalar spectrum power-law ($k_0 = 0.05$ Mpc^{-1})	n_s	0.9667 ± 0.0040
Expansion rate in km s^{-1} Mpc^{-1}	H_0	67.74 ± 0.46
Dark energy density	Ω_Λ	0.6911 ± 0.0062
Total matter density (inc. massive neutrinos)	Ω_m	0.3089 ± 0.0062
RMS matter fluctuations in linear theory	σ_8	0.8159 ± 0.0086
Reionization redshift	$z_{\rm re}$	$8.8^{+1.2}_{-1.1}$
$10^9\times$ curvature power spectrum at $k_0 = 0.05$ Mpc^{-1}	$10^9 A_s$	2.142 ± 0.049
Age of the Universe (in Gyr)	t_0	13.799 ± 0.021
Redshift at which the optical depth equals unity	z_*	1089.90 ± 0.23
Comoving size of the sound horizon at $z = z_*$	r_*	144.81 ± 0.24
$100\times$ angular size of sound horizon at $z = z_*$	$100\,\theta_*$	1.04112 ± 0.00029
Redshift at which baryon-drag optical depth equals unity	$z_{\rm drag}$	1059.68 ± 0.29
Comoving size of the sound horizon at $z = z_{\rm drag}$	$r_{\rm drag}$	147.50 ± 0.24
Characteristic damping comoving wavenumber (in Mpc^{-1})	k_D	0.14038 ± 0.00029
Redshift of matter-radiation equality	$z_{\rm eq}$	3371 ± 23
Comoving horizon angular size at equality	$100\,\theta_{\rm s,eq}$	0.4523 ± 0.0023

Table 2.1: Cosmological parameters derived by the PLANCK mission (Table 4. from Planck Collaboration et al. 2015b)

a large variety of structures appears, such as walls, filaments, clusters and superclusters of galaxies, as well as voids between them; all these structures are connected into the so-called "cosmic web", as suggested both by the results of N-body simulations (Springel, Frenk & White, 2006), and observationally by the measurements of galaxy and quasar clustering (e.g. Davis et al. 1982; Platen et al. 2011; Tempel et al. 2014).

In the standard picture, the formation of all these structures is seeded by microscopic perturbations in the primordial quantum field, expanded during Inflation to cosmological scales at $t \sim 10^{-34}$ s after the Big Bang. After the perturbations exit the cosmological horizon, no causal physics can affect them and they remain with constant amplitude until they re-enter the horizon at a later time. The fluctuations associated with cosmological structures re-enter the horizon when the universe was about 10^5 years old, shortly before the decoupling of the CMB photons (Baumann, 2009); inflation predicts scale-free primordial density fluctuations described by a power spectrum of the form $P(k) \propto k^n$ with $n \simeq 1$. Once inside the horizon, causal physics affect the perturbation amplitudes leading to the acoustic peak structure of the CMB and to the collapse of high-density fluctuations by gravitational instability. Since the physics of the perturbations' evolution inside the horizon is well-understood, CMB and LSS observations can in principle be used to probe the physical conditions of the Universe at time $t \sim 10^{-34}$ s after the Big Bang, at energies far beyond the ones reached by the most powerful particle colliders such as LHC ($\sim$ 13 TeV), in a regime where the Standard Model of Particle Physics is not valid anymore.

When the density fluctuations are small they are in the so-called "linear regime", and the evolution of the fluctuations can be studied analytically expanding at first order the equations of motion. For structure formation studies the non-relativistic Newtonian approximation

provides an accurate description of the early evolution of these perturbations if the scales are much smaller than the Hubble radius ($r_H \equiv c/H_0$), for small values of the gravitational potential ϕ and for non-relativistic matter, which is true in the case of CDM and baryons after decoupling[1].

Once the perturbations become non-linear, their evolution is significantly more complicated and cannot be treated analytically except for simple symmetric cases (e.g. the spherical top-hat model, Gunn 1977; Padmanabhan 1993), which however provide some insight into the basic behavior of the collapse. The most useful tool to study collapse in the non-linear regime are N-body simulations (e.g. Klypin & Shandarin 1983; Springel et al. 2008; Heitmann et al. 2008), which predict that the final result of the non-linear evolution of a dark matter density perturbation is the formation of a dark matter halo, an approximately stable, near-equilibrium state supported against its own self-gravity by the random motions of its constituent particles. According to the ΛCDM paradigm, halos grow in time hierarchically from the first small ones to the later generations formed through merging of the earlier generations of halos (for further details see classic textbooks such as Peebles 1993; Longair 2008).

We will now describe the two regimes, providing both analytical and numerical results.

2.2.1 Linear regime

The linear theory, describing the evolution of density perturbations in the linear regime, is based on the equations of motion of a self-gravitating cosmological fluid with density ρ, pressure p, velocity distribution $\mathbf{v}$, internal energy per unit mass u and gravitational potential Φ. For studying the evolution of cosmological fluids, it is convenient to write the equations in the so-called Lagrangian coordinates, following the motion of a particular fluid element, for which the time derivate is written

$$\frac{d}{dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla). \tag{2.16}$$

The equations describing the fluid are the continuity and Euler equations, the conservation of energy (first law of thermodynamics), and the Poisson equation

$$\begin{aligned} \frac{d\rho}{dt} &= -\rho \nabla \cdot \mathbf{v} \\ \frac{d\mathbf{v}}{dt} &= -\frac{1}{\rho}\nabla p - \nabla\phi \\ \frac{du}{dt} &= -\frac{p}{\rho}\nabla \cdot \mathbf{v} - \frac{\Lambda(u,\rho)}{\rho} \\ \nabla^2\phi &= 4\pi G\rho \end{aligned} \tag{2.17}$$

where $\Lambda(u, \rho)$ is the cooling function (relevant only for the gas component, see sec. 2.3.2).

After perturbing the fluid equations at first order and defining the density contrast δ

$$\delta = \frac{\rho - \bar{\rho}}{\bar{\rho}} \tag{2.18}$$

[1]However, outside r_H or for relativistic fluids (such as photons and neutrinos) the correct description of the evolution of fluctuations is given only by a full general-relativistic treatment (see e.g. reviews by Mukhanov 2005; Miedema 2011).

(where $\bar{\rho}$ is the mean density of the fluid), we can obtain the evolution of the density contrast in time; in the case of a static background (i.e. $\dot{a} = 0$) the density perturbations are separated into two regimes by the *Jeans' length*

$$\lambda_J = \frac{2\pi}{k_J} = c_s \left(\frac{\pi}{G\bar{\rho}} \right)^{1/2}. \tag{2.19}$$

Density perturbations on scales $\lambda < \lambda_J$ result in stable oscillations, while for $\lambda > \lambda_J$ the perturbations collapse. Another important solution is in the case of the Einstein-de Sitter Universe, namely a Universe with cosmological parameters $\{k = 0,\, \Omega_m = 1, \Omega_\Lambda = 0\}$ (which is an approximate description of the early stages of the matter-dominated phase), for which the perturbations grow as a power of the time $\delta \propto t^{2/3} (\propto a)$. Models with arbitrary Ω_m and Ω_Λ (including the ΛCDM model) have perturbations growing as

$$\delta \propto \frac{\dot{a}}{a} \int_0^a \frac{da}{\dot{a}^3} \tag{2.20}$$

and several analytical approximations of this integral can be found in the literature (see e.g. Longair 2008).

2.2.2 Non-linear regime

When the density contrast reaches $\delta \gtrsim 1$ local self-gravitation becomes important and structure formation enters in the non-linear regime, where linear theory is not able to describe the evolution of perturbations anymore; dark matter halos form as final result of gravitational collapse. Although even in the non-linear regime analytical solutions based on simple arguments applied to idealized cases are able to give some insights into the halo formation process, the problem is usually studied using N-body numerical simulations, that can follow the gravitational collapse of dark matter particles in the general non-symmetric case. Here we review the basis of the N-body technique.

N-body simulations

Numerical N-body simulations are able to follow the non-linear evolution of density perturbations attempting to directly solve the equations governing the gravitational dynamic of a N-body system. Since collisionless dark matter interacts only gravitationally, its dynamics is relatively simple to model using the Newtonian theory which is valid in the regime of structure formation for non-relativistic particles.

The use of N-body simulations to study structure formation started already in the 60's, to investigate the formation of elliptical galaxies from the collapse of a (cold) top-hat perturbation (van Albada, 1961; Peebles, 1970). Later, this technique has been used to follow the collapse of a spherical overdensity trying to reproduce the properties of galaxy clusters (White, 1976), and there were the first attempts to simulate the formation of structures in a CDM scenario (Peebles, 1983; Frenk et al., 1985; Quinn, Salmon & Zurek, 1986), in some cases with enough resolution to resolve the halo inner structure (e.g. Dubinski & Carlberg 1991). Recent projects focused on studying the evolution of CDM perturbations inside big cosmological boxes using a large number of particles ($\sim 10^{10}$ for the MILLENIUM simulation, Springel et al. 2005), starting from realistic and well constrained initial conditions derived from the CMB power spectrum, and have been able to reproduce the observed Large Scale Structure of the Universe. Such simulations provided several information about the properties of dark matter halos, such as the mass function (Jenkins et al., 2001; Springel et al., 2008), the clustering at different redshifts (Kravtsov, Gnedin & Klypin, 2004; Reed et al.,

2009), their density profiles (Navarro, Frenk & White, 1997), angular momenta distribution (van den Bosch et al., 2002; Bett et al., 2007), concentrations (Neto et al., 2007) and internal sub-structures (Springel et al., 2008).

Mathematically, cold dark matter particles can be described as a collisionless, non-relativistic fluid of particles with mass m; the distribution function $f(\mathbf{r}, \mathbf{p}, t)$ evolves in the phase-space following the collisionless Boltzmann equation. The basic idea of N-body simulation is to replace $f(\mathbf{r}, \mathbf{p}, t)$ by a discrete set of particles (each representing a δ-function in the phase-space) evolving under gravity; for large number of particles N the set provides a good approximation for the evolution of the distribution function.

In an expanding background where $\mathbf{r}$ and $\mathbf{p} = ma^2\dot{\mathbf{x}}$ are respectively the comoving position and momentum, $f(\mathbf{r}, \mathbf{p}, t)$ evolves as

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{ma^2} \cdot \nabla f - m\nabla\Phi \cdot \frac{\partial f}{\partial \mathbf{p}} = 0 \tag{2.21}$$

in which the gravitational potential Φ is related to the density according to

$$\begin{aligned} \nabla^2\Phi(\mathbf{r}, t) &= 4\pi G a^2[\rho(\mathbf{r}, t) - \bar{\rho}(t)] \\ \rho(\mathbf{r}, t) &= \int f(\mathbf{r}, \mathbf{p}, t) d^3p \end{aligned} \tag{2.22}$$

for a mean background density $\bar{\rho}$.

Sampling the phase-space distribution function by a finite number N of particles, each of them obeys the equations of motion (in comoving coordinates)

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= -m\nabla\Phi \\ \frac{d\mathbf{x}}{dt} &= \frac{\mathbf{p}}{ma^2} \end{aligned} \tag{2.23}$$

which can be rewritten, after introducing the peculiar velocity $\mathbf{v} = a\dot{\mathbf{x}}$ as

$$\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{\dot{a}}{a} = -\frac{\nabla\Phi}{a}. \tag{2.24}$$

The time derivative of the scale factor a is given by eq. (2.15) in terms of the cosmological parameters:

$$\dot{a} = H_0\sqrt{\Omega_r / a^2 + \Omega_m / a + \Omega_\Lambda a^2} \tag{2.25}$$

while the gravitational potential Φ can be obtained summing the contribution of all N particles

$$\Phi(r) = -G\sum_{j=0}^{N} \frac{m_j}{(|\mathbf{r} - \mathbf{r}_j|^2 + \epsilon^2)^{1/2}} \tag{2.26}$$

where the *gravitational softening* ϵ is normally introduced to smooth the gravitational force between two particles in close encounters, and is typically chosen as $1/20 - 1/50$ of the mean inter-particle separation within the simulation, determining also the smallest physical length scale accurately resolved in the simulation (Dolag et al., 2008; Barnes, 2012). According to this formulation, called the "Particle-Particle" (or PP) method, the total gravitational potential is calculated summing the contribution of each particle. Although it is the most accurate way to compute the gravitational potential generated by a distribution of N particles, it is quite computationally expensive since it scales as $\propto N^2$.

In order to handle larger number of particles, more computationally efficient algorithms have been developed. One of them is the "Particle-Mesh" (PM) method; it is based on discretizing the space introducing a regular spatial grid, and assigning to each grid element a density computed from the particle positions; the density field is Fourier transformed, and the Poisson's equation is then solved in Fourier space. Since periodic boundary conditions are usually assumed, the Fast Fourier Transform (FFT) technique can be applied increasing the computational speed to $\propto N \log N$. The most relevant shortcoming of the PM method is poor force resolution on small scales due to the grid size; a possible solution is to introduce smaller subgrids, in the so-called "Adaptive Mesh Refinement" (AMR) scheme (Kravtsov, Klypin & Khokhlov, 1997). Another popular scheme, that combines the advantages of both the PP and the PM methods, is the "Particle-Particle-Particle-Mesh" (P^3M, Efstathiou et al. 1985); P^3M applies the efficient PM algorithm to compute forces on larger scales and the more accurate PP scheme on smaller scales. The "Tree-Particle-Mesh" (TreePM) method is similar to P^3M in the long-range force computation (handled by a PM algorithm), but it uses a tree algorithm to calculate the short-range forces, treating clumps of particles as a single mass point increasing the size of the clumps with the distance. The gravitational potential is the sum over the multipole expansion of the gravitational fields generated by the clumps of particles (Springel, 2005).

One important issue in fast algorithms such as AMR, P^3M or TreePM is the domain decomposition when large parallel computers with distributed memory are used. Since gravity acts over long scales and gravitational collapse creates highly inhomogeneous distributions of matter, a bad balance of the computing load between processors and excessive communication among them could significantly reduce the computational speed; therefore, efficient domain decomposition schemes are usually implemented in modern parallelized N-body codes (Springel, 2005; Gottloeber & Klypin, 2008; Wadsley, Stadel & Quinn, 2004).

Properties of dark matter halos

Most of the information about the properties of dark matter halos has been derived from results of N-body simulations; however, when dealing with dark matter halos in simulations, the first problem is the definition of the virial radius $r_{\rm vir}$ which identifies the material belonging to a certain halo. The common definition r_{200} and $M_{200} = M(< r_{200})$, by which $r_{\rm vir}$ is the radius where $\rho > 200\,\rho_{\rm crit}$, is somehow justified by analytical calculations (see Coles & Lucchin 1995; Longair 2008) and is the most used in simulation studies since spherical overdensities are easy to measure, although some authors have shown that ΛCDM halos form very differently than assumed in the spherical top-hat model, in particular $r_{\rm vir}$ is a larger fraction of the turnaround radius (Diemand, Kuhlen & Madau, 2007), and the conventional definition r_{200} may be too small and underestimate the extent and mass of dark matter halos significantly (Prada et al., 2006).

In the next sections we summarize the properties of ΛCDM halos as derived by N-body simulations and analytic calculations.

- **Press-Schechter mass function**

  The mass function $n(M)$ describes the abundance of halos at a given mass and can be easily calculated at different redshifts from N-body simulations. In addition, the Press & Schechter (1974) theory (PS hereafter) provides an analytic expression for the mass function from basic physical considerations. Although the PS approach has several limitations, the results of simulations match surprisingly well with the prediction of the PS theory at $z = 0$, though with some discrepancies in the number of small halos

(Tinker et al., 2008). At $z = 1 - 2$ the agreement is less impressive, and the PS mass function lies above the one derived from N-body simulations (Cuesta et al., 2008).

According to the PS theory, halos are associated with peaks in the Gaussian density field in the Early Universe, where the phases of the waves were randomly distributed. The probability distribution $p(\delta_M, M)$ of the amplitudes of the perturbations on a scale R, corresponding to a mass $M \propto \rho R^3$ is described by a Gaussian function:

$$p(\delta_M, M) = \frac{1}{\sqrt{2\pi\sigma^2(M)}} \exp\left[-\frac{\delta_M^2}{2\sigma^2(M)}\right] \tag{2.27}$$

where $\sigma^2(M)$ is the mean-squared fluctuation of mass M

$$\langle \delta_M^2 \rangle = \sigma^2(M). \tag{2.28}$$

The PS analysis assumes an Einstein-de Sitter cosmological background, ($k = 0, \Omega_m = 1, \Omega_\Lambda = 0$) and a power-law spectrum of initial perturbations $P(k) \propto k^n$. The basic idea is that, when a perturbation reaches an amplitude greater than some critical value ($\delta_M > \delta_c$), it collapses into a bound object with mass M. The fraction $F(M)$ of perturbations that become bound is easily computed from (2.27)

$$F(M) = \frac{1}{\sqrt{2\pi\sigma^2(M)}} \int_{\delta_c}^{\infty} \exp\left[-\frac{\delta_M^2}{2\sigma^2(M)}\right] d\delta_M = \frac{1}{2}[1 - \Phi(t_c)] \tag{2.29}$$

where $t_c = \delta_c/\sqrt{2}\sigma$ and $\Phi(x)$ is the integral

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{2.30}$$

for which is valid

$$\frac{d\Phi}{dx} = \frac{2}{\sqrt{\pi}} e^{-x^2}. \tag{2.31}$$

Since dF/dM is the fraction of independent regions that collapse with mass between $[M, M + dM]$, the integral (2.29) is related to the mass function $n(M)$ as

$$n(M) = \frac{2}{M}\bar{\rho}\left|\frac{dF(M)}{dM}\right|. \tag{2.32}$$

The additional factor two in $n(M)$, given by Press & Schechter (1974), has been the source of a long debate in the so-called "cloud-in-cloud" problem. The problem arises in the non-linear stage, when mass is accreted from the vicinity of the perturbation; according to the results of N-body simulations most of the mass is already condensed into discrete structures. Press and Schechter were aware of the problem and argued that the mass spectrum should be multiplied by a factor two to take into account the accretion of mass during the non-linear evolution. However, although this argument does not justify the reason for multiplying the mass function by a value of exactly two, its validity has been confirmed by comparing the PS mass function with the results of N-body simulations (see Peacock & Heavens 1990; Bond et al. 1991 for a full discussion of this topic). The problem is solved analytically in a more rigorous way by the excursion set theory (Bond et al., 1991) which is sometimes referred to as the "extended Press-Schechter theory" (EPS).

The mean-square $\sigma(M)$ of the mass perturbation on mass scale M can be related to the exponent of the primordial power spectrum $P(k) \propto k^n$ as (see Longair 2008):

$$\sigma^2(M) = AM^{-(3+n)/3} \tag{2.33}$$

for some $A(t)$. We can now write the term t_c in eq. 2.29 as a function of the mass:

$$t_c = \frac{\delta_c}{\sqrt{2}\sigma(M)} = \left(\frac{M}{M_*}\right)^{(3+n)/6} \tag{2.34}$$

after introducing a reference mass $M_* = (2A/\delta_c^2)^{3/(3+n)}$ which changes in time. Since in an Einstein-de Sitter Universe the perturbations grow as $\delta \propto t^{2/3}$ it follows that $\sigma^2(M) \propto t^{4/3}$ or $A \propto t^{4/3}$, then

$$M_* \propto A^{3/(3+n)} \propto t^{4/(3+n)} \tag{2.35}$$

Inserting into eq. (2.32), we obtain the following expression for the mass function

$$n(M) = \frac{1}{\sqrt{\pi}}\left(1+\frac{n}{3}\right)\frac{\bar{\rho}}{M^2}\left(\frac{M}{M_*}\right)^{(3+n)/6}\exp\left[-\left(\frac{M}{M_*}\right)^{(3+n)/3}\right] \tag{2.36}$$

(note that $n(M)$ depends on time through the factor M_*).

The PS formalism provides a remarkably effective description of the form and evolution of the mass function of dark matter halos with cosmic epoch and has been the starting point for more detailed analysis based on simulations (see e.g. Yano, Nagashima & Gouda 1996; Sheth & Tormen 1999); however, when high precision measurements of the halo mass function as given by modern N-body simulations are needed, the PS analysis has been proven to be less accurate compared the numerical methods (Tinker et al., 2008).

- **Halo density profiles**

  Early models of collapse predicted the formation of virialized halos with almost-isothermal density profile $\rho \propto r^{-2}$ (e.g. Gunn & Gott 1972). Later, N-body simulations have found halo profiles steeper than $\rho \propto r^{-2}$ in the outer parts, and shallower in the inner regions, following the so-called Navarro-Frenk-White (NFW) profile (Navarro, Frenk & White, 1997) over a large mass range and for a wide variety of cosmological scenarios

  $$\rho_{\rm NFW}(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}. \tag{2.37}$$

  r_s and ρ_s are respectively a characteristic radius and density, which vary from halo to halo. The density goes as $\rho \propto r^{-1}$ near the center ($r \ll r_s$) and $\rho \propto r^{-3}$ at large radii; r_s is related to the radius at which the circular velocity has a peak by $r_{(\rm V_{max})} = 2.163\, r_s$, and defines the halo concentration parameter as $c \equiv r_{\rm vir}/r_s$. The concentration parameter is a function of halo mass and has higher values for smaller halos (Navarro, Frenk & White, 1997); it is also related to the halo formation time, since older halos have higher c at low-redshift, although with large scatter depending on the mass (Macciò, Dutton & van den Bosch, 2008).

The NFW function was suggested as a fit to the CDM density profiles in the range $[0.01, 1] \times r_{200}$. Higher resolution simulations have shown that in the inner part of the halo the density profile becomes significantly steeper, and is better described by the Einasto profile (Einasto & Haud, 1989)

$$\rho(r) = \rho_s \exp\left\{-\frac{2}{\alpha}\left[\left(\frac{r}{r_s}\right)^{\alpha} - 1\right]\right\} \tag{2.38}$$

where α is a free parameter. The Einasto profile is very similar to NFW, and fits simulated CDM halos equally well, but the two functions differ significantly at radius below ~ 500 pc (Navarro et al., 2004), at a resolution however currently reached only by very few simulations.

An important open question related to the shape of dark matter density profiles is the so-called "core-cusp problem", namely the discrepancy between the observed dark matter density profiles of low-massive galaxies and those predicted by N-body simulations (de Blok, 2010). On one hand, simulations already in the 90's predicted steep power-law density distribution at small radii, the so-called "cusps", described by the function $\rho \propto r^{\alpha}$ with $\alpha = -1$ (Dubinski & Carlberg, 1991; Navarro, Eke & Frenk, 1996; Navarro, Frenk & White, 1997). On the other hand, the rotation curves of most of the observed dwarf galaxies reveal nearly constant dark matter density in the inner region (a "core"), well-fitted by $\alpha = 0$ (Moore, 1994; Oh et al., 2015; Navarro, Frenk & White, 1996, 1997; Donato et al., 2009; Oh et al., 2008, 2011). Several possible solutions have been proposed, in particular pointing out the effect of baryonic processes such as SNe and AGN feedback and their dynamical interactions with the dark matter, since outflows due to the feedback can move large amounts of low-angular-momentum gas from the central parts "flattening out" the core of a galaxy's dark matter profile (Navarro, Eke & Frenk, 1996; Pontzen & Governato, 2012; Governato et al., 2010). Other works tried to solve the problem with alternative cosmological paradigms, for instance warm or self-interacting dark matter, claiming that they are able to remove "cusps" in simulated galaxies' density profiles (Lovell et al., 2012; Elbert et al., 2015; McGaugh et al., 2007).

- **Halos shapes and spins**

Dark matter halos posses angular momentum, which may have an impact on different aspects of galaxy formation and evolution. For instance, it has been shown that angular momentum affects the gravitational collapse with detectable consequence on the mass function (Del Popolo, 2006; Macciò et al., 2007; Jimenez et al., 1998), and has an influence in regulating the formation and evolution of central accretion disks in AGNs (Peebles, 1971; Fall & Efstathiou, 1980; Mo, Mao & White, 1998).

Early theoretical investigations predicted that the spin parameter λ of dark matter halos, a dimensionless quantity defined as a function of the total angular momentum $J \equiv |\mathbf{J}|$, total energy $E \equiv |E_{\rm kin} + E_{\rm grav}|$ and dark matter halo mass M as

$$\lambda = \frac{JE^{1/2}}{GM^{5/2}} \tag{2.39}$$

should have an approximately log-normal distribution (Peebles, 1969)

$$P(\lambda) = \frac{1}{\lambda\sqrt{2\pi\sigma_\lambda^2}} \exp\left[-\frac{\ln^2(\lambda/\bar{\lambda})}{2\sigma_\lambda^2}\right]. \tag{2.40}$$

This prediction was later found to be valid also when higher order effects were taken into account (Catelan & Theuns, 1996a,b), and according to the results of numerical simulations the free parameters have values $\bar{\lambda} = 0.04$ and $\sigma_\lambda = 0.56$ (van den Bosch et al., 2002). In general, the values of the spin parameters found both in simulations and observations are in the range $0.02 - 0.11$ (Steinmetz & Bartelmann, 1995; Navarro, Frenk & White, 1997; Burkert & D'Onghia, 2004), although the measurements are difficult and somehow inaccurate in both cases. Bett et al. (2007), after analysing halos in the Millennium simulation, gave a different fitting function for the spin parameter distribution, claiming that it may also depend on the numerical definition of the halos (e.g. number of particles, definition of r_{vir}):

$$P(\lambda) \propto \left(\frac{\lambda}{\lambda_0}\right)^3 \exp\left[-\alpha\left(\frac{\lambda}{\lambda_0}\right)^{3/\alpha}\right]. \tag{2.41}$$

The function is similar to the log-normal distribution, except for both very low and high values of the spin.

Several theoretical and numerical studies have aimed to understand the origin of the log-normal functional form of $P(\lambda)$, which may arise from small asymmetries in the velocity distribution, acquired by an asymmetric halo through tidal torques (Chiueh, Lee & Lin, 2002; Knebe & Power, 2008; Gao & White, 2007; Davis & Natarajan, 2009; Antonuccio-Delogu et al., 2010).

- **Dark matter substructures**

  In the ΛCDM scenario structures grow via hierarchical merging and accretion of smaller progenitors; when halos merge they virialize, redistributing energy among the collisionless mass components. Early low-resolution numerical simulations showed that the results of a "bottom-up" formation scenario were smooth and featureless triaxial halos, while later higher-resolution simulations predicted a vast amount of surviving substructure on all resolved mass scales, both gravitationally bound (subhalos) and unbound (streams) (Klypin et al., 1999; Moore et al., 1999). This was confirmed by subsequent simulations such as AQUARIUS (Springel et al., 2008), VIA LACTEA II (Diemand, Kuhlen & Madau, 2007) and GHALO (Stadel et al., 2009). They have also shown that the merging of progenitors is not always a complete process, and the cores of accreted halos often survive as gravitationally bound sub-halos orbiting within a larger host system (Springel et al., 2008).

  Observationally, we should expect to see many satellites around the Milky Way or M31 in the Local Group, corresponding to the sub-halos found in simulations. However, only a small number of dwarfs has been found and there is a considerable discrepancy between the number of satellites in the Milky Way and the number of halos predicted by CDM simulations, although the problem has been recently somehow mitigated by the discovery of the ultra-faint satellites (Willman et al., 2005; Belokurov et al., 2006; Irwin et al., 2007). The mismatch between the number of subhalos in simulations and satellites in the Local Group is often referred to as the "missing satellite"problem (Madau, Diemand & Kuhlen, 2008; Kravtsov, 2010). Careful investigations about the

possible mechanisms that are able to suppress star formation in small sub-halos have been carefully investigated; some authors claimed that the stellar content of dwarf galaxies can be stripped by strong tidal forces (Kravtsov, Gnedin & Klypin, 2004). Another possible solution involves some sort of feedback that can make star formation in small halos very inefficient, such as SNe (Sawala et al., 2010), photoionization (Okamoto, Gao & Theuns, 2008; Brooks et al., 2013) or reioniziation (Bullock, Kravtsov & Weinberg, 2000); in particular reionization is able to prevent the acquisition of gas by dark matter halos of small mass at high redshift (Ricotti & Gnedin, 2005; Moore et al., 2006). Other possibilities are based on non-standard cosmological paradigms, such as alternative dark matter models (Sommer-Larsen & Dolgov, 2001; Peebles, 2000; Kaplinghat, Knox & Turner, 2000), extended versions of the General Relativity (Starobinsky, 1980; Ferraro, 2012; Milgrom, 1983b,a), or some exotic form of inflation in the Early Universe that was able to modify the power spectrum at small scales (e.g. Kamionkowski & Liddle 2000; Zentner & Bullock 2003).

Another tension exists between the spatial distribution of satellites in simulations and observations. According to the results of simulations in a ΛCDM cosmological context, the sub-halo distribution around Milky Way-sized halos should be roughly isotropic; however, recent studies have found a strong correlation in the distribution of satellites in the Milky Way and M31, which lie aligned within a relatively thin disk (Ibata et al., 2013). Some studies suggested that the ambient shear field plays a fundamental role in the alignment of satellite systems, and a strongly-correlated spatial distribution does not contradict the ΛCDM paradigm (Libeskind et al., 2015).

2.3 Baryonic physics

The interplay between dark matter and baryons during gravitational collapse is crucial for understanding galaxy formation and evolution. In the early stages of the collapse baryons trace the dark matter distribution and play a minor or indirect role in driving the evolution of the linear perturbations (Arons & Silk, 1968); the baryons later concentrate towards the potential well generated by dark matter halos, and cooling removes the pressure support settling the gas in a disk, eventually leading to the beginning of the star formation processes. During these later stages the perturbations are in the non-linear regime, and baryonic physics becomes a relevant aspect in the formation of galaxies.

To follow the joint evolution of baryonic and dark matter density perturbations the most accurate way is by using hydrodynamical codes[2], that are able to describe the dynamics of gas (see e.g. reviews by Rosswog 2009; Bertschinger 1998). However, most of the important physics relevant in galaxy formation, such as star formation, feedback from SuperNovae (SNe) and Active Galactic Nuclei (AGNs) accretion and feedback, is well beyond the spatial resolutions usually reached by hydrodynamical codes, and must be described at the "subgrid" level using phenomenological recipes (e.g. Booth & Schaye 2009; Okamoto 2008; Oppenheimer & Davé 2008; Scannapieco et al. 2005; Ciotti, Ostriker & Proga 2009; Stinson

[2]An alternative approach is the so-called "semi-analytic modelling" (see review by Baugh 2006), which is based on implementing into the dark matter halos' merger trees, extracted from large N-body simulations, some physical recipes to describe baryonic processes inside dark matter halos. These recipes are usually derived from observed relations or from the results of simulations. One of the advantages of the method is that is computationally faster compared to hydrodynamical simulations, and thus suited for exploring a large range of physical models and for determining their free parameters (Henriques et al., 2009), however with a larger degree of approximation (Benson, 2010).

et al. 2013a), which somehow requires model-adjustment comparing the results of simulations with observations (Tormen, 1996; Dolag et al., 2008).

In the next sections we will describe the main baryonic processes involved in galaxy formation.

2.3.1 Accretion of gas

After the formation of a dark matter halo, surrounding gas falls into the central potential well. Since the accreted gas provides the fuel source for star formation, which is a crucial factor in driving a galaxy's evolution, the way galaxies accrete gas has received a great deal of attention in galaxy formation studies.

According to the standard picture (Rees & Ostriker, 1977; Silk, 1977; White & Rees, 1978; Fall & Efstathiou, 1980) the gas falling into the dark matter potential well gains kinetic energy, and is shock-heated to the halo virial temperature $T_{\rm vir}$ settling into quasi-hydrostatic equilibrium. $T_{\rm vir}$ has the following expression, given the halo virial mass $M_{\rm vir}$, virial radius $R_{\rm vir}$, Boltzmann constant k_B, proton mass m_p and mean molecular weight μ (Benson, 2010):

$$T_{\rm vir} = \frac{2}{3}\frac{GM_{\rm vir}}{R_{\rm vir}}\frac{\mu m_p}{k_B}. \tag{2.42}$$

Gas is able to cool radiatively and so it will eventually lose energy and pressure support, leading to cooling flows onto the central region of the halo.

However, different works based on hydrodynamical simulations (e.g. Kereš et al. 2005) suggested that this picture is somehow oversimplified; in reality only $\sim 1/2$ of the gas is shock-heated to $T_{\rm vir}$ ("hot mode"), while the other half radiates its acquired gravitational energy at a much lower temperatures $T \lesssim 10^5$ K ("cold mode"). The hot mode of accretion is likely to dominate the growth of high mass systems, while the cold mode is prevalent for low mass galaxies. Numerical simulations suggest a value of $M_{\rm halo}$ for the transition from cold to hot accretion of a few times $10^{11} M_\odot$ (Birnboim & Dekel, 2003), although the exact value depends on the resolution and on the numerical schemes implemented in the simulations. Furthermore, the transition mass is not sharp, and massive halos may still experience some cold accretion, with the relative contribution decreasing with increasing halo mass. Observationally this picture has not been yet fully confirmed; the Milky Way contains a significant mass of cold, neutral gas in the form of high velocity clouds (Putman et al., 2003), which may indicate that it is in the transition mass range between purely cold and purely hot accretion (Crain et al., 2009), while in extragalactic observations cold flows are rarely observed (Faucher-Giguère et al., 2010). Some indirect evidence of the existence of cold accretion is however given by studies of star formation in dwarf galaxies (Tolstoy, Hill & Tosi, 2009).

Accretion in these two regimes may result in very different spatial and spectral distributions of the cooling radiation, since cold accretion allows much of the radiation to emerge in the Lyα line instead of the X-ray continuum (Fardal et al., 2001). The different accretion may also have consequences on the properties of the galaxies, with cold accretion resulting in earlier star formation compared to the hot mode (Brooks et al., 2009).

2.3.2 Gas cooling and heating

The gas is able to radiate its internal energy through radiative cooling, while it can also gain energy from the incident radiation field by radiative heating; the rate at which gas cools is a

crucial ingredient in galaxy formation studies, and has a strong influence on star formation and on the phase structure of the ISM (Sutherland & Dopita, 1993).

In the simplest picture, gas which cools sufficiently below $T_{\rm vir}$ loses pressure support and flows smoothly towards the minimum of the gravitational potential. However, several studies have shown that cooling gas fragments into cold ($T \approx 10^4$ K), dense clouds in pressure equilibrium with a hot ($T \sim T_{\rm vir}$) diffuse component with much longer cooling time scales (Maller & Bullock, 2004), which may reduce the rate of gas supply, particularly in massive halos (Kaufmann et al., 2009).

Since the dominant cooling processes result from two-body interactions, the cooling rate of (optically-thin) gas per unit volume C can be expressed as

$$C = n_e^2 \, \Lambda(T, Z) \tag{2.43}$$

where n_e^2 is the electron number density and $\Lambda(T, Z)$ is the *cooling function*. Several calculations of the cooling function are given in the literature (e.g. Sutherland & Dopita 1993; Wiersma, Schaye & Smith 2009; Lykins et al. 2013), normally computed assuming collisional ionization equilibrium, which however is not always valid in the ISM (Wiersma, Schaye & Smith, 2009); for instance, in dense and extremely metal-enriched regions such as around the central AGNs (that can reach $\rho \sim 10^{14}$ cm^{-3}, $Z \sim 30\, Z_\odot$) collisional ionization equilibrium assumption breaks down (Osterbrock & Ferland, 2006; Wang et al., 2014). The cooling function can be also modified by the incident radiation field (Gnedin & Hollon, 2012).

Cooling occurs mainly through four mechanisms, depending on the ionization state of the gas (determined by the temperature), on the chemical composition and on the density:

- **Inverse Compton scattering**: this mechanism involves the scattering of low-energy photons to higher energies by ultrarelativistic electrons, so that the electrons lose energy in favour of the photons. CMB photons are efficiently scattered by electrons in the hot halos at high redshift ($z > 6$), while at lower redshift this cooling channel is not efficient anymore (Longair, 2008; Benson, 2010).

- $\mathbf{T_{vir} < 10^4\ K}$: for low temperatures the gas is usually neutral, and cooling is strongly suppressed. However, in presence of molecular hydrogen or heavy elements, cooling is possible through the collisional excitation/de-excitation of the rotational and vibrational energy levels in molecules, or transitions between fine (or hyperfine) structure lines in heavy elements.

- $\mathbf{10^4\ K < T_{vir} < 10^7\ K}$: in this temperature range cooling occurs mainly through recombination of electrons with ions, or by atomic transitions between energy levels of collisionally-excited atoms. Since different atomic species have different excitation energies, the cooling rates in this range depends strongly on the chemical composition of the gas, which cools faster at higher metal abundances.

- $\mathbf{T_{vir} > 10^7\ K}$: at this temperatures, typically reached in massive halos, the gas is almost fully collisionally ionized, and it cools mainly through Bremstrahlung emission from free-free transitions in electron-ion collisions.

The following processes are instead able to heat the gas:

- **Photoionization heating**: photoionization can modify the cooling function as well as heat the gas by converting the energy of a photon with frequency ν into kinetic energy $\Delta K = h(\nu - \nu_i)$ of an atom with threshold frequency ν_i. Photoionizing photons

are then able to heat the InterGalactic Medium (IGM) increasing the thermal pressure and inhibiting or reducing the rate of accretion onto dark matter halos. This effect is expected to be particularly relevant in the Early Universe during the Epoch of Reionization at $z \sim 6.5$ (Fan, Carilli & Keating, 2006), when the first generation of galaxies fully ionized hydrogen (Zaroubi, 2013). Several works (e.g. Doroshkevich, Zel'dovich & Novikov 1967; Benson et al. 2002; Couchman & Rees 1986; Efstathiou 1992) have shown that high energy background photons are able to suppress (or inhibit) galaxy formation in small halos, reducing the abundance of low-mass galaxies in the local Universe.

- **Heating from SNe and AGNs feedback**: the energy released by SNe explosions ($\sim 10^{51}$ erg) heats up the surrounding material, leading to disruption of cold gas clouds and reducing star formation (Scannapieco et al., 2006). The energy released by the explosions is converted both into thermal and kinetic energy of the gas; however the fraction of the released energy that is thermalized depends on the time scale of thermalization relative to the cooling time, and is still debated (Cole et al., 1994; Mo, van den Bosch & White, 2010). SNe feeedback in particular affects smaller systems because of their shallower potential (Dekel & Silk, 1986; White & Frenk, 1991), while for massive halos feedback from AGNs dominates (Di Matteo, Springel & Hernquist, 2005), since the amount of energy released is up to a factor $\sim 20-50$ higher than for SNe (Benson & Madau, 2003). Galaxy formation models which include feedback from AGNs are given for instance by Croton et al. (2006); Bower et al. (2006); Somerville et al. (2008); however, the details of the mechanism by which the energy output from the AGNs is coupled to the surrounding gas is still under discussion (Ruszkowski, Brüggen & Begelman, 2004; Ruszkowski et al., 2008; King, 2009; Brüggen & Scannapieco, 2009).

- **Photoelectric effect on grain surface**: this mechanism, by which electrons from interstellar dust grains are photo-ejected, is a heating mechanism relevant for a neutral ISM. The energy of an ejected electron is

$$E = h\nu - W \tag{2.44}$$

  where $h\nu$ is the average photon energy for the diffuse interstellar radiation field and W is the grain work function (Pollack et al., 1994).

2.3.3 Star formation

The problem of star formation is an important aspect of galaxy formation studies, since stars are one the main factor in driving the structure and evolution of galaxies and their chemical evolution. Observationally, the sites of stars formation are large condensations of dust and molecular gas (giant molecular clouds) which collapse because of gravitational instability, leading to episodes of star formation. In these regions stars form in the mass range $0.08\,M_\odot \lesssim M \lesssim 100\,M_\odot$, as for smaller masses they cannot ignite nuclear fusions, while for larger masses they are unstable to their own radiation pressure.

The theory of star formation describes how stars form inside giant molecular clouds (see e.g. the review by McKee & Ostriker 2007); however the theory is incomplete in several fundamental aspects, in particular many details of the physical processes acting in the clouds during the collapse, such as turbulence, magnetic fields or fragmentation, are not fully understood. For that reason, in galaxy formation studies observationally-based empirical rules

are usually adopted to describe star formation. The most important one is the Schmidt-Kennicutt law (Schmidt, 1959; Kennicutt, 1998b), which states that the star formation rate per unit surface area Σ_{SFR} depends on the surface density of gas Σ_{gas} as

$$\Sigma_{SFR} \propto (\Sigma_{gas})^n \tag{2.45}$$

with $n = (1.4 \pm 0.15)$ (Kennicutt, 1998b). Since this law is empirical its validity is limited to the regimes where it was originally measured; for example, it does not fit well for dwarf galaxies, that are much less efficient in forming stars than larger galaxies (Benson & Bower, 2010).

An open question in the star formation theory is the observational evidence that the collapse of molecular clouds is a very inefficient process, transforming only a few percent of the available gas into stars and leading to the observed low global star formation rate (SFR). For instance, the star formation rate of the Milky Way is about $2 M_\odot$ yr^{-1} (Chomiuk & Povich, 2011), while with the estimated molecular gas available the star formation rate should be $1-2$ orders of magnitudes higher (Zuckerman & Evans, 1974); the same issue has been found in spiral galaxies in the nearby Universe (Kennicutt, 1998a). The star formation efficiency α is defined as

$$\dot{\rho}_* = \alpha \frac{\rho_{\rm gas}}{t_{\rm dyn}} \tag{2.46}$$

where $\dot{\rho}_*$ and $\rho_{\rm gas}$ are respectively the local star formation rate and gas density, and $t_{\rm dyn}$ the dynamical time, usually estimated as $t_{\rm dyn} = (4\pi G\rho)^{-1/2}$. The inefficiency of the star formation process is translated into a small value of the star formation efficiency, $\alpha \approx 0.02$ according to Silk (1997); Elmegreen (1997). Some physical mechanisms must prevent molecular clouds from collapsing; possible solutions involve magnetic fields (Mestel & Spitzer, 1956; Mouschovias, 1976; Shu, 1983), ionization feedback from massive stars (Vazquez-Semadeni, 2015) or turbulence motion in the clouds (Mac Low & Klessen, 2004; Elmegreen & Scalo, 2004; McKee & Ostriker, 2007).

Another open question in the star formation process, particularly relevant for galaxies, is the mass spectrum with which stars form, i.e. the Initial Mass Function (IMF) which determines the relative number of stars born with masses in the range $[m, m + dm]$. Based on observations in the solar neighborhood, Salpeter (1955) first estimated the IMF $F(m)$ as

$$F(m) \propto m^{-\alpha} \tag{2.47}$$

with $\alpha = 2.35$. Other IMFs, commonly used in the literature, include Miller & Scalo (1979), Scalo (1986), Kroupa (2002), and Chabrier (2003); all these are similar to Salpeter at $m \gtrsim 1\, M_\odot$, but are in general flatter at lower masses. These more recent IMFs are based on numerous observational studies that have been carried out in systems with a wide range of properties, in order to establish whether the IMF is universal, or whether it depends on parameters such as the metallicity (e.g. Scalo 1998; von Hippel et al. 1996). In general, there is no clear evidence in the nearby Universe of any dependence of the IMF on the local conditions of the ISM; however, some indirect evidence suggested that the IMF may vary for starburst and high-redshift galaxies (Larson, 1998), although this has not been yet fully confirmed.

2.3.4 Chemical evolution

Stellar evolution changes the abundance of elements in the Universe in time by stellar nucleosynthesis and during SNe explosions, releasing the synthesized elements back into the ISM

by stellar or SNe-driven winds; subsequent generations of stars are then born with different chemical compositions. The first generation of stars (also called "Population III") formed from primordial gas, which consists primarily of hydrogen and helium created during the epoch of primordial nucleosynthesis, and then are expected to have very different properties compared to stars created at later time (see review by Yoshida, Hosokawa & Omukai 2012).

Metals are released into the ISM mainly by the following processes:

- **SNe Type II explosions**: this type of SNe explosions occur when massive stars ($M_* > 8\,M_\odot$) experience core-collapse, and they are primarily responsible for the production of α-elements [O, Ne, Mg, Si, S, Ar, Ca, Ti] (Arnett, 1978; Woosley & Weaver, 1995).

- **SNe Type Ia explosions**: they are thought to originate from the collapse of a white dwarf after accretion of material from a binary companion. During the collapse it is believed that almost the entire star will be changed into iron-peak elements, so it is generally assumed SNe Type Ia explosions synthesized most of the iron present in the Universe (Nomoto, Thielemann & Yokoi, 1984; Thielemann, Nomoto & Yokoi, 1986).

- **AGB stars stellar winds**: stars in the Asymptotic Gian Branch (AGB) phase, when H and He burn intermittently in two shells surrounding an inert CO core, are the major site of production of several chemical elements and fundamental contributors to the stellar nucleosynthesis yields (Tosi, 2007). Such stars are the main producers of *s*-elements and they synthesize large amounts of He, N, C, O, F (Goriely & Siess, 2005). During the thermal pulses, material from the core may be mixed into the outer layers, changing the surface composition in a process referred to as "dredge-up". AGB stars are long-period variables, and suffer strong mass loss in the form of stellar winds during thermal pulses, that are even able to detach their external shells and release into the ISM up to $50 - 70\%$ of their mass (Wood, Olivier & Kawaler, 2004).

In galaxy formation studies the relative abundances of metals can provide information about the time-scale of star formation because of the different lifetimes of the progenitor stars. For instance, the [α/Fe] ratio can be used to constrain the star formation history of galaxies, since α-elements are mainly produced on short time-scales by SNe II explosions, while Fe is released at later time by SNe Ia that explode with a delay compared to SNe II (Thomas, 1999; Prantzos, 2008).

2.3.5 Feedback

Feedback processes have great impact on galaxy evolution; as first pointed out by White & Rees (1978); Larson (1974); Dekel & Silk (1986), the luminosity function of galaxies is very different from the distribution of dark matter halo masses (which is well described by the Schechter function), dropping both towards lower and higher masses (Martizzi et al., 2014; Tinker et al., 2008; Benson et al., 2003); some baryonic feedback mechanism, that prevents gas from cooling or reheats and expels the existing cold gas, is required to make galaxy formation 'inefficient' both for low and high mass galaxies. The most plausible feedback processes acting in galaxies are thought to be SNe explosions and energy injection from AGNs accretion (Efstathiou, 2000).

Observational evidence for feedback from SNe explosions (see review by Veilleux & Rupke 2005) ranges from the galactic fountains observed in the Milky Way halo to the outflows in local edge-on starburst galaxies, also observed at higher redshift (Heckman & Lehnert, 2000;

Pettini et al., 2001; Steidel et al., 2010; Martin et al., 2012). It has long been understood that the cumulative effects of multiple supernovae (each of them releasing $\sim 10^{51}$ erg in the form of both thermal and kinetic energy) can produce supernova-driven winds (Ott, Walter & Brinks, 2005) that can eject gas from the galaxy, powering a galactic outflow or wind (Mathews & Baker, 1971; Larson, 1974). The ejected material, which consists of a multiphase mixture of gas at different temperatures, dust and magnetized relativistic plasma, is transported by the winds at a rate that depends on a number of unknown physical processes (Mac Low & Ferrara, 1999). SNe feedback has a strong influence on galaxies with masses up to the Milky-Way, since for more massive galaxies the deeper potential well makes the SNe-driven winds inefficient in removing the material (Powell, Slyz & Devriendt, 2011). For galaxies with masses similar to the Milky-Way, SNe feedback affects the amount of baryons that collect into the central part of the halo (Piontek & Steinmetz, 2011); in smaller galaxies, SNe explosions can efficiently eject the ISM out of the galaxy, suppressing star formation (which may explain the behaviour of the luminosity function at low masses). Another possible candidate for powering strong galaxy-scale winds is the energy input from massive stars, which can reach a speed ~ 2000 km/s around OB stars (Murray, Quataert & Thompson, 2005; Krumholz & Thompson, 2013).

For massive galaxies a more efficient feedback mechanism is provided by accretion onto the supermassive black hole in the central Active Galactic Nucleus (AGN) region. The AGN responds to the accretion of gas by feeding back energy to the surroundings in both radiative and kinetic forms; the interaction between AGN feedback and the surrounding gas may heat the IGM, suppressing cooling and star formation, and sweeping out matter via jets and lobes, as is normally observed in AGN-hosting galaxies at different wavelengths. However, it is still unclear how the energy and momentum resulting from the accretion into the AGN is coupled and interacts with the surrounding medium, either involving some mechanical or radiative processes (Ruszkowski, Brüggen & Begelman, 2004; Ruszkowski et al., 2008; King, 2009; Brüggen & Scannapieco, 2009), although it is believed that AGN feedback plays a role in shaping the galaxy luminosity function at the high masses. The physical connection between the central black hole and the hosting galaxy is reflected into a number of empirical relations; the most important one is a correlation between the black hole mass and the bulge stellar velocity dispersion $M_{\rm BH} \propto \sigma^4$, known as the $M_{\rm BH} - \sigma$ relation (Gebhardt et al., 2000; Ferrarese & Merritt, 2000); a similar relation holds between the black hole mass and the bulge stellar mass (Marconi & Hunt, 2003; Häring & Rix, 2004).

2.3.6 Environmental effects

In the ΛCDM scenario, structures form hierarchically by merging dark matter halos into larger systems; when this occurs, the galaxies in the center of the halos interact with the environment, both gravitationally or hydrodynamically with the surrounding IntraCluster Medium (ICM), such as in galaxy clusters. In the observed Universe interactions are relatively frequent and galaxies are rarely found in isolation; most of them are surrounded by satellites or embedded in larger groups or clusters. The galaxy number density strongly depends on the large structure environment, spanning three orders of magnitude from the low-density voids to the huge galaxy clusters (Geller & Huchra, 1989).

According to the results of simulations, subhalos may survive when orbiting within the host halo (see sec. 2.2.2), and each orbit can move the satellite galaxies into a region where collisions or close encounters may happen. This can affect the properties of the satellite, in particular distorting the morphology creating new structures such as warps or bars, or even tidal tails that extend well beyond the main body of the galaxy, and inducing new

bursts of star formation (Hernandez-Jimenez et al., 2013; Patton et al., 2013; Ellison et al., 2013). When halos merge, simulations have shown that two disk galaxies with similar mass can produce ellipticals as merger remnants (Toomre & Toomre, 1972; Barnes & Hernquist, 1996) and that accretion of small satellites onto spirals can transform the host spiral into an S0 type (Walker, Mihos & Hernquist, 1996).

Although the role played by the environment in driving galaxy evolution is an open issue (the so-called "nature versus nurture" debate), the properties of galaxies in clusters are in general statistically different compared to the isolated ones. For example, galaxies in isolation are bluer (Butcher & Oemler, 1978; Dressler, 1980) and have on average different star formation rates and gas content at a given redshift (Brodwin et al., 2013). Several works established that morphological type and local density are not independent quantities, and the fraction of spirals decreases from the field to the center of rich galaxy clusters (Dressler, 1980; Whitmore, Gilmore & Jones, 1993); high-redshift observations have also shown that the Universe had in the past a higher fraction of spirals in galaxy clusters (Butcher & Oemler, 1978, 1984; Dressler et al., 1997; Couch et al., 1998; Treu et al., 2003).

Here we review the main mechanisms through which the environment affects galaxy properties (see also Boselli & Gavazzi 2006):

- **Interaction with the intracluster medium**

  Ram pressure stripping

  When a satellite galaxy is orbiting through the hot atmosphere of a host halo it experiences a large ram pressure (Gunn & Gott, 1972) from the ICM and the hydrodynamical forces can strongly affect the gas content of the satellite, removing the hot atmospheres. If the ram pressure is strong and greater than the binding force, the dynamical pressure from the ICM may even sweep the cold gas out (Abadi, Moore & Bower, 1999), removing a large fraction of gas and truncating star formation on a short timescale ($\sim 10^7$ yr); some authors claimed that the ram-pressure can also generate significant compression ahead of the galaxy, temporarily triggering star formation (Bekki & Couch, 2003). Ram-pressure stripping is considered to be effective in the central region of clusters, where the density of the ICM is higher.

  Starvation

  This effect is similar to ram pressure stripping, although the loss of gas is limited to the diffuse hot gas reservoir that is confined in the satellite galaxy halo, since is less bound than the cold gas in the central part (Bekki, Couch & Shioya, 2002). Without gas supply, a satellite galaxy exhausts its cold gas in ~ 1 Gyr, and the star formation rate declines on a time scale $t \gtrsim 1$ Gyr, longer than for ram-pressure stripping. This effect is expected to happen for satellites in the outer regions of the host halo, resulting in a slow decline of star formation rate and progressive reddening (Larson, Tinsley & Caldwell, 1980).

  Viscous stripping

  When a satellite galaxy travels into the ICM, the outer layers of its ISM may experience a viscosity momentum transfer, that could be sufficient for dragging out part of its gas (Nulsen, 1982). Viscous stripping needs high-density ICM, and thus is likely to happen only in the central region of galaxy clusters.

- **Gravitational interaction**

  Tidal Effects

Tidal effects arise when the external gravitational force acting on a satellite galaxy is not constant across its diameter. The net effect works as an external force which, in the extreme cases, may entirely disrupt the galaxy resulting in a stellar stream (as in the Sagittarius dwarf galaxy, Belokurov et al. 2006). Usually the tidal radius r_t is defined as the distance beyond which stars within a satellite are stripped by the tidal forces and become unbound (Read et al., 2006; Gajda & Lokas, 2015). This effect works both on the gas component, by removing gas supply, and on stars, leading to morphology transformation and stellar mass loss. Tidal interactions with the cluster gravitational potential can also have opposite effects compressing the gas, enhancing both gas and stellar surface densities in the satellite (Merritt, 1984; Miller, 1986; Byrd & Valtonen, 1990; Nehlig, Vollmer & Braine, 2016).

Harassment

Galaxy harassment is characterized by the combined effect of tidal interactions between a satellite and the host halo potential, together with close high-speed encounters with other satellites (Moore, Katz & Lake, 1996). In this process the rapidly changing tidal fields can make the stars in the satellite gaining energy in the form of random motion. When tidal interactions are relatively moderate, the transfer of energy into internal motions of stars effectively heats the ordered structures in the satellite (Gnedin, 2003; Mastropietro et al., 2005), while several close encounters are able to fully destroy the disk (Moore, Katz & Lake, 1996; Mastropietro et al., 2005; Aguerri & González-García, 2009).

Mergers

Mergers have strong effects on galaxies, and are able to change their morphology (Kormendy & Kennicutt, 2004). They are usually classified according to the mass ratio M_1/M_2 between the involved galaxies (where M_1 is the bigger mass) as major mergers ($M_1/M_2 < 4$), minor mergers ($M_1/M_2 > 10$) and intermediate mergers ($4 < M_1/M_2 < 10$). Mergers can be also classified as "wet merger", between gas-rich galaxies (which are able to induce star formation bursts) or "dry merger" between gas-poor galaxies (Hopkins et al., 2005; Li et al., 2007). Since the work of Toomre & Toomre (1972) there has been increasingly theoretical and observational effort to account for the effects of mergers in galaxies (e.g. Toomre 1976; Fall 1979; Farouki & Shapiro 1982; Barnes 1992; Naab & Burkert 2003; Mihos & Hernquist 1996). Major mergers are able to strongly modify the morphology; for instance mergers between two spirals are thought to be the main formation mechanism for intermediate mass ellipticals (e.g. Naab & Burkert 2003; Bournaud, Jog & Combes 2005; Cox et al. 2006), while two early-type galaxies can generate a more massive elliptical when merging (Naab, Khochfar & Burkert, 2006; González-García et al., 2009) with a wider fundamental plane (González-García & van Albada, 2003). Minor mergers produce less effects on large spiral galaxies (Walker, Mihos & Hernquist, 1996; Naab, Johansson & Ostriker, 2009; Velazquez & White, 1999; Abadi et al., 2003), although they can trigger bursts of star formation and AGN activity (Mihos & Hernquist, 1994), form bulges (Wyse, Gilmore & Franx, 1997), and dynamical heat the stellar disc (Tapia, Balcells & Eliche-Moral, 2010; Villalobos, Kazantzidis & Helmi, 2010; Di Matteo et al., 2011). On elliptical galaxies minor mergers are able to generate counter-rotating cores (Balcells & Quinn, 1990) and perhaps play a role in the formation of the slowest rotators (Emsellem et al., 2007).

2.3.7 Hydrodynamical simulations

While N-body simulations can follow the gravitational evolution of dark matter density perturbations (sec. 2.2.2), hydrodynamical codes are able to describe the coupled evolution of dark matter and baryons, and consistently simulate the formation and evolution of galaxies in a cosmological context from primordial perturbations. A variety of hydrodynamical numerical schemes have been developed over the years, divided into two broad categories: particle (Lagrangian) methods (Smoothed-particle hydrodynamics, Gingold & Monaghan 1977; Lucy 1977) which discretize the mass, and grid-based (Eulerian) methods, which discretize the space (for a comparison among the two methods see Agertz et al. 2007). Recently a new particularly promising method has also been developed, the so-called "moving-mesh" technique, based on an unstructured moving grid generated from Voronoi tessellation (Vogelsberger et al., 2012).

In general, the grid-based codes (e.g. RAMSES, Teyssier 2002; ENZO, O'Shea et al. 2004; ART, Kravtsov 1999; FLASH, Fryxell et al. 2000) work well in both low- and high-density regions and in shocks, which are common in the cosmological environment. However, they suffer from limited spatial resolution and lack of Galilean-invariance, as well as from the presence of overmixing (Dolag et al., 2008). Particle-based (or SPH) codes (e.g. HYDRA, Couchman, Thomas & Pearce 1995; GRAPESPH, Steinmetz 1996; ASPH, Owen et al. 1998; GADGET, Springel 2005; GASOLINE, Wadsley, Stadel & Quinn 2004; VINE, Wetzstein et al. 2009) can achieve good spatial resolutions in high-density regions, but work poorly at low densities, and tend to artificially suppress the fluid instabilities (Dolag et al., 2008).

We here review the basis of the SPH approach.

Smoothed Particle Hydrodynamics

In the Smoothed Particle Hydrodynamics (SPH) scheme (for a review see Monaghan 1992; Dolag et al. 2008) the continuous fluid quantity $A(\mathbf{x}_i)$ in $\mathbf{x}_i$ is written, discretizing the mass by a set of individual particles m_j at position x_j, as

$$\langle A_i \rangle = \langle A(\mathbf{x}_i) \rangle = \sum_j \frac{m_j}{\rho_j} A_j W(\mathbf{x}_i - \mathbf{x}_j, h) \tag{2.48}$$

which is the discrete version of the general definition of a kernel smoothing method

$$\langle A(\mathbf{x}) \rangle = \int W(\mathbf{x} - \mathbf{x}', h) A(\mathbf{x}') dx' \tag{2.49}$$

where the kernel $W(\mathbf{x}, h)$ satisfies these requirements:

$$\begin{aligned} \int W(\mathbf{x}, h) d\mathbf{x} &= 1 \\ W(\mathbf{x}, h) &\rightarrow \delta(\mathbf{x}) \qquad \text{for } h \rightarrow 0 \end{aligned} \tag{2.50}$$

i.e. the kernel is normalized to one and goes to the delta function as the smoothing length h approaches zero. Note that we replace the volume element of the integration $d\mathbf{x} = d^3x$ with the ratio of the mass and density m_j/ρ_j of the particles in (2.48)[3]. For kernels with compact support, i.e. $W(\mathbf{x}, h) = 0$ for $|\mathbf{x}| > h$, the summation is done only over the neighbour particles within a sphere of radius h from the particle i under consideration.

If the quantity A_i is the density ρ_i, we are able to obtain the density estimate:

[3]See Hopkins (2013) for a different formulation of the SPH equations.

$$\langle \rho_i \rangle = \sum_j m_j W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{2.51}$$

The derivative of A_i can be calculated as

$$\nabla \langle A_i \rangle = \sum_j \frac{m_j}{\rho_j} A_j \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h) \tag{2.52}$$

where ∇_i is the derivative with respect to x_i. Using these identities, the Euler equation in (2.17) can be written in the SPH formulation as

$$\frac{d\mathbf{v}_i}{dt} = -\sum_j m_j \left(\frac{P_j}{\rho_j^2} + \frac{P_i}{\rho_i^2} + \Pi_{ij} \right) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h) \tag{2.53}$$

and the conservation of energy (without cooling) as

$$\frac{du_i}{dt} = \frac{1}{2} \sum_j m_j \left(\frac{P_j}{\rho_j^2} + \frac{P_i}{\rho_i^2} + \Pi_{ij} \right) (\mathbf{v}_j - \mathbf{v}_i) \nabla_i W(\mathbf{x}_i - \mathbf{x}_j, h). \tag{2.54}$$

The artificial viscosity term Π_{ij} is added in order to capture shocks; several forms for it have been proposed, by e.g. Monaghan & Gingold (1983); Balsara (1995); Steinmetz (1996); Monaghan (1997); Morris & Monaghan (1997). The continuity equation in (2.17) does not have to be evolved explicitly, as it is automatically satisfied due to the Lagrangian nature of SPH.

Additional physics

In addition to gravity and hydrodynamics, several other physical processes are relevant in galaxy formation (see previous sections); modern hydrodynamical codes usually implement recipes to describe these additional processes, in most of the cases at the sub-resolution level (Dolag et al., 2008).

- **Radiative cooling**

  Including radiative cooling into the hydrodynamical scheme is straightforward achieved by simply adding the cooling function $\Lambda(u, \rho)$ into eq. (2.54); this is usually accomplished using tabulated cooling functions for optical thin gas (e.g. Sutherland & Dopita 1993). When calculating cooling rates it is also important to account for photoionisation due to the UV background (as done e.g. by Wiersma, Schaye & Smith 2009; Smith, Sigurdsson & Abel 2008), because it may affect both the thermal and ionization state of the plasma (Efstathiou, 1992; Gnedin & Hollon, 2012). The optical-thin gas approximation however breaks down for gas densities $\rho \gtrsim 10^{-3}$ cm^{-3}, where the gas efficiently attenuates the UV background radiation field; some simple prescriptions to account for self-shielding from UV background are given e.g. in Vogelsberger et al. (2013).

- **Star formation**

  For modelling star formation, the recipe by Katz, Weinberg & Hernquist (1996) is often used. According to this prescription, for a gas particle to be eligible to form a star, it must have a convergent flow

$$\nabla \cdot \mathbf{v} < 0 \tag{2.55}$$

and have density above some threshold value $\rho > \rho_*$ which is usually assumed of the order $\rho_* \approx 0.1 - 10$ atoms cm^{-3} (e.g. Vogelsberger et al. 2013; Wang et al. 2015; Stinson et al. 2006). In addition the gas must be Jeans unstable (sec. 2.2.1)

$$h > c_s \frac{1}{\sqrt{4\pi G\rho}} \tag{2.56}$$

where h is either the SPH smoothing length or the mesh size for Eulerian codes, and c_s is the local sound speed. Note that this condition is often automatically fulfilled when $\rho > \rho_*$ holds (Dolag et al., 2008).

Once a gas particle is eligible to form a star, its star formation rate $\dot{\rho}_*$ is calculated as

$$\dot{\rho}_* = c\,\frac{\rho}{t} \tag{2.57}$$

where c is the dimensionless star formation efficiency parameter and t the characteristic timescale for star formation, usually taken to be the dynamical time $t_{\rm dyn} = 1/\sqrt{4\pi G\rho}$. The conversion of gas into star particles is usually implemented in hydrodynamical codes in a stochastic way (see Springel & Hernquist 2003; Lia, Portinari & Carraro 2002; Dolag et al. 2008); when a star particle is generated, it has usually a mass of the order $\sim 10^4 - 10^5 M_\odot$ and it represents a Simple Stellar Population (SSP, Bruzual & Charlot 2003) of coeval stars with mass spectrum given by the assumed IMF.

- **Multiphase ISM**

  The multiphase nature of the ISM (McKee & Ostriker, 1977; Efstathiou, 2000) is difficult to describe with the standard SPH formulation, in particular the co-existence of a hot and a cold phase gas is poorly represented. Multiphase models of the ISM embedded in hydrodynamical codes (such as Springel & Hernquist 2003; Scannapieco et al. 2006) are usually sub-resolution; they allow a more realistic description of star formation, and can follow the energy and metal injection from SNe explosions to the distinct phases of the ISM.

- **Feedback from SNe**

  The number of SNe type II explosions is calculated in hydrodynamical codes from the chosen IMF as the number of stars with mass larger than $8M_\odot$. It is often assumed that the typical lifetime of massive stars does not exceed the integration time-step of the simulation, and the feedback energy (typically 10^{51} erg per SN) is deposited into the surrounding gas in the same time step; lifetimes of SNII progenitors depend on mass and metallicity, and are usually considered in current simulations (see e.g. Raiteri, Villata & Navarro 1996). On the other hand, the rate of the SNIa explosions is still poorly understood (Howell, 2011), and several ways to calculate it stochastically are often used (e.g. Greggio & Renzini 1983; Thielemann, Nomoto & Hashimoto 1993; Maoz & Mannucci 2012; Madau 1998). Due to the resolution limitation, two approaches are mainly followed to couple SNe explosions to the surrounding ISM: injecting the SN energy into kinetic form (e.g. Navarro & White 1993; Mihos & Hernquist 1994; Kay et al. 2002; Oppenheimer & Davé 2006; Dubois & Teyssier 2008; Dalla Vecchia & Schaye 2008; Hopkins et al. 2012), into thermal energy (e.g. Thacker & Couchman 2000a; Kawata & Gibson 2003; Sommer-Larsen, Götz & Portinari 2003; Brook et al. 2004; Stinson et al. 2006; Piontek & Steinmetz 2011; Dalla Vecchia & Schaye 2012; Stinson et al. 2013b; Scannapieco et al. 2006) or both (e.g. Agertz et al. 2013; Aumer

et al. 2013). Alternative SNe feedback recipes can be found in the literature, such as the blast-wave formalism (Stinson et al., 2006) or the super-bubble feedback (Keller et al., 2014; Keller, Wadsley & Couchman, 2015).

- **Feedback from massive stars**

  As suggested by Hopkins, Quataert & Murray (2011) using high-resolution simulations, feedback from young massive stars may play a critical role in the galaxy evolution process and has a comparable effect on the ISM than SNe feedback. The exact form of feedback from massive stars is uncertain, although some models exist in the literature (e.g. Agertz et al. 2013; Hopkins, Quataert & Murray 2011; Stinson et al. 2013b) which have been already implemented in hydrodynamical codes by e.g. Aumer et al. (2013).

- **Black holes accretion and feedback**

  The details of accretion into central AGNs are currently impossible to resolve in cosmological simulations, since their Schwarzschild radius is typically much smaller than the spatial resolution. Subgrid models that describe AGN growth and feedback have been implemented by e.g. Springel et al. (2005); Di Matteo, Springel & Hernquist (2005); Thacker, Scannapieco & Couchman (2006); Sijacki et al. (2007); Di Matteo et al. (2008); Okamoto (2008); Booth & Schaye (2009); however, such models typically depend on various free parameters, which are not well constrained. AGN feedback schemes are based on the observational evidence of two distinct phases of AGN activity, namely the "radio-mode" characterized by low accretion rate and radio jets (Mezcua & Prieto, 2014), and the "quasar-mode" with high accretion rate and emission dominated by the accretion disc (Prieto et al., 2010); this is usually implemented applying two different types of feedback (respectively mechanical and thermal) chosen according to the BH accretion rate (Springel et al., 2005; Sijacki et al., 2007). Some models also include the effect of the AGN radiation field (e.g. Vogelsberger et al. 2013) that can alter the photo-ionization and photo-heating rates of the nearby plasma (Gnedin & Hollon, 2012; Efstathiou, 1992; Hambrick et al., 2011).

- **Chemical enrichment**

  After the first approaches to include chemical evolution in SPH by Steinmetz & Mueller (1994); Steinmetz & Muller (1995); Raiteri, Villata & Navarro (1996), chemical enrichment is currently tracked by several galaxy formation codes (e.g. Scannapieco et al. 2005; Pontzen et al. 2008; Crain et al. 2015; Vogelsberger et al. 2014). After the formation of a star particle, the metals are released to the ISM assuming some metal-dependent yields for SNe II and SNe Ia explosions (such as Portinari, Chiosi & Bressan 1998; Woosley & Weaver 1995; Chieffi & Limongi 2004; Thielemann et al. 2003; Marigo 2001), and in some codes including also the chemical enrichment from stars in the AGB phase. The problem of distributing the metals to the surrounding ISM is complex, since simulations do not reach enough resolution to properly treat the small-scale mixing process; a numerical scheme usually applied in SPH codes is the kernel interpolation method, where gas particles close to the exploding stars receive a fraction of the ejected elements according to their kernel weight (Mosconi et al., 2001). However, this may lead to situations where particles with similar thermodynamical properties but very different metallicities are close to each other (Aumer et al., 2013). As suggested by Wiersma, Schaye & Smith (2009), smoothing the metallicity between neighbour particles may improve the modelling of chemical enrichment in SPH; some

numerical schemes are given by Martínez-Serrano et al. (2008); Greif et al. (2009); Shen, Wadsley & Stinson (2010).

- **Magnetic fields**

  Magnetohydrodynamics (MHD) can be included in galaxy formation codes to describe the effect of cosmological magnetic fields on structure formation (Durrer & Neronov, 2013), and it has been already implemented in SPH (e.g. Dolag & Stasyszyn 2009) and in grid-based codes (e.g. ATHENA, Stone et al. 2008). MHD is important in galaxy cluster studies, as various radio observations have shown that the hot atmospheres of the clusters are magnetized (Govoni & Feretti, 2004; Ferrari et al., 2008). Since the understanding of the origin of cosmological magnetic fields is particularly limited, as well as their evolution and implications for structure formation, simulations including MHD may help to investigate the physics behind the observed magnetic field in galaxy clusters, providing also theoretical support to interpret the results of the new generation of radio telescopes, such as LOFAR, (van Haarlem et al., 2013), SKA (Huynh & Lazio, 2013) and VLASS (Hales, 2013).

- **Thermal conduction**

  Thermal conduction has been proposed by Narayan & Medvedev (2001) as a possible heating mechanism for offsetting cooling in the center of rich galaxy clusters, as observed from X-ray spectra (David et al., 2001). Thermal conduction has been introduced in SPH codes by Jubelgas, Springel & Dolag (2004) to study its effects on galaxy clusters, finding that it may lead to substantial changes in the thermodynamical properties of rich clusters (Dolag et al., 2004).

Zoom-in technique

N-body simulations that follow the evolution of dark matter perturbations are usually run in large, periodic volumes in order to provide a fair representation of the real Universe including the long-scale modes in the power spectrum (Power & Knebe, 2006). However, when hydrodynamics is added in order to simulate for instance the formation of galaxies in cosmological context, running simulations in these big cosmological volumes is time-consuming and requires a large amount of computer memory.

The method that allows to self-consistently embed a high-resolution region into a large cosmological volume at lower resolution, considering the effect of the cosmological environment on large scale and simultaneously reaching sufficiently high resolution to describe the physics at galactic scales, is the so-called "zoom-in technique" (Navarro & White, 1994; Frenk et al., 1996; Tormen, Bouchet & White, 1997; Thacker & Couchman, 2000b; Bertschinger, 2001; Springel et al., 2008; Klypin et al., 2001); for galaxy formation studies, the high-resolution area of interest is usually the virialized region of a dark matter halo.

Generating a "zoom-in" simulation normally requires the following steps:

- Low-resolution initial conditions for a dark matter-only simulation with a single refinement level are created, choosing a cosmological model, box size, initial redshift and creating the matter power spectrum at that redshift.

- The low-resolution dark-matter only simulation is run up to the desired redshift.

- From the low-resolution simulation interesting volumes are identified, usually dark matter halos with specific properties. Particles inside these volumes are traced back to the initial conditions, and the Lagrange volume is defined by the initial positions of those particles.

- New initial conditions with higher mass resolution and including baryons are created in the Lagrange volume, while boundary regions around the Lagrange volume may be initialized at an intermediate resolution.

- The zoom-in with these new initial conditions is run including hydrodynamics and the additional subgrid physics specific of the particular hydrodynamical code used.

It should be noticed that a possible issue of having multiple resolution levels as with the zoom-in technique is when low-resolution particles end up in the high-resolution region (referred to as "contamination", see Oñorbe et al. 2014).

The zoom-in technique is now a standard method to study galaxy formation in cosmological context, and it has been used in several works such as Scannapieco et al. (2008); Wang et al. (2015); Stinson et al. (2010, 2012); Anglés-Alcázar et al. (2014); Hopkins et al. (2014); Governato et al. (2007); Aumer et al. (2013); Governato et al. (2010).

Chapter 3

Biases and systematics in the observational derivation of galaxy properties

In this Chapter we begin our study of the possibilities offered by connecting simulations and observations in the study of galaxies.

In observations, complex algorithms are usually applied to analyze the shape and features of the galaxies' Spectral Energy Distributions (SEDs) recovered by the observing instruments, in order to estimate the different physical properties of the target galaxies such as the stellar, gas and metal content, the conditions of the InterStellar Medium (ISM) and the characteristics of the hosted Active Galactic Nuclei (AGNs). Recently, several galaxy surveys such as SDSS (Abazajian et al., 2003), 2dFGRS (Colless, 1999), HDUF (Beckwith et al., 2006) and 2MASS (Skrutskie et al., 2006) have obtained spectroscopic and photometric information for a large number of galaxies at different redshifts, allowing to statistically study the formation and evolution of galaxies in the Universe. On the other hand, numerical simulations are a useful and complementary tool in galaxy formation studies, and also an important aspect in the interpretation of observational results, since they allow to investigate the link between a galaxy's formation, merger and accretion history and its properties at different times. Progress on the simulation of realistic galaxies has been presented in many recent works (e.g. Governato et al. 2007; Scannapieco et al. 2008; Aumer et al. 2013; Vogelsberger et al. 2014; Nelson et al. 2015; Schaye et al. 2015) that, together with advances in computational resources, are starting to allow to simulate relatively large cosmological volumes, recreating a virtual universe where galaxies are naturally diverse as the result of their individual evolution.

Despite the individual progress in observations and simulations, much uncertainty still remains in the comparison between them. Previous works using simulations have shown that, as the methods usually applied to derive the properties of simulated galaxies are very different compared to the observational techniques, several biases might be introduced making the comparisons between simulations and observations unreliable (e.g. Abadi et al., 2003; Governato et al., 2009; Scannapieco et al., 2010; Snyder et al., 2011; Munshi et al., 2013; Christensen et al., 2014). In addition, each observational method has its own uncertainties and biases, mostly related to the detector noise and to the simplified modelling of the fitted SEDs that are used to estimate the galaxies' properties; these effects have a strong impact on the physical quantities recovered in observations, which are often quite discrepant when the same object is analyzed using different methods, as shown for instance by Rosa-González, Terlevich & Terlevich (2002); Kewley & Ellison (2008); Pforr, Maraston & Tonini

(2012, 2013); Michałowski et al. (2014). In order to better understand the reliability of observational analysis, as well as to properly judge the agreement between simulations and observations, it is of primary importance that the effects of these uncertainties are considered.

This is the main focus of this Chapter, in which we estimate the biases and uncertainties in the derivation of several galaxies' properties in observations (stellar masses, star formation rates, gas/stellar metallicities, stellar ages and magnitudes/colors) by applying observational methods on synthetic SEDs created from hydrodynamical cosmological simulations of galaxy formation. As the properties of the simulated galaxies are known, we can quantify how close the physical quantities obtained observationally are to the real ones, which in turn allows to more consistently compare observations and simulations, as well as to better interpret the results of observations.

3.1 The simulations

In this work, we use three sets of galaxy simulations consisting in total of fifteen galaxies formed in a ΛCDM universe. Each set comprises the same five galaxies but adopts a different modelling of chemical enrichment and feedback, which is known to introduce differences in the final properties of galaxies and in their evolution. The focus of this work is to identify whether the analysis techniques used to extract galaxy properties from simulations and observations introduces important biases that make the comparisons unreliable. Our set of fifteen galaxies is ideal for this purpose as they all have similar total mass but span a wide range in gas/stellar metallicities, stellar ages and star formation rates.

As the goal of this work is not to decide how realistic the galaxies are, covering a variety of galaxy properties allows one to test the biases properly, unaffected by particular details of a given implementation. We use our sample to test minimum and maximum biases introduced in the conversion of simulations into observables, as this conversion is expected to primarily depend on the age of the stellar populations, the stellar masses, the metallicities and the galaxy morphologies. The latter is important as, on one side, dust will influence differently observations of face-on and edge-on galaxies and, on the other hand, the presence of gradients in galaxy properties might affect the derivation of their global properties from the observationally-obtained fiber quantities, since fiber spectrographs such as those used for SDSS sample only the inner region of galaxies.

The initial conditions correspond to five galaxies which are the hydrodynamical counterparts of (a subset of) the Aquarius halos (Springel et al., 2008). These are similar in mass to the Milky Way and formed in isolated environments (no neighbour exceeding half their mass within 1.4 Mpc at redshift $z = 0$), but have different merger and accretion histories, as discussed in Scannapieco et al. (2009). The galaxies have virial masses between 0.7 and $1.7 \times 10^{12} \mathrm{M}_{\odot}$ (calculated within the radius where the density contrast is 200 times the critical density), stellar masses of $1 - 10 \times 10^{10} \mathrm{M}_{\odot}$, and gas masses of $3 - 10 \times 10^{10} \mathrm{M}_{\odot}$.

For our first set of simulations, we have used the extended version of the Tree-PM SPH code Gadget-3 (Springel, 2005), which includes star formation, chemical enrichment, supernova Type Ia and TypeII feedback, metal-dependent cooling and a multiphase model for the gas component which allows the coexistence of dense and diffuse phases (Scannapieco et al., 2005, 2006). These simulations have been first presented in Scannapieco et al. (2009) and further analysed in Scannapieco et al. (2010) and Scannapieco et al. (2011), and will be called throughout this chapter *A(-E)-CS* or *CS* galaxies. The Scannapieco et al. code has been extensively used for simulating galaxies of a wide range of total/stellar masses, and shown to be successful in reproducing the formation of galaxy discs from cosmological initial

conditions of Milky Way-mass galaxies (Scannapieco et al. 2008, see also Sawala et al. 2010; Scannapieco et al. 2012; Nuza et al. 2014; Creasey et al. 2015; Scannapieco et al. 2015).

For our second simulation set, we used an updated version of the model of Scannapieco et al., in relation to the treatment of chemical enrichment (Poulhazan et al., in prep). These simulations will be referred to as *A(-E)-CS*$^{+}$ or *CS*$^{+}$. The updated code implements a different Initial Mass Function (Chabrier instead of Salpeter), chemical yields from Portinari, Chiosi & Bressan 1998 (while the CS model uses the Woosley & Weaver 1995 yields), and the treatment of feedback from stars in the AGB phase, which contribute significant amounts of given chemical elements, such as carbon and nitrogen. For these reasons, the CS^{+} galaxies have systematically higher chemical abundances compared to those in the CS sample. The modelling of energy feedback is the same as in the standard Scannapieco et al. code.

The final set of simulations, which will be referred to as *A(-E)-MA* or the MA galaxies, have used the Aumer et al. (2013) code, which is an independent update to the Scannapieco et al. (2006) model. This code has a different set of chemical choices in relation to the initial mass function and chemical yields, includes also stars in the AGB phase, and assumes a different cooling function. More importantly, it has a different treatment of supernova energy feedback: unlike in the Scannapieco et al. code, where feedback is purely thermal, in Aumer et al. the supernova energy is divided into a thermal and a kinetic part, and also includes feedback from radiation pressure. This results in stronger feedback effects compared to the Scannapieco et al. model and, as a consequence, produces galaxies that are more disk dominated, younger and more metal rich compared to the rest of our simulations. We refer the interested reader to Aumer et al. (2013) for full details on this implementation.

The assumed cosmological parameters of the simulations are as follows: $\Omega_{\rm m} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.04$, $\sigma_8 = 0.9$ and $\rm H_0 = 100\,\mathit{h}\,km\ s^{-1}Mpc^{-1}$, with $h = 0.73$. All simulations have similar mass resolution ($2 - 5 \times 10^5 \rm M_\odot$ for stellar/gas particles and $1 - 2 \times 10^6 \rm M_\odot$ for dark matter particles) and adopt similar gravitational softenings ($300 - 700$ pc).

Fig. 5.2 shows the color-composite images of the fifteen simulated galaxies, in face-on and edge-on views, obtained with the code SUNRISE (see below). The composite images in the (u, r, z)-bands are generated using the algorithm described in Lupton et al. (2004). The edge-on and face-on views are defined such that the total angular momentum of the stars in the galaxy is aligned with the z-direction.

3.2 Creating synthetic spectra of the simulated galaxies

To derive the observables of our simulated galaxies and investigate the biases introduced in the process, we follow three approaches. First, we use different Stellar Population Synthesis (SPS) models, which provide the resulting spectrum due to the emission of the stars, as well as different information about the stellar populations at a given metallicity and age (e.g. mass loss, ionizing photon flux, Lick indices, number of black holes and neutron stars). Second, we add a simple dust model to the prediction of the SPSs, in order to better compare the resulting spectra from those of observed galaxies. Finally, we postprocess the simulations with the radiative transfer code SUNRISE (Jonsson, 2006; Jonsson, Groves & Cox, 2010), which is computationally slower but more consistent with the underlying hydrodynamical simulation. This gives the full SED including stellar absorption features, nebular emission, and extinction due to dust as light travels out through the interstellar medium.

The SPS models are commonly used as a postprocessing of hydrodynamical simulations as they provides a fast and easy way to estimate the light distribution of an ensemble of stars.

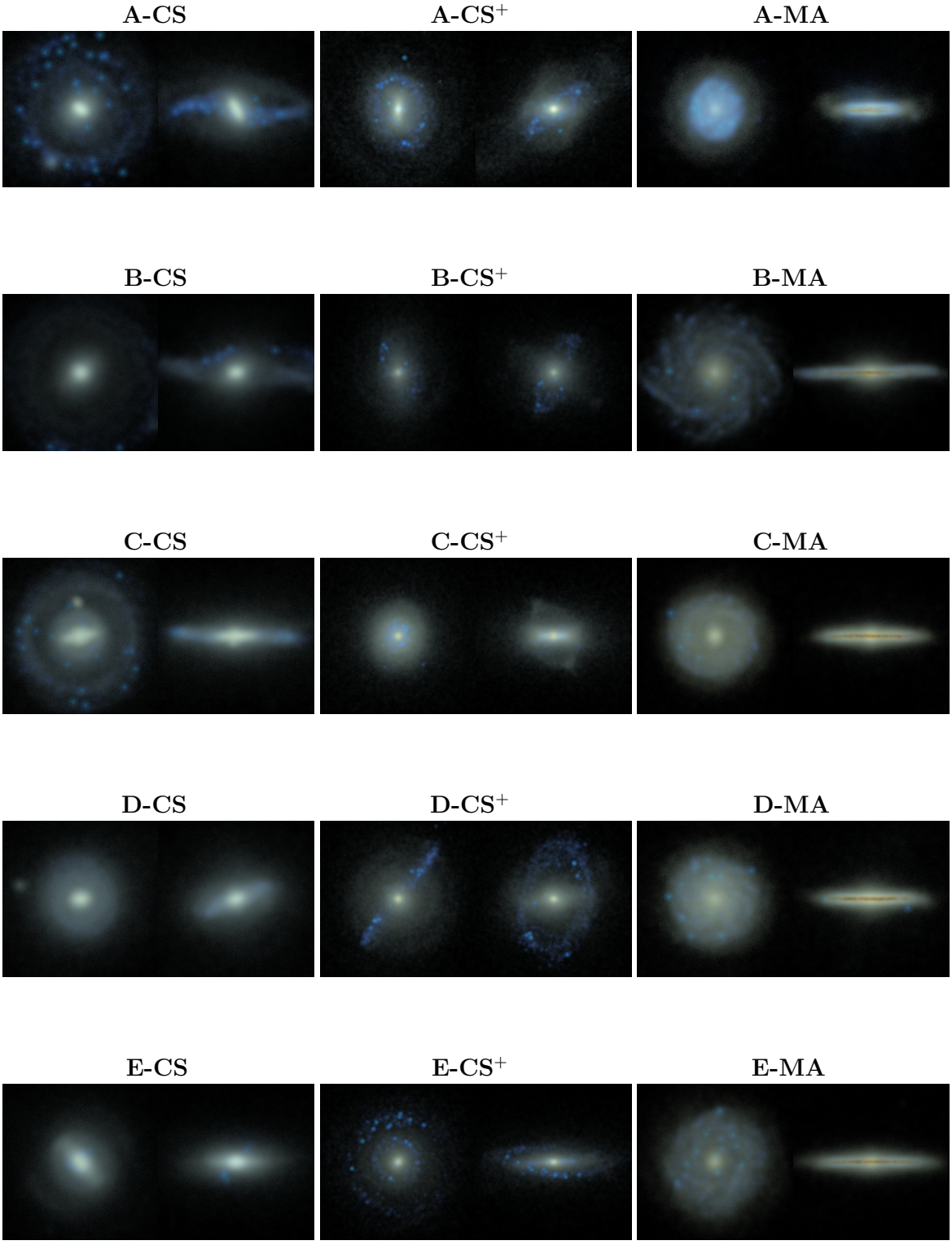

Figure 3.1: Multi-band (u, r, z) images of our fifteen galaxies, for face-on and edge-on views.

The use of radiative transfer codes has also been applied to hydrodynamical simulations (e.g. Governato et al. 2009; Scannapieco et al. 2010; Hayward et al. 2013; Christensen et al. 2014). This approach has the advantage that can consistently consider the distribution of dust (traced by the metals) although it also needs to introduce certain assumptions and simplifications.

In the following subsections, we describe the main aspects of these models and the assumptions that are relevant for our work, as well as discussing the main uncertainties and sources of biases in the creation of the synthetic SEDs.

3.2.1 Stellar population synthesis Models

We use various SPS models to create synthetic spectra of our simulated galaxies. The spectra are obtained assuming that each star particle represents a Simple Stellar Population (SSP) parametrized by the age and metallicity of the star, and normalized by the mass of the particle. In the simulations, the mass of a star particle changes with time owing to the supernova ejecta and the mass loss during the AGB phase; the mass normalization should be made with the mass of the stars *at the formation time* (note that this is important in order to avoid double-counting the effects of stellar mass loss). The spectrum of the galaxy is obtained by summing up the spectra of individual stars. The galaxy spectra can be convolved with given photometric bands and integrated to get the magnitude of the galaxy in the bands.

We use 5 different SPS models to obtain spectra, magnitudes and colors of our simulated galaxies, which allows us to identify systematic effects due to their varying assumptions. Although moderate, some differences between these models are expected, as uncertainties in modelling the stellar evolution still exist (e.g. different treatment of convection, rotation, mass-loss, thermal pulses during AGB evolution, close binary interaction), as well as in the empirical and theoretical stellar spectral libraries (see the reviews by Walcher et al. 2011; Conroy 2013). The main characteristics of the SPS models are described in the following, and in Table 1 we give a summary of the input parameters we chose, taken as homogeneous as possible to make the interpretation of our results more clear.

- **BC03** (Bruzual & Charlot, 2003) computes, using different evolutionary tracks, the spectral evolution of stellar populations at a resolution of 3 Å FWHM in the optical and at lower resolution in the full wavelength range.

- **SB99** ("STARBURST99", Leitherer et al. 1999; Vázquez & Leitherer 2005) is a web-based platform that allows users to run customized SPS models for a wide range of IMF, isochrones, model atmospheres, ages and metallicities. The spectral resolution reaches ~ 1 Å FWHM in the optical wavelength range for the fully theoretical spectra.

- **PE** ("PEGASE", Fioc & Rocca-Volmerange 1997, 1999) uses an algorithm that can accurately follow the stellar tracks of very rapid evolutionary phases such as red supergiants or TP-AGB. The stellar library includes also cold star parameters, and a simple model for nebular emission (continuum + lines) can be added to the stellar spectrum.

- **FSPS** ("Flexible Stellar Population Synthesis 2.3", Conroy, Gunn & White 2009) is a flexible SPS package that can compute spectra at resolving power $\lambda/\Delta\lambda \approx 200 - 500$. In addition to the choice of IMF, metallicities, ages, the user can select a variety of assumptions on Horizontal Branch (HB) morphology, blue straggler population, location of the TP-AGB phase in the HR-diagram and post-AGB phase.

Table 3.1: Summary of the characteristics and choices of parameters for the 5 SPS models used in our work. The input stellar model in SUNRISE is SB99 with the same parameters.

Model	IMF	Age range (yr)	Metallicity range	Wavelength range	Stellar tracks	Stellar library
BC03	Chabrier(1)	$10^5 - 2 \cdot 10^{10}$	0.0001 - 0.05	$91\,\AA - 160\,\mu m$	Padova 1994	BaSeL3.1 / STELIB
M05	Kroupa(2)	$10^3 - 1.5 \cdot 10^{10}$	0.0001 - 0.07	$91\,\AA - 160\,\mu m$	Cassisi / Schaller	BaSeL3.1 / Lançon & Mouchine
SB99	Kroupa(2)	$10^4 - 1.5 \cdot 10^{10}$	0.0004 - 0.05	$91\,\AA - 160\,\mu m$	Padova 1994	Pauldrach / Hillier
PE	Kroupa(2)	$0 - 2 \cdot 10^{10}$	0.0001 - 0.05	$91\,\AA - 160\,\mu m$	Padova 1994	BaSeL3.1 / ELODIE
FSPS	Kroupa(2)	$3 \cdot 10^5 - 1.5 \cdot 10^{10}$	0.0002 - 0.03	$91\,\AA - 10\,mm$	Padova 2007	BaSeL3.1 / Lançon & Mouchine

NOTES: (1) Mass range: m = 0.1-100 $M_\odot$, $\alpha = 2.3$ for $m > 1$ $M_\odot$; (2) $\alpha = 1.3$ for m =0.1-0.5 $M_\odot$, $\alpha = 2.3$ for m = 0.5-100 $M_\odot$.

- **M05** (Maraston, 2005) is different from the other SPS models in the treatment of Post Main Sequence stars, namely using the Fuel Consumption theorem (Renzini & Buzzoni, 1986) to evaluate the energetics. It gives spectra at a resolution $5 - 10$ Å in the visual region and at $20 - 100$ Å from the NUV to the near-IR for either blue or red populations on the horizontal branch.

3.2.2 Dust

Dust extinction is an important ingredient in the estimation of observables from the simulations, as in observed galaxies dust effects can be large, especially for edge-on systems. Different authors have modelled dust extinction curves of the Milky Way (e.g. Seaton, 1979; Cardelli, Clayton & Mathis, 1989), of the Large and Small Magellanic clouds (e.g. Fitzpatrick, 1986; Bouchet et al., 1985), and also of external galaxies (e.g. Silva et al., 1998; Charlot & Fall, 2000; Calzetti et al., 2000; Fischera & Dopita, 2005), which are often used to correct for dust extinction.

In our calculations of the dust-corrected magnitudes and colors, we use the model of Charlot & Fall (2000, CF00 hereafter) in the slightly different formulation given in da Cunha, Charlot & Elbaz (2008, dC08 hereafter), that consistently extend the CF00 model to include dust emission, in order to allow the interpretation of the full UV-far infrared galaxy SEDs.

CF00 is an angle-averaged time-dependent model, with extinction curve depending on the wavelength, and on the Stellar Population (SP) age. This model has been proven to work reasonably well for a wide class of galaxies, and it has been already implemented in SPS models (Bruzual & Charlot, 2003). In CF00, stars are assumed to born in "birth clouds" that disperse after a certain amount of time (see also Silva et al., 1998); the transmission function of the SP is the product of the transmission function in the birth cloud (which depends on the SP age t) and in the ISM:

$$T_\lambda^{\rm BC}(t) = \begin{cases} \exp\left[-(1-\mu)\,\tilde{\tau}_V \left(\frac{\lambda}{5500\,{\rm A}}\right)^{-N_{\rm BC}}\right] & \text{for } t \le t_{\rm BC} \\ 1 & \text{for } t > t_{\rm BC} \end{cases}$$

$$T_\lambda^{\rm ISM} = \exp\left[-\mu\,\tilde{\tau}_V \left(\frac{\lambda}{5500\,{\rm A}}\right)^{-N_{\rm ISM}}\right]$$

where $t_{\rm BC}$ is the birth cloud life-time (see CF00 and dC08 for details). The resulting attenuated luminosity L_λ given the intrinsic luminosity S_λ is then

$$L_\lambda(t) = S_\lambda\, T_\lambda^{BC}(t) T_\lambda^{ISM}.$$

We set the free parameters to the values given in dC08 for normal star-forming galaxies, namely: $t_{\rm BC} = 10$ Myr, $\tilde{\tau}_V = 1.5$, $\mu = 0.3$, $N_{\rm BC} = 1.3$ and $N_{\rm ISM} = 0.7$.

3.2.3 Radiative Transfer

To calculate the full far-UV to submillimeter SED of our simulated galaxies we use the Monte Carlo Radiative Transfer code SUNRISE in the post-processing phase. SUNRISE (Jonsson, 2006; Jonsson, Groves & Cox, 2010) is a 3D adaptive grid polychromatic Monte Carlo radiative transfer code, suited to process hydrodynamical simulations. SUNRISE assigns a spectrum to each star particle in the simulation, and then propagates photon "packets" from these sources through the dusty ISM using a Monte Carlo approach, assuming a constant dust-to-metals mass ratio, which we fixed to 0.4 according to Dwek (1998).

In the standard SUNRISE implementation, each stellar particle older than 10 Myr is assigned a spectrum corresponding to its age and metallicity from the input SPS model, in our case SB99 (see Table 1). Star particles younger than 10 Myr are assumed to be located in their birth clouds of molecular gas, and are given a modified spectrum which accounts for the effects of HII and photo-dissociation regions (PDRs). The evolution of HII regions and PDRs are described by the photo-ionization code MAPPINGS III (Groves, Dopita & Sutherland, 2004; Groves et al., 2008), which is used to calculate the propagation of the source spectrum through its nebula. The HII regions absorb effectively all ionizing radiation and are the sources of hydrogen recombination lines, as well as hot-dust emission. The only MAPPINGS III parameter not constrained by the hydrodynamical simulation is the PDR clearing timescale; in SUNRISE this free parameter has been changed to the time-averaged fraction of stellar cluster solid angle covered by the PDR ($f_{\rm PDR}$), for which we use the fiducidal value of $f_{\rm PDR} = 0.2$ as in Jonsson, Groves & Cox (2010).

Dust extinction in SUNRISE is described by a Milky Way-like extinction curve normalized to $R_V = 3.1$ (Cardelli, Clayton & Mathis, 1989; Draine, 2003) which includes also the $2175\mathring{A}$ bump observed in our galaxy (Czerny, 2007). Model cameras are placed around the simulated galaxies, to sample a range of viewing angles. The emergent flux is determined by the number of photons that flow from the galaxy unscattered in a given camera's direction, as well as those scattered into the line of sight or reemitted in the infrared by dust into the camera. In the calculations presented here we use two cameras, one looking face-on and the other edge-on to the galaxy, to take into account the two extreme cases in terms of optical depth. Finally, the flux in the cameras is convolved with the bandpass filters and integrated to get the broadband images and magnitudes in the chosen photometric bands.

Fig. 3.2 is an example of the spectra obtained for one of our simulated galaxies, C-MA, showing in the upper panel the SB99 spectrum both without dust (pink line) and corrected with the CF00 model (blue line), as well as the dust-free stellar spectrum with SUNRISE (black dashed line). The lower panel shows again the SUNRISE stellar spectrum without dust, the full nebular+stellar spectrum without dust (red), and the full spectrum with dust from the edge-on camera (green).

3.2.4 Uncertainties and biases

In addition to the different assumptions of the SPS models and the treatment of dust and radiative transfer, there are a number of choices on the process of constructing the synthetic SEDs of simulated galaxies. In some cases, these constitute important sources of biases that need to be kept in mind when results are interpreted, while others do not strongly affect

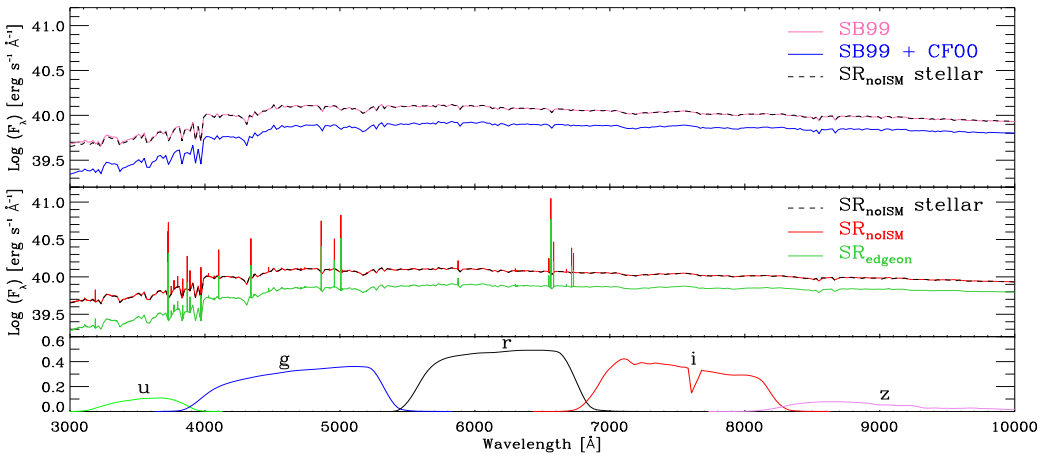

Figure 3.2: Examples of the SEDs obtained for one of our simulated galaxies (C-MA) using different methods, including the stellar spectrum only (SB99), the stellar spectrum and the CF00 simple dust model (SB99 + CF00), the stellar SUNRISE dust-free spectrum (SR_{noISM} stellar) and total SUNRISE spectra in the absence (SR_{noISM}) and presence of dust, for the edge-on view (SR_{edgeon}). The lower panel shows the SDSS filter transmission functions for the u, g, r, i and z-bands.

the final SEDs. When appropriate we will test and quantify these effects, which we describe below.

Aperture bias

While simulations give the full information on the phase space of particles that constitute a galaxy, observations usually have limitations related to the region that can be observed/measured, and the different techniques used to define this region in observations and simulations can introduce some systematics (Stevens et al., 2014). In SDSS, the properties derived from spectra (e.g. stellar age and metallicity, gas metallicity, SFRs) are affected by the small aperture (3" arcsec) of the SDSS spectrograph, which samples only the inner part of the target galaxies. The presence of metallicity gradients in galaxy properties and colors (Bell & de Jong, 2000; Pilkington et al., 2012; Welikala & Kneib, 2012; Sánchez et al., 2015) can therefore lead to substantial uncertainties in the observed measurements when one wants to estimate the total quantities. The implications of this bias have been discussed by several authors (Bell & de Jong, 2000; Kochanek, Pahre & Falco, 2000; Baldry et al., 2002; Gómez et al., 2003; Brinchmann et al., 2004; MacArthur et al., 2004, e.g.); in some cases methods for aperture correction which exploit the spatially-resolved color information can be used to extrapolate to global quantities (e.g. SFRs, Brinchmann et al. 2004, Salim et al. 2007).

We used our simulations to test the effects of using a single fiber in the derivation of properties such as metallicities and stellar ages, by including in the calculations both all particles in a 60x60 kpc field of view (FoV) in the face-on projection (full FoV), or only those within a small region that mimics the size sampled by the SDSS fiber spectrograph, i.e. a circular region of 4 kpc radius with the galaxy in the center observed face-on (fiber FoV)[1], which according to the cosmology adopted in our simulations corresponds to redshift

[1]In the z-direction we count all stars in the halo as identified with SUNFIND as bound to the halo.

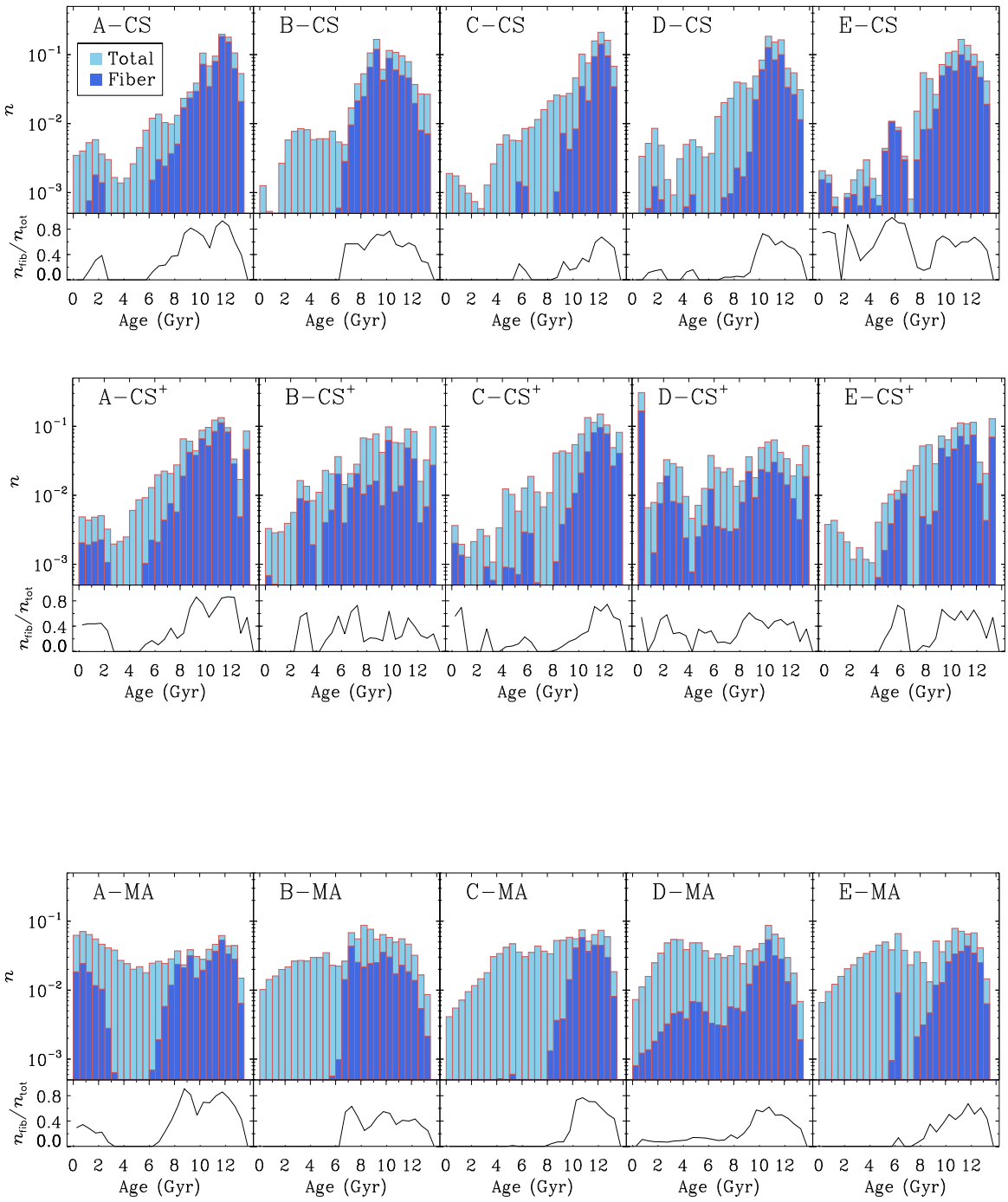

Figure 3.3: Number-density histograms of stellar ages for the 15 simulated galaxies, considering all star particles in the galaxies as well as only those within the fiber. For each galaxy, we also show the fraction between the number of stars within the fiber to the total number of stars in each bin, $n_{\rm fib}/n_{\rm tot}$, that we refer to as the sampling function.

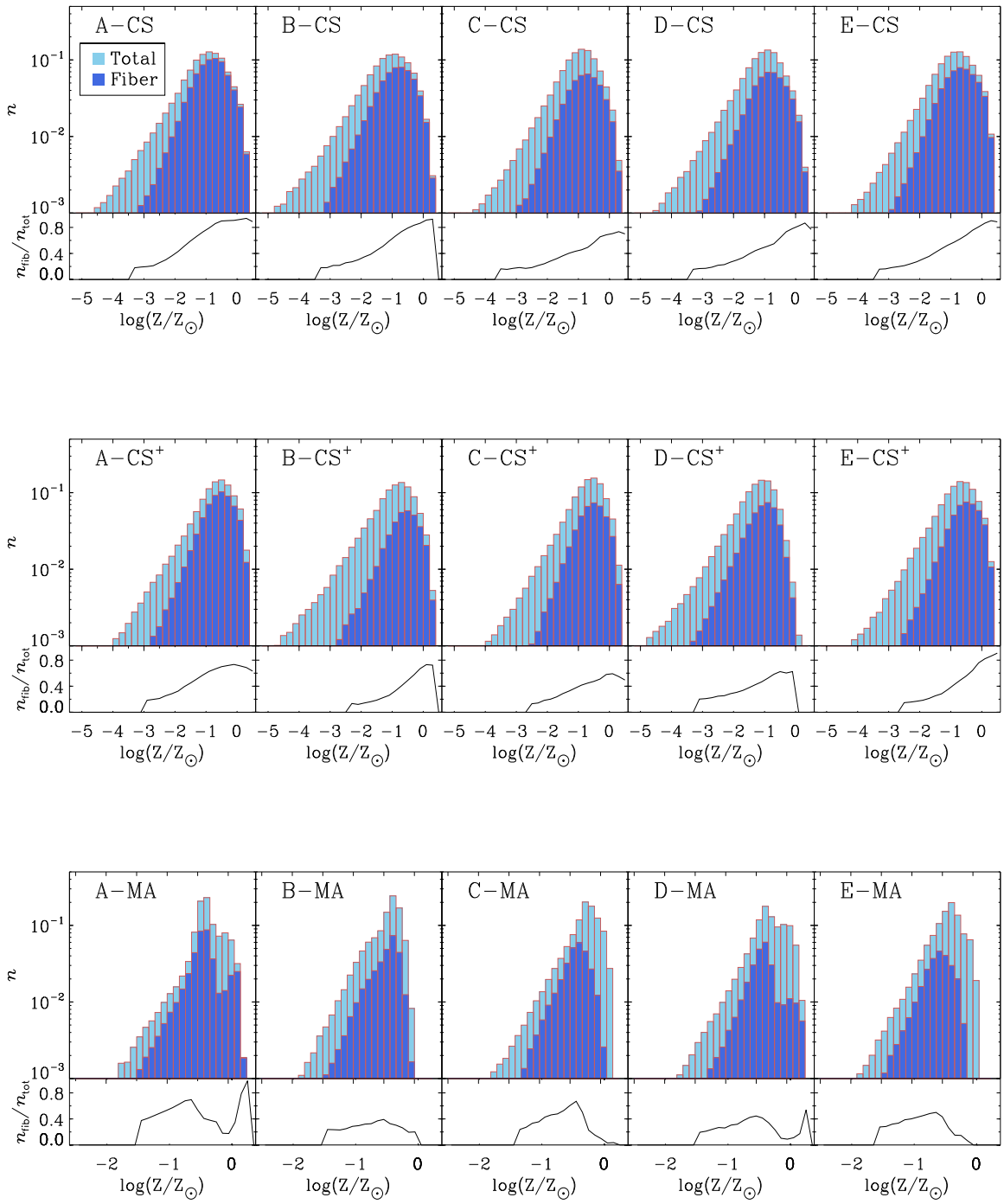

Figure 3.4: Number-density histograms of stellar metallicity (in solar units, assuming $Z_\odot = 0.02$), both for all star particles and inside the fiber field of view. For each galaxy, we also show the fraction between the number of stars within the fiber to the total number of stars in each bin, $n_{\rm fib}/n_{\rm tot}$, that we refer to as the sampling function.

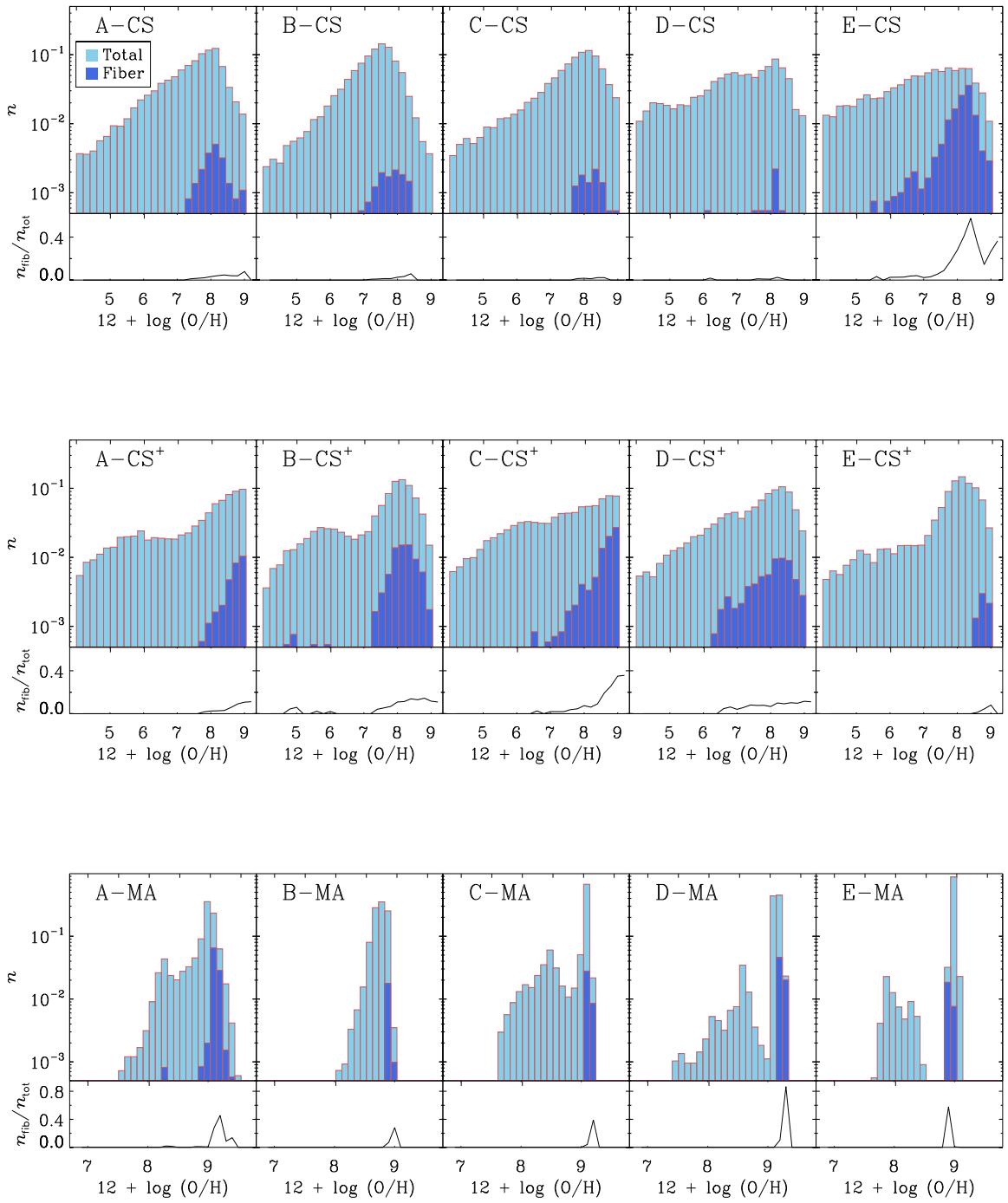

Figure 3.5: Number-density histograms of gas metallicity, both for all star particles and inside the fiber field of view. For each galaxy, we also show the fraction between the number of stars within the fiber to the total number of stars in each bin, $n_{\rm fib}/n_{\rm tot}$, that we refer to as the sampling function. Note the different x-range plotted for the CS/CS$^+$ and MA samples.

$z \sim 0.15$ for the SDSS instrument aperture. Note that we are limited to further decrease the size of the fiber field of view by having enough particles to reach a good statistical sample, in particular in the case of gas particles.

Figs. 3.3, 3.4 and 3.5 show the number-density histograms of stellar ages and stellar/gas metallicities for the 15 simulated galaxies, when we consider all particles in the galaxy and only those within the fiber. For each galaxy, we also show the fraction between the number of star/gas particles within the fiber to the total number of star/gas particles in each bin, $n_{\rm fib}/n_{\rm tot}$, that we refer to as the sampling function. In the case of stellar ages we find that, in general, fiber quantities preferentially sample the older populations, with sampling functions typically higher than 50%. In contrast, the young populations are sampled in various ways, depending on the age profile of the galaxy. In some cases, young populations are not at all sampled and will be completely missed in the calculations using particles within the fiber. It is clear, in any case, that the *sampling within the fiber is not constant or even similar for all galaxies, which means that it is not possible to reliable estimate the mean stellar age of a galaxy from the fiber quantities, or even estimate the fiber bias without having additional information (e.g. importance of age/metallicity profiles) of the regions outside the fiber.* We note that the relative contribution of old and young populations in a galaxy will in general reflect the relative importance of bulges and disks, but with fiber quantities this information might not be a reliable reflection of the real relative contribution of the different stellar components of a galaxy.

Similar considerations can be made in the case of the stellar metallicities, shown in Fig. 3.4. The metallicity distributions of the galaxies in the CS sample is broader compared to those in the MA sample, with the CS$^+$ galaxies lying in between. This results from the different implementation of chemical enrichment (chemical yields increase from the CS galaxies, to CS$^+$ and then to MA), chemical diffusion (only included in the MA simulations) and feedback (from weaker in CS to stronger in MA, with CS$^+$ in between). In particular, the MA galaxies have more peaked distributions, and have less strong metallicity gradients (see also Aumer et al. 2013) compared to the CS and CS$^+$ galaxies (see also Tissera, White & Scannapieco 2012). Unlike for the stellar ages, *for the stellar metallicities we find that the sampling shows much less variation from galaxy to galaxy, as in all cases there is a preferential sampling of the more metal-rich populations.*

The distributions of gas metallicities are more complex, and show important differences depending on the details of the implementation of chemical enrichment (IMF, chemical yields, etc.) and on the absence/presence of chemical diffusion, which is included only in the MA sample and leads to a much less broad distribution of gas metallicities (note the different x-scales in the lower panels). From Fig. 3.5, it is clear that the sampling of gas metallicities varies significantly from galaxy to galaxy. While in general there is a preferential sampling of the metal-rich regions, gas particles with intermediate metallicities can also contribute significantly to the final metallicity when fiber quantities are considered. For the MA sample, we find that, within the fiber, only a narrow range of gas metallicities are sampled, which reflects the effects of metal diffusion that tends to give a smoother metallicity distribution. In the case of the D-CS galaxy, we also find that although the metallicity distribution is broad, only a very narrow range of metallicities are sampled within the fiber. According to our findings, *for estimating the gas metallicy of a galaxy it is of primary importance that the fiber bias can be quantified; otherwise it is not possible to derive the mean gas metallicity of the whole galaxy reliably.*

In summary, our results show that *the bias due to the fiber is in general reflected in the tendency to sample older and more metal-rich stellar populations and metal-enriched regions of the ISM, although the shape of the sampling function highly depends on the significance*

of gradients in ages/metallicities (which also somewhat depend on the modelling of physical processes in the simulations), *affecting in particular galaxies with stronger gradients.*

Variation of the assumed Initial Mass Function

The choice of the Initial Mass Function (IMF) can strongly bias the derivation of many observables of galaxies (see Salpeter, 1955; Miller & Scalo, 1979; Scalo, 1986; Kroupa, 2002), for example, the assumed IMF has a direct influence on the M/L ratio depending on the age, composition and past star formation history (Baldry & Glazebrook, 2003; Chabrier, 2003), which in turn influences the stellar mass derivation. The IMF is measured in many environments, still there is not yet consensus on a number of important aspects, such as whether it is universal or not (e.g. Hoversten & Glazebrook, 2008; Bastian, Covey & Meyer, 2010).

We assume in general a Kroupa IMF with slope $\alpha = 1.3$ for masses 0.1-0.5 $M_{\odot}$ and $\alpha = 2.3$ in the mass range 0.5-100 $M_{\odot}$ when we apply the different SPS models, except in the case of BC03, where we instead use a Chabrier IMF (Table 1). The difference between these two IMFs is moderate, however, due to their very similar shape, particularly at the high-mass end (see Fig. 17 in Ellis 2008). Note that our three galaxy samples use different IMFs to calculate feedback and chemical enrichment: Salpeter (CS), Chabrier (CS^{+}) and Kroupa (MA).

3.3 Results

In this section we describe the methods we use to derive the galaxy properties – magnitudes and colors, stellar masses, stellar ages, stellar and gas metallicities and star formation rates – from their synthetic spectra, and make a detailed comparison between results obtained with the different methods, in order to identify biases and systematics. For each property, different techniques are used in the derivation, the selection of which has been made taking into account the questions we want to answer: (*i*) what is the range of variation in observationally-derived quantities when different techniques are used, (*ii*) do they agree with the 'real' quantity; and (*iii*) how do observational definitions of galaxy properties effect their measured values. In the case of observations, we particularly focus on the techniques employed in SDSS, which will be used in the next chapter to compare simulated and observed galaxies.

In order to avoid additional biases, we always use the same field of view defined as a 60 kpc×60 kpc region with the galaxy in the center both when we derive the properties directly from the simulations and from the SPS/SUNRISE spectra.

3.3.1 Magnitudes and colors

We first compare the colors and the absolute magnitudes of our simulated galaxies, in the 5 SDSS photometric bands (u, g, r, i, z, see Fig. 3.2), obtained using different methods, namely:

- **BC03, SB99, FSPS, PE, M05**: these refer to the magnitudes calculated by applying the five different SPS models described in Section 3.2.1. As explained in the previous section, for each star particle the magnitudes are obtained via a linear interpolation of the SPS tables according to its age and metallicity, normalizing with the particle mass *at the formation time.*

- **CF00**: we include dust effects to BC03 using the model of CF00/dC08 described in Section 3.2.2.

- **SR$_{\rm noISM}$, SR$_{\rm faceon}$, SR$_{\rm edgeon}$**: results from the radiative transfer code SUNRISE as explained in Section 3.2.3, in the absence of dust[2], and for the face-on and edge-on views including dust, respectively. Note that the input SPS model for SUNRISE is SB99.

The left-hand panel of Fig. 3.6 compares the magnitudes obtained using the different SPS methods and SUNRISE, when we ignore the effects of dust. We show the differences with respect to BC03[3] as a function of the stellar metallicity of the galaxy[4] (in solar units, assuming $Z_{\odot} = 0.02$), as systematics in SPS models are expected to increase both at low/high metallicity (as well as at younger stellar ages, e.g., Conroy, Gunn & White 2009; Conroy & Gunn 2010). The different SPS models show in general very good agreement, with magnitude differences of $\lesssim 0.1$. We detect however some systematics: models FSPS and M05 usually predict lower magnitudes (i.e. brighter galaxies) compared to BC03, while models SB99 and PE give systematically higher magnitudes than BC03. As expected, differences are somewhat larger (but still moderate) for the most metal-rich galaxies, which are also those that exhibit a higher fraction of young (ages < 10 Myr) and intermediate (ages in the range $[0.1 - 2]$ Gyr) stellar populations. For these ages the uncertainties in the treatment of young stars, post red-giant branch and TP-AGB phase stars are larger. The galaxy which exhibits the largest ΔM (in all bands) is A-MA, which has $\log(Z/Z_{\odot}) \approx -0.26$ the most extreme SFR (Section 3.3.5) and the youngest mean stellar age in our sample (see next sections).

The results of SR$_{\rm noISM}$ also agree well with the rest of the models, particularly with those of SB99, as SB99 is the input SPS model of SUNRISE. The remaining differences are because in SUNRISE the spectrum for young stars is calculated with MAPPINGS III, which includes the contribution from nebular continuum and emission as well as dust absorption and IR emission that is modeled as sub-grid physics. The largest differences are found for A-MA, the galaxy with the highest number of young star particles. In fact, on one side the ionizing photons are reprocessed as nebular emission lines and continuum increasing the total flux in the optical and, on the other hand, the young particles are also more extincted due to the sub-grid treatment of dust absorption in MAPPINGS III, giving total magnitudes in general lower (brighter) by < 0.1 mag than modelling the SED including only stellar emission.

The right-hand panel of Fig. 3.6 shows our results in the case of including the effects of dust, either with the CF00/dC08 correction or, more consistently, with the SUNRISE code. For comparison, the magnitude differences are calculated with respect to the BC03 model. When the CF00 simple dust model is included we find, as expected, fainter galaxies, with differences between $\sim 0.6 - 0.8$ mag for the u and g-bands and of $\sim 0.4 - 0.5$ mag for i, r and z. As the CF00 model is angle-averaged, there is no difference between face-on and edge-on views. When we consider radiative transfer effects in SUNRISE, we detect almost no difference if we see the galaxies face-on (SR$_{\rm faceon}$) or if dust is ignored (SR$_{\rm noISM}$ method, included here for comparison).

As expected, larger differences are detected when galaxies are seen edge-on, as dust effects are maximal in this case. However, we find significant differences only for metal-rich galaxies, while for galaxies that are metal-poor ($\log(Z/Z_{\odot}) \lesssim$-0.45) dust effects are unimportant. This is because the amount of dust in SUNRISE is directly proportional to the gas metallicity (galaxies with low stellar metallicity also have low gas metallicity). Observations

[2]We note, however, that dust is included through the sub-grid dust model around young stars in MAPPINGS.

[3]We note that, even if BC03 assumes a different IMF (Chabrier) compared to the other SPS models, this does not introduce any significant systematic effect in the derived magnitudes/colors, since for the more luminous stars ($M_* > 1 M_{\odot}$) the IMF slope is the same as Kroupa (see Table 1).

[4]This is calculated as the average mass-weighted metallicity over all stellar particles in the simulated galaxy (method SIM in Section 3.3.3).

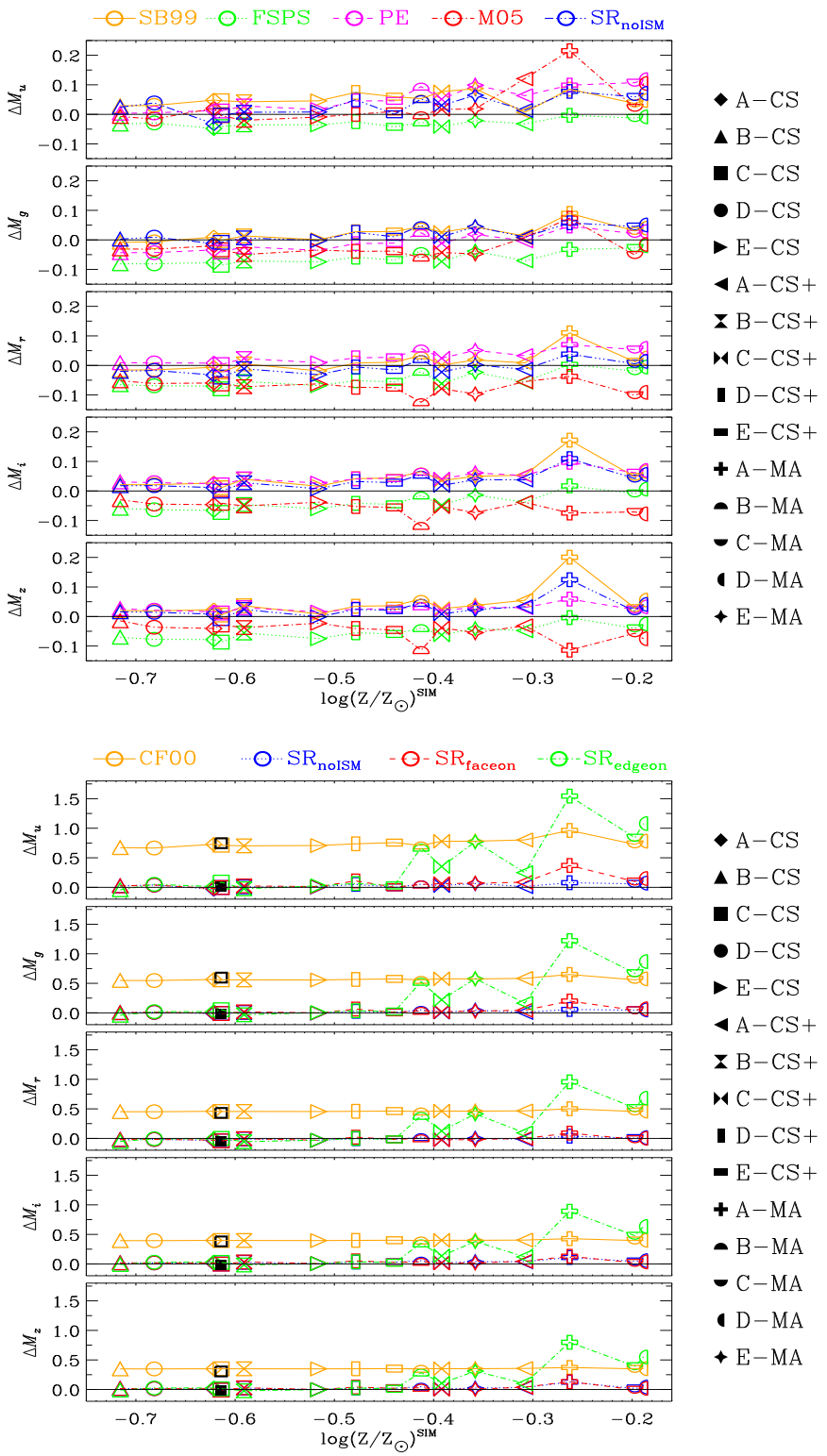

Figure 3.6: Difference between the absolute magnitudes of the simulated galaxies obtained using different methods and those predicted by the dust-free BC03 model, in the 5 SDSS bands and as a function of stellar metallicity. The *left*-hand panel shows results for dust-free models and the *right*-hand panel compares results of the CF00 simple dust model and of the SUNRISE code. The open/filled black squares are results for C-CS, when a higher dust-to-metals ratio is assumed in the SUNRISE calculations in edge-on and face-on views, respectively.

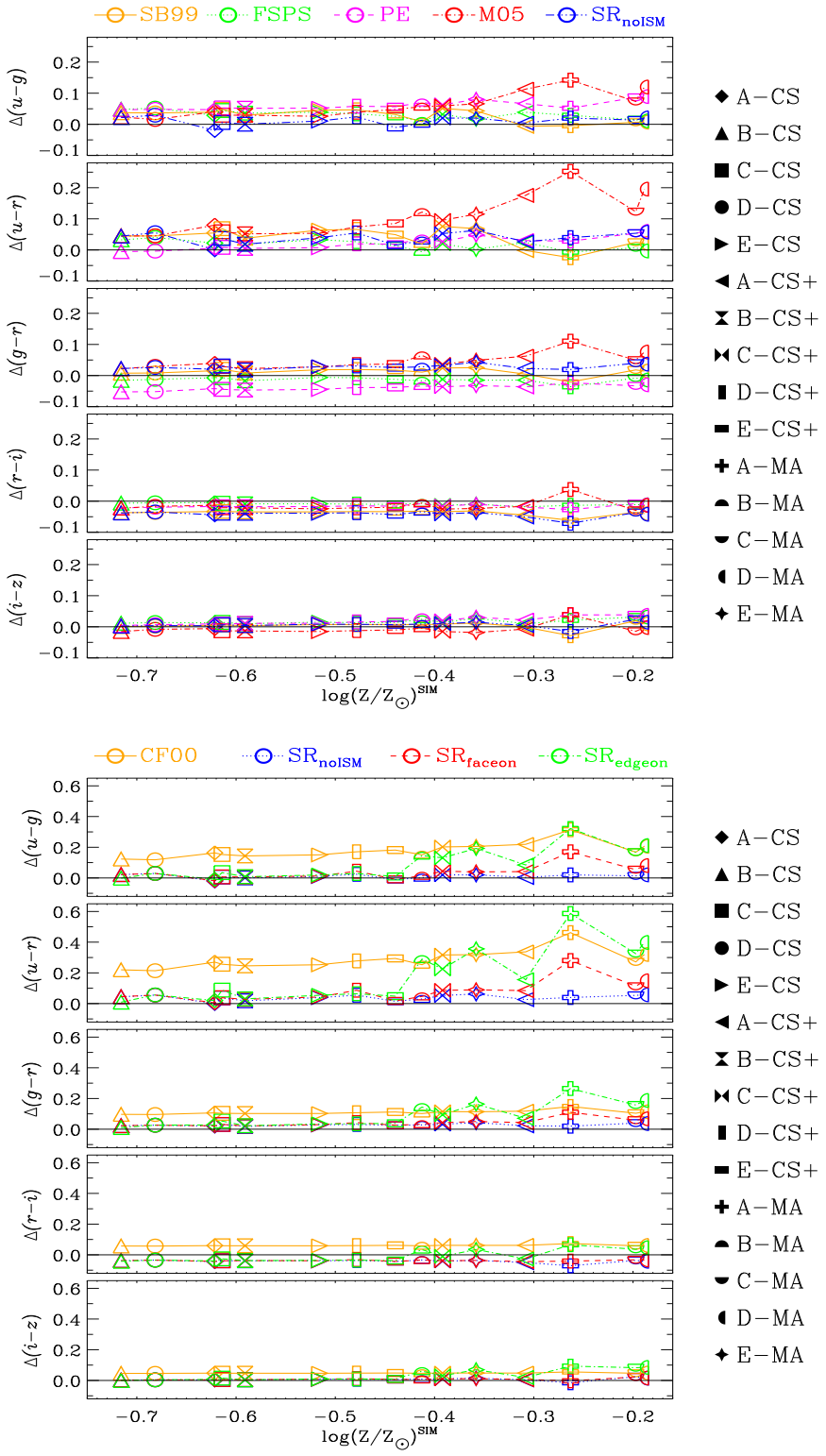

Figure 3.7: Difference between the colors of the simulated galaxies obtained using different methods and those predicted by the dust-free BC03 model, as a function of stellar metallicity. The *left*-hand panel shows results for dust-free models and the *right*-hand panel compares results of the CF00 dust model and of the SUNRISE code.

also confirm that the dust-to-metals ratio is nearly constant over a large range of metallicities and redshifts, and dust effects are small in metal-poor galaxies (Zafar & Watson 2013, Mattsson et al. 2014). For the metal-rich galaxies, the inclusion of dust makes them fainter by factors of $\sim 0.8-1$ dex in the u-band and of $\sim 0.3-0.5$ dex for the z-band. Interestingly, for the most metal-rich galaxies the magnitudes derived from the $\mathrm{SR_{edgeon}}$ and CF00 methods are very similar. Note also that the $\mathrm{CS^{+}}$ galaxies have a systematically lower reddenning compared to the MA ones, even at high metallicity. This is explained by the lower amount of metals in the ISM, and therefore lower amount of dust of the galaxies in the $\mathrm{CS^{+}}$ sample compared to those in MA.

To explore the dependence of dust effects on metallicity for the metal-poor sample, we rerun SUNRISE for one galaxy (C-CS), using a higher dust-to-metals ratio of 4 (10 times larger than the standard value of 0.4 assumed in our SUNRISE calculations, Dwek 1998). The results are shown as open and filled squares, respectively for the edge-on and face-on views. In this case, we indeed find little effects in the case of the face-on view and larger differences in the edge-on case, with results very similar to those of CF00.

The differences in magnitudes, when different methods are applied, are translated into differences in the colors, as shown in Fig. 3.7. Both in the absence (left-hand panel) and in the presence (right-hand panel) of dust, differences are larger for the bluer colors ($u-g$, $u-r$ and $g-r$), and for the most metal-rich galaxies.

Our results show that *applying different SPS to simulations gives visual magnitudes of galaxies with spread* $0.05-0.1$ *dex for metal-poor galaxies and of* $0.1-0.2$ *for metal-rich ones*, depending on the model, *with BC03 appearing as intermediate among the 5 SPS models tested.* Furthermore, *while for face-on galaxies or edge-on galaxies with low metallicities* ($\log(Z/Z_{\odot}) \lesssim -0.45$) *the effects of dust can be in first approximation ignored, when galaxies are seen edge-on and have a significant amount of dust (i.e. are metal-rich), errors in the magnitudes can be up to* $0.7-1.5$ *dex if dust is ignored*, depending on the band. This means that for edge-on galaxies, particularly metal-rich ones, reliable magnitudes will not be possible to obtain without a proper modelling of the dust extinction. Similar considerations can be made for the colors, that can not be calculated with precision better than $0.1-0.5$ dex.

The previous figures showed the dependence of the galaxies' magnitudes on different models and assumptions. In simulation studies, it is common to use SPS models to convert masses into luminosities, and to compare models with observational data (e.g. light profiles, magnitudes, etc). We have shown that the use of different SPS models introduces some changes in the predicted magnitudes, that are however moderate. On the other hand, when dust effects are included or a proper radiative transfer treatment is considered, larger differences might appear. In the left-hand panel of Fig. 3.8 we show the position of our simulated galaxies in the color-magnitude diagram; in this case in the M_r vs $(u-r)$ plane. We show the results for BC03, the spread found for all SPS models in both plotted quantities, and the results for the BC03 model with the CF00 correction. We also show results for $\mathrm{SR_{noISM}}$, $\mathrm{SR_{faceon}}$ and $\mathrm{SR_{edgeon}}$. These cover the commonly used ways to post-process simulation results, and allow to understand what are the possible offsets expected when more sophisticated methods are used to calculate the magnitudes. Most of our simulated galaxies have M_r between -22 and -19, and $(u-r)$ color between ~ 1.5 (except for A-MA) and ~ 2.3. The A-MA galaxy lies at a different position compared to the other galaxies, particularly when dust effects are ignored or when it is included but the galaxy is seen face-on. Note that this is a young galaxy with strong emission lines, and the Hα line falls in the r-filter at $z=0$. A-MA however moves photometrically closer to the other galaxies if dust extinction is included.

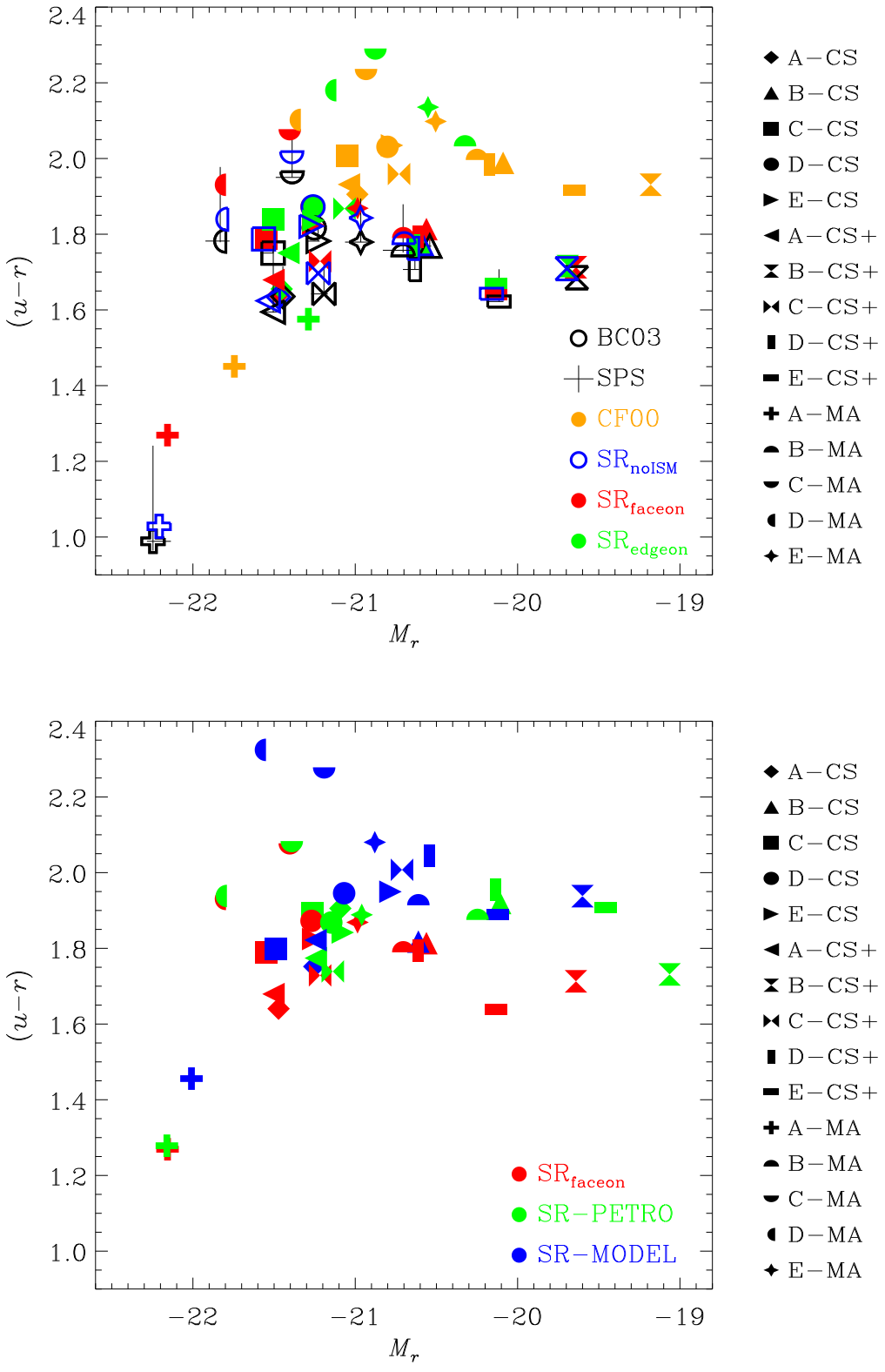

Figure 3.8: Color-magnitude diagram of the simulated galaxies, obtained using different methods to derive the synthetic spectra. The left-hand panel show results for the SPS, CF00 and SUNRISE methods and the right-hand panel compares results obtained following the SDSS techniques.

In observations, measuring the total magnitudes of galaxies can be plagued by different observational problems (e.g. low signal-to-noise, sky brightness, bad calibrations and sky subtraction), which affect particularly the outer part of the profiles where the S/N is lower. For this reason, definitions on how to measure the total light are used in all galaxy surveys, although they can differ substantially among authors and collaborations (e.g. Kron 1980, Blanton et al. 2001, Norberg et al. 2002). We have calculated the magnitudes and colors of our simulated galaxies following the techniques used in SDSS, in particular calculating the so-called Petrosian and Model magnitudes (Petrosian, 1976; Blanton et al., 2001; Yasuda et al., 2001; Kauffmann et al., 2003; Salim et al., 2007). These were derived from the images obtained with SUNRISE, including dust and in the face-on view (SR_{faceon}), as follows (Blanton et al., 2001):

- **PETRO**: we calculate the Petrosian Radius (Petrosian, 1976) in the r-band, and we derive the Petrosian magnitudes in all bands taking the flux inside $N_P = 2$ Petrosian radii; as in SDSS, we assume a Petrosian Ratio $R_P = 0.2$. Note that, in general, this method samples only part of the total flux, the fraction depending on the luminosity profile (see Graham et al. 2005 for detailed calculations).

- **MODEL**: in this case, each image is matched to different luminosity profiles, and the magnitudes are calculated from the profile which gives the best fit. We used the code GALFIT to perform the fit (Peng et al., 2010), assuming arbitrary axis ratio and position angle, and used two different profiles: Exponential and DeVaucouleur. Depending on the galaxy, one of these profiles provided the best fit: for 5 galaxies this was an exponential, for 10 a DeVaucouleur profile.

The right-hand panel of Fig. 3.8 compares the results obtained with the PETRO and MODEL methods to those of SR_{faceon}. We find in general similar values of magnitudes for the PETRO and SR_{faceon} methods, with PETRO magnitudes always higher (i.e. fainter) than those of SR_{faceon} as expected. In the case of the MODEL method, we also find fainter galaxies compared to the results of SR_{faceon}. In both cases, differences are lower than 0.3 dex for 9/15 galaxies, with a maximum of 0.6 dex for the remaining 6 systems. Colors are similar in the three models, with differences in $(u-r) \lesssim 0.1-0.2$ dex. Our results show that, for our simulated galaxies the *PETRO and MODEL methods can not recover the real magnitudes/colors of galaxies with a precision better than* $0.2-0.3$ *dex.*

3.3.2 Stellar mass

The stellar mass of galaxies is an important proxy of how the galaxy populations evolved over cosmic time (e.g. Bell et al., 2003). In this section, we compare the stellar mass of our simulated galaxies, obtained in different ways, including the simple sum over the mass of stellar particles (as done in simulation studies) and those obtained using different post-processing techniques that mimic observations, in particular the ones used in SDSS. These are described below:

- **SIM**: the total mass of star particles in the simulated galaxy. We include all stars in the 60 kpc×60 kpc field of view (see sec. 3.3), i.e., the same field of view of SUNRISE.

- **BC03**: the final mass of star particles is calculated with the BC03 model, which considers the mass lost by a stellar population since it was formed, and until the present time. The BC03 final mass of each star particle is obtained normalizing with its initial mass. The total stellar mass of the galaxy is the sum over the final masses of the star particles within the field of view.

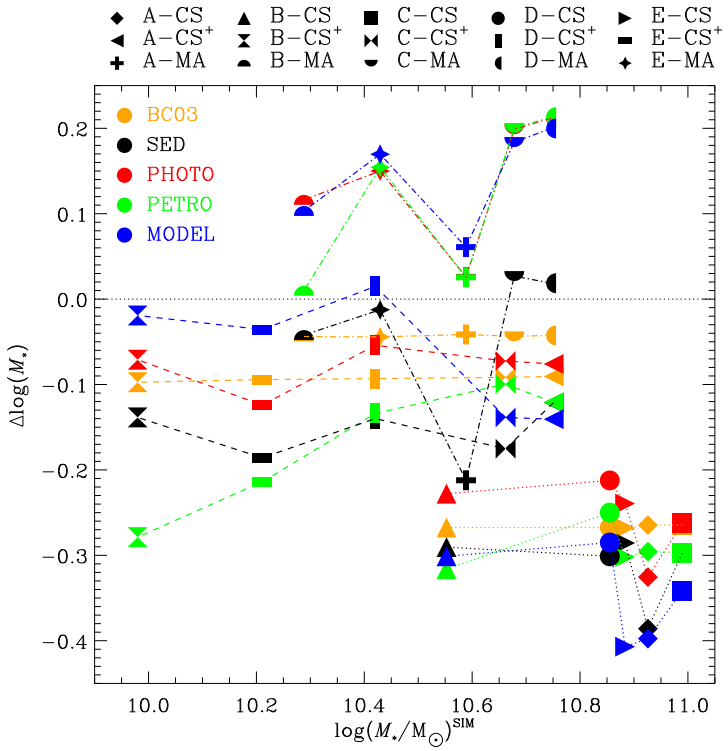

Figure 3.9: Comparison between different estimators of the stellar mass of galaxies, with respect to the real stellar mass, calculated as the simple sum of the mass of star particles ($M_*^{\rm SIM}$). Dashed and dot-dashed lines indicate simulations with a proper accounting of stellar mass loss, while dotted lines are for simulations where mass loss is not consistently followed.

- **SED**: the total stellar mass is estimated by fitting the SUNRISE face-on optical galaxy spectrum ($\rm SR_{faceon}$ method) over the range $3800 - 9000$ Å using the STARLIGHT code (Cid Fernandes et al., 2005, 2009) with the BC03 SPS model. The spectra include nebular emission (masked during the fit) and are dust-extincted.

- **PHOTO**: we fit the photometric (u, g, r, i, z)-band magnitudes obtained from the $\rm SR_{faceon}$ spectrum to a grid of models as updated from BC03 in 2007 (CB07, Charlot & Bruzual, priv. comm.); the grid spans a large range in galaxy star formation histories, ages and metallicities (see Walcher et al. 2008 for a description of the code). To remove the nebular contribution from the broad-band magnitudes (the fitted model considers only stellar light), we calculate the relative contribution of nebular emission within the fiber in each photometric band fitting the fiber spectrum with the STARLIGHT code, and assume that the relative contribution of nebular emission for the total galaxy is the same as in the fiber. This procedure allows one to mimic the stellar mass estimation of the Garching SDSS DR7.

- **PETRO**: Petrosian stellar masses are derived as in PHOTO-MASS, but using the Petrosian magnitudes (Section 3.3.1) as input for the fit (i.e. within 2 times the Petrosian radius). The same procedure to remove nebular emission as in PHOTO-MASS was applied in this case.

- **MODEL**: The same as PETRO-MASS, but using the Model magnitudes (Section 3.3.1).

In order to assess the effects of observational biases on the determination of the stellar mass of galaxies, it is important to consider the effects of mass loss of stellar particles in the simulations, as for all observationally-derived quantities, the fitting of spectra/magnitudes is done using the SPS models that include a mass loss prescription. Our 15 simulations cover the cases where mass loss is not followed, and where it is properly treated, allowing one to assess how important this effect can be. Moreover, simulations assume sometimes choices for the IMF that are different than those used in the SPS models. It is for these reasons that we calculated the stellar mass of simulated galaxies using the BC03 method additionally to the direct result of the simulation (SIM method).

In Fig. 3.9 we compare the stellar masses of simulated galaxies, obtained with the different methods, to the stellar mass obtained directly from the simulations (SIM). Also shown is the 1-to-1 relation (solid black line). In the lower panel, dotted, dashed and dot-dashed lines indicate respectively galaxies in the CS/CS$^+$/MA samples, that differ in the treatment of mass loss of stars. We find that for galaxies where mass loss was not properly treated (CS), the stellar masses obtained with all methods are systematically lower compared to the direct result of the simulations. Furthermore, the offset is very similar for all galaxies, of the order of 0.3 dex in $\log(M_*)$. The results of the PHOTO, PETRO and MODEL methods are similar, with the former giving systematically higher stellar masses, and the latter predicting the lowest values of stellar masses.

In contrast, for galaxies in the simulations with a proper mass loss treatment (CS$^+$/MA), the stellar masses obtained with the BC03 shows better agreement with the direct result (SIM), with differences up to ~ 0.1 dex. The SED method predicts slightly lower stellar masses, with typical differences of the order of $0.05 - 0.2$ dex. In the case of the PHOTO, PETRO and MODEL methods, stellar masses have some scatter up to $\sim 0.1 - 0.2$ dex in $\log(M_*)$, and are systematically lower for CS$^+$ (dashed lines) and higher for MA (dot-dashed lines) compared to the direct result of simulations.

The results of this section show that *the observational estimators are able to recover the mass up to 0.1-0.2 dex in logarithmic scale, only in simulations where the mass loss of stars is consistently included. In contrast, simulations with no treatment of mass loss can still be compared with observations in a meaningful manner after a fixed correction (at least for the mass range of our sample) is applied.*

3.3.3 Stellar ages and metallicities

The determination of stellar metallicities and ages of galaxies from observational data is usually accomplished by fitting either selected sensitive absorption line indices (e.g. Lick indices, Worthey et al. 1994, Trager et al. 1998) derived from high S/N spectra to a grid of SPS models (see Trager et al. 2000a,b; Gallazzi et al. 2005, 2006) or the full available spectrum (λ-to-λ method, e.g. Cid Fernandes et al. 2005; Tojeiro et al. 2007, 2009; Chen et al. 2010), which also requires high quality and S/N spectra for a reliable fit. Some authors also exploit colors to estimate ages and metallicities, especially at high redshift (e.g. Lee et al., 2010; Pforr, Maraston & Tonini, 2012; Li, 2013).

In this Section, we compare the stellar ages and metallicities of the simulated galaxies obtained in different ways. In order to assess the effects of the fiber bias, for all methods we consider stars within the same field of view than in the previous section, i.e. in a projected area of $(60\text{kpc})^2$, and also only those within the fiber. The different methods are as follows:

- **SIM**: the direct result of the simulations, i.e., the mean stellar age/metallicity of a galaxy is calculated as the corresponding mass-weighted mean over the stellar particles.

This requires no post-processing or additional assumptions, and is the most common way to derive ages/metallicities from simulations.

- **SIM-LUM**: we weight with the luminosity of stellar particles to calculate the average ages/metallicities, which more closely reflects what would be obtained in an observation; we use the r-band particle luminosity for stellar ages and the particles' luminosity in all visible bands for the metallicities (both obtained with the BC03 model) to mimic the SDSS analysis (Gallazzi et al. 2005).

- **SED-FIT**: we fit the spectra of simulated galaxies obtained with SUNRISE using the STARLIGHT code (Cid Fernandes et al., 2005), selecting the best mixture of ~ 300 instantaneous-burst SSPs with different ages and metallicities. The fitted spectra correspond to the face-on views and include dust and nebular emission ($\mathrm{SR_{faceon}}$ method). From the fitted mixture of SSPs we calculate the mean (optical) light-weighted ages and metallicities.

- **LICK-IND**: the mean ages and metallicities were computed by Anna Gallazzi (priv. comm.) for a subsample of 10/15 galaxies using the method described in Gallazzi et al. (2005). The method is based on fitting five absorption features (Lick indices D4000n, Hβ, Hδ_A+Hγ_A, [$\mathrm{Mg_2}$Fe], [MgFe]) to libraries constructed with the BC03 model with different star formation history, metallicity and velocity dispersion, deriving the (optical luminosity-weighted) stellar metallicities and (r-band weighted) mean ages. As the spectra generated with SUNRISE have the resolution of the input stellar model (SB99) with spacing 20 Å (which is too low to compute meaningful absorption indices), we thus explore an alternative way to measure the Lick indices of the simulated galaxies directly from the BC03 tables, after interpolating with age and metallicity for each star particle and averaging weighting with the particle mass. However, we note that these may differ from the indices measured directly from the spectrum, as done in observations and in the model library of Gallazzi et al. used to interpret observations. For this reason, this method is not fully consistent with the one used in SDSS. We will explore this point in subsequent analysis with higher-resolution spectra.

The upper panel of Fig. 3.10 shows the galaxies' stellar ages obtained with the various methods described above as a function of the age obtained directly from the simulations (i.e. with the SIM method, $\mathrm{Age^{SIM}}$), together with the 1:1 relation, while the lower panel shows the corresponding differences. Results are shown in the cases of considering all stars in the field of view and stars within the fiber.

In general, when all stars in the field of view are considered (filled symbols and solid lines), the ages obtained using the SIM-LUM and SED-FIT methods are similar, and are systematically lower than the simple average of the age of stellar particles. This results from the fact that young stellar populations emit more light than old ones, therefore have a comparatively higher weight in SIM-LUM and SED-FIT. Typical variations are of the order of $\lesssim 2$ Gyr, but the discrepancies are larger for some systems, with a maximum of about ~ 5 Gyr for the youngest galaxy. The estimation with the (mass-weighted) Lick indices gives systematically older ages for young galaxies ($\mathrm{Age^{SIM}} < 8$ Gyr) compared to SIM-LUM and SED-FIT, and slightly older ages compared to SIM, while it is in the range of the other observational estimators for the older galaxies.

Contrary to what we found for the ages estimated from all stars, in the case of the fiber quantities (open symbols and dotted lines) we find systematically higher ages compare to $\mathrm{Age^{SIM}}$, with differences up to ~ 4 Gyr. Note that the differences depend sensitively on

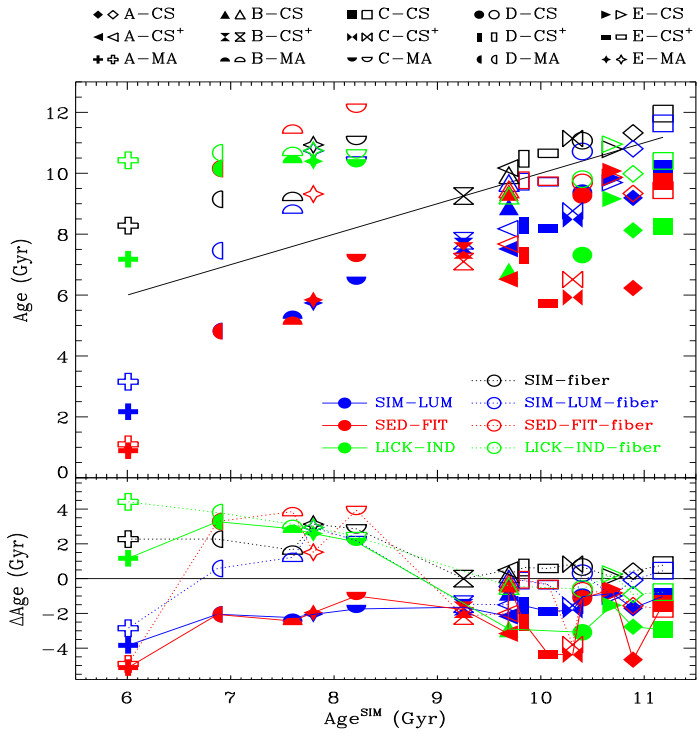

Figure 3.10: Difference between the mean mass-weighted stellar age in simulations and the ages calculated with the observational methods. The galaxies are sorted by increasing mean age to the right.

the age but also on the presence/absence of age gradients which, as shown in Fig. 3.3, can vary significantly from galaxy to galaxy. For old galaxies, all methods, in general, agree better with each other. We note that, for quantities within the fiber, the two sources of bias (preferential sampling of inner populations and dependence on age gradients) can lead to both positive and negative differences with respect to the direct result of the simulation. The LICK-IND method taking only particles in the fiber (LICK-IND fiber) gives in general slightly older ages compared to LICK-IND.

We also applied the three methods and calculated total and fiber quantities for the stellar metallicities. Fig. 3.11 shows a comparison of results. Compared to the direct result of the simulation, the SIM-LUM and SED-FIT methods give in general lower metallicities for metal-poor (old) galaxies and higher metallicities for more metal-rich (younger) ones. This is again explained by the different relative weight of old and young stars to the average metallicity (note that the spread in stellar metallicities is much smaller than that for stellar ages). Differences are however moderate, always of the order of ± 0.3 dex.

For quantities within the fiber, we also find moderate differences; in this case, metallicities tend to be higher compared to total quantities, as the contribution of very low metallicity stellar populations gets smaller (Fig. 3.4). When we apply the LICK-IND and LICK-IND fiber methods, we find, for the low-metal sample ($\log(Z/Z_{\odot}) \lesssim -0.45$), systematically higher metallicities, while for metal-rich galaxies we find better agreement with the other indicators. We detect small differences between LICK-IND and LICK-IND fiber, the latter giving in general slightly higher metallicities. It is notable that for the metal-poor galaxies, most observational methods disagree with the direct result of the simulations by an approximately constant factor, instead of scaling with the metallicity.

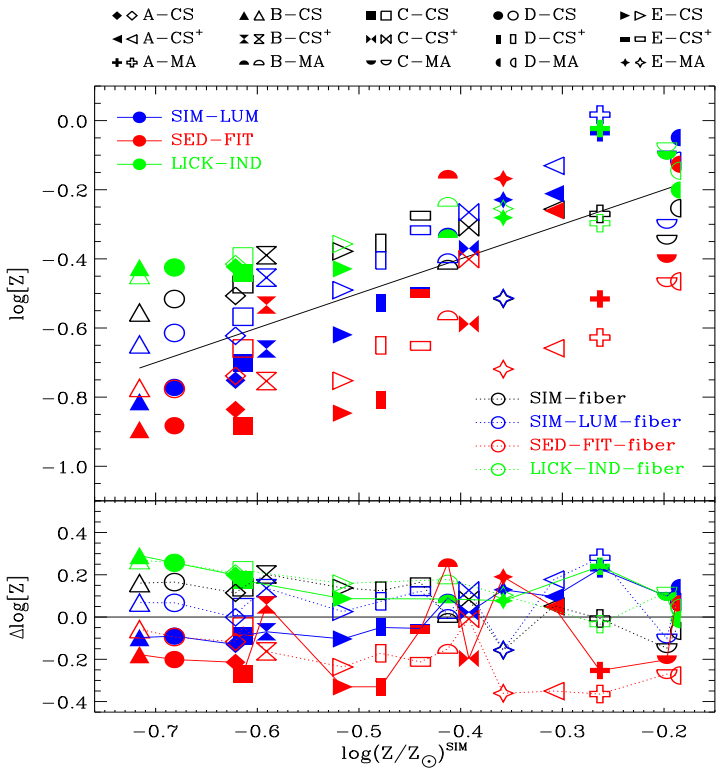

Figure 3.11: Difference between the mean mass-weighted stellar metallicity and the observational values, ordering the galaxies by increasing stellar metallicity to the right.

Our results show that in our simulations it is not possible to estimate the mean stellar age with accuracy better than ~ 3 Gyr, and the stellar metallicities better than ~ 0.3 dex, depending on the method and on the way averaged quantities are calculated. The fiber bias has a strong effect on the derived ages/metallicities, due to the preferential sampling of old and metal-rich stars; our 15 galaxies show wide variety of gradients, and therefore we find that the fiber bias when applying different methods can significantly change from galaxy to galaxy.

3.3.4 Gas metallicity

In this section, we compare the gas oxygen abundance of our simulated galaxies, using different estimators. Observationally, the determination of the chemical composition of the gas in galaxies is based on metallicity-sensitive emission line ratios. A number of different metallicity calibrations are used to estimate the O/H ratio in nebulae and galaxies, and comparisons among the results of various calibrations reveal large discrepancy with systematic offsets that can reach ~ 0.7 dex (Pilyugin 2001, see also Appendix 3.5).

As in the previous sections, we calculate the gas chemical composition following the most common methods from simulations and observations, and, when appropriate, considering all gas particles and only those within the fiber (labelled as "fiber"). When we use (face-on) SUNRISE spectra for the calculation (T04, KK04, T_e), we first Balmer-correct the emission line ratios for dust extinction using the Calzetti law (Calzetti, Kinney & Storchi-Bergmann, 1994). The methods are described in the following.

- **SIM**: the gas metallicity of a galaxy is calculated as the mean (O/H) of gas particles,

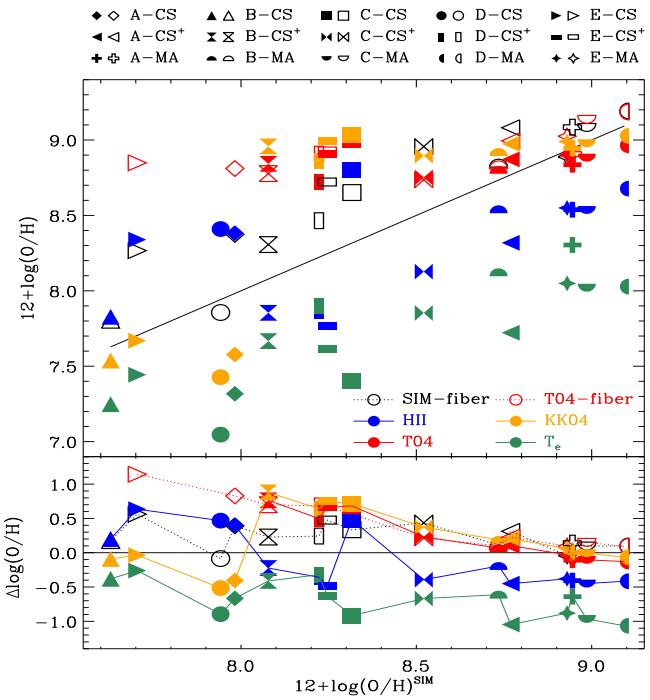

Figure 3.12: Comparison between the different estimators for the the galaxies' oxygen gas metallicities, as a function of that obtained with the SIM method. For reference, in the upper panel we include the 1:1 relation as a solid line. As explained in the text, T04 calibration is valid only for a subsample of the galaxies.

weighted by their mass.

- **HII**: is the mean mass-weighted O/H of gas particles, but only considering particles around stars younger than 20 Myr and inside 1 kpc radius. In this way, we mimic the preferential bias to young stellar populations in nebulae-based measurements. Note that although the HII regions have sizes which can vary in a wide range of $0.1 - 200$ pc from ultra-compact to giant extragalactic HII regions, we are limited to define smaller sizes of the HII regions by the resolution of the simulations, which is of the order of ~ 1 kpc.

- **T04** (Tremonti et al., 2004): the gas metallicity is computed using the calibration of the R_{23}-upper branch given in Tremonti et al. (2004). According to Kewley & Ellison (2008) we use [NII]/[OII] to remove the degeneracy in the R_{23}-metallicity relation, and we define the upper branch as Log ([NII]/[OII]) ≥ -1.2 (see also Appendix 3.5).

- **KK04** (Kobulnicky & Kewley, 2004): is a widely used metallicity calibration, based on an iterative method that allows to solve both the oxygen abundance and ionization parameter using the O_{32} and R_{23} line ratios.

- $\mathbf{T_e}$: the electron-temperature calibration (also referred to as "direct" method) is based on the ratio between the auroral line [OIII] $\lambda 4363$ and [OIII] $\lambda 4959, 5007$; it is commonly used to determine gas metallicities when the weak auroral line can be detected. Here we follow the procedure outlined in Izotov et al. (2006).

Fig. 3.12 shows our results for the oxygen abundances of the simulated galaxies obtained using the different methods. We show results as a function of the real gas metallicity, i.e. $12+\log(\mathrm{O/H})^{\mathrm{SIM}}$, the differences are also calculated with respect to this estimator. As expected, the metallicities predicted using the SIM-fiber and T04-fiber methods are systematically higher compared to SIM as the most metal-enriched regions in the bulge are preferentially sampled in these cases. This happens however only for the metal-poor galaxies, while for the metal-rich ones these estimators agree very well. Note that, in our simulations, the galaxies with low/moderate metal content (CS and CS^+ samples) are also the ones with stronger metallicity gradients, while galaxies in the MA sample do not show significant metallicity gradients, masking a possible bias due to preferential sampling (see also Fig. 3.5). The HII estimator, which samples the gas near young star particles, shows ± 0.5 dex scatter around the SIM value, in general higher for the galaxies with $12+\log(\mathrm{O/H}) \lesssim 8.3$ and lower for the more oxygen-rich ones.

The gas metallicities obtained with the emission-line calibration of T04 are systematically higher, with large discrepancies only for galaxies with low metal content. However, note that only for 11/15 galaxies this calculation is possible[5]. Also for the KK04 method, we find good agreement with SIM for the metal-rich systems, with some offsets for metal-poor galaxies (i.e. $12 + \log(\mathrm{O/H}) \lesssim 8.3$).

The largest discrepancies in gas metallicities, both with respect to the real value and to the other observational estimators, are found for the T_e method, which gives the lowest metallicity values. Unlike the other methods, the systematically lower metallicities are found even for the most metal rich galaxies where the rest of the methods agree well. The largest differences are of the order of ~ 1 dex, and only a weak trend with the metallicity is detected.

It is worth noting that the emission line intensities in the spectra calculated from SUNRISE rely on the MAPPINGS III photoionization code. The line ratios are then affected by uncertainties, assumptions and approximations in the model that describe the photodissociation regions, e.g. on-the-spot approximation (Stasińska, 2002, 2007), assumptions on the geometry of the nebula, dust grain composition, shocks, etc. (Groves, Dopita & Sutherland, 2004), which will in turn affect the derivation of the gas metallicities. In any case, our results show that wide differences in metallicity can appear due to the use of various calibrations even on the same spectrum modeled with a given photoionization code.

In summary, our analysis shows that *the fiber bias has a strong effect on the mean gas metallicity of galaxies, particularly for metal-poor systems and for systems with strong metallicity gradients, with metallicities systematically higher, up to* ~ 1 *dex, compared to the total mean metallicities. Deriving metallicities from the emission-line calibrations shows also large offsets, that can reach* $0.5-0.8$ *dex compared to the value in the simulations, but in this case, the results are more diverse, with both positive and negative differences. Finally, the calibration based on electron temperature predicts systematically lower metallicities, by* $0.7-1$ *dex, compared to the rest of the methods and to the metallicities derived directly from the simulations.*

3.3.5 Star formation rate

In observations, the SFR of galaxies is estimated using different methods at different redshifts, e.g. the luminosity of the Hα line, of the [OII]λ3727 line and the UV continuum for low, intermediate and high redshifts, respectively. Each SFR indicator is affected by theoretical biases and observational uncertainties (e.g. different timescale, sparse wavelength

[5]Note that, when the T04 calibration is used inside the fiber, the estimation is possible for 12 out of the 15 galaxies.

sampling, contribution from old stars, AGN contamination) which can influence the final SFR value (Calzetti, 2008). Furthermore, the use of different SFR proxies at different redshifts may introduce systematic errors that can bias the comparison between simulations and observations (Rosa-González, Terlevich & Terlevich, 2002).

In this section, we calculate the SFRs of our simulated galaxies with different methods. We use the calibrations given in Kennicutt (1998a), applying the correction factor $f_{\rm IMF} = 1.5$ (Calzetti et al., 2009) to account for the use of a different IMF (Kennicutt 1998a assumes a Salpeter IMF while we assume a Kroupa IMF in SUNRISE and Chabrier in BC03). We note that, in SDSS, the fiber-derived SFRs (Brinchmann et al., 2004) are corrected to total SFR with the Salim et al. (2007) method. For this reason, we use the full face-on spectra obtained with SUNRISE, dust-corrected with the Calzetti law, to derive SFRs, instead of the spectra within the fiber. The various methods to derive SFRs are described below.

- **SIM**: the real SFR extracted directly from the simulations (stellar mass formed over time interval), averaged over the past 0.2 Gyr.

- **BC03**: we convert the rate of ionizing photons calculated with BC03 into SFR:

$$\mathrm{SFR}(\mathrm{M}_\odot\,\mathrm{yr}^{-1}) = 1.08\, f_{\rm IMF}^{-1} \times 10^{-53} Q(H^0)\,(\mathrm{s}^{-1}).$$

- **Hα**: we extract the Hα-luminosity from the SUNRISE spectrum and convert into SFR according to:

$$\mathrm{SFR}(\mathrm{M}_\odot\,\mathrm{yr}^{-1}) = 7.9\, f_{\rm IMF}^{-1} \times 10^{-42} L(H\alpha)\,(\mathrm{erg\ s}^{-1}).$$

- **UV**: we calculate the flux from SUNRISE spectra of the nearly-flat region between $1500 - 2800\,\mathring{A}$ and calculate the SFR as:

$$\mathrm{SFR}(\mathrm{M}_\odot\,\mathrm{yr}^{-1}) = 1.4\, f_{\rm IMF}^{-1} \times 10^{-28} L_\nu\,(\mathrm{erg\ s}^{-1}\mathrm{Hz}^{-1}).$$

- **[OII]**: we estimate the SFR using the luminosity of the forbidden-line doublet [OII]$\lambda 3726, 3729$ in SUNRISE:

$$\mathrm{SFR}(\mathrm{M}_\odot\,\mathrm{yr}^{-1}) = 1.4\, f_{\rm IMF}^{-1} \times 10^{-41} L(\mathrm{OII})\,(\mathrm{erg}\ s^{-1}).$$

- **FIR**: the SFR is estimated from the SUNRISE FIR luminosity integrated over the range $8 - 1000\,\mu$m as:

$$\mathrm{SFR}(\mathrm{M}_\odot\,\mathrm{yr}^{-1}) = 4.5\, f_{\rm IMF}^{-1} \times 10^{-44} L_{FIR}\,(\mathrm{erg}\ s^{-1}).$$

All observational methods are sensitive to the emission from young massive stars, although the UV indicator has 10 times longer timescale (~ 100 Myr), because massive stars stay luminous for longer time in the UV with small production of ionizing photons (Calzetti, 2008).

Fig. 3.13 shows the differences between the SFR obtained with the different estimators and that obtained directly from the simulations ($\mathrm{SFR}^{\rm SIM}$). Except for the galaxy with the lowest SFR, the SFRs obtained with BC03, H_α and UV methods agree well, with differences of $\lesssim 0.2 - 0.4$ dex in logarithmic scale with respect to the real SFR of the simulations. Note that BC03 and H_α are based on the same method (conversion of the number of ionizing photons into SFR), but use two different models.

The other observational indicators ([OII] and FIR) predict in general lower SFRs compared to $\mathrm{SFR}^{\rm SIM}$, with maximum differences of about 1 dex. In particular, the [OII]-SFRs

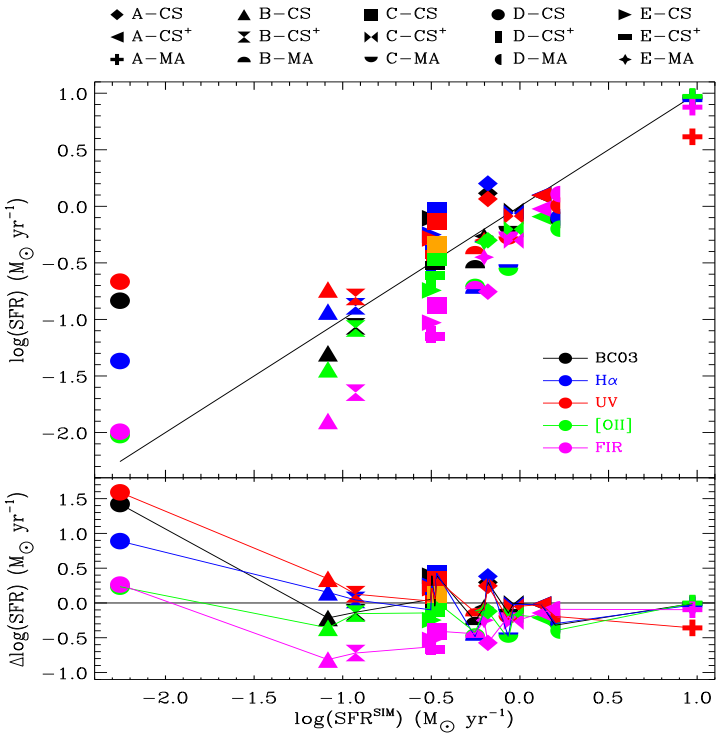

Figure 3.13: Comparison between the SFRs obtained using various methods as a function of the real SFR of the simulation (method SIM). In the upper box, we also include the 1:1 relation for reference.

of metal-poor galaxies are significantly lower than SFR^{SIM}, by ~ 0.5 dex. Note that the [OII] line intensity and [OII]/Hα ratio strongly depend on the metallicity (Kewley, Geller & Jansen, 2004) and has been calibrated by Kennicutt (1998a) for around-solar metallicity; this calibration is not accurate for the metal-poor galaxies of our sample. Also in the case of the SFR obtained with the FIR method differences are expected, as the Kennicutt (1998a) calibration is valid only for dusty metal-rich starburst galaxies (note that calibrations that follow both dust-obscured and unobscured star formation can be found in the literature, see Calzetti 2008); for the metal-poor sample, since the amount of dust in SUNRISE is proportional to the metallicity, we have only small dust absorption (Sec. 3.3.4) and reemission in the infrared (see also Hayward et al. 2014). To further investigate this effect, we plot in Fig. 3.13 the result obtained using the FIR method of one of our metal-poor galaxies (C-CS, orange square) where the assumed dust-to-metal ratio in SUNRISE was increased 10 times (as in Section 3.3.1). In this case, the SFR raises by ~ 0.6 dex, although it is still moderately lower compared to the results of the other estimators.

Finally, note that SFRs derived observationally and in simulations assume different timescales, of about $10 - 100$ Myr for the former compared to $0.2 - 0.5$ Gyr for the latter. This can result in significant discrepancies among the two estimations if we are in the presence of recent starbursts. In our sample, however, we observe such a systematic difference only in the galaxy with the lowest SFR (D-CS).

The results of this section show that the SFRs indicators exhibit differences of $\lesssim 0.4$ *dex in log(SFR), which are caused both by the difference in methods and the different characteristic time-scales to which the methods are sensitive. The* H_α *and UV methods predict similar SFRs compared to the direct results of the simulations, except for the galaxy with the lowest SFR.*

For galaxies with low metallicities the [OII] method gives systematically lower SFRs, with absolute differences lower than 0.5 *in logarithmic scale. For the FIR method the differences are larger in metal-poor galaxies, with SFRs systematically lower by* $0.5-1$ *dex in log(SFR), while the FIR estimation agrees with the SFR of the simulations up to* $0.1-0.2$ *dex for metal-rich galaxies.*

3.4 Discussion and conclusions

We have used a set of 15 simulated galaxies, of similar mass to the Milky Way, to study biases and systematics in the derivation of observables, and in the comparison between them and observational data. The aim of this study is twofold; first, help simulators to be aware of systematics and be able to reliably judge the agreement between simulated and observed galaxies and, second, help observers to interpret observational results by being able to better quantify the differences between the observationally-obtained galaxy properties and the real ones (that are known in the simulations).

Our simulations comprise 15 galaxies with a variety of merger, formation and accretion histories, which results in a variety of final morphologies, star formation rates, gas fractions and metallicities. As these properties somewhat depend on the modelling of feedback, the same 5 galaxies were simulated using three different models for chemical/energy feedback; the strength of feedback and the amount of chemical yields varies from moderate to strong. For moderate feedback we find galaxies that are more metal-poor and form their stars earlier, while for strong feedback we find younger galaxies with a higher metal content. This diversity, typical of real galaxies, makes the sample ideal to test biases and systematics in the derivation of the synthetic spectra, and to study the dependence of such effects on galaxy properties (with the caveat that with our sample we do not represent the more metal rich galaxies observed).

For our study we have computed the synthetic spectra of the simulated galaxies at $z = 0$, as at low-redshift many databases of galaxy properties derived from large galaxy surveys (e.g. SDSS, 2dFGRS and 6dFGS) are available. For this purpose, we followed three different approaches: (*i*) Stellar Population Synthesis (SPS) models, which give the spectra coming from stars; (*ii*) SPS models including dust extinction with a simple recipe; and (*iii*) a full radiative transfer calculation that gives the spectra including stellar and nebular emission, as well as the effects of dust which are parametrized using the metallicity of the interstellar medium. We have used the synthetic spectra to derive the observables (magnitudes/colors, stellar masses and ages, stellar/gas metallicities and star formation rates) in various ways, as to mimic real observations, and we have compared the results with the direct outputs of the simulations, i.e. the real galaxy properties.

Biases and systematics appear at various stages in the process of obtaining the observables from the simulations, due to:
- assumptions and parameters of SPS models,
- dust/radiative transfer/projection effects,
- weighting with the mass instead of luminosity for mean quantities,
- observational biases, such as extrapolation to external regions of galaxies where no spectral/photometric data are available (e.g. fiber bias, Petrosian/Model magnitudes),
- different parametrization of the star formation history and dust extinction when quantities are derived fitting a pre-constructed grid of models,
- in the particular case of gas metallicities and SFRs, the use of different calibrations.

We tested the effects of such biases on the magnitudes and colors of simulated galaxies,

their stellar masses, ages, stellar/gas metallicities and star formation rates. Our results can be summarized as follows:

- *Magnitudes*

  - The galaxies' magnitudes in the (*u,g,r,i,z*) SDSS bands derived from the five different SPS models used in our work show good agreement, with differences lower than 0.1 dex. Differences are larger for the most metal-rich galaxies, which also have a higher contribution of young and intermediate age stars, for which model uncertainties are the largest.

  - Dust effects can be important if galaxies are seen edge-on. A simple angle-averaged dust model predicts galaxies that are $\sim 0.4 - 0.8$ dex fainter compared to the no-dust case. If full radiative transfer is considered, where dust is traced by the metals, edge-on galaxies appear in general $\sim 0.3 - 1$ dex fainter, but only if these are relatively metal-rich.

  - Estimating the magnitudes using a more observational approach (as for instance the Petrosian and Model magnitudes of SDSS) exhibits some differences compared to the real magnitudes of the simulated galaxies, with offsets $\sim 0.2 - 0.3$ dex for 60% of the systems and up to 0.6 dex for the remaining 40%.

- *Stellar masses:*

  When observational techniques are applied to the simulated galaxies to estimate their stellar masses, the results vary depending on the treatment of mass loss in the simulations:

  - If mass loss includes only SNe (which is a small effect), all observational estimators give systematically lower stellar masses compared to the direct result of the simulation, as the fitted models include the full mass loss of Stellar Populations at each stage of evolution. The offset is similar for all galaxies, of about a factor 0.3 in $\log(M_*)$ or, equivalently, of 50% in M_*.

  - If mass loss is properly treated in the simulations (e.g. adding AGB stars), the observational techniques recover the real stellar masses with differences $0.1 - 0.2$ dex in $\log(M_*)$.

- *Stellar ages:*

  - If stellar ages are estimated mimicking observational techniques weighting with the luminosity (but ignoring the fiber bias), galaxies appear younger, typically by ~ 2 Gyr, compared to the direct result of simulations (mean age of stellar particles). Some galaxies exhibit larger differences, up to a maximum of 5 Gyr; among them, we find both very young and very old galaxies.

  - If only stars in the nuclear part of galaxies are included (as to mimic single-fiber surveys such as SDSS), the observationally-estimated galaxy ages are in general higher compared to the direct result of the simulations with differences of about $2 - 4$ Gyr for young systems and of $\lesssim 1$ Gyr for old ones. However, the results depend sensitively on the particular properties of galaxies, namely the presence of significant/insignificant age gradients and the resulting preferential sampling of old stars within the fiber.

- *Stellar metallicities:*

- Observationally-estimated stellar metallicities (ignoring the fiber bias) are in general higher compared to the direct result of the simulations, in particular if galaxies are metal-poor, with differences up to ~ 0.3 in logarithmic scale. For more metal-rich systems, differences are lower and closer to the direct result of the simulations. The differences in the behaviour of metal-rich/metal-poor galaxies originates in the different weight of old/metal-poor and young/metal-rich stars in the different methods, in particular depending on the way the mean metallicity is calculated (mass/luminosity-weighted).

  - If only stars within the fiber are considered, stellar metallicities tend to be higher compared to the direct result of the simulations with typical differences of $\sim 0.1 - 0.3$ dex and no strong dependence with the real stellar metallicity.

- *Gas metallicities:*

  - Deriving gas metallicities using different emission-line calibrations show large spread in the resulting metallicity values, up to $0.5 - 0.8$ dex. In general, metallicities lower than the real value are predicted for very metal-poor systems, while for more metal-rich galaxies they show a better agreement. Significant differences are found when the calibration based on electron temperature is used; in this case the derived metallicities are systematically lower by differences of the order of $0.7 - 1$ dex.

  - Gas metallicities are systematically higher than the direct result of simulations when the fiber bias is included, due to a preferential sampling of metal-rich regions in the galaxy. Differences can be up to 1 dex for the most metal-poor galaxies, while the for metal-rich sample differences are always smaller than 0.1 dex.

- *Star formation rates:*

  - Observationally-derived SFRs using different indicators present in general offsets of the order of ± 0.4 dex in log(SFR), mainly due to the differences in methods and time-scales to which they are sensitive. The largest differences are found for the estimations based on the FIR, which gives systematically lower SFR values up to 1 dex in logarithmic scale for the metal-poor galaxies.

In summary, we have shown that it is important to properly take into account different observational biases when galaxy observations are interpreted and linked to different underlying physical processes. Furthermore, a meaningful comparison between observations and simulations of galaxies requires a good understanding of the systematics; if this is not considered it is not possible to properly judge agreement/disagreement between simulations and observations, and to decide which of the physical processes included in the simulations are more relevant in the context of the formation and evolution of galaxies in a cosmological context.

Finally, we note that other possible sources of biases and systematics have not been explored here, e.g. related to the inability to resolve the typical height of gaseous discs (which could affect the comparison between face-on and edge-on projections). Moreover, how well the simulated galaxies reproduce the observed sizes/concentrations/luminosity profiles can affect our findings related to the fiber bias and Petrosian/Model Magnitudes (although in previous works we have shown that our models produce galaxies in broad agreement with observational results, e.g. Scannapieco et al. 2010, Scannapieco et al. 2012, Aumer et al. 2013).

In the next chapter we will compare the properties of our simulated galaxies to observations of the SDSS survey, in order to gain more insight in their agreement/disagreement in

Method	Class	Reference
T04	theoretical	Tremonti et al. (2004)
KK04	theoretical	Kobulnicky & Kewley (2004)
T_e	direct	Izotov et al. (2006)
Z94	theoretical	Zaritsky, Kennicutt & Huchra (1994)
KD02	theoretical	Kewley et al. (2002)
M91	theoretical	McGaugh (1991)
D02	combined	Denicoló, Terlevich & Terlevich (2002)
PP04	empirical	Pettini & Pagel (2004)
PP04(O3N2)	empirical	Pettini & Pagel (2004)
P05	empirical	Pilyugin, Vílchez & Thuan (2010)

Table 3.2: Main characteristics of the different metallicity calibrations (see Kewley & Ellison 2008 for more details).

relation to the treatment of feedback, star formation, mass loss and chemical enrichment in the simulations. We will also quantity differences from the observationally-derived quantities and the direct result of the simulations to provide the relevant scalings for a meaningful comparison, in the case of the different galaxy properties.

3.5 Appendix: gas metallicity calibrations

The observational derivation of the gas metallicity of galaxies requires a calibration of emission lines to recover the (O/H) ratio; different ones are used in various observational campaigns (e.g Zaritsky, Kennicutt & Huchra 1994; Kobulnicky & Kewley 2004; Kewley et al. 2002; Pettini & Pagel 2004). Previous work has already shown that large discrepancies, as large as ~ 0.7 dex, arise when different calibrations are used (Pilyugin, 2001). The calibrations can be broadly classified as empirical (further subdivided into "direct" and "statistical") and theoretical. The former derives metal abundances directly from electron temperature-sensitive lines or relations between temperature-estimated metallicities and strong emission lines, while the latter relies on photoionization models to calibrate the metallicity indicators (see the reviews by Ferland 2003; Stasińska 2007 and Kewley & Ellison 2008).

In general, theoretical calibrations give, for a given galaxy, a higher metallicity compared to electron temperature-based estimators (Liang et al., 2006; Kewley & Ellison, 2008). The causes of these discrepancies are still unclear, although some authors conjecture either problems in the photoionization models (Kennicutt, Bresolin & Garnett, 2003), or temperature gradient fluctuations in the nebulae (Stasińska, 2005, 2007; Bresolin, 2008). The effects of these disagreements can also bias the determination of the shape and the zero point in the mass-metallicity relation (Andrews & Martini, 2013).

Fig. 3.14 presents a comparison between the oxygen abundance of our simulated galaxies assuming various calibrations, that we summarize in Table 3.2. All results are derived after Balmer-correcting the emission line ratios for dust extinction using the Calzetti law (Calzetti, Kinney & Storchi-Bergmann, 1994). Some calibrations have limited ranges of validity, and here we only show results for galaxies that satisfy these conditions (see Kewley & Ellison 2008 for details).

It is clear that the metallicities for the various emission-line calibrations show larger differences, up to $\sim \pm 1$ dex. The use of different calibrations is certainly a source of concern, although in general the discrepancies are similar for all galaxies regardless their metallicities,

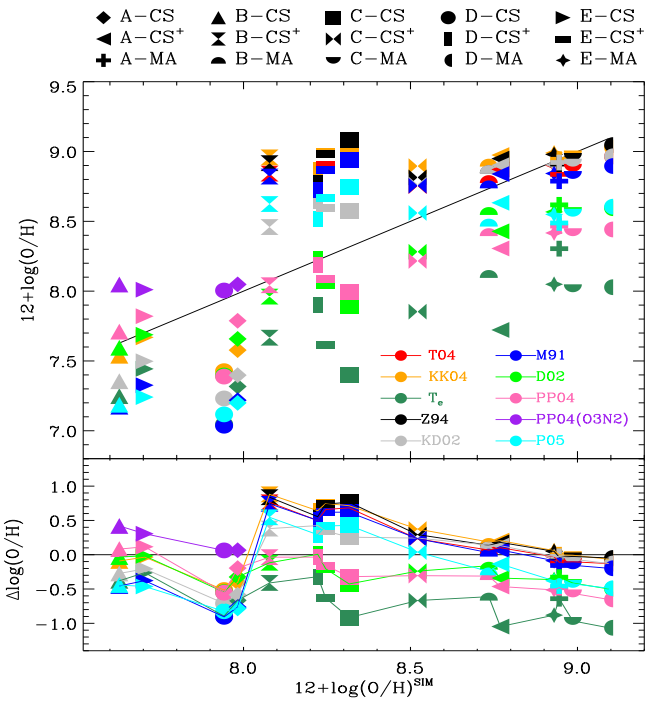

Figure 3.14: Comparison of the gas metallicity of simulated galaxies using different calibrations, as a function of the direct result of the simulation.

although variations are smaller for the more metal-rich galaxies. In particular, the calibration based on electron temperature is in general lower than the others (see Sec. 3.3.4), manifesting itself as a clear offset, particularly for galaxies with (direct) metallicities higher than $12 + \log(O/H) \sim 8.2$. On the other hand, T04 and Z94 give the highest metallicity values, with the other calibrations lying in between. For the most metal-rich galaxies, a number of indicators agree relatively well, particularly KK04, KD02, M91, and even T04 and Z94, with a small spread of $\sim 0.1 - 0.2$ dex (note that all of these are theoretical calibrations). For the most metal-poor galaxies, we also find that various of the calibrations agree relatively well with each other, still with variations of ± 0.5 dex.

Chapter 4

Toward a consistent comparison between hydrodynamical simulations and SDSS

Here we follow the approach illustrated in Chapter 3, but focusing on making an unbiased comparison between the SDSS dataset and our simulations. While in Chapter 3 we aimed to evaluate the accuracy of different observational techniques in recovering the galaxies' properties, in this work we study the comparison between simulations and observations. We derive the properties of the simulations following the simple derivations usually done in simulation studies, and we extract the same quantities mimicking the observational techniques applied in the SDSS survey. In this way we estimate the differences between the two derivations, and we accurately compare the results of our simulations with the SDSS observational dataset in an unbiased manner, which is a fundamental step to test the recipes for star formation, feedback and metal enrichment implemented in our hydrodynamical simulation codes.

In the previous Chapter we have shown that the simulations' properties derived following an observational approach may be significantly different from the ones calculated directly from the simulations snapshots. We have also shown that, in order to perform a reliable comparison with an observational dataset, the simulations must be converted into full synthetic observations, considering that each galaxy survey suffers from observational biases related to the observational setup and strategy, as well as to the different assumptions in the pipelines used to derive the physical properties (Walcher et al., 2011). In fact, we found that these observational effects and biases have a strong influence on the derived properties of galaxies, that may show large variations using different methods (as shown also by e.g. Scannapieco et al. 2010; Michałowski et al. 2014; Hayward & Smith 2015; Smith & Hayward 2015; Kewley & Ellison 2008).

In this Chapter we apply the SDSS analysis procedures to a sample of fifteen galaxies, simulated up to redshift $z = 0$ in a cosmological context using three different feedback and chemical enrichment models. In order to provide simulators a way to reliably compare their galaxies with SDSS data, for each physical property that we studied we give scaling relations, that can be easily used to convert the values extracted from the simulations into quantities more consistent with the SDSS dataset.

4.1 Methodology

4.1.1 The simulations

We use in this work cosmological hydrodynamical simulations of galaxy formation that are based on the dark-matter only Aquarius simulations (Springel et al., 2008). In particular, we use five galaxy halos with present-day virial mass similar to the Milky Way, i.e. $0.7 \times 10^{12}\mathrm{M}_{\odot} < M_{200} < 1.7{\times}10^{12}\mathrm{M}_{\odot}$ (see Scannapieco et al. 2009 for details). Each of the five halos is then re-simulated up to the present time including a baryonic component using the zoom-in technique (Tormen, Bouchet & White, 1997), with three different hydrodynamical codes based on Gadget-3 (Springel, 2005) which assume various recipes for star formation, chemical enrichment, metal-dependent cooling and SNe feedback, composing a total of fifteen galaxies. As shown in Scannapieco et al. (2012), differences in the implementation of SNe feedback have strong effects on the properties of simulated galaxies, and different hydrodynamical codes can produce galaxies with a large range of physical properties (e.g. morphologies, sizes, metallicities, ages, star formation rates) even for the same dark-matter halo.

In order to identify the different galaxies we assign letters from A to E for the five dark-matter halos, adding a label for the hydrodynamical code with which the galaxies have been simulated: either CS, CS$^+$ or MA. Simulations CS are run with the model described in Scannapieco et al. (2005, 2006), which includes star formation, chemical enrichment, metal-dependent cooling, feedback from supernova Type Ia and TypeII, and a multiphase model for the ISM. The second set of five simulations labelled CS$^+$ is generated with an updated version of the Scannapieco et al. (2005, 2006) model by Poulhazan et al. (in prep.), which adopts different choices for the chemical yields, a Chabrier Initial Mass Function (IMF), and includes chemical feedback from AGB stars. The third set, referred to as MA, is simulated with the update to the Scannapieco et al. code by Aumer et al. (2013); the main changes are a different set of chemical yields (which also include AGB stars), a different metal-dependent cooling function and a Kroupa IMF. In addition, the code has a different implementation of energy feedback from SNe, which is divided into a thermal and a kinetic part, and it includes feedback from radiation pressure coming from massive young stars. The MA model has in general stronger feedback compared to the CS/CS$^+$ models, resulting in more disky, younger galaxies. All the fifteen galaxies of our sample have, at redshift $z = 0$, total stellar masses between $1 - 10 \times 10^{10}\mathrm{M}_{\odot}$, gas masses in the range $3 - 10 \times 10^{10}\mathrm{M}_{\odot}$, stellar/gas mass resolution of $2 - 5 \times 10^{5}\mathrm{M}_{\odot}$, dark matter particle mass of $1 - 2 \times 10^{6}\mathrm{M}_{\odot}$, and gravitational softening of $300 - 700$ pc. The cosmological parameters assumed are: $\Omega_{\mathrm{m}} = 0.25$, $\Omega_{\Lambda} = 0.75$, $\Omega_{\mathrm{b}} = 0.04$, $\sigma_8 = 0.9$ and $H_0 = 100\,h$ km s^{-1} Mpc^{-1} with $h = 0.73$.

4.1.2 Creating the mock observations

The hydrodynamical simulations at redshift $z = 0$ have been post-processed with the radiative transfer code SUNRISE (Jonsson, 2006; Jonsson, Groves & Cox, 2010), which simulates the propagation of light through a dusty ISM using Monte Carlo techniques, and self-consistently derives the spectra of the simulated galaxies from different observing positions, including stellar/nebular emission, dust absorption and IR-emission. In a first stage, SUNRISE assigns each star particle a spectrum. For star particles older than 10 Myr, the stellar spectrum is selected according to the age, metallicity and mass of the particle from a template of spectra generated with the stellar population synthesis code STARBURST99, choosing the Padova 1994 stellar tracks (Fagotto et al., 1994a,b), a Kroupa IMF (Kroupa, 2002) (with $\alpha = 1.3$ for $m_{\mathrm{star}} = 0.1 - 0.5$ M$_{\odot}$ and $\alpha = 2.3$ for $m_{\mathrm{star}} = 0.5 - 100$ M$_{\odot}$) and Paul-

drach/Hillier stellar atmospheres. On the other hand, for star particles younger than 10 Myr, SUNRISE assigns a nebular spectrum that takes into account the effects of photo-dissociation and recombination of the surrounding gas. The nebular spectra are pre-computed with the photo-ionization code MAPPINGS III (Groves, Dopita & Sutherland, 2004; Groves et al., 2008), and depend on the metallicity of the star particle and the gas around it, on the ISM pressure[1], and on the Photo-Dissociation Region (PDR) covering fraction $f_{\rm PDR}$. The MAPPINGS III parameters not constrained by the underlying hydrodynamical simulation, $f_{\rm PDR}$ and $\rm M_{\it cl}$, have been set respectively to $f_{\rm PDR} = 0.2$ and $\rm M_{\it cl} = 10^5\, M_\odot$, following Jonsson, Groves & Cox (2010).

Once a spectrum is assigned to each star particle, SUNRISE enters the radiative transfer stage, where random-generated photon "packets" (rays) are propagated from these sources through the ISM using a Monte Carlo approach (we use $\sim 10^7$ Monte Carlo rays in our case. It is assumed that the dust is traced by the metals with a constant dust-to-metals ratio of 0.4 (Dwek, 1998), and that dust extinction is described by a Milky Way-like extinction curve normalized to $R_V = 3.1$ (Cardelli, Clayton & Mathis, 1989; Draine, 2003). In our model, unlike in other hydrodynamical codes where gas particles represent a mix of gas/stellar phases, (e.g. Springel & Hernquist 2003), each gas particle has a single temperature, density and entropy (see Scannapieco et al. 2006 for details), and so the amount of dust is linked to the total amount of metals in the gas particle (see also Hayward et al. 2011; Snyder et al. 2013; Lanz et al. 2014).

The tracing of the rays is done on an adaptive grid, which for our simulations is represented by a number of cells between $\sim 30.000 - 400.000$, and covers a box with side 120 kpc, with minimum cell size of $\sim 220 - 460$ pc. To compute the grid, we have assumed a value of tolerance $\rm tol_{\it met} = 0.1$ and metals opacity $\kappa = 3 \times 10^{-5}$ kpc^2 $\rm M_\odot^{-1}$ following Jonsson (2006).

We have also followed a simple approach by applying Stellar Population Synthesis (SPS) models to derive some of the galaxy properties, which we discuss here when appropriate.

4.2 Observational data

We compare the properties of our simulated galaxies, at redshift zero, with the Sloan Digital Sky Survey (SDSS) dataset. The SDSS camera (Abazajian et al., 2003) is designed to collect multi-color images and spectra of a large number of objects in an area of a third of the sky, at median redshift (for galaxies) of $z \sim 0.1$. The images are taken in 5 different photometric bands (*u, g, r, i, z*) of increasing effective wavelength, with filter curves defined for an airmass of 1.3 at the Apache Point Observatory, pixel size of 0.396" and exposure time of 53.9 s (Gunn et al. 1998, Gunn et al. 2006).

SDSS magnitudes are based on the AB photometric system (Oke, 1965; Oke & Gunn, 1983), which allows immediate conversion from magnitudes to physical fluxes (Fukugita et al., 1996). The spectrographic survey observes spectra of ~ 640 target objects simultaneously, and the light of each object is collected with a single optical fibre of diameter 3 arcsec in the sky pointing at the center of the object (see York et al. 2000; Smee et al. 2013 for a technical description, see also Chapter 3 and Stevens et al. 2014 for a discussion of the effects of the limited fibre size on simulation properties). The wavelength covering is between 3800 and 9200Å at resolution $R = 1800 - 2200$, and S/N > 4 at g-mag $= 20.2$. In a PLANCK cosmology (Planck Collaboration et al., 2015a), at $z \sim 0.1$, the fibre encloses a circular region of ~ 5.8 kpc in diameter, sampling only $\approx 1/3$ of the total light for a typical spiral

[1]The ISM pressure enters the MAPPINGS computation through the compactness parameter C, which is also related to the assumed cluster mass $\rm M_{\it cl}$ (see Groves et al. 2008; Jonsson, Groves & Cox 2010).

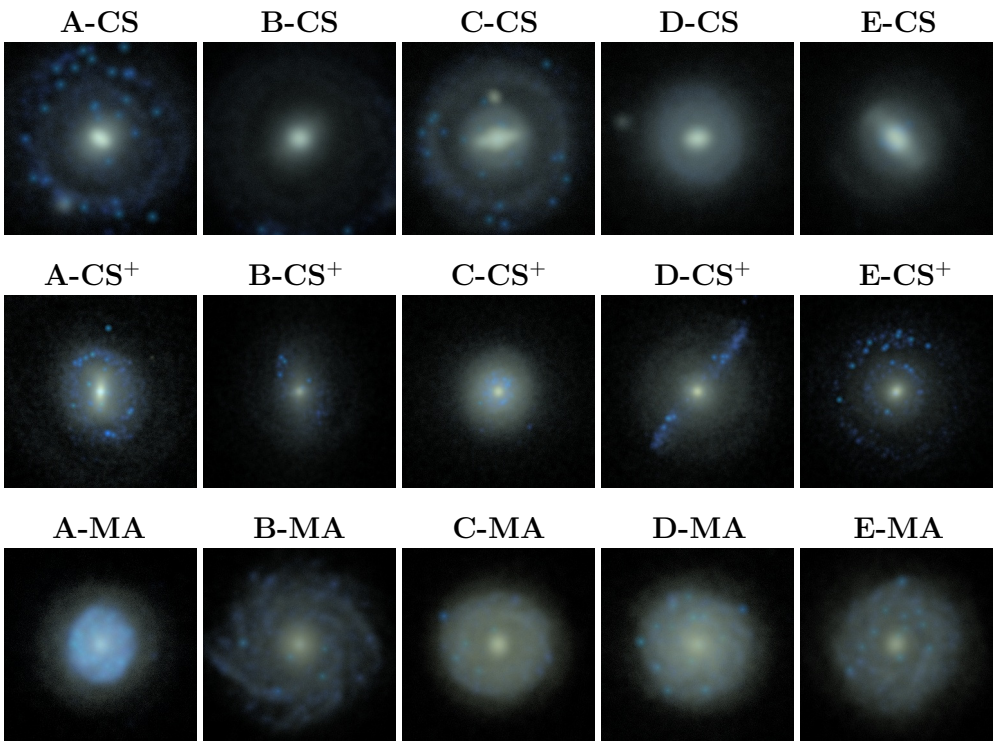

Figure 4.1: (u, r, z) multi-band face-on images of the fifteen simulated galaxies, as predicted from the radiative transfer calculation of SUNRISE, for a 60x60 kpc Field of View.

galaxy (Brinchmann et al., 2004). The SDSS Data Release 4 (DR4, Adelman-McCarthy et al. 2006) includes an imaging catalogue of about 180 million objects and spectroscopic data of $\sim$ 850.000 objects, of which $\sim$ 565.000 are galaxies. The DR7 (Abazajian et al., 2009) contains photometric information of more than 350 million objects and spectra of $\sim$ 930.000 galaxies.

In this work, we use datasets of derived galaxy properties from the MPA-JHU analysis of SDSS DR4 (for stellar masses/ages/metallicities) and DR7 (magnitudes/colors, stellar masses, gas metallicities, SFRs). From the MPA-JHU datasets we first select galaxies in the local universe ($z < 0.3$), and we separate the sample into early and late types according to visual classification. For the DR4 data we use the Nair & Abraham (2010) catalogue, which includes $\sim$ 14.000 galaxies, while for DR7 we select early/late-type galaxies according to the Galaxy Zoo classification (Lintott et al., 2008, 2011), a crowdsourcing-based project of morphological classification of $\approx 1/3$ of the SDSS DR7 galaxies. In order to better estimate how close/far from real spirals our simulated galaxies are, we further split the samples (both for DR4 and DR7) in *green valley* galaxies (Martin et al., 2007) defined according to the Salim (2014) condition on the specific Star Formation Rate (sSFR):

$$-11.8 < \log(sSFR) < -10.8 \qquad \text{(green valley)}$$

independently of the visual classification; in the figures we will plot the spiral, green valley and elliptical galaxies respectively in blue, green and red (notice that the division into separated sequences of *star-forming*, *intermediate* and *passive* galaxies is still under debate, see e.g. Casado et al. 2015). After selecting these subsets of galaxies from SDSS, the final galaxy sample that we will use consists of $\sim$ 7200 spirals, $\sim$ 700 ellipticals and $\sim$ 5700 green valley galaxies for DR4, and $\sim$ 145.000 spirals, $\sim$ 45.000 ellipticals and $\sim$ 63.000 green valley

galaxies for DR7. For the sake of clarity, in the figures we will show only $\sim 10\%$ of these galaxies, but we will plot the contours enclosing 50% and 80% of objects of each type.

4.3 Galaxy properties

In the next sections we will compare the different physical properties of the simulated galaxies (magnitudes/colors, concentrations, Sérsic indices, stellar masses, mean stellar ages/metallicities, gas metallicities, SFRs) with SDSS data. For most of the galaxy properties, we will show the effects of including the observational biases in the calculation mimicking SDSS (OBS) when the galaxies are observed in the face-on projection, and the results derived directly from the simulations or with little post-processing (SIM) as commonly used in simulation studies. Furthermore, we will show the results derived using an intermediate approach, where we apply simple refinements to the SIM method that are more directly comparable to the OBS results. We have also derived the properties following the OBS method but using the edge-on projections (OBS-edge). For the different galaxy properties, we will provide linear best-fit formulae of the relation between the OBS and SIM values, as well as between OBS and the other methods, using linear regression.

The techniques and models used to generate mock spectra from our simulated galaxies and to extract their physical properties have been described in Chapter 3, where we also discussed the effects of the SDSS small-aperture spectrograph (fibre bias)[2]. As explained below, these are based on the conversion of the simulation's outputs into mock SEDs obtained either using a SPS model or the radiative transfer code SUNRISE. In the next sections we will focus on showing for which properties it is more important to apply observational techniques when simulations are compared to observations, and in particular whether mimicking the biases of SDSS makes the simulations look closer to real spiral/elliptical/green valley SDSS galaxies.

4.3.1 Magnitudes, colours and stellar masses

In this section we compare the position of our simulated galaxies in the colour-magnitude/colour-mass diagrams with SDSS data, showing also the changes due to the different techniques applied to calculate these quantities. In the figures we will use the $(u-r)$ colour, as for this colour the SDSS galaxies show a clear bimodality in the colour-magnitude diagram (Strateva et al. 2001; Baldry et al. 2004). We apply the following methods to calculate the magnitudes (see also the previous chapter):

- **OBS [PETRO]**[3]: the magnitudes are derived from SUNRISE face-on images (edge-on for the OBS-edge method) with a procedure that mimics the SDSS Petrosian magnitudes calculation (Blanton et al., 2001; Yasuda et al., 2001), i.e. extracting the Petrosian Radius (Petrosian, 1976) in the r-band, and taking the flux inside two Petrosian Radii in all bands to calculate the magnitudes. The Petrosian radius R_P is the radius that for a galaxy with luminosity profile $I(r')$ satisfies:

$$\frac{\int_{0.8\,R_P}^{1.25\,R_P} dr' 2\pi r' I(r') / [\pi (1.25^2 - 0.8^2)\, R_P^2]}{\int_0^{R_P} dr' 2\pi r' I(r') / [\pi R_P^2]} \equiv 0.2$$

[2]The global properties are derived considering always a Field of View (FoV) of 60x60 kpc, see the previous chapter.

[3]In the following, we give in brackets the reference labels used in Chapter 3 when different.

The flux inside two Petrosian radii recovers nearly 98% of the light for an exponential profile and $\sim$ 80% for a DeVacouleur profile (Shen et al., 2003). For our galaxies, the fraction of $r-$band flux inside two Petrosian radii varies from $\sim$ 52% (E-CS$^+$) to $\sim$ 100% (C-MA), and is on average $\sim$ 80%.

- **SIM [BC03]**: the magnitudes of the simulations in the (u, g, r, i, z)-bands have been calculated with the Bruzual & Charlot (2003, BC03 hereafter) SPS model, assigning each stellar particle a spectrum according to its age, metallicity and mass. This is one of the most common and fast ways to calculate the spectra/magnitudes of simulated galaxies.

- **BC03-dust [CF00]**: we add to the spectra calculated with BC03 the effects of dust extinction assuming the Charlot & Fall (2000) model (CF00). CF00 uses different extinction curves for young stellar populations (supposed to be born in dusty molecular clouds) and old ones (extincted only by dust in the ISM), without any dependence on the inclination of the galaxy[4]. The CF00 free parameters are set according to the values given in da Cunha, Charlot & Elbaz (2008), which are slightly different from the ones derived by CF00 fitting a set of star-forming galaxies; the results of the BC03-dust model then depend somehow on the values assumed for the free parameters, and cannot be considered fully-predictive (see also Chapter 3).

For the calculation of the total stellar masses we follow these procedures:

- **OBS [PETRO]**: the masses are derived fitting the Petrosian magnitudes in all five photometric (u, g, r, i, z)-bands in the face-on view (edge-on in OBS-edge) to the grid of models described in Walcher et al. (2008), after subtracting nebular emission, as the fitted models include only stellar light[5].

- **SIM**: the total stellar mass is calculated summing the mass of star particles (within the 60x60 kpc FoV), extracted directly from the simulations' snapshots. Note that the mass in stellar particles inside two Petrosian radii in our simulations is on average $\sim$ 85% of the total stellar mass in the FoV, ranging from $\sim$ 60% (B-CS$^+$) to $\sim$ 100% (C-MA).

In Fig. 4.2 we show the one-to-one relation of the SIM/BC03-dust/OBS-edge versus OBS (r-band) magnitudes, and the results of linear fits to the relations. From the figure we can see that the SIM method gives in general lower (brighter) magnitudes compared to OBS, in particular for the fainter galaxies (r-mag $\gtrsim -21$), while for the brighter objects (r-mag $\lesssim -21$) it agrees better with the OBS estimation. These discrepancies are the result of the composite effect of dust extinction – dust is not considered in SIM and hence gives brighter magnitudes than OBS – and cutting the luminosity profile at $2\,R_P$ – OBS uses the Petrosian magnitudes, further reducing the total flux (as it misses the external part of the galaxy profile, the effect being larger for galaxies with non-exponential profiles). Note that these results strongly depend on the amount of dust, the orientation (face-on/edge-on), and the

[4]We note that the BC03-dust method shifts the results of BC03 approximately by a constant factor, as the offset between different galaxies can only (slightly) change due to the different number of young star particles.

[5]To remove the nebular contribution from the broad-band magnitudes we calculate the relative contribution of nebular emission within the fiber in each photometric band fitting the fiber spectrum with the STARLIGHT code, and assume that the relative contribution of nebular emission for the total galaxy is the same as in the fiber (see Chapter 3 for details).

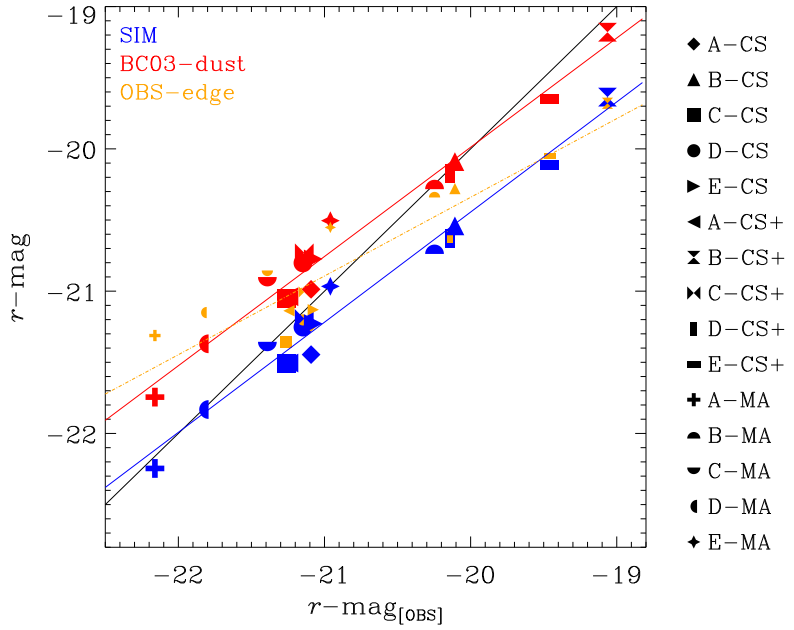

Figure 4.2: r-band absolute magnitudes of the simulated galaxies using the SIM, BC03-dust and OBS-edge methods, plotted against the observational value OBS in the 1-to-1 relation (black solid line). The blue, red and orange lines are, respectively, the linear fits of the SIM, BC03-dust and OBS-edge points.

luminosity profile of each galaxy. In fact, when the galaxies are seen edge-on (OBS-edge), the derived magnitudes are different compared to those found from the face-on images. The differences are due to the effect of dust which in general affects more the edge-on projections for dust-rich galaxies, and to the different values of the Petrosian radii, which are much higher when we see the galaxies edge-on (in particular for B-CS$^+$, D-CS$^+$ and E-CS$^+$), resulting in most of the cases in higher (i.e. fainter) magnitudes compared to OBS.

The BC03-dust model gives in most of the cases galaxies with fainter r-band magnitudes compared to OBS. In this case, it is important to note that BC03-dust is based on an angle-averaged dust model, from which we expect fainter magnitudes compared to the face-on magnitudes from OBS (at least for galaxies with sufficiently large R_P). For galaxies of intermediate brightness, the OBS values lie in between those given by the SIM and BC03-dust methods. Note that, as explained above, the SIM and BC03-dust models give a similar slope, indicating that the shift caused by the inclusion of dust in the magnitudes of galaxies (in the range analysed here) is approximately the same for all of them.

We have performed a linear fit to the relations between the SIM/BC03-dust/OBS-edge and the OBS r-magnitudes, and found the following values for the slope, zero point and correlation coefficient R:

$$r - \mathrm{mag}_{[\mathrm{SIM}]} = 0.77 \times \{r - \mathrm{mag}_{[\mathrm{OBS}]}\} - 4.96 \ \ [\mathrm{mag}]$$
$$R_{[SIM]} = 0.986$$

$$r - \mathrm{mag}_{[\mathrm{BC03\text{-}dust}]} = 0.77 \times \{r - \mathrm{mag}_{[\mathrm{OBS}]}\} - 4.63 \ \ [\mathrm{mag}]$$
$$R_{[\mathrm{BC03\text{-}dust}]} = 0.986$$

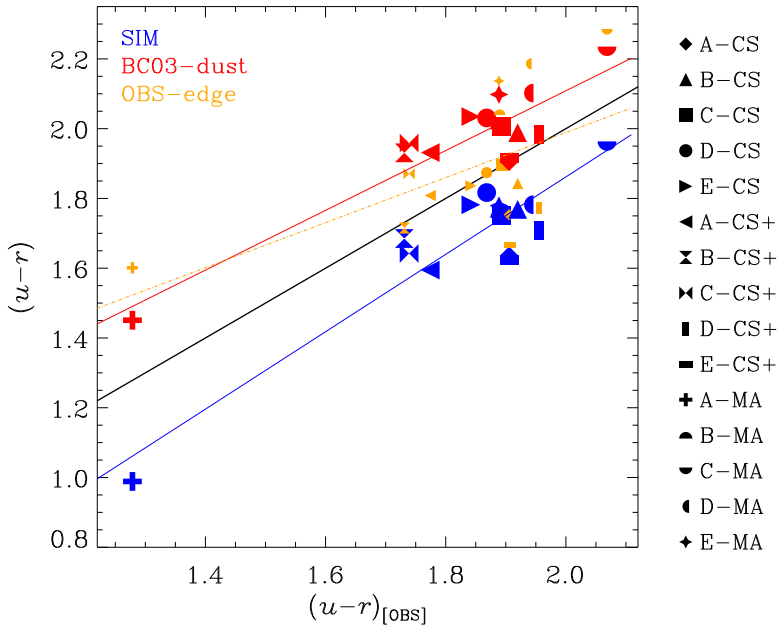

Figure 4.3: $(u-r)$ colours of the simulated galaxies using the SIM, BC03-dust and OBS-edge methods, as a function of the corresponding values obtained with OBS. We also show the 1−to−1 relation (black line) and linear fits of SIM, BC03-dust and OBS-edge (blue, red and orange lines respectively).

$$r - \mathrm{mag}_{[\mathrm{OBS\text{-}edge}]} = 0.55 \times \{r - \mathrm{mag}_{[\mathrm{OBS}]}\} - 9.28 \ \ [\mathrm{mag}]$$
$$R_{[\mathrm{OBS\text{-}edge}]} = 0.911$$

As can be seen in Fig. 4.2, a linear fit is a good approximation for these relations, as quantitatively indicated by the high values of R. Note, however, that these relations (as well as the ones that we discuss below) somehow depend on the specific simulation code and sub-resolution model adopted, and the dependence of these relations on the specific implementation of hydrodynamics has not been fully explored yet.

From Fig. 4.3 we see a similar behaviour for the $(u-r)$ colours of the SIM, BC03-dust and OBS-edge methods against the OBS values. The SIM method gives, as the effects of dust are ignored, bluer colours, in particular for A-MA (this is a very young and star forming galaxy, see Sections 4.3.3 and 4.3.5). In contrast, the BC03-dust method predicts redder colours compared to OBS, as the reddening is estimated angle-averaged. Note that the use of the Petrosian magnitudes can also have an impact on the estimation of the colours, in some cases with differences reaching $\sim 0.1 - 0.2$ mag (galaxies have in general different luminosity profiles/scalelengths in the different photometric bands, see e.g. Fathi et al. 2010), although the effects are strongly galaxy-dependent. The OBS-edge colours are in general redder compared to OBS, although in some cases they appear slightly bluer. This is due to the combined effect of the low amount of dust extinction and the changes in the $(u-r)$ colour due to the use of the Petrosian magnitudes.

The linear fitting functions (with respective goodness-of-fit indicators) obtained for the

SIM, BC03-dust and OBS-edge results are:

$$(u-r)_{[\mathrm{SIM}]} = 1.11 \times (u-r)_{[\mathrm{OBS}]} - 0.36 \ \ [\mathrm{mag}]$$
$$R_{[\mathrm{SIM}]} = 0.925$$

$$(u-r)_{[\mathrm{BC03\text{-}dust}]} = 0.86 \times (u-r)_{[\mathrm{OBS}]} + 0.40 \ \ [\mathrm{mag}]$$
$$R_{[\mathrm{BC03\text{-}dust}]} = 0.912$$

$$(u-r)_{[\mathrm{OBS\text{-}edge}]} = 0.65 \times (u-r)_{[\mathrm{OBS}]} + 0.70 \ \ [\mathrm{mag}]$$
$$R_{[\mathrm{OBS\text{-}edge}]} = 0.587$$

Note that the slope of the fit is close to one for the SIM method, but the correlation is slightly worse compared to that found for the $r-$magnitudes[6].

In Fig. 4.4 we show the colour-magnitude diagram of SDSS galaxies and the results obtained for our simulated galaxies using the OBS method. From the SDSS data, we select the Petrosian colours/magnitudes, making it consistent with our calculation in OBS. We also show contours for blue, red and green valley galaxies that enclose 50% and 80% of the corresponding datapoints. We find that most of the simulated galaxies are consistent with the photometrical properties of SDSS blue/green valley galaxies. A-MA is in the bluer outer part of the blue sequence, while E-CS$^+$ and B-CS$^+$ are outside of the region where most of the data are located. For comparison, we also show in this figure the results obtained with the SIM, BC03-dust and OBS-edge methods. As discussed above, using the SIM method moves the galaxies slightly down, more into the region of the blue sequence. Applying the BC03-dust model to calculate the magnitudes, the simulated galaxies move to the right (i.e. fainter magnitudes) and up (i.e. redder colours) towards the green valley region. With the OBS-edge method the galaxies look in general slightly redder and fainter, although for most of them the position in the diagram does not change significantly.

Fig. 4.5 shows the colour-mass diagram, including the SDSS datapoints and the results of the simulations using the four methods. The results are similar to those found in Fig. 4.4, with most of the simulated galaxies both in the blue sequence and its intersection with the green valley when the OBS method is applied, while they move slightly down (up) using the SIM (BC03-dust/OBS-edge) technique. Note that the shift in stellar mass obtained applying the OBS method is significant for the CS sample (~ 0.3 dex), as in the CS code the mass loss of stellar particles due to stellar evolution is not well described (see Chapter 3). When the code includes stellar mass loss by stars in the AGB phase (CS$^+$/MA samples), the shift in stellar mass is less important, of the order of $\sim 0.1 - 0.2$ dex. Note that the main uncertainties in the observational derivation are related to the use of the Petrosian magnitudes and to the simplified procedure to construct the grid of models used in the fit, in particular in the assumptions of star formation history (SFH) and dust attenuation (see e.g. Michałowski et al. 2014; Mitchell et al. 2013; Wuyts et al. 2009).

In summary, we find that *including the observational biases has limited influence on the position of the simulated galaxies in the colour-magnitude and colour-mass diagrams.* As the differences between the direct results of the simulations and those obtained mimicking the biases of SDSS data are small, *the values of magnitudes and colors derived applying SPS*

[6]The A-MA $(u-r)$ colours are far from the range covered by the rest of the galaxy sample. If we ignore this galaxy for the fits, we obtain lower correlation factors ($R \approx 0.55 - 0.65$) in all cases, with SIM slope and zero-point of 0.68 and 0.45, for the BC03-dust method a slope 0.63 and zero-point 0.84, and in the case of OBS-edge a values of 1.10 and -0.17 for the slope and zero-point respectively.

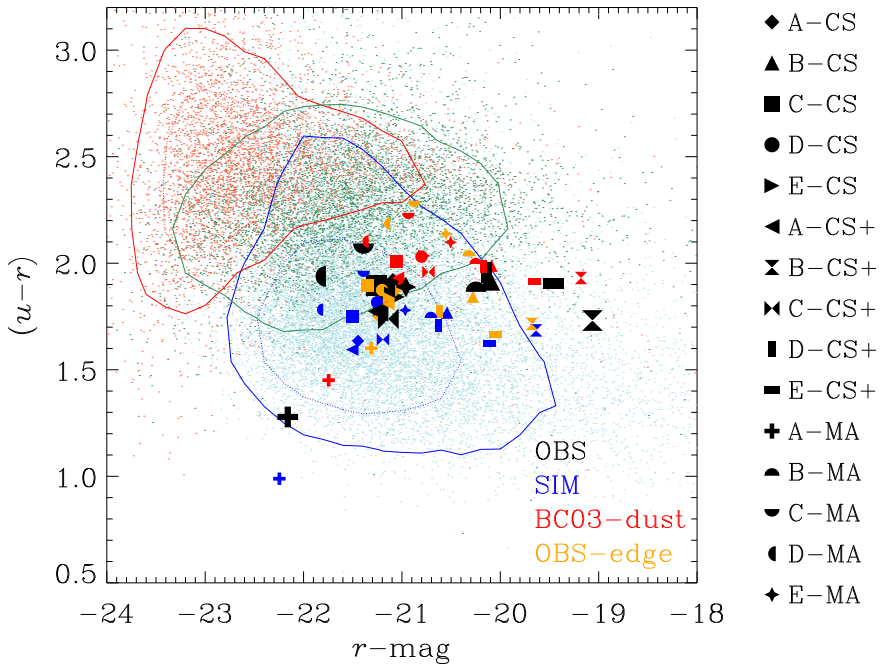

Figure 4.4: Colour-magnitude diagram of SDSS galaxies and simulations, using different methods to calculate the magnitudes of the simulated galaxies. In blue, green and red are the SDSS galaxies classified as spirals, green valley and ellipticals with their respective contours enclosing 50% (dotted lines) and 80% (solid lines) of the datapoints.

models to simulations can be compared with observations at a good approximation (at least for old enough galaxies). Including a simple dust model to the direct result of simulations does not seem to be useful to improve the comparison with observations. The use of the different projections (face-on/edge-on) has small influence on the position of the galaxies in the colour-magnitude and colour-mass diagrams, resulting in (slightly) redder and fainter galaxies for edge-on views. Our galaxy sample looks photometrically similar to galaxies in the blue sequence/green valley, and in most of the cases is inside the range of real galaxies in the colour-magnitude and colour-mass diagrams.

The accuracy of the stellar mass determination in observations, which affects the positions of the galaxies in the colour-mass diagram, has been also investigated by several previous studies. For example, Wuyts et al. (2009) have shown that, if more filters are used when fitting the photometry, the precision in the derived stellar mass can increase up to $\sim 0.03 - 0.13$ dex. On the other hand, stellar mass estimations based on fitting the SED have an accuracy of a factor of $\sim$ 2 for normal star-forming galaxies (Hayward & Smith, 2015; Torrey et al., 2015), which can however be improved assuming double-component SFHs in the fitted templates (Michałowski et al., 2014). In the case of SDSS, the masses derived from the SED (correcting for the limited fiber size using the $z-$band luminosity) and from photometry have been shown to agree within ~ 0.2 dex over the range $10^8 - 10^{12} M_\odot$ (Drory, Bender & Hopp, 2004).

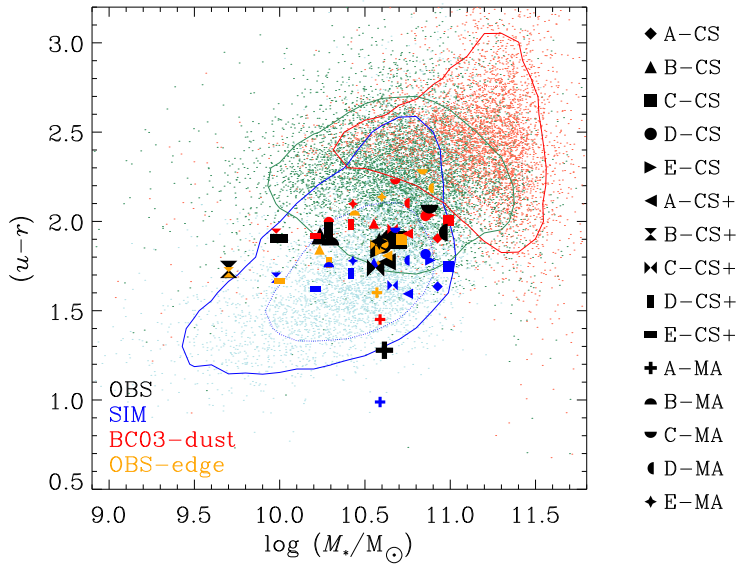

Figure 4.5: Colour-mass diagram of SDSS and simulated galaxies. The SIM/BC03-dust methods use the data of stellar particles of the simulation snapshots, while OBS and OBS-edge mimic the biases in the SDSS derivation of stellar mass and colours from face-on and edge-on photometric images. In blue, green and red are the SDSS galaxies classified as spirals, green valley and ellipticals with their respective 50% (dotted) and 80% (solid) contours.

4.3.2 Concentration and Sérsic index

The concentration index c (Fraser, 1972; Abraham et al., 1994) is defined in SDSS as the ratio (Shen et al., 2003):

$$c = \frac{R_{90}}{R_{50}}$$

where R_{90} and R_{50} are respectively the radii including 90% and 50% of the total Petrosian light. The concentration index has been shown to correlate with the morphological type (Shimasaku et al., 2001), and hence is a useful tool for morphological classification in large galaxy surveys.

In Fig. 4.6 we show the concentration index c, calculated in the r-band, as a function of the Petrosian r-band absolute magnitude, of our simulated galaxies. As these are purely observational properties, we only show the results extracted from the SUNRISE face-on (OBS) and edge-on (OBS-edge) images. From this figure we see that the concentrations obtained using the face-on images (OBS) of the CS/CS$^+$ samples are different compared to those of the MA galaxies, the former being more concentrated than the latter – with concentration indices larger by $\sim 0.5 - 1.5$. In fact, most of the CS/CS$^+$ simulations have concentrations consistent with SDSS green valley galaxies; however, some them (A-CS$^+$, B-CS$^+$, E-CS$^+$) are outside the region covered by the majority of the datapoints (note that two of these galaxies have concentrations in the range of observations but disagree in the magnitudes). In particular, the concentration index of A-CS$^+$ is larger than 3.7; only $\approx 0.07\%$ of the galaxies have concentrations above this value. On the other hand, the MA galaxies have low concentration indices and lie at the bottom of the range covered by SDSS data – only 6.5%

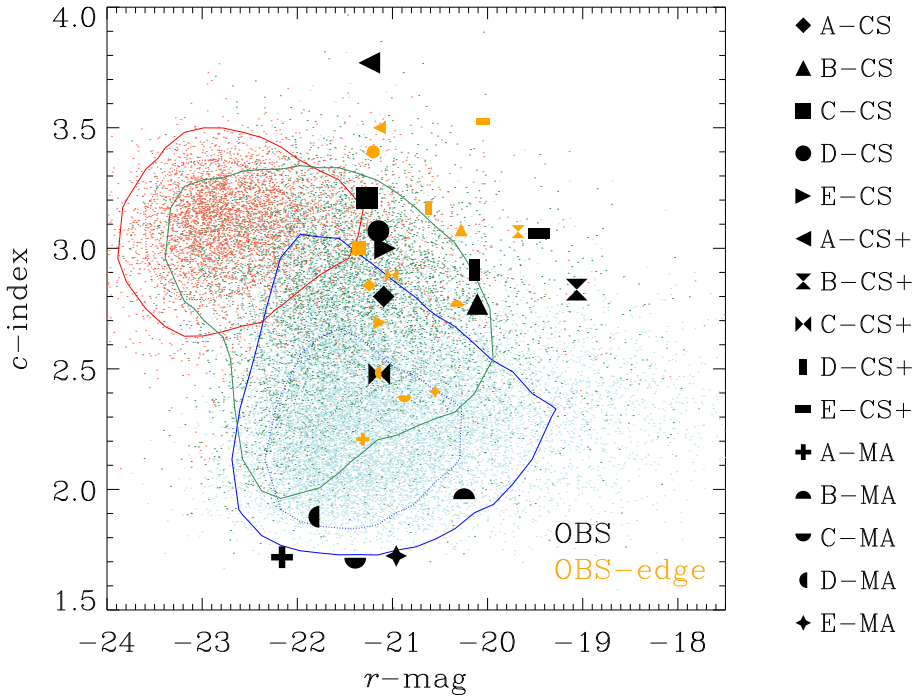

Figure 4.6: Concentration index c (in the r-band) for SDSS galaxies and simulations, plotted versus the r-band Petrosian absolute magnitude, calculated in face-on (OBS) and edge-on (OBS-edge) projections. The dotted (solid) contours enclose 50% (80%) of the blue, green and red SDSS sample.

of the SDSS galaxies have concentrations below 2 where the 5 MA galaxies appear, and only $\approx 0.4\%$ below 1.8 where 3 of the MA galaxies lie.

The effect of the orientation on the estimated concentration indices is significant: when we use the edge-on images, the concentrations of the three samples get closer, and most of them lie in the region of the green valley/blue sequence galaxies, although the MA concentrations are in general lower than those of the CS/CS$^+$ samples. Note that the concentrations obtained from the simulations are derived considering the two extreme orientations (face-on/edge-on), while the observational sample is not corrected for inclination effects.

In Fig. 4.7 we plot the r-band Sérsic indices of the SDSS galaxies and simulations as a function of the Petrosian r-band magnitudes. The observational data are taken from the NYU-VAGC catalogue (Blanton et al., 2005)[7], while the Sérsic indices of the simulations have been calculated fitting a single Sérsic profile to the r-band face-on (OBS) and edge-on (OBS-edge) images generated with SUNRISE, using the GALFIT code (Peng et al., 2002, 2010) and assuming arbitrary axis ratio, central pixels positions and angle in the fit. In the figure we see that the samples have different Sérsic indices; when the CS/CS$^+$ galaxies are observed face-on they have in general indices $2 \lesssim n_s \lesssim 5$ and they lie in the region of spirals/green valley galaxies, even though some of them are somewhat outside the area

[7]Note that the Sérsic fit in the catalogue is performed on the photometric images without considering the inclination.

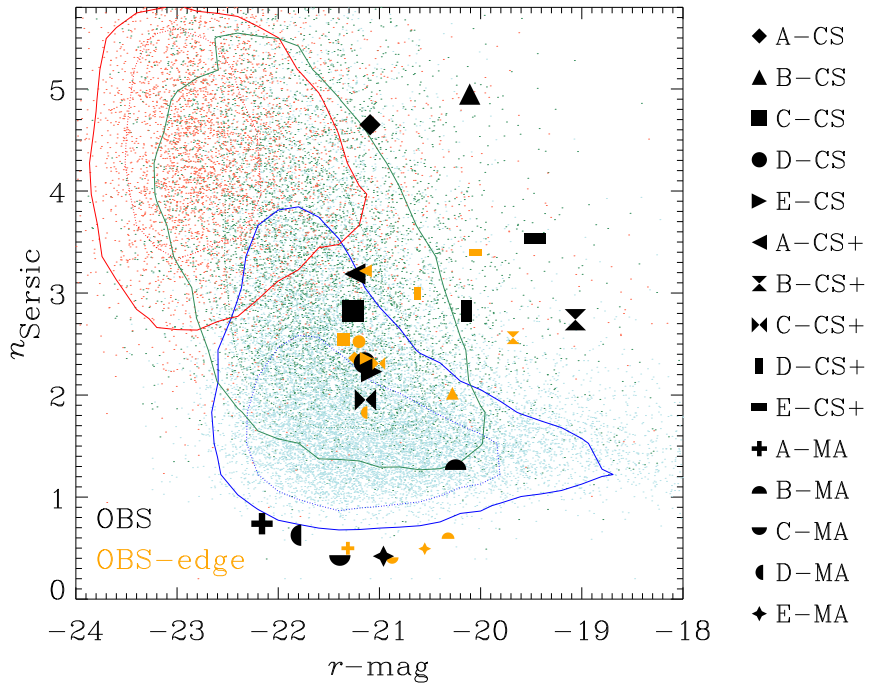

Figure 4.7: Sérsic indices versus absolute (Petrosian) magnitudes in the r-band for SDSS galaxies (from Blanton et al. 2005) and simulations in face-on and edge-on views (OBS/OBS-edge respectively), together with the 50% (dotted) and 80% (solid) contours.

covered by the data (B-CS, B-CS$^+$, E-CS$^+$), mainly in terms of the $r-$magnitudes. The MA sample has face-on indices $n_s \lesssim 1$ (hence close to an exponential profile $n_s = 1$), below the contour that includes 80% of the data points of spiral galaxies. In fact, the five MA galaxies have Sérsic index $n_s < 1.4$, where less than 10% of the observational datapoints are. Furthermore, four MA galaxies have $n_s < 0.8$, which corresponds to only 1.1% of the SDSS sample.

Similarly to what we found for the concentration indices, we find that projection has an impact on the derivation of the Sérsic indices. The use of the edge-one views causes the CS/CS$^+$ galaxies to have lower values compared to those obtained from the face-on projections; as a consequence they lie closer to the region of the green valley/blue sequence galaxies. In the case of the MA sample, the OBS-edge method predicts higher values for n_S, but the galaxies are in most of the cases still outside the 80% contour of the SDSS spirals (for a discussion about the origin of these trends see e.g. Maller et al. 2009; Pastrav et al. 2012).

Figs. 4.6 and Fig. 4.7 show that *the CS/CS$^+$ and MA samples are morphologically different*, with the MA galaxies lying at the low extreme of the range of spirals both in concentration and Sérsic index when these quantities are calculated face-on, consistent with only a small fraction of the SDSS galaxies, while galaxies of the CS/CS$^+$ sample lie in the region where spirals and green valley galaxies overlap. The use of the different projections has however a significant influence on the position of the galaxies both in the concentration-magnitude and Sérsic index-magnitude diagram.

4.3.3 Stellar ages and stellar metallicities

In this section we compare the stellar ages and metallicities of the simulated galaxies with SDSS data. The ages/metallicities in the SDSS-Garching DR4 are derived using the method described in Gallazzi et al. (2005, 2006), which is based on simultaneously fitting different absorption features adopting a Bayesian inference approach. To calculate the mean ages/metallicities in our simulations we follow these procedures:

- **OBS [LICK-IND-fibre]**: we run SUNRISE without nebular emission and using BC03 as input stellar model[8]; we select the spectra inside a circular region of 4 kpc radius from the center of the galaxies both in face-on (OBS) and edge-on (OBS-edge) projections, mimicking the fibre size of the SDSS spectrograph at $z \sim 0.15$ (fibre FoV). From the spectra, we measure the strength of the D4000n, Hβ, Hδ_A+Hγ_A, [Mg_2Fe] and [MgFe]' absorption features (Worthey et al., 1994; Worthey & Ottaviani, 1997; Balogh et al., 1999) and we compute the mean ages and metallicities fitting these indices with the method described in Gallazzi et al. (2005). As the Gallazzi et al. method is also sensitive to the estimation of the errors on the indices, from each noiseless galaxy spectrum obtained with SUNRISE we produce 1000 different spectra, adding Gaussian-distributed random noise with S/N = 10, and we measure 1000 times the strength of the absorption features for each galaxy in the two projections, using as final value for the indices and related errors, respectively, the average and standard deviation of the measurements. Note that this method to calculate the indices is different from the one presented in the previous chapter, as now the indices are more consistently extracted directly from the total stellar spectra, and the errors on the measurements are better estimated.

- **SIM [SIM-fibre]**: the mean ages/metallicities have been calculated averaging the (linear) ages/metallicities of stellar particles in the fibre FoV, weighted by mass as is common in simulations studies.

- **Lum-W-fibre [SIM-LUM-fibre]**: we compute the mean ages and metallicites weighting, respectively, with the stellar particle's luminosity in the r-band and in all SDSS bands, calculated both with BC03, and considering only particles in the region sampled by the fibre.

We plot in Fig. 4.8 the different estimations of the mean stellar ages of the simulated galaxies in the one-to-one relation with the value derived mimicking the SDSS observational biases, for the face-on projection (OBS). It is evident from the figure that the SIM values are systematically higher than the OBS ones, giving older stellar ages even by $\sim 2-4$ Gyr. When the mean age is estimated weighting with the luminosity (Lum-W-fibre) we obtain younger ages compared to SIM and in better agreement with OBS, although the majority of galaxies remains older with respect to OBS by $\sim 1-2$ Gyr. The discrepancy among the SIM and Lum-W-fibre methods tends to increase at younger ages ($\lesssim 4$ Gyr); furthermore, both relations exhibit a different slope. According to the OBS age estimations, the galaxies in general appear much younger. In the case of the OBS-edge method, we obtain similar values compared to OBS, although the oldest galaxies appear slightly younger (note that the edge-on

[8]Since the input stellar model commonly used in SUNRISE is SB99, which has sampling $\Delta\lambda \sim 20$ Å in the optical, the spectral resolution of the SED when the SB99 input spectra is used is too low to reliably measure the Lick indices. Note also that if nebular emission is neglected and the BC03 stellar model is used, the interpretation of results is more direct, as the Gallazzi et al. method assumes BC03 in the fit, and requires the subtraction of nebular emission for the calculation of the indices.

projections sample also part of the disks, which have in general a younger stellar content than the bulge). The Lum-W-fibre method gives systematically lower ages compared to SIM, as weighting with the luminosity gives more weight to the younger stellar populations, which in general emit more light than the old ones. The luminosity-weighted ages calculated with the OBS/OBS-edge methods are in general younger than Lum-W-fibre, due to the uncertainties and simplified assumptions in the procedure to construct the grid of models used to fit the Lick indices. Also note that the scatter, in particular for the SIM estimation, is relatively large, evidencing the variety of star formation histories of the galaxies. It should be noted that all methods consistently consider the same set of stellar particles in the central part of the galaxies sampled by the fibre (apart from projection effects for the OBS-edge method), so the differences are not caused by the presence/strength of age/metallicity gradients (fibre bias, see Chapter 3), but are purely related to the different techniques applied to derive the properties.

We have made linear fits (blue/red/orange lines) to the SIM/Lum-W-fibre/OBS-edge datapoints obtaining the following relations and correlation factors:

$$\mathrm{Age}_{[\mathrm{SIM}]} = 0.30 \times \mathrm{Age}_{[\mathrm{OBS}]} + 8.28 \ \ [\mathrm{Gyr}]$$
$$R_{[\mathrm{SIM}]} = 0.551$$

$$\mathrm{Age}_{[\mathrm{Lum\text{-}W\text{-}fibre}]} = 0.97 \times \mathrm{Age}_{[\mathrm{OBS}]} + 2.37 \ \ [\mathrm{Gyr}]$$
$$R_{[\mathrm{Lum\text{-}W\text{-}fibre}]} = 0.884$$

$$\mathrm{Age}_{[\mathrm{OBS\text{-}edge}]} = 0.75 \times \mathrm{Age}_{[\mathrm{OBS}]} + 1.25 \ \ [\mathrm{Gyr}]$$
$$R_{[\mathrm{OBS\text{-}edge}]} = 0.850$$

The scatter in the SIM datapoints is reflected in the low value of the correlation coefficient $R_{[\mathrm{SIM}]}$. As expected, a much larger correlation factor is found for the Lum-W-fibre method.

In Fig. 4.9 we compare the simulations' stellar ages obtained with the OBS method with mean ages from SDSS, in the stellar age-stellar mass diagram (note that the stellar mass of the OBS method is the PETRO mass estimation, sec. 4.3.1). Our results show that about half of the simulated galaxies look older compared to the observations, and the rest is close or inside the contours corresponding to green valley galaxies (CS sample) or well inside the blue sequence (A-CS$^+$, C-CS$^+$, A-MA, D-MA) where it intersects with green valley and red galaxies. When the OBS method is applied to edge-on spectra in the fibre (OBS-edge), we obtain similar results compared to OBS, while some simulations move into the region covered by the observational data (B-MA, C-MA, E-MA). For reference, we also include in the figure results for the SIM and Lum-W-fibre methods; note that in these cases the stellar mass is calculated as the sum of the mass of stellar particles of the simulated galaxies. As discussed above, using the SIM method the galaxies appear too old compared to the observations, while weighting with the luminosity (Lum-W-fibre) moves the points down towards the range of observations, but still most of the galaxies are too old compared to the real ones.

We make a similar analysis for the mean stellar metallicities, showing in Fig. 4.10 the comparison among the different methods. From the figure we see that both the SIM and Lum-W-fibre methods give systematically higher metallicities compared to OBS, with the offsets increasing (up to $\sim 0.4 - 0.5$ dex) for metal poor galaxies. The discrepancy between SIM and Lum-W-fibre is explained by the different weight of old and young stars when the average metallicity is calculated weighting with the luminosity, while OBS (OBS-edge) results (for which the mean metallicity is computed luminosity-weighted) suffer from the

uncertainties intrinsic in the method. When the observational method is applied to edge-on spectra we obtain very similar results compared to OBS over the full range of metallicities.

The values obtained for the linear fit and correlation coefficient are:

$$\log(Z/Z_{\odot})_{[\mathrm{SIM}]} = 0.27 \times \log(Z/Z_{\odot})_{[\mathrm{OBS}]} - 0.15 \ \ [\mathrm{dex}]$$
$$R_{[\mathrm{SIM}]} = 0.525$$

$$\log(Z/Z_{\odot})_{[\mathrm{Lum\text{-}W\text{-}fibre}]} = 0.69 \times \log(Z/Z_{\odot})_{[\mathrm{OBS}]} + 0.23 \ \ [\mathrm{dex}]$$
$$R_{[\mathrm{Lum\text{-}W\text{-}fibre}]} = 0.701$$

$$\log(Z/Z_{\odot})_{[\mathrm{OBS\text{-}edge}]} = 1.1 \times \log(Z/Z_{\odot})_{[\mathrm{OBS}]} + 0.08 \ \ [\mathrm{dex}]$$
$$R_{[\mathrm{OBS\text{-}edge}]} = 0.987$$

The correlations are similar to the ones found for the stellar ages, with the Lum-W-fibre method increasing the value of R, and the OBS-edge method having $R \sim 1$.

In Fig. 4.11 we show a comparison of the stellar metallicities obtained with the OBS method and observational results. We find that some of the galaxies are in the area of metal-poor spirals, while most lie outside of the region where most of the observations are (in particular the CS sample), with $\log(Z/Z_{\odot}) < -0.9$. Note that only $\approx 1.6\%$ of the SDSS galaxies have metallicities lower than this value. In the case of stellar metallicities derived using the SIM or Lum-W-fibre methods, galaxies appear slightly more metal-rich, with most of them lying in the blue sequence, although the CS sample is again outside the region covered by SDSS galaxies.

In conclusion, *we find that the effects of using different methods to calculate the mean stellar ages/metallicities are strong and affect significantly the comparison of simulations with observations*. When simple derivations of stellar ages and metallicities done in simulation studies are used, the galaxies appear older and more metal-rich compared to results obtained following observational techniques. In the case of stellar ages, the discrepancies between SIM and OBS values are large, in particular at younger ages. Weighting with the luminosity to obtain the stellar ages also affects the results, which are closer to the observational values but shifted compared to OBS by almost a constant factor. In the case of stellar metallicities, the discrepancy between SIM and OBS increases at lower metallicity, while the offset is almost constant when the mean metallicity is calculated weighting with the luminosity. The effect of the projection, estimated applying the observational method to edge-on spectra, is secondary compared to the differences arising from the use of different derivation methods. Our results show that the majority of our simulated galaxies appear older than real spirals, and with metallicities similar or lower than the most metal-poor spirals in SDSS.

It should be noted that comparing the Lum-W-fibre and OBS results (both providing the luminosity-weighted ages/metallicities) the observational method shows a bias to systematic younger ages and lower metallicities. Notice also that our Lum-W-fibre results have a scattering with respect to the Lum-W-fibre/OBS relation of ~ 0.3 dex in ages and 0.2 dex in metallicities, similar to the errors of the method estimated by Gallazzi et al. (2005) in case of good signal-to-noise (i.e. S/N$>$ 20), namely ~ 0.2 dex for the ages and ~ 0.3 dex for the metallicities. In the previous chapter we have shown that SED fitting methods are able to reach a higher accuracy in stellar age and metallicity determination, constraining the ages by ~ 0.06 dex and the metallicities by $\sim 0.15 - 0.25$ dex, although still with some trends (see Chapter 3); other SED fitting studies (e.g. Cid Fernandes et al. 2005) claim results similar to our findings (in the case of good S/N), with ages constrained by ~ 0.08 dex and metallicities by ~ 0.1 dex.

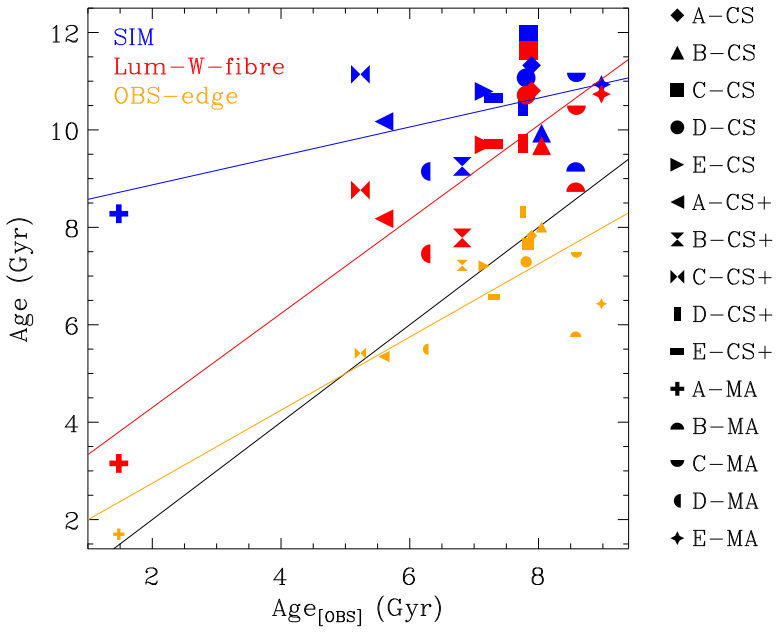

Figure 4.8: Mean stellar ages estimated with the different methods, plotted against the observational estimations for the face-on projection, in the one-to-one relation (black solid line).

4.3.4 Gas metallicities

In the SDSS-Garching DR7 dataset, the gas oxygen abundances (metallicities) are extracted using the method described in Tremonti et al. (2004, T04) (see also Brinchmann et al. 2004), which is based on the simultaneous fit of different emission lines according to the Charlot & Longhetti (2001) model (CL01). The authors also give a calibration of the R_{23}-metallicity relation – so-called T04 calibration – which is valid on the upper branch. In this work we derive the gas metallicities of the simulated galaxies in the following ways:

- **OBS [T04-fibre]**: the gas metallicity is calculated applying the T04 calibration to SUNRISE face-on spectra (edge-on for OBS-edge) extracted inside a circular region of 4 kpc radius at the center of the galaxy (fibre FoV, sec. 4.3.3), after correcting for dust extinction with the Calzetti law (Calzetti, Kinney & Storchi-Bergmann, 1994). Since the T04 calibration is only valid in the upper branch of the relation, from our sample of fifteen galaxies we are able to include twelve objects (eleven for OBS-edge), selected according to the [NII]/[OII] ratio (Kewley & Ellison, 2008).

- **SIM [Mass-W]**: we calculate the mean $12 + \log(O/H)$ abundance from the oxygen/hydrogen ratio of each gas particles, weighted by the particle's mass.

- **Sim-fibre**: the same as SIM, but only considering gas particles inside the fibre FoV in the face-on orientation.

The results of these different techniques are shown in Fig. 4.12, plotted in the one-to-one relation with the OBS method (T04-fibre). The plot reveals a large scatter among the

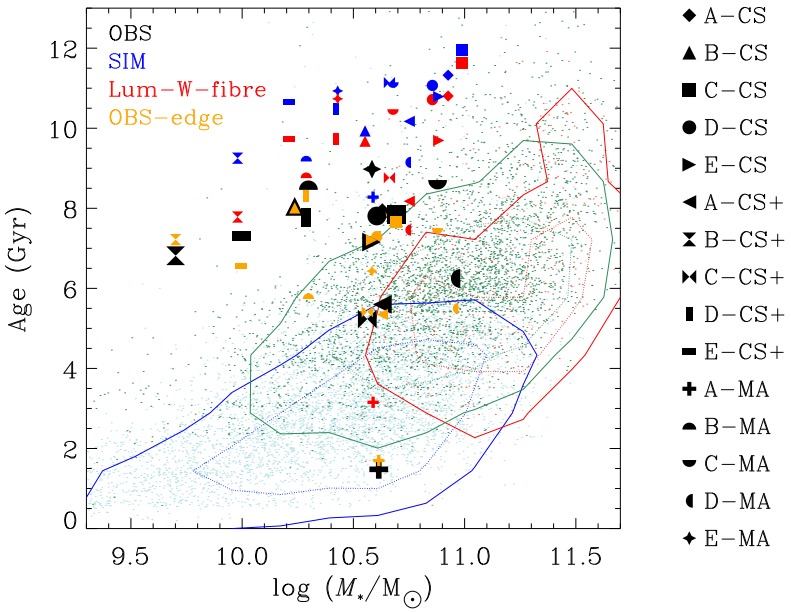

Figure 4.9: Stellar ages of the simulated galaxies, plotted together with SDSS data and the contours that enclose 50% and 80% of the datapoints shown respectively as dotted and solid lines.

methods, particularly for galaxies with 12+log$(O/H) \lesssim 9$. Deriving the metallicities with the SIM method gives in most of the cases lower values compared to OBS/Sim-fibre/OBS-edge (note that the OBS/Sim-fibre/OBS-edge methods measure the metallicity in the metal-enriched central part of the galaxies, and that our CS and CS$^+$ samples have stronger metallicity gradients compared to MA, see Chapter 3).

Using the Sim-fibre method moves the metallicities closer to OBS, even though with some scatter and with the tendency to underestimate the metallicity of metal-poor galaxies. Although the Sim-fibre and OBS methods sample the same central region of a galaxy, a discrepancy among the two methods is somehow expected as the emission line ratio from which the OBS values are extracted are based on the MAPPINGS III code, and the several uncertainties and assumptions on modelling nebular emission in the photoionization code may affect the derivation of the gas metallicities (see Groves, Dopita & Sutherland 2004).

The OBS-edge method gives similar results compared to OBS; however, the gas metallicity of the most metal-rich galaxies is systematically lower by $0.1 - 0.2$ dex compared to the results when the galaxies are observed face-on (OBS). These differences may indicate that the use of a different orientation will affect the region sampled by the fiber due to projection effects, which on its turn will affect the metallicity estimation. Note also that the uncertainties in the dust corrections for edge-on/face-on galaxies may also influence the determination of the gas metallicities.

We fitted the relations between the SIM/Sim-fibre/OBS-edge and the OBS methods with linear functions (blue, red and orange lines, respectively), and obtained the following

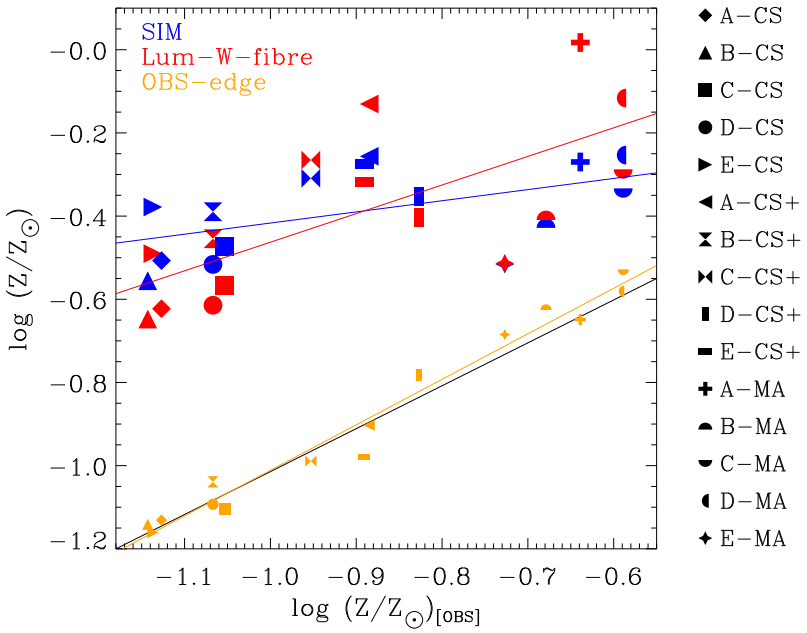

Figure 4.10: One-to-one relation of the different mean stellar metallicity estimations (together with the best-fit models in blue, red and orange lines) versus the method closest to SDSS applied to face-on spectra (black line).

parameters:

$$12 + \log(O/H)_{[\mathrm{SIM}]} = 2.22 \times \{12 + \log(O/H)_{[\mathrm{OBS}]}\} - 11.33$$
$$R_{[\mathrm{SIM}]} = 0.674$$

$$12 + \log(O/H)_{[\mathrm{Sim\text{-}fibre}]} = 1.50 \times \{12 + \log(O/H)_{[\mathrm{OBS}]}\} - 4.65$$
$$R_{[\mathrm{Sim\text{-}fibre}]} = 0.621$$

$$12 + \log(O/H)_{[\mathrm{OBS\text{-}edge}]} = 0.54 \times \{12 + \log(O/H)_{[\mathrm{OBS}]}\} + 3.97$$
$$R_{[\mathrm{OBS\text{-}edge}]} = 0.804$$

Note the different slopes of the relations, particularly in the case of the SIM and Sim-fibre methods, which also have a similarly low correlation factor. As expected, the relation between the OBS-edge and OBS datapoints has a higher correlation factor.

In Fig. 4.13 we compare the gas metallicities of the simulated galaxies obtained with the OBS method and the SDSS dataset. We find for face-on views that most of the sample is in good agreement with the observations and inside the area covered by the data, even though the majority of the galaxies have metallicities slightly below the T04 analytical relation (dashed line), and the A/E-CS and C-CS$^+$ galaxies are outside the contour containing 80% of the data. For metallicities derived from edge-on spectra, all the galaxies are below the analytic relation, and only four galaxies are inside the 80% contour. Note that, in the gas metallicity-stellar mass plane, following observational techniques makes the galaxies more

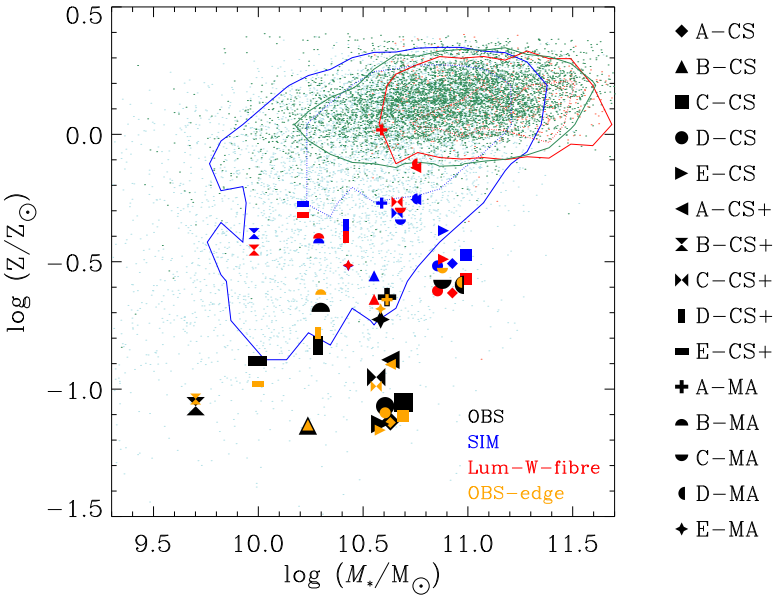

Figure 4.11: Mean stellar metallicity of SDSS galaxies and simulations, estimated using different techniques. The simulated galaxies appear in the area of young spirals (the blue contours contain 50% and 80% of SDSS spirals), but some of them are outside the range of real galaxies.

consistent with observations, by shifting them both in the metallicity and stellar mass values, compared to the common estimations done in simulation studies. For reference we also show results for the SIM and Sim-fiber; for SIM and Sim-fiber galaxies appear too metal-poor compared to observational data. The discrepancies are large, in some cases even by more than 1 dex. Note also that when we take into account only particles inside the fibre FoV (Sim-fibre), the metallicities of the CS/CS$^+$ samples significantly increase (while those of the MA galaxies remain similar, since these galaxies have flatter metallicity gradients as shown in the previous chapter).

Our results show that, *to compare gas metallicities of simulated and real galaxies it is important to apply to the simulations the same methods and calibrations than in observations, in order to make the comparisons reliable.* An intermediate step of obtaining a more comparable but simple gas metallicity estimation from the simulations is to mimic the most relevant biases of the survey, in particular the SDSS fibre size. The use of face-on or edge-on views has also an influence on the gas metallicity estimation, as the fiber may sample different regions in a galaxy due to projection effects. We have shown that *simple calculations obtained directly from the simulations, that neglect all the observational biases, can not be properly compared to SDSS observations.* Mimicking the SDSS derivation, our galaxies are close to the gas metallicities of spirals in SDSS, although with the trend of having oxygen abundances slightly lower than observational results. This is related to the particular calibration adopted (T04), as different studies (e.g. Kewley & Ellison 2008) have shown that the effects of the metallicity calibration used to determine the oxygen abundance are strong, affecting the determination with offsets up to ~ 0.7 dex comparing theoretical (such as T04) and empirical calibrations, with the theoretical calibrations giving in general higher

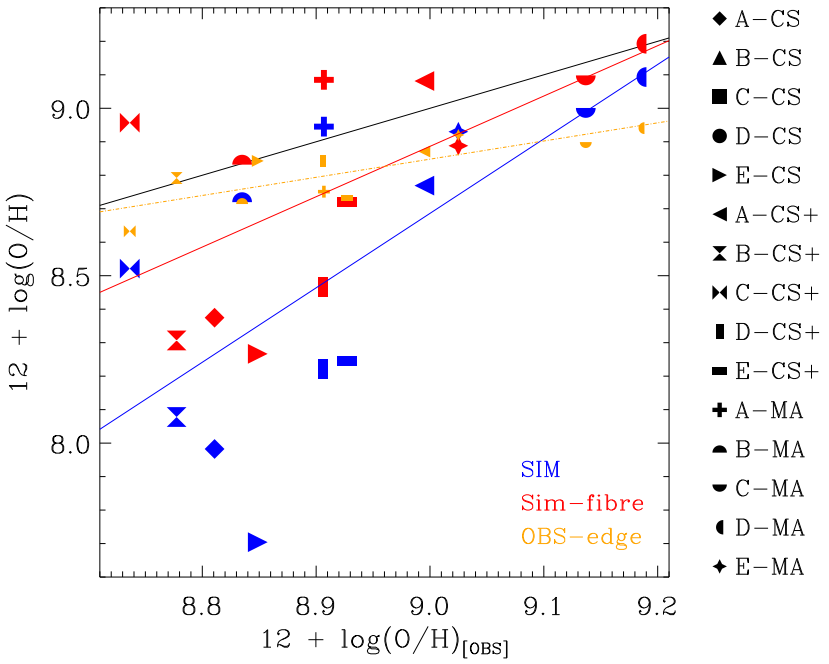

Figure 4.12: Comparison of the gas metallicities extracted using the different methods, in the one-to-one relation (black solid line) with the metallicity derived applying the T04 calibration to the emission lines ratios of the simulated SUNRISE face-on (OBS) and edge-on (OBS-edge) spectra.

metallicities (see also Chapter 3)

4.3.5 Star formation rates

We analyse in this section the SFRs of our simulated galaxies, and compare them to the SDSS-Garching data. The method used to calculate the total SFRs from the SDSS spectra is described in Brinchmann et al. (2004) and is based on the CL01 model, correcting for the limited fibre size of the spectrograph with the technique described in Salim et al. (2007). For our simulated galaxies, we estimate the SFRs following these procedures:

- **OBS [Hα]**: we extract the Hα-luminosity L(Hα) from the SUNRISE face-on (edge-on for the OBS-edge method) spectra that, after correcting for dust extinction with the Calzetti law using the Hα/Hβ ratio, we convert into SFR according to the Kennicutt calibration, taking into account with the factor $f_{\rm IMF} = 1.5$ the use of Kroupa/Chabrier IMF (Calzetti et al., 2009):

$$\mathrm{SFR}\ (\mathrm{M}_\odot\,/\,\mathrm{yr}) = 7.9/f_{\rm IMF} \times 10^{-42}\, L(H\alpha)\,(\mathrm{erg}\,/\,\mathrm{s})$$

  Note that this method, although different from the one used in SDSS analysis, has been shown in Brinchmann et al. (2004) to be in good agreement with it, at least for our range of stellar masses. Note also that, in addition, the method is sensitive (together with the BC03-ionizing flux) to the emission from young massive stars with lifetimes

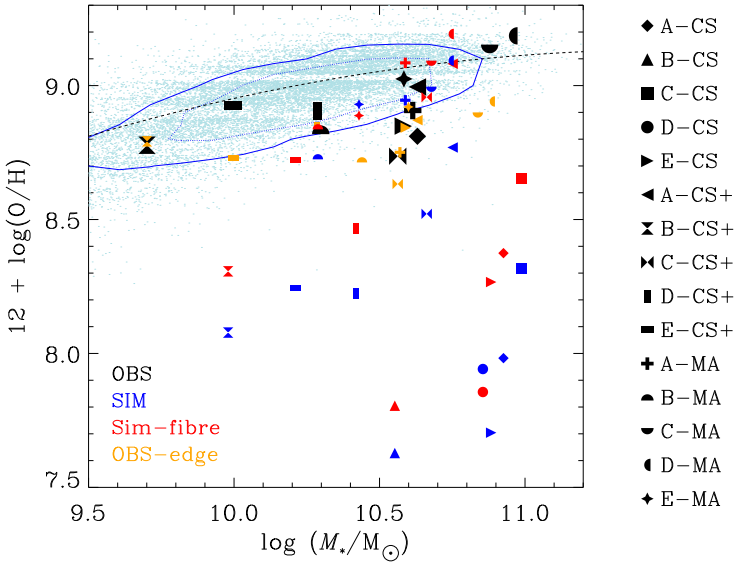

Figure 4.13: Mass-metallicity diagram for SDSS galaxies and simulations. The dashed line is the analytic mass-metallicity relation from Tremonti et al. (2004), while the blue contours encloses respectively 50% (dotted) and 80% (solid) of the datapoints. OBS and OBS-edge methods closely mimic SDSS biases, using the T04 calibration applied to the spectra inside the fibre FoV respectively in face-on and edge-on projections.

$\lesssim$ 10 Myr (Calzetti, 2008), while the SFR derived with the SIM method is averaged over a larger timescale (0.2 Gyr).

- **SIM**: we calculate the SFR directly from the simulation's snapshots, considering the amount of total stellar mass formed over a certain time interval, that we set to the last 0.2 Gyr.

- **BC03**: we convert the rate of ionizing photons $Q(H^0)$ calculated with BC03 into SFR according to the calibrations given in Kennicutt (1998a):

$$\mathrm{SFR}\ (\mathrm{M}_\odot\,/\,\mathrm{yr}) = 1.08/f_{\mathrm{IMF}} \times 10^{-53}\, Q(H^0)\ (\mathrm{s}^{-1})$$

In Fig. 4.14 we show the estimations of the SFR using the different methods in the one-to-one relation with the OBS results. We find in general a tight agreement among them, with scatter of the order of $\lesssim 0.2 - 0.4$ dex, and only the lowest-SFR galaxy (D-CS) has a significantly different SIM value compared to the OBS estimator. Projection effects do not strongly affect the derived SFRs, as evidenced by the similar relation found in the case of the OBS-edge method (note that both OBS and OBS-edge are corrected for dust extinction, and sample the full field of view of 60 × 60 kpc).

The linear functions that best-fit the SIM/BC03/OBS-edge datapoints, and the values of the correlation factors R, are:

$$\mathrm{SFR}_{[\mathrm{SIM}]} = 1.11 \times \mathrm{SFR}_{[\mathrm{OBS}]} - 0.01 \quad [\mathrm{M}_\odot/\mathrm{yr}]$$
$$R_{\mathrm{SIM}} = 0.871$$

$$\mathrm{SFR}_{[\mathrm{BC03}]} = 0.90 \times \mathrm{SFR}_{[\mathrm{OBS}]} + 0.01 \quad [\mathrm{M}_\odot/\mathrm{yr}]$$
$$R_{\mathrm{BC03}} = 0.921$$

$$\mathrm{SFR}_{[\mathrm{OBS\text{-}edge}]} = 0.90 \times \mathrm{SFR}_{[\mathrm{OBS}]} - 0.07 \quad [\mathrm{M}_\odot/\mathrm{yr}]$$
$$R_{\mathrm{OBS\text{-}edge}} = 0.987$$

Note that the BC03 points appear to be in a better agreement with OBS compared to SIM, with slightly smaller scatter (i.e. higher correlation coefficient).

It is worth noting that both BC03 and OBS are based on the conversion of the rate of ionizing photons into SFR and hence sample the same timescale of star formation, while for the SIM method we consider a much longer timescale, as star formation is treated stochastically in the simulations. This however may result in significant discrepancies with the observational estimators, in particular in the presence of recent starbursts (Sparre et al., 2015). We have tested this effect using a timescale of 10 Myr for the SIM method; in this case we obtain a similar fit, with slope and zero points of 1.103 and -0.061, respectively, but a higher correlation factor of $R_{\mathrm{SIM}} = 0.973$.

In Fig. 4.14, we additionally show the SFRs derived from Hα without correcting for dust, to provide a visual impression of the amount of dust extinction in our simulations, both in the face-on and edge-on projections. We see that that for most of our galaxies the effects of dust are small on the face-on spectra, while dust affects the edge-on views more significantly, particularly for the galaxies with higher metal content (MA sample).

In Fig. 4.15 we compare the SFRs of simulated galaxies obtained with the OBS method to observations. The majority of the simulated galaxies have SFRs consistent with those of green valley galaxies, and only A-MA and E-CS$^+$ are in the area of the actively star-forming spirals. On the contrary, B-CS and D-CS have low SFRs, closer to the ones of red ellipticals. As shown in the previous figure, the position of the galaxies in the SFR-stellar mass diagram when different methods are applied are similar, although in some cases the use of the OBS method (which includes the Petrosian masses) moves the simulations towards the range of real galaxies (D-CS), or from the red sequence to the green valley (e.g. B-CS) due to the different timescales over which the SFR is derived with the OBS and SIM methods.

The Specific Star Formation Rate (sSFR) diagram (Fig. 4.16) confirms the trends of Fig. 4.15, with most of the galaxies in the transition between the green valley and the blue sequence, although now more galaxies appear in the blue sequence area (note that the diagram is divided in three regions by the definition of green valley galaxies, see Sec. 4.2).

We conclude that *the estimation of the SFR in simulations is not strongly affected by which method is used, and only mildly by the projection if the spectra are corrected for dust extinction*; the values extracted directly from the simulations can be meaningfully compared with observations, at least for normal star-forming galaxies. Most of our simulated galaxies have SFRs at the transition between the green valley/blue sequence of SDSS galaxies. The good agreement between the star formation rate in simulations and the one extracted from an observational indicator such as the Hα flux is in general found for other star formation rate proxies, such as the [OII] line intensity (Jonsson, Groves & Cox, 2010) or the IR luminosity (Hayward et al., 2014; Hayward & Smith, 2015), at least for quiescent star-forming galaxies.

4.4 Discussion and conclusions

We have made an unbiased comparison between simulated and observed galaxies, converting the simulation outputs into synthetic observations and applying observational techniques

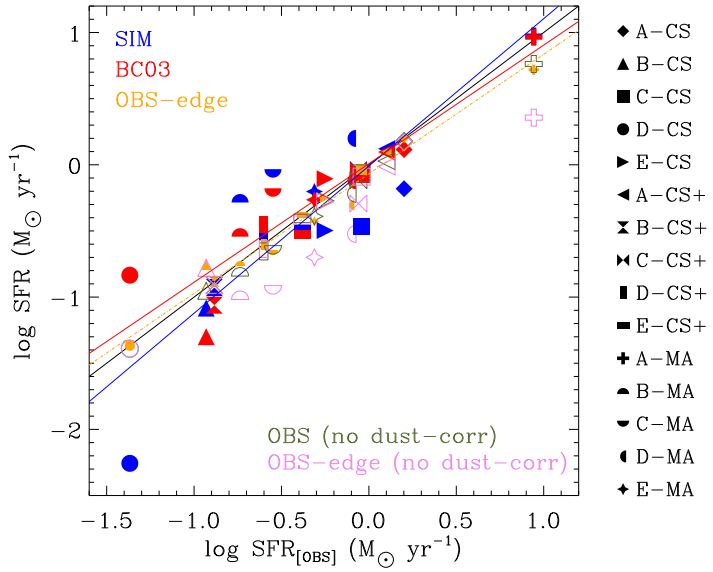

Figure 4.14: Comparison of the different SFR estimators, in the one-to-one relation (black line) with the observational method (OBS) based on Hα luminosity. The best-fits of both SIM and BC03-ionizing flux (solid blue and red lines) are in good agreement with the relation, as well as the values derived applying the observational method in the edge-on projection (OBS-edge). The results not corrected for dust show the effect of dust extinction on the spectra for these simulations.

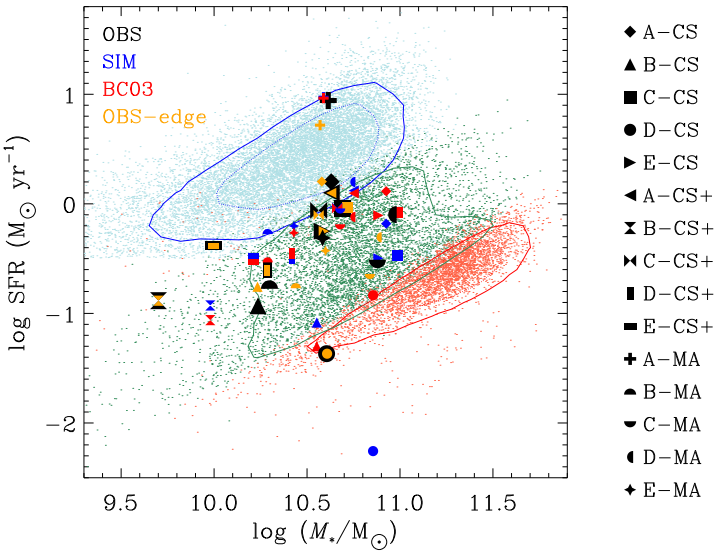

Figure 4.15: SFR-stellar mass diagram for SDSS and simulated galaxies, where the SDSS galaxies are shown in blue, green and red according to morphological classification, with their respective 50% (dotted) and 80% (solid) contours. For each simulation we plot the results of the different methods.

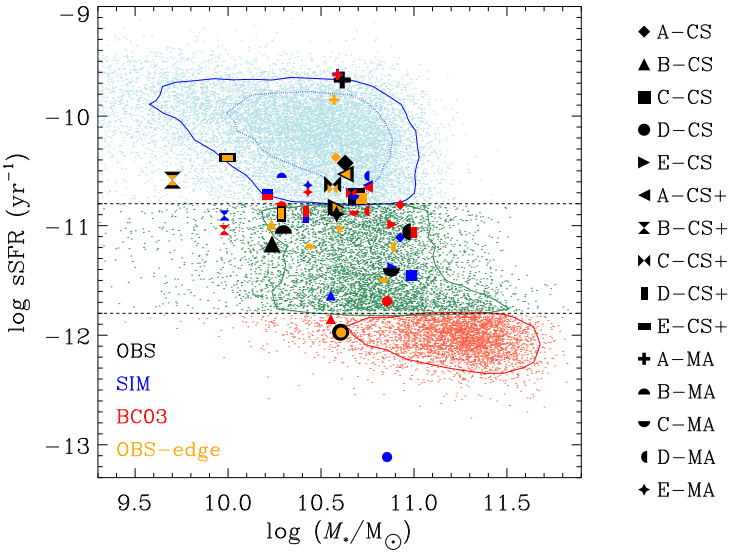

Figure 4.16: Specific SFR−stellar mass diagram showing SDSS data and the 50% and 80% contours of each morphological type (spirals, green valley galaxies and ellipticals respectively in blue, green and red), together with the values extracted from the simulated galaxies using different techniques.

to derive the galaxies' magnitudes, colours, stellar masses, mean stellar ages, stellar and gas metallicities and star formation rates. We have used 15 hydrodynamical simulations of galaxies formed in a cosmological context adopting three different models for chemical enrichment and feedback, and we compared their properties with data from the Sloan Digital Sky Survey (SDSS). In order to extract the physical properties of the simulated galaxies, we first created synthetic spectra, using both a SPS model (BC03) and a full radiative-transfer code (SUNRISE) to post-process the snapshots.

In Chapter 3 we compared the properties of the galaxies obtained applying different observational methods, and we have shown that large variations can appear, most notably in the case of the galaxies' ages and metallicities. Here we focused on the methods that mimic the SDSS techniques, in order to make an unbiased comparison with the observational dataset. In particular, we studied which physical properties are more affected by the observational biases that, in the case of SDSS, are mainly originated by the limited fiber size, by the methods applied to recover ages and metallicities, and by the use of the Petrosian quantities that affect both the magnitudes and stellar mass estimation.

We have given simple scalings to convert the direct results of simulations into values that can be compared with the SDSS dataset in a reliably manner, although in some cases (most notably in the mean stellar ages of galaxies) the correlation has a large scatter. Moreover, the scalings we found might depend somehow on the particular hydrodynamical code we used to simulate the galaxies, which we partially tested by applying three different versions of chemical enrichment and feedback. In addition, our results are also sensitive to some choices in the derivation of the observables (e.g. inclination, as shown by the results derived in the edge-on projections), and these caveats should be taken into account when the provided scaling relations are used.

We found that the biases that appear when observational techniques are applied affect differently the various galaxy properties that we studied here. In particular, for the colours and magnitudes, mimicking observational techniques has a small effect, and the direct results of simulations can be reliably compared with SDSS data. Stellar masses derived fitting the photometry show some discrepancies with respect to the stellar masses in simulations, although they are small except in the cases where mass loss due to stellar evolution is not properly modelled in the hydrodynamical codes. In contrast, in the case of stellar ages and stellar and gas metallicities the effects are stronger. For stellar ages and metallicities, the values of the simulations following observational techniques predict younger and more metal-poor galaxies compared to their mass-weighted values. The discrepancy among the methods increases both for young and for metal-poor galaxies. Refining the direct calculation by weighting with the luminosity of the stellar particles improves the mean age and metallicity determination, although strong trends with respect to the observational method appear.

For the mean oxygen abundance of the gas, we find that applying the SDSS metallicity calibration (T04) to the spectra in the fiber makes the simulated galaxies to be in much better agreement with observations compared to a direct calculation. Ignoring the fibre bias makes the galaxies appear more metal-poor, the effects being stronger for galaxies with steeper metallicity gradients. This has been already investigated in several studies (e.g. Kewley & Ellison 2008; Tremonti et al. 2004), which found that the limited fiber size strongly affects the determination of the gas metallicity and the shape of the mass-metallicity relation, particularly for galaxies with masses $M_* > 10^{10} M_\odot$. Our findings show that, when simulated spectra are not available, the comparison of simulations' gas metallicity with SDSS data can be improved mimicking the fibre size of the SDSS spectrograph.

The determination of the star formation rates is the quantity less affected by the method used, and the values extracted from the simulations and the Hα flux can be meaningfully compared (if dust extinction is correctly taken into account). We also found that the effects of the projection are significant for quantities such as the concentration and Sérsic indices, while projection has smaller effects when physical quantities are corrected for dust extinction such as in the case of gas metallicity and SFR.

Our results show that, when an unbiased comparison with the SDSS data is performed, our simulated galaxies:

(i) look photometrically similar to SDSS blue/green valley galaxies,
(ii) have concentrations and Sérsic indices mostly in the range of SDSS galaxies, even though the different feedback codes give different results and in some cases outside the observed range,
(iii) are in good agreement with SDSS ages, although most of them appear older compared to SDSS spirals,
(iv) have stellar metallicities consistent with metal-poor spirals,
(v) show good agreement with observations of the gas oxygen abundances, even if they remain slightly more metal-poor,
(vi) have Hα-based SFRs in the region between the SDSS green valley galaxies and the blue sequence, although there are objects with Hα-based SFRs both in the region of strongly star-forming spirals and in the red sequence of passive ellipticals.

In summary, we have shown that *a reliable comparison between observations and simulations requires in general the conversion of the direct results of simulations into observationally-derived quantities* taking into account the biases of the survey and mimicking its algorithms. A consistent comparison of the galaxy properties is the only possible way to reliable test the recipes for star formation, feedback and metal enrichment in simulation codes.

Chapter 5

Generating synthetic observations of CALIFA galaxies from hydrodynamical simulations

In this Chapter we consider a different way to combine observations and simulations in galaxy formation studies, by converting hydrodynamical simulations of galaxies in cosmological context into a set of reliable synthetic Integral Field Spectroscopy (IFS) observations, reproducing the properties of the Calar Alto Legacy Integral Field Area (CALIFA) survey (Sánchez et al., 2012). IFS is now becoming an important discipline in observational astronomy (see the review by Allington-Smith 2007), as it allows to recover spatially-resolved spectra of objects in the sky over a two-dimensional field, substantially improving previous techniques such as long-slit or fiber spectroscopy, in which the spatial resolution is obtained only in the dimension along the slit or inside the region covered by the fiber. Studying the properties of galaxies with the additional spatial information is important in order to obtain a more comprehensive view of a galaxy's formation and evolution, as the presence of gradients or regions with different spectral features is missed using spectrographs that are able to recover only a single spectrum for each object.

However, the technical possibilities offered by the new Integral Field Unit spectrographs such as MUSE, (Bacon et al., 2004), WEAVE, (Dalton et al., 2014) and PMAS (Roth et al., 2005) require also a higher level of complexity in the observational algorithms used in the data analysis, as well as additional data processing, for instance the spatial segmentation (i.e. binning) of the spectra to achieve a sufficient S/N. In order to provide observers a tool to reliably test the accuracy of this new class of observational algorithms (in particular the ones used in CALIFA), we create from the hydrodynamical simulations realistic synthetic spatially-resolved spectra (3-dimensional data or datacubes) mimicking the most relevant observational effects of the CALIFA survey, as well as resolved maps of the properties of our simulations, that can be used to compare the results of a given IFS analysis algorithm applied to our synthetic dataset.

5.1 Hydrodynamical simulations

To produce our mock data sample we use three hydrodynamical simulations of galaxies in a ΛCDM Universe, generated from a dark-matter simulation with the zoom-in technique (Tormen, Bouchet & White, 1997). The initial conditions for the hydrodynamical simulations are taken from the Aquarius dark-matter only simulation (Springel et al., 2008), identifying

at redshift zero halos (as defined by the SUBFIND halo finder algorithm, Springel et al. 2001) that are possible candidates for the formation of galaxies with properties similar to the Milky Way, with virial mass (calculated as the mass within the radius where the density is $200\,\rho_c$) between 0.7 and 1.7×10^{12} $M_\odot$, and a quiet merger history in the recent past, excluding halos with neighbours more massive than half of their mass within a spherical region of 1.4 Mpc radius at $z = 0$ (see Scannapieco et al. 2009 for details). The cosmological parameters assumed are the following: $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_b = 0.04$, $\sigma_8 = 0.9$, and $H_0 = 100\,h$ km s^{-1} Mpc^{-1} with $h = 0.73$. The simulations have, at redshift $z = 0$, mass resolution of $1 - 2 \times 10^6$ $M_\odot$ for dark matter particles and of $2 - 5 \times 10^5$ $M_\odot$ for stellar/gas particles, and gravitational softening of $300 - 700$ pc.

Two halos, that we name C-CS$^+$ and E-CS$^+$ (where the first letter identifies the Aquarius halos according to the Springel et al. convention) have been simulated with a new version of Scannapieco et al. (2005, 2006) model (CS hereafter), which implements chemical enrichment and Supernovae (SNe) feedback in the Tree-PM SPH code Gadget-3 (Springel, 2005). The main changes in the updated version (CS$^+$ model, Poulhazan et al., in prep.) concern the use of new metal yields (from Portinari, Chiosi & Bressan 1998) including chemical enrichment from AGB stars by Portinari et al. (1998); Marigo (2001), a new IMF (Chabrier 2003), while the assumed cooling function is the one by Sutherland & Dopita (1993) as in the original Scannapieco et al. model.

The third halo (D-MA) is simulated with the Aumer et al. (2013) independent update (MA hereafter) to the CS model. MA model is different from CS in the chemical yields (which also include AGB stars contribution) with the additional modelling of metal diffusion in the ISM, in the use of a Kroupa IMF, and in the cooling function, which is taken from Wiersma, Schaye & Smith (2009). More important, the energy feedback from SNe is, unlike in the CS$^+$ model where feedback is purely thermal, divided into a thermal and a kinetic part, and the code also includes the feedback on the ISM of the radiation pressure due to massive young stars. The MA model gives in general stronger feedback compared to CS/CS$^+$, and hence younger, more metal rich and disk-dominated galaxies (for details on this model we refer the reader to Aumer et al. 2013).

5.1.1 Properties of the simulated galaxies

Measurements of some of the integrated properties of the simulated objects are listed in Table 5.1. The global properties of these galaxies (already derived in Guidi, Scannapieco & Walcher 2015; Guidi et al. 2016) have been computed considering the particles belonging to the main halo, in a 60 kpc×60 kpc region with the galaxy in the centre. The galaxies are oriented face-on according to the direction of the total angular momentum.

We calculate the following global properties:

- **Total stellar mass:** the mass in stars (in units of $\log[M_*/M_\odot]$) is computed considering the stellar particles in the simulations inside the 60 kpc×60 kpc central region.
- **$r-$band absolute magnitude:** we calculate the absolute magnitudes of each galaxy in the $r-$band convolving the total (face-on) spectrum generated with the SUNRISE radiative transfer code (Sec. 5.2) with the $r-$band filter (Gunn et al., 1998, 2006).
- **Mean stellar age:** we derive the global mean stellar age weighting both by the mass of the stellar particles ($\log\langle age\rangle_M$), and by the luminosity in the r-band calculated with the Bruzual & Charlot (2003) SPS model ($\log\langle age\rangle_L$); the units are $\log$ [yr][1].

[1]Notice that in this work we use arithmetic means both for the ages and metalliticies (Asari et al.,

Name	Total mass	Absolute magnitude	Stellar age (log [yr])		Stellar metallicity (log $[Z/Z_\odot]$)		v_{disp}	Gas metallicity
	$\log(M_*/M_\odot)$	($r-$band)	$\log\langle age\rangle_M$	$\log\langle age\rangle_L$	$\log\langle Z\rangle_M$	$\log\langle Z\rangle_L$	[km/s]	[12+log (O/H)]
C-CS$^+$	10.66	-21.22	10.01	9.93	-0.39	-0.37	94.2	8.52
E-CS$^+$	10.21	-20.13	10.00	9.91	-0.44	-0.49	62.4	8.24
D-MA	10.75	-21.83	9.84	9.68	-0.19	-0.05	65.8	9.09

Table 5.1: Global properties of the simulated galaxies used to generate the CALIFA mock datacubes. These properties have been calculated in a 60 kpc×60 kpc region and for face-on orientation. Edge-on values differ from the ones presented here, and can be found in Guidi, Scannapieco & Walcher (2015) together with several other physical properties, while in Guidi et al. (2016) these galaxies have been compared with the Sloan Digital Sky Survey dataset (Abazajian et al., 2009).

$$\log\langle age\rangle_M = \log\left[\frac{\sum_i M_i \cdot age_i}{\sum_i M_i}\right] \quad (5.1)$$

$$\log\langle age\rangle_L = \log\left[\frac{\sum_i L_i \cdot age_i}{\sum_i L_i}\right] \quad (5.2)$$

- **Mean stellar metallicity:** the global mean stellar metallicity is calculated weighting by the mass ($\log\langle Z\rangle_M$) and by the luminosity in the r-band ($\log\langle Z\rangle_L$); both are in logarithmic solar units with $Z_\odot = 0.02$.

$$\log\langle Z\rangle_M = \log\left[\frac{\sum_i M_i \cdot Z_i}{\sum_i M_i}\right] \quad (5.3)$$

$$\log\langle Z\rangle_L = \log\left[\frac{\sum_i L_i \cdot Z_i}{\sum_i L_i}\right] \quad (5.4)$$

- **Velocity dispersion:** we compute the velocity dispersion in the face-on projection as

$$V_{\rm disp} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(v_i - \bar{V})^2} \quad (5.5)$$

  where $\bar{V}$ is the mean velocity. The units are km $\cdot$ s^{-1}.

- **Mean gas metallicity:** we derive the mean oxygen abundance of the gas 12 + log(O/H) as the mean of the (O/H) ratio of each gas particle.

Together with the global properties described above we also calculate the spatially-resolved ones, i.e. considering the particles enclosed in the physical size covered by each spaxel of our "virtual" CALIFA observations (see Section 5.3.1). We refer to these maps of directly measured properties (listed in tab. 5.2) as *product datacubes* (some examples can

2007; Cid Fernandes et al., 2013). A different definition often found in the literature is the geometric mean $\langle \log age\rangle_M = \frac{\sum_i M_i \cdot \log age_i}{\sum_i M_i}$ (e.g. Gallazzi et al. 2005; González Delgado et al. 2014, 2015; Sánchez et al. 2016c). Since we will provide these quantities smoothed by the CALIFA spatial PSF (Sec. 5.3.3), we choose to weight the linear quantities in order to avoid biases in the calculation of the smoothed properties.

be seen in Figure 5.1), that represent the 'solutions' to be recovered by the observational algorithms.

We describe now these calculations:

- **Stellar mass/Stellar mass density:** the stellar mass and stellar mass density maps have been derived from the simulation snapshots, considering the amount of stellar mass in the region corresponding to each spaxel; the units of the maps are $\log[M_*/M_\odot]$ and $\log[M_*/(M_\odot\,\mathrm{pc}^2)]$ respectively.

- **Mass-weighted mean stellar age:** we compute the mass-weighted mean stellar ages as the logarithm of the mean of the ages in each spaxel ($\log\langle age\rangle_M$, eq. 5.1); the units are $\log[\mathrm{yr}]$.

- **Mass-weighted mean stellar metallicity:** we derive the mean mass-weighted stellar metallicities ($\log\langle Z\rangle_M$, eq. 5.3), where Z_i is the metalllicty of a stellar particle in solar units (with $Z_\odot = 0.02$) and we provide the logarithm of the mean metallicity in each spaxel. The units are $\log[Z/Z_\odot]$.

- **Luminosity-weighted mean stellar age:** we compute the luminosity-weighted mean stellar age weighting the ages by the flux L_i of each star particle at $\lambda = 5635$ Å calculated with STARBURST99[2] (eq. 5.2). We store the logarithm of the mean age in units $\log[\mathrm{yr}]$.

- **Luminosity-weighted mean stellar metallicity:** to derive the mean luminosity-weighted stellar metallicity ($\log\langle Z\rangle_L$, eq. 5.4), we weight the metalllicty of a stellar particle Z_i (in solar units assuming $Z_\odot = 0.02$) by the luminosity L_i calculated at 5635 Å using STARBURST99 SPS model. It is in units $\log[Z/Z_\odot]$.

- **Mean velocity/Velocity dispersion:** the mean velocity v_{mean} and velocity dispersion v_{disp} maps (both in units of $\mathrm{km\cdot s^{-1}}$) have been derived weighting the line-of-sight velocity of the stellar particles inside the region sampled by each spaxel by their luminosity at 5635 Å, calculated with STARBURST99.

$$v_{\mathrm{mean}} = \frac{\sum_i L_i \cdot v_i}{\sum_i L_i} \tag{5.6}$$

$$v_{\mathrm{disp}} = \sqrt{\frac{\sum_i L_i \cdot (v_i - v_{\mathrm{mean}})^2}{\sum_i L_i}} \tag{5.7}$$

- **Star formation rate:** the maps of the spatially-resolved SFRs are generated from the simulation snapshots, considering the amount of stellar mass formed in the previous 10 Myr, on a timescale similar to the one sampled by most of the observational indicators (Kennicutt, 1998a); it is in units of $[M_\odot\ \mathrm{yr}^{-1}]$.

[2]To calculate luminosity-weighted quantities in the product datacubes we use the flux at $\lambda = 5635$ Å following the choice done in several studies of the CALIFA galaxies, e.g. González Delgado et al. (2014); Sánchez et al. (2016b); Ruiz-Lara et al. (2016).

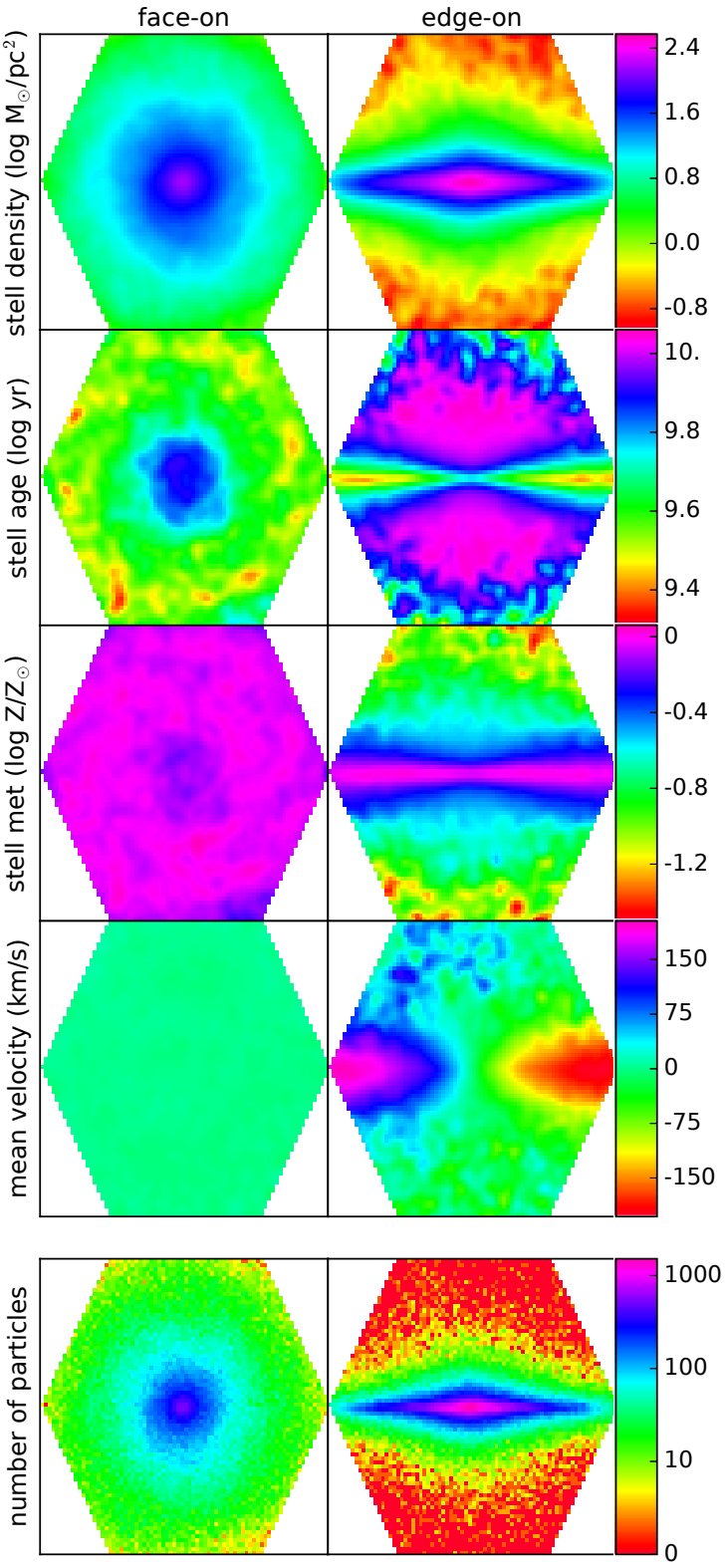

Figure 5.1: Spatially-resolved stellar properties of two simulated galaxies, D-MA_0 (face-on) on the left and D-MA_2 (edge-on) on the right. These maps show, from top to bottom, the stellar mass density, the mean luminosity-weighted ages and metallicities, the mean velocity along the line of sight, and the number of stellar particles in each spaxel (in logarithmic colour scale).

Stellar property	Units
Mass	$\log(M_\odot)$
Mass density	$\log(M_\odot/pc^2)$
Mean age mass-weighted	log(yr)
Mean metallicity mass-weighted	$\log(Z/Z_\odot)$
Mean age luminosity-weighted	log(yr)
Mean metallicity luminosity-weighted	$\log(Z/Z_\odot)$
Mean velocity	km/s
Velocity dispersion	km/s
Star formation rate	$M_\odot$/yr

Table 5.2: List of the spatially-resolved stellar properties provided in the product datacubes.

5.2 Simulated spectra

In this section we describe the procedure followed to generate the spatially-resolved spectral energy distribution of our simulated galaxies. To this end, we post-process the simulation snapshots at redshift zero with the Monte Carlo Radiative Transfer code SUNRISE (Jonsson, 2006; Jonsson, Groves & Cox, 2010), considering as input for the radiative transfer post-processing only the stellar particles belonging to the main halo in the hydrodynamical simulations, defined according to the SUBFIND algorithm. SUNRISE is a 3-D polychromatic Monte Carlo radiative transfer code, which is able to self-consistently simulate the emission and propagation of light in a dusty InterStellar Medium (ISM) from hydrodynamical snapshots, to obtain the full spatially-resolved UV-to-submillimetre Spectral Energy Distributions (SEDs). The resulting SEDs include the contribution of stellar and nebular emission, dust absorption and scattering, and hence show stellar absorption features, emission lines, as well as the effects of kinematics.

The procedure followed to obtain the spectra of the hydrodynamical simulations consists mainly in three distinct steps:

1. SUNRISE assigns a specific spectrum to every stellar particle depending on the age, metallicity, normalized by the mass of the particle. In particular, according to the age of the particle, two different model spectra are considered.

   - age > 10 Myr: spectra from the STARBURST99 Stellar Population Synthesis (SPS) model (SB99, Leitherer et al. 1999) are assigned to the stellar particles. To create the input stellar model we have selected the Padova 1994 stellar tracks (Fagotto et al., 1994a,b) assuming a Kroupa IMF (Kroupa, 2002) with $\alpha = 1.3$ for $m_{\rm star} = 0.1 - 0.5$ $M_\odot$ and $\alpha = 2.3$ for $m_{\rm star} = 0.5 - 100$ $M_\odot$. The low-resolution spectra (which have sampling ~ 20 Å) have been computed choosing the Pauldrach/Hillier stellar atmospheres, while for the high-resolution region of the STARBURST99 spectra (available only in the range $3000 - 7000$ Å with sampling of 0.3 Å) we have used the fully theoretical atmospheres by Martins et al. (2005). The final input stellar model is the combination of the low-resolution spectra for wavelengths $\lambda < 3000$ Å, $\lambda > 7000$ Å, and the high-resolution spectra in the range 3000 Å $\leq \lambda \leq 7000$ Å.
   - age ≤ 10 Myr: young stellar particles are assumed to be the source of significant amount of ionizing photons, which are efficiently absorbed by the surrounding gas

producing recombination lines and forming an HII region. To these young stellar particles a modified spectrum is assigned, that takes into account the effects of photo-dissociation and recombination of the gas. The spectra coming from the HII regions are pre-computed with the 1D photo-ionization code MAPPINGS III (Groves, Dopita & Sutherland, 2004; Groves et al., 2008) assuming spherical geometry for the surrounding gas, and depend on the metallicity of the stellar particle and of the gas around it, on the compactness parameter C (which in turn depends on the ISM pressure and on the chosen value of the cluster mass M_{cl} by equation 13 in Groves et al. 2008), and on the covering fraction f_{PDR}, which is the time-averaged fraction of stellar cluster solid angle covered by the Photo-Dissociation Region (PDR). We set the MAPPINGS free parameters f_{PDR} and M_{cl} to the fiducial values of $f_{\mathrm{PDR}} = 0.2$ and $\mathrm{M}_{cl} = 10^5\, M_\odot$ given by Jonsson, Groves & Cox (2010).

2. After SUNRISE assigns stellar or nebular spectra to all stellar particles, in the radiative transfer stage randomly-generated photon packets are propagated through the dusty ISM (assuming a constant dust-to-metals ratio of 0.4 according to Dwek 1998) with a Monte Carlo approach. Since in the multi-phase model of the ISM implemented in our hydrodynamical code (Scannapieco et al., 2006) each gas particle has a single temperature, density and entropy (while other ISM models may have cold/hot phases in a given gas particle, e.g. Springel & Hernquist 2003) the amount of dust is directly linked to the total amount of metals in each gas particle (for a discussion of the effects of the ISM sub-resolution structure on radiative transfer calculations see Hayward et al. 2011; Snyder et al. 2013; Lanz et al. 2014). Dust extinction is described by a Milky Way-like curve with $R_V = 3.1$ (Cardelli, Clayton & Mathis, 1989; Draine, 2003), while for dust scattering the phase function by Henyey & Greenstein (1941) is adopted. The $\sim 10^7$ Monte Carlo rays generated by the SUNRISE algorithm are traced on an adaptive grid made by $\sim 30\,000 - 150\,000$ cells covering a region of $(120\ \mathrm{kpc})^3$ with minimum cell size of ~ 250 pc, calculated assuming in SUNRISE a value of tolerance $\mathrm{tol}_{met} = 0.1$ and V-band metals opacity for unit mass of metals $\kappa = 3 \times 10^{-5}\ \mathrm{kpc}^2\ \mathrm{M}_\odot^{-1}$ (see Jonsson 2006 for details).

3. The model cameras placed around the simulated galaxies obtain the SED in each pixel. In our calculations we place cameras with three different orientations (defined according to the alignment of the total angular momentum of the stars with the z direction) for each galaxy, respectively face-on, 45° and edge-on. The flux in the cameras may be convolved with bandpass filters to get broadband magnitudes and images as in Fig. 5.2, which shows the (u, r, z)-band colour-composite images of the simulated galaxies in the three orientations; the region of the simulations observed by the CALIFA hexagonal field of view is in red in this figure.

In order to reduce the random noise introduced by the SUNRISE Monte Carlo algorithm, the radiative transfer process described above is run ten times for each galaxy, changing only the random seeds, and the resulting spectra are averaged over the ten different random realizations In this way we are able to reach a 'signal-to-noise' S/N (where N is the standard deviation over the ten realizations) of $\sim 300 - 400$ in the central spaxels and S/N$\sim 8 - 10$ in the outskirt regions, which is negligible compared to the typical values of the S/N in the CALIFA spectrograph that we aim to mimic (Sánchez et al., 2012).

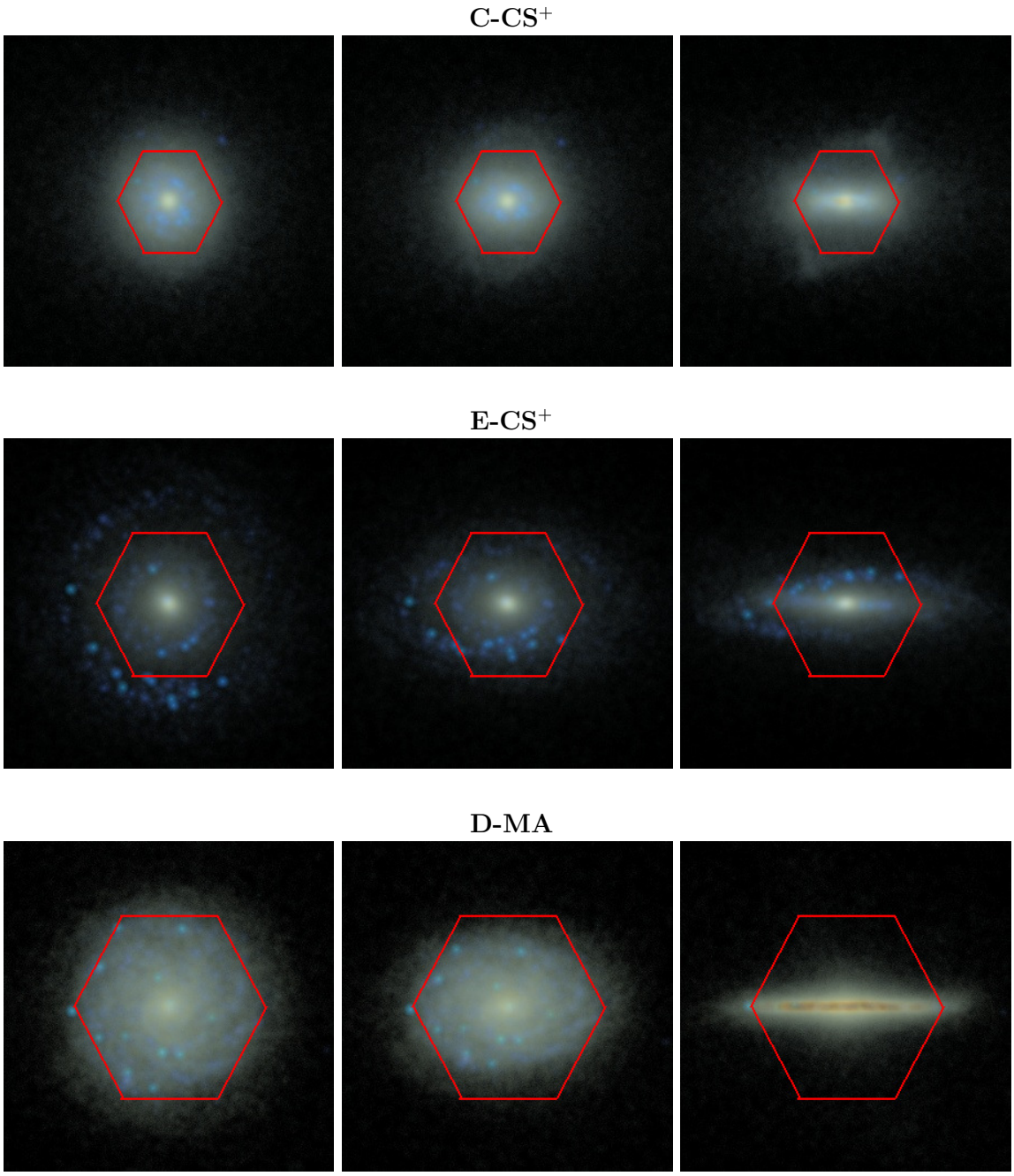

Figure 5.2: Composite synthetic broadband images created in the (u, r, z)-bands using the Lupton et al. (2004) composition algorithm for our three simulated galaxies (from top to bottom, C-CS$^+$, E-CS$^+$, and D-MA) in a field of view of 60×60 kpc with 300×300 pixels. The orientations are, from left to right, face-on, 45° and edge-on, labelled in the synthetic datacubes as _0, _1 and _2 respectively. The red hexagon is the region of the simulations sampled by the CALIFA field of view, with physical sizes of $\sim 19, 27, 35$ kpc respectively for the C-CS$^+$, E-CS$^+$ and D-MA galaxies.

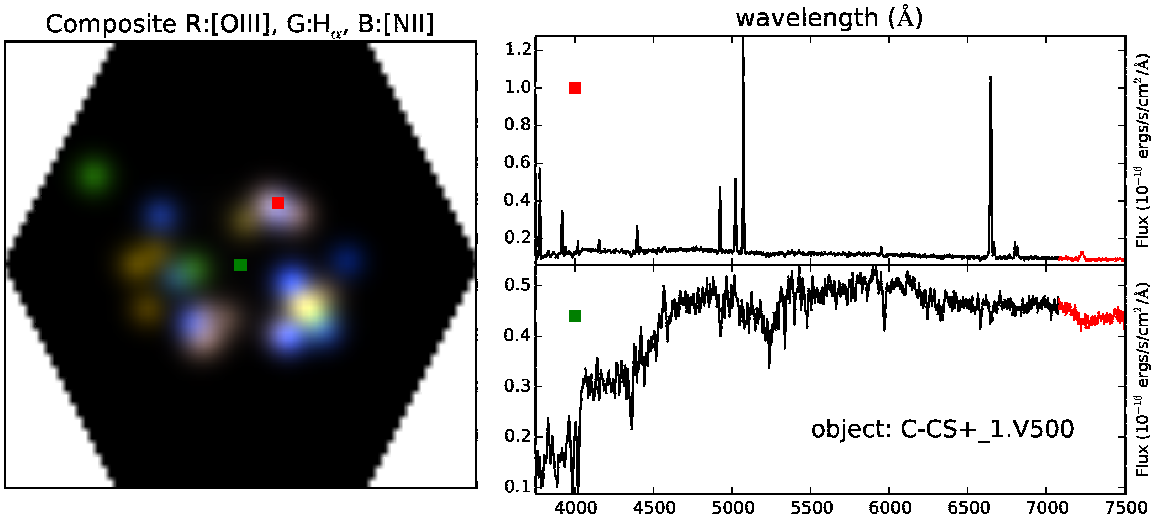

Figure 5.3: Left panel: RGB image of the [OIII]5007, Hα and [NII]6584 emission lines for the galaxy C-CS$^+$_1. Right panels: synthetic spectra in the spaxels corresponding to the red and green squares in the RGB image. In the upper right panel, the spaxel samples a nebular region (red square), while the lower right panel shows a V500 spectrum containing only stellar emission (green square). The part of the spectrum generated with the low-resolution stellar model (Sec. 5.2) is marked in red in the plot.

5.2.1 Measurements on the simulated spectra

As part of the product datacubes we also provide resolved maps of some spectral features as derived from noiseless datacubes (i.e. prior to the addition of the detector noise, see Section 5.3.4). To obtain measurements of the properties of the stellar and nebular spectra without using any observational algorithm to separate the two spectral components (which may introduce many caveats and uncertainties), we additionally generate *stellar-only* synthetic datacubes following the same procedure described in Sec. 5.2, but switching off the nebular contribution. These stellar-only synthetic datacubes are then subtracted from the full datacubes, to obtain the spatially-resolved *nebular-only* spectra.

- **Lick indices:** we derive the strength of the Lick stellar absorption features from the stellar-only datacubes. The spectra have 2 Å sampling (mimicking the V500 setup, Section 5.3.2), while the CALIFA spectral PSF has been not included in this calculation. The list of the absorption features provided in the product datacubes is given in Table 5.3[3].

- **Nebular emission line intensities:** the fluxes of the emission lines listed in Table 5.4 are measured from the *nebular-only* synthetic datacubes (see above). We compute the flux of each emission line as the total flux in the *nebular-only* spectra between the lower and upper bounds of each line (given in Table 5.4). Note that the *nebular-only*

[3]Notice that the Lick indices depend also on the velocity dispersion at which they are measured (see e.g. Sánchez-Blázquez et al. 2006; Oliva-Altamirano et al. 2015). In IFS observational studies the spectra in each spaxel are usually broadened to a single velocity dispersion prior to the measurement of the Lick indices, in order to consistently compare them with models with the same dispersion (e.g. Wild et al., 2014). In our product datacubes we do not change the broadening of the absorption lines, since this procedure introduces additionally uncertainties in the analysis. The spaxel-by-spaxel velocity dispersion is provided in the GALNAME.stellar.fits files (Sec. 5.4.1) and can be used to tune the fitted models.

Name	Index Bandpass (Å)	Blue continuum bandpass (Å)	Red continuum bandpass (Å)	Units	Reference
CN_1	4142.125 - 4177.125	4080.125 - 4117.625	4244.125 - 4284.125	mag	Worthey et al. (1994)
CN_2	4142.125 - 4177.125	4083.875 - 4096.375	4244.125 - 4284.125	mag	Worthey et al. (1994)
Ca4227	4222.250 - 4234.750	4211.000 - 4219.750	4241.000 - 4251.000	Å	Worthey et al. (1994)
G4300	4281.375 - 4316.375	4266.375 - 4282.625	4318.875 - 4335.125	Å	Worthey et al. (1994)
Fe4383	4369.125 - 4420.375	4359.125 - 4370.375	4442.875 - 4455.375	Å	Worthey et al. (1994)
Ca4455	4452.125 - 4474.625	4445.875 - 4454.625	4477.125 - 4492.125	Å	Worthey et al. (1994)
Fe4531	4514.250 - 4559.250	4504.250 - 4514.250	4560.500 - 4579.250	Å	Worthey et al. (1994)
Fe4668	4634.000 - 4720.250	4611.500 - 4630.250	4742.750 - 4756.500	Å	Worthey et al. (1994)
$H\beta$	4847.875 - 4876.625	4827.875 - 4847.875	4876.625 - 4891.625	Å	Worthey et al. (1994)
Fe5015	4977.750 - 5054.000	4946.500 - 4977.750	5054.000 - 5065.250	Å	Worthey et al. (1994)
Mg_1	5069.125 - 5134.125	4895.125 - 4957.625	5301.125 - 5366.125	mag	Worthey et al. (1994)
Mg_2	5154.125 - 5196.625	4895.125 - 4957.625	5301.125 - 5366.125	mag	Worthey et al. (1994)
Mgb	5160.125 - 5192.625	5142.625 - 5161.375	5191.375 - 5206.375	Å	Worthey et al. (1994)
Fe5270	5245.650 - 5285.650	5233.150 - 5248.150	5285.650 - 5318.150	Å	Worthey et al. (1994)
Fe5335	5312.125 - 5352.125	5304.625 - 5315.875	5353.375 - 5363.375	Å	Worthey et al. (1994)
Fe5406	5387.500 - 5415.000	5376.250 - 5387.500	5415.000 - 5425.000	Å	Worthey et al. (1994)
Fe5709	5696.625 - 5720.375	5672.875 - 5696.625	5722.875 - 5736.625	Å	Worthey et al. (1994)
Fe5782	5776.625 - 5796.625	5765.375 - 5775.375	5797.875 - 5811.625	Å	Worthey et al. (1994)
Na D	5876.875 - 5909.375	5860.625 - 5875.625	5922.125 - 5948.125	Å	Worthey et al. (1994)
TiO_1	5936.625 - 5994.125	5816.625 - 5849.125	6038.625 - 6103.625	mag	Worthey et al. (1994)
TiO_2	6189.625 - 6272.125	6066.625 - 6141.625	6372.625 - 6415.125	mag	Worthey et al. (1994)
$H\delta_A$	4083.500 - 4122.250	4041.600 - 4079.750	4128.500 - 4161.000	Å	Worthey & Ottaviani (1997)
$H\gamma_A$	4319.750 - 4363.500	4283.500 - 4319.750	4367.250 - 4419.750	Å	Worthey & Ottaviani (1997)
$H\delta_F$	4091.000 - 4112.250	4057.250 - 4088.500	4114.750 - 4137.250	Å	Worthey & Ottaviani (1997)
$H\gamma_F$	4331.250 - 4352.250	4283.500 - 4319.750	4354.750 - 4384.750	Å	Worthey & Ottaviani (1997)
D4000_n		3850.000 - 3950.000	4000.000 - 4100.000		Balogh et al. (1999)

Table 5.3: List of the absorption line indices for which the strength in each spaxel is provided, together with the definition of the continuum and bandpass wavelength ranges.

datacubes have spectral sampling of 2 Å, and do not consider the effect of the spectral PSF. It is also important to emphasize here that the nebular emission in the datacubes is limited to the stellar particles younger than 10 Myr (HII regions), and we do not count on any other sources of ionizing photons. The line intensities are stored in units of 10^{-16} erg $\cdot$ s^{-1} $\cdot$ cm^{-2} (Section 5.4.1).

In Figure 5.3 we show an RGB image of the intensities (derived from the nebular maps) of the [OIII]5007, H_α and [NII]6584 emission lines, together with spectra in two different spaxels in the synthetic datacubes, one containing nebular emission and the other only stellar light. An example of these maps is given in Figure 5.4, where we show for one of our simulated galaxies the intensities of the BPT (Baldwin, Phillips & Terlevich, 1981) emission lines (H_α, H_β, [OIII]5007, [NII]6584), and the corresponding signal-to-noise maps (see Section 5.3.4 for the discussion of the detector noise implementation in the synthetic datacubes).

5.3 CALIFA mock observations[4]

In this section we describe how we convert the output of the SUNRISE radiative transfer algorithm into synthetic IFS observations mimicking the CALIFA survey (Sánchez et al., 2012; García-Benito et al., 2015b). The CALIFA observations were taken with the Potsdam Multi Aperture Spectrograph (PMAS, Roth et al., 2005), mounted on the Calar Alto 3.5 m telescope, utilizing the large hexagonal Field-Of-View offered by the PPak fiber bundle (Verheijen et al., 2004; Kelz et al., 2006). The final CALIFA Public Data Release (DR3, Sánchez et al. 2016a) consists of 667 galaxies[5].

[4]This section was written by Javier Casado.

[5]The DR3 is available at the URL http://califa.caha.es/DR3.

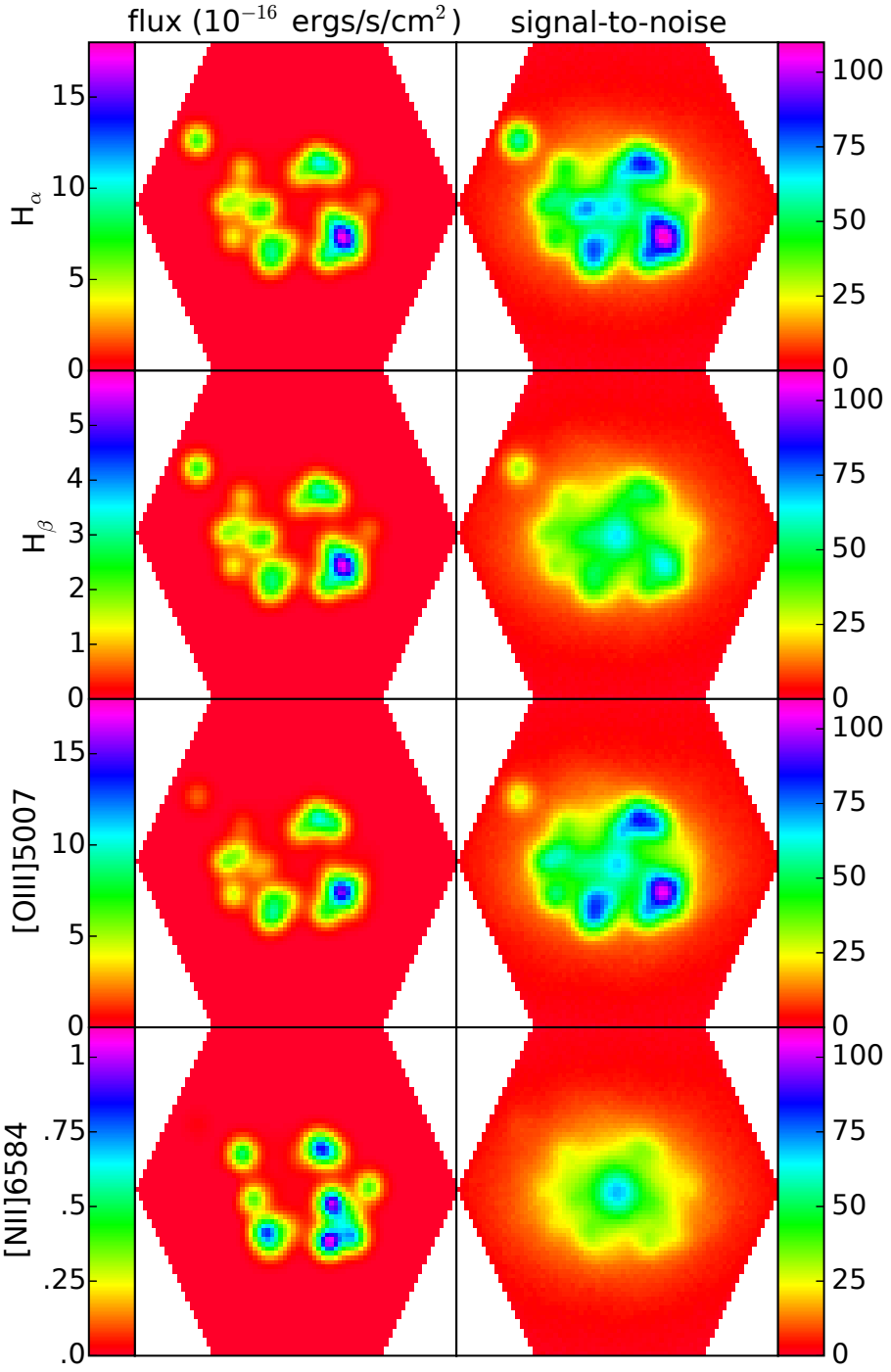

Figure 5.4: Line intensity (left) and Signal-to-Noise (right) maps of the four BPT lines (Baldwin, Phillips & Terlevich, 1981) for the galaxy C-CS$^+$_1. The S/N in every spaxel is obtained as the ratio between the mean signal and noise in the wavelength range of the corresponding emission line given in Table 5.4.

Species	Line center (Å)	Lower/ upper bounds (Å)
[Ne III]3869	3869.060	3859 – 3879
Hδ	4101.734	4092 – 4111
Hγ	4340.464	4330 – 4350
[O III]4363	4363.210	4350 – 4378
Hβ	4861.325	4851 – 4871
[O III]4959	4958.911	4949 – 4969
[O III]5007	5006.843	4997 – 5017
HeI 5876	5875.670	5866 – 5886
[N II]6548	6548.040	6533 – 6553
Hα	6562.800	6553 – 6573
[N II]6584	6583.460	6573 – 6593
[S II]6717	6716.440	6704 – 6724
[S II]6731	6730.810	6724 – 6744

Table 5.4: List of the emission line intensities provided in the product datacubes. Line centers, lower and upper bounds are taken from the Sloan Digital Sky Survey-Garching DR7 analysis (available at the URL http://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/).

In this section we present a summary of the technical properties of the CALIFA survey, and we describe how we reproduce the CALIFA observations in our synthetic datacubes.

5.3.1 Field-of-View and spaxel size

CALIFA sample selection criteria (Walcher, Wisotzki & Bekeraité, 2014) and the observing strategy (three points dither pattern, Sánchez et al. 2012) were carefully conceived to reach a filling factor of 100% across the Field-of-View (FoV) and guarantee to cover the entire optical extent of the galaxies up to ~ 2.5 effective radii R_e. The CALIFA datacubes (i.e. three-dimensional data) present a typical spaxel physical size of ~ 1 kpc, with a $77'' \times 73''$ 2D distribution[6] of spectra with $1'' \times 1''$ spatial sampling and 2.5" Full Width at Half Maximum (FWHM) spatial resolution (Sánchez et al., 2012). To reproduce in our simulated observations the CALIFA FoV size, we first derive for each object the half-light radius R_{50} in the r-band from low-resolution radiative transfer simulations with a FoV of 60×60 kpc, and we calculate the physical size of the spaxels with the cosmological calculator by Wright (2006), assuming that a FoV of $78'' \times 78''$ with spatial sampling $1'' \times 1''$ covers a region up to $2\,R_{50}$ of each simulated galaxy. We also compute the redshift and luminosity distance corresponding to the physical size of the spaxels for the CALIFA $1'' \times 1''$ aperture. These quantities are given in Table 5.5.

5.3.2 Spectral properties

In the CALIFA datasets two different overlapping spectral setups are available, the V500 low-resolution mode covering the range $3749 - 7501$ Å with sampling $\mathrm{d}_{\lambda 500} = 2.0$ Å and

[6]The number of spaxels is different among the objects due to the observing conditions and the disposition of the dithering pattern; usually it is in the range $76''-78''$ and $71''-73''$ in the right ascension and declination axes respectively. We have chosen a $77'' \times 73''$ scheme for our simulated datacubes.

Object	Redshift	Luminosity distance [Mpc]	Physical size of the spaxels [kpc]
C-CS$^+$	0.013	58.1	0.25
E-CS$^+$	0.018	80.8	0.35
D-MA	0.024	108.2	0.45

Table 5.5: Redshift, luminosity distance and physical size covered by the spaxels in our synthetic CALIFA observations.

FWHM spectral resolution $\delta_{\lambda 500} = 6.0$ Å, and the blue mid-resolution setup V1200 that covers the range $3650-4840$ Å with spectral sampling $d_{\lambda 1200} = 0.7$ Å and FWHM resolution $\delta_{\lambda 1200} = 2.3$ Å (Sánchez et al., 2012; García-Benito et al., 2015b). In order to generate the synthetic datacubes in both setups we must cover their respective spectral ranges. The high resolution stellar model (see Sec. 5.2) is available only in the wavelength range $3000-7000$ Å. Therefore, to cover the full spectral range of the CALIFA V500 configuration ($3749-7501$ Å) we generate for each galaxy two different sets of radiative transfer simulations, one at higher resolution including the kinematics between $3000-7000$ Å, the other at lower resolution from $\lambda = 7000$ Å to 7600 Å without kinematics, and we paste the SEDs together at $\lambda = 7000$ Å. In this way we are able to exactly match the CALIFA V500 wavelength range (while the V1200 range is fully covered by the high-resolution stellar model), although the redder part of the spectrum has a spectral resolution lower than the CALIFA sampling. The regions of the spectra generated with the low-resolution stellar model are then flagged as bad pixels in the datacubes (see Sec. 5.4.2) and should not be used for SED fitting analysis[7].

After we create these datacubes that cover the V500 and V1200 spectral range we redshift the spectra to their corresponding redshift, and we resample them to a spacing of 2 Å and 0.7 Å according to the spectral sampling of the V500 and V1200 configurations respectively. In the last steps, we remove the wavelengths outside the V500 and V1200 spectral range, as well as some of the pixels to obtain a $77'' \times 73''$ hexagonal FoV configuration.

5.3.3 Point Spread Functions

In our datacubes we account for both the spatial and spectral Point Spread Functions (PSFs) affecting CALIFA observations. We convolve the two spatial dimensions of our synthetic 3-dimensional datacubes with a two dimensional Gaussian PSF, to account for the 2.5" FWHM spatial resolution in CALIFA, and we also include the effect of the PSF in the spectral dimension by convolving it with a one-dimensional Gaussian kernel, reproducing in our synthetic datacubes the known FWHM values for the PMAS/PPak spectral resolution of $\delta_{\lambda 500} = 6.0$ Å and $\delta_{\lambda 1200} = 2.3$ Å for the V500 and V1200 setups respectively.

5.3.4 CALIFA detector noise

We include in our simulated CALIFA dataset the charateristic noise of the detector by adding random Gaussian noise to the synthetic datacubes. In order to characterize the noise associated to the PMAS/PPak instrument we have considered a sample of 20 objects from the CALIFA survey and performed a spatial and spectral analysis of the noise in both the V500 and V1200 setups. The analysed sample covers a set of different visually-classified

[7]Notice that when we redshift our synthetic spectra we reduce the range of bad pixels, starting from the wavelength $\sim 7090 - 7170$ Å depending on the redshift of the object.

Setup	N_α	N_δ	N_λ	d_λ	δ_λ	σ_0	I_0
V500	77	73	1877	2.0 Å	6.0 Å	0.23	12.10
V1200	77	73	1701	0.7 Å	2.3 Å	0.6	11.26

Table 5.6: Sizes of the simulated datacubes in the spatial and spectral dimensions $(N_\alpha, N_\delta, N_\lambda)$, spectral sampling and spectral resolution $(d_\lambda, \delta_\lambda)$, and the best-fitted noise parameters of equation 5.8 (σ_0, I_0).

morphological types (Walcher, Wisotzki & Bekeraité, 2014)[8] consisting of 9 spiral galaxies (3 face-on, 3 edge-on and 3 with intermediate inclination $\sim 45^o$), 7 ellipticals and 4 objects in the process of merging. The results of the analysis show that the dependency of the noise with the intensity of the signal, characteristic of charge-coupled devices, can be modelled with a simple parametric formula:

$$\sigma_n = \sigma_0 \sqrt{1 + \frac{I_n}{I_0}} \tag{5.8}$$

I_n and σ_n refer to the intensity and the noise provided in the CALIFA datacube normalized to the median value of the intensity in the given datacube, i.e. $\sigma_n = \sigma \setminus I_{50}$ and $I_n = I \setminus I_{50}$. We use normalized errors and fluxes in order to obtain a uniform object-independent unit-free characterization of the noise. σ_0 and I_0 are the parameters we aim to fit. The first term σ_0 in the equation is associated to the 'white noise' in the detector, while the second one $\propto \sqrt{I_n/I_0}$ accounts for the Poisson noise (or shot noise), known to be characteristic of photon counting detectors.

Our analysis also shows that the detector noise does not depend on wavelength, except for the expected edge effects. Some spatial artifacts are also found, related to specific observational errors (e.g. missalignments of the different pointings). None of them are modelled or included in our synthetic datacubes. Besides, we must comment on the well-known noise correlation caused by the CALIFA three-point dithering scheme (Sánchez et al., 2007), as already shown by Husemann, Jahnke & Sánchez (2013); García-Benito et al. (2015b). Since our goal is to set the basis in order to generate state-of-the-art synthetic IFS data, and not to approach the specific problem of combining the PMAS/PPak observations and the CALIFA observational strategy (other instruments/surveys do not use dithering techniques), we have not considered this source of error.

We fit our data to equation 5.8 and average our results over the 20 objects considered in the analysis. The values obtained (for both setups) are summarized in Table 5.6, together with the main properties of the simulated datacubes: spatial and spectral dimensions (N_α, N_δ, N_λ), spectral sampling (d_λ) and spectral resolution (δ_λ). Fig. 5.5 shows an example of the fitted noise-intensity relation for three of the galaxies. White solid line corresponds to the best fitting of our modelled detector noise.

5.4 The SELGIFS data challenge

The main goal of this work is to provide the scientific community with a reliable set of synthetic IFS observations, and with the corresponding maps of directly measured properties,

[8]The morphological classification is part of DR3 and can be found at the URL http://www.caha.es/CALIFA/public_html/?q=content/dr3-tables

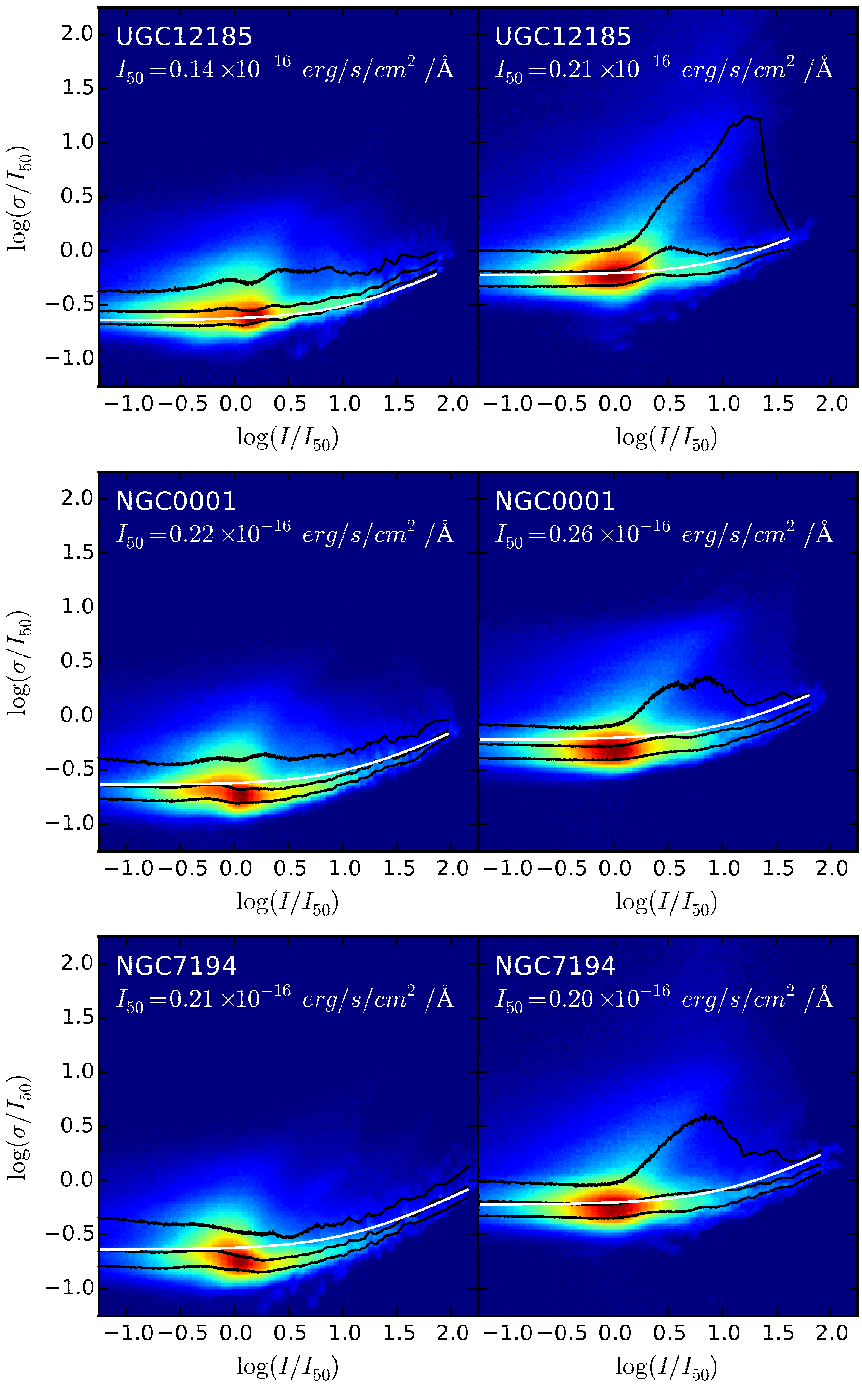

Figure 5.5: Colour maps showing the distribution of the noise σ relative to the signal I, normalized to the median value of the signal I_{50} (in logarithmic scale) for three out of the twenty galaxies considered to characterize the properties of the noise. Left (right) panel displays the values corresponding to the V500 (V1200) setup. Black solid lines correspond to the median and $\pm1\sigma$ value of the distributed noise at a given intensity. White solid lines show the best-fit curves, obtained averaging the values fitted for each galaxy over the full sample of 20 objects; the parameters of the white curve σ_0 and I_0 are given in Table 5.6. In these plots the linear colour scale corresponds to the number of pixels in the CALIFA datacubes ($\sim 10^7$ in total both for the V500 and V1200 configuration) at a given signal and noise.

that allows to test existing (and future) dedicated analysis tools, as well as to create a benchmark for verifying hypothesis and/or preparing observations.

The data are distributed through a web page[9] hosted by the Universidad Autónoma de Madrid. The description of the different files and their data format is presented in the following sections.

5.4.1 Product datacubes

The direct calculation of the resolved (spaxel-by-spaxel) galaxy properties described in Sections 5.1 and 5.2 are provided in separate files. These maps have been calculated directly from the simulations' output, or from the noiseless synthetic spectra prior to the addition of any observational effect.

The name of the files and data format are listed below:

- **GALNAME.stellar.fits:** the file contains the resolved maps obtained directly from the hydrodynamical simulations as described in Section 5.1.1. These FITS files have a single Header Data Unit (HDU) holding a 9-layer matrix, containing the nine 77×73 maps of the stellar properties in the order given in Table 5.2. The header includes the information about the physical property stored in every layer (DESC_*) and its units (UNITS_*), where * refers to the layer number.

- **GALNAME.Lick_indices.fits:** this file stores the resolved maps for the 26 Lick indices measured from the noiseless stellar-only datacube (see Section 5.2.1). Each file consists of a single HDU unit with a 26-layer matrix that contains the twenty-six 77×73 maps of the different absorption features listed in Table 5.3. The header provides for each layer the Lick index name (DESC_*) and its measured units (UNITS_*), with * indicating the layer number.

- **GALNAME.nebular.fits:** it encloses the resolved maps for the 13 nebular line intensities measured from the noiseless nebular-only datacube (Sec. 5.2.1). The data are stored in a single HDU unit with a 13-layer matrix, containing all the thirteen 77×73 maps of the nebular lines given in Table 5.4. The header stores the line names (DESC_*), rest frame wavelengths (LAMBDA_*) and units (UNITS_*) for each layer * in the file.

- **GALNAME.particle_number.fits:** these files provide the number of stellar particles in the region covered by each spaxel of the simulated datacubes.

In order to provide results directly comparable with the ones generated by the observational algorithms applied to the syntetic datacubes, maps at the same spatial resolution of the synthetic datacubes are additionally available. These have been obtained convolving the stellar maps with a 2.5 FWHM Gaussian kernel, and the synthetic spectra with a 2.5" FWHM PSF before extracting the Lick indices and the nebular line intensities as described in Section 5.2.1. Notice that when we compute the logarithmic quantities in the stellar maps the PSF is added prior to the calculation of the logarithm.

5.4.2 Synthetic observations

Our synthetic CALIFA datacubes in the two V500 and V1200 setups (Section 5.3) are provided in different files, identified following the CALIFA DR2 naming convention GAL-

[9] http://astro.ft.uam.es/selgifs/data_challenge/

HDU	Extension name	Format	Content
0	PRIMARY	32-bit float	flux density in units 10^{-16} erg/s/cm^2/Å
1	ERROR	32-bit float	1σ error on the flux density
2	ERRWEIGHT	32-bit float	error weighting factor
3	BADPIX	8-bit integer	bad pixel flags (1=bad, 0=good)
4	FIBCOVER	8-bit integer	number of fibers used to fill each spaxel

Table 5.7: Structure of the CALIFA FITS files in DR2 (from García-Benito et al. 2015b).

NAME.V500.rscube.fits.gz and GALNAME.V1200.rscube.fits.gz for the V500 and V1200 respectively. The data structure of these simulated data closely follows the one adopted in CALIFA, namely datacubes in the standard FITS file format.

The FITS header of the simulated datacubes stores only the most relevant keywords available in the DR2 header. Most of the DR2 keywords containing information about the pointing, the reduction pipeline, Galactic extinction, sky brightness, etc. have been removed. In addition to the mandatory FITS keywords, we store in the primary HDU the information about declination and right ascension of the object (according to the Greisen & Calabretta 2002 standard). We give arbitrary values to these two parameters to avoid problems with visualization tools. The flux unit has been stored under the keyword PIPE UNITS as in the CALIFA datacubes, and also under the keyword BUNIT following the most recent FITS file keywords definition by Pence et al. (2010).

Each FITS file contains the data for a single galaxy stored in five HDU (see Table 5.7), every one of them providing different information according to the data format of the pipeline V1.5 used in DR2 (García-Benito et al., 2015b). The first two axes in the datacubes (N_α, N_δ) correspond to the spatial dimensions (along the right ascension and declination) with a 1" $\times$ 1" sampling. The third dimension (N_λ) represents the wavelength axis, with ranges and samplings described in Section 5.3.2 and Table 5.6.

Here we summarize the content of each HDU:

0) Primary (PRIMARY)

The primary HDU contains the measured flux densities in CALIFA units of 10^{-16} erg $\cdot$ s^{-1} $\cdot$ cm^{-2} $\cdot$ Å^{-1}.

1) Error (ERROR)

This extension provides the values of the 1σ noise level in each pixel, calculated according to Eq. 5.8. In the case of bad pixels, we store a value of 10^{10} following the CALIFA data structure.

2) Error weight (ERRWEIGHT)

In the CALIFA datacubes, this HDU gives the error scaling factor for each pixel, in the case that all valid pixels of the cube are co-added; in our case we set all the values to 1.

3) Bad pixel (BADPIX)

This extension stores a flag advising on potential problems in a pixel; in the CALIFA dataset this may occur for instance due to cosmic rays contamination, bad CCD columns, or the effect of vignetting. In our datacubes we flag as bad pixel (i.e. equal to 1) the regions in the spectra that are generated with the lower-resolution stellar model (see Section 5.3.2).

4) Fiber coverage (FIBCOVER)
This HDU, available only from DR2, accounts for the number of fibers used to recover the flux, and is set to 3 in our mock datacubes.

5.5 Summary

We use hydrodynamical simulations of galaxies formed in a cosmological context to generate mock data mimicking the Integral Field Spectroscopy (IFS) survey CALIFA (Sánchez et al., 2012). The hydrodynamical code follows, in addition to gravity and hydrodynamics, many other relevant galactic-scale physical processes, such as energy feedback and chemical enrichment from SNe explosions, multi-phase InterStellar Medium (ISM), and metal-dependent cooling of the gas. Our hydrodynamical simulations have been post-processed with the radiative transfer code SUNRISE, in order to obtain their spatially-resolved spectral energy distributions. These spectra contain the light emitted by the stars and the nebulae (young stars) in the simulations, and include the broadening of the absorption and emission lines due to kinematics, as well as the extinction and scattering by the dust in the ISM.

The input parameters in SUNRISE have been tuned to reproduce the properties of the CALIFA instrument in terms of field of view size, number of spaxels and spectral range. After we obtain the results of the radiative transfer with SUNRISE, we redshift the simulated spectra to match the physical size covered by the spaxels in the radiative transfer stage with the angular resolution of the PMAS instrument used in CALIFA, and we resample and cut these spectra according to the sampling and wavelength range of the low-resolution V500 and blue mid-resolution V1200 CALIFA setups. We convert our spatially-resolved spectra into the V500 and V1200 data format of the CALIFA DR2, and we convolve these 3-dimensional datasets with Gaussian Point Spread Functions both in the spatial and spectral dimensions, mimicking the properties of the CALIFA observations in terms of spatial and spectral resolution. Finally, after we parametrize the properties of the noise in a sample of 20 galaxies both from the CALIFA V500 and V1200 datasets, we add similar noise to the simulated V500 and V1200 data.

Our final sample of 18 datacubes (3 objects with 3 inclinations both in the V500 and V1200 setups) provide observers with a powerful benchmark to test the accuracy and calibration of their analysis tools and set the basis for a reliable comparison between simulations and IFS observational data. To this purpose we generate, together with the synthetic IFS observations, a corresponding set of *product datacubes*, i.e. resolved maps of several properties computed directly from the simulations and/or simulated noiseless datacubes.

Although this work is specifically designed to reproduce the properties of the CALIFA observations, the method illustrated in this Chapter can be easily extended to mimic other integral field spectrographs such as MUSE (Bacon et al., 2004), WEAVE (Dalton et al., 2014), MaNGA (Bundy et al., 2015) or SAMI (Allen et al., 2015) by changing some of the input parameters in the radiative transfer stage and performing a similar study of the detector noise. Hence, this procedure can be easily applied to generate synthetic observations for different IFS instruments, or for studying a specific science case prior to applying for observing time. The present project can also be extended to use other hydrodynamical simulations, which will be very important in order to enlarge the given dataset and consider a more complete sample of galaxies in terms of morphology, total mass, stellar age and metallicity, gas content and merger history.

Chapter 6

Summary and conclusions

In this dissertation we have explored the possibility to profitably connect the results of two different approaches in the study of galaxies in the Universe, namely the analysis of observational data from extragalactic surveys and cosmological hydrodynamical simulations.

Galaxy formation and evolution has been a topic of intense research in the last years, and many progresses have been made thanks to both the design and development of new observing facilities, that have allowed to create large datasets of galaxies' spectral and photometric properties over a wide range of redshifts and wavelengths, and the improvements in the numerical techniques applied in galaxy formation codes, that now include physically-motivated descriptions of many processes relevant for the formation and evolution of galaxies in cosmological context.

These two approaches to the study of galaxies have been often considered as separated fields with few interaction and overlap, as the methods, the techniques and the information that they provide are significantly different. Observations aim to obtain accurate measurements of the spectra and photometry of galaxies, and observers analyze the light emitted by the galaxies to recover their physical properties, trying to disentangle the processes acting at the various stages of the galaxies' formation and evolution. On the other hand, cosmological hydrodynamical simulations follow the formation and evolution of galaxies in the Universe from primordial perturbations, taking into account the different physics acting at the various scales over cosmic time, being able to connect a galaxy's merger, accretion and interaction history with its properties at various times. Despite the differences in these two approaches, observers often rely on simulations to connect given observational results to given physical processes acting on galaxies, while simulators use observations to test whether the implemented physical modules included in galaxy formation codes yield results that are consistent with the available observations.

A fruitful way to study galaxy formation is then to combine these two approaches; as the main source of information in observations comes from the galaxies' photometry and spectra, the first step to connect simulations and observations is to calculate synthetic spectra of simulated galaxies, although at the cost of having additional modelling together with the corresponding uncertainties and assumptions. For a reliable comparison between observations and simulations, it is also necessary to use similar methods and techniques to analyse their properties. Moreover, it must be taken into account that each of the two approaches suffers from several limitations; in the case of observations mainly related to instrumental effects (e.g. point spread functions, low signal-to-noise, limited spatial sampling), and in the case of simulations related to the limited spatial and mass resolution, implying that many important physical processes act at the sub-resolution level and are implemented using sub-grid recipes.

In this thesis we aimed to fill the gap between these two approaches by developing techniques that allow to more consistently compare simulations and observations, helping simulators to better judge their results against real galaxies; on the other side, we have used hydrodynamical simulations to generate mock data that can be employed to test observational algorithms.

Here we will summarize the main results of our work, and we will formulate a few more general conclusions and future prospects.

6.1 Testing biases and systematics in the observational derivation of galaxy properties

The work presented in Chapter 3 is a study of the ability of several observational algorithms to recover the physical properties of galaxies. To this end, we create synthetic SEDs from hydrodynamical simulations of galaxies in cosmological context, and compare the real properties of the galaxies, that are known in the simulations, with the observationally-derived quantities. In this way, we estimate the accuracy of the algorithms and study the biases and systematics intrinsic in the observational derivation of the galaxy properties[1].

The synthetic SEDs have been created from a set of 15 zoom-in hydrodynamical simulations of galaxies similar in mass to the Milky-Way, generated using the same hydrodynamical code GADGET-3 with three different implementations of chemical and SNe energy feedback. To create the synthetic SEDs, we have applied to the simulations three methods with increasing complexity:

1. Stellar Population Synthesis (SPS) models, that provide only the stellar spectra of the simulated galaxies.

2. SPS models with the addition of a simple parametric model of dust extinction.

3. Radiative transfer calculations with the 3D polychromatic Monte Carlo radiative transfer code SUNRISE, self-consistently simulating the transfer of the light generated by young (i.e. nebulae) and old star particles through a dusty ISM.

Comparing the direct outputs of the simulations with the quantities derived mimicking the observational procedures we have studied:

- the accuracy, systematics and biases of the observational algorithms in recovering the galaxy properties,

- the effects of the assumptions in the creation of the synthetic SEDs on the derived galaxy properties, and the differences arising due to the different procedures in the calculation of the physical quantities,

- for each quantity studied, the reliability of a direct comparison between simulated and observed galaxies without post-processing the simulations.

We have found that the derivation of the properties of galaxies is influenced by several factors, that we list here:

[1]We note that this is valid provided the simulated galaxies are not too dissimilar from real galaxies. In previous works we have shown that this is indeed the case for the galaxies studied in this Chapter (see e.g. Scannapieco et al. 2009, 2011, 2012; Aumer et al. 2013; Nuza et al. 2014; see also Chapter 4).

– when creating the synthetic SEDs, assumptions intrinsic to the SPS models (such as the choice of stellar tracks, stellar atmospheres, IMF), and assumptions into the dust curve (e.g. grain composition),

– orientation of a galaxy (e.g. fainter magnitudes for edge-on views) and projection effects,

– biases related to a specific instrumental setup (for instance the limited spatial sampling of the SDSS single-fiber spectrograph),

– different ways to calculate mean quantities such as age or metallicity, e.g. weighting by the mass emphasizing the properties of the most massive stellar populations, or by luminosity giving more weight to the most luminous ones,

– when performing a fit to the spectra/broadband magnitudes, simplified assumptions in the construction of the template of the fitted models,

– the use of different calibrations when physical quantities are calculated converting values derived from single (or few) spectral features, such as in the case of the SFRs or the gas oxygen abundances.

All of these factors influence the physical quantities recovered in different ways; assuming different SPSs to model the SED gives small offsets ($\lesssim 0.1$ dex) in the optical magnitudes and colours, while if dust extinction is taken into account the effects are larger, in particular for metal-rich (i.e. dust-rich) galaxies observed edge-on (fainter by $\sim 0.5-1$ mag and redder by $\sim 0.1-0.3$ mag). Following an observational approach in the calculation of the magnitudes (such as the Petrosian or Model Magnitudes in the Sloan Digital Sky Survey) the differences compared to the magnitudes obtained with SPS models are of the order $\sim 0.2-0.3$ dex.

The limited spatial sampling of single-fiber spectrographs (such as in SDSS) strongly affects the determination of stellar ages/metallicities and gas metallicities; mimicking the sampling of the SDSS spectrograph gives in general older ages and higher metallicities compared to the global ones, as well as higher gas oxygen abundances due to the preferential sampling of the metal-rich regions in the bulge; however, the strength of this effect depends on the presence of gradients in the galaxy properties, which are expected to vary from galaxy to galaxy.

The use of luminosity or mass to calculate the mean stellar ages and metallicities has also a strong influence, giving in general younger ages and lower metallicities for luminosity-weighted quantities (as is common in observations) compared to the ones calculated weighting by the mass, which is more common in simulations.

Considering the quantities derived fitting the SED or the photometry to a pre-constructed template of models, the stellar ages and metallicities recovered from the SEDs are in good agreement with the ones calculated directly from the simulations weighting with the luminosity. The stellar masses estimated from the photometry are within $\sim 0.1-0.2$ dex compared to the stellar masses in simulations if mass loss is properly treated in the simulations (following the evolution of stars in the different stages including the AGB phase), while if mass loss is not taken into account the observationally-derived stellar masses are systematically lower by ~ 0.3 dex.

Among the physical quantities based on calibrations, the gas metallicity is the one with the stronger dependence on the emission-line calibration used (up to $0.5-0.8$ dex), while the SFR is only mildly affected by the choice of the calibration applied, giving similar values within $\sim 0.2-0.4$ dex.

In summary, we found that the most uncertain quantities in observations are the stellar ages/metallicities and gas metallicities, for which their recovered values depend on several factors and assumptions; this must be taken into account also when simulations are compared with observations, as the values extracted directly from the simulations can be very different compared to the ones derived observationally.

6.2 Unbiased comparison between hydrodynamical simulations and SDSS

In Chapter 4 we have compared our set of hydrodynamical simulations of galaxies at redshift zero with observations from the Sloan Digital Sky Survey (SDSS), converting the outputs of the simulations into synthetic observations and deriving their properties applying the same observational techniques used in SDSS.

In this chapter we extend the work done in Chapter 3, and focus on the techniques used in SDSS and adopting a simulators' point of view, in order to test the reliability of the comparison between the properties of simulated galaxies and observation. As shown in Chapter 3, the biases related to the observational setup of a specific galaxy survey and the uncertainties intrinsic to the algorithms used to derive the properties of galaxies have a significant impact on the recovered quantities, and must be taken into account when simulations and observations are compared.

We found that mimicking SDSS has the following effects on the simulated galaxies' properties studied in this work:

- *Colour and magnitude*: the effect is small ($0.1-0.2$ dex), and the magnitudes calculated from the simulations in a simple way (for instance post-processing the snapshots with a SPS model) can be reliably compared with SDSS data.

- *Stellar mass*: the masses derived fitting the SDSS photometry show some discrepancies with respect to the stellar masses in simulations; the offsets are small ($0.1 - 0.2$ dex) except in the cases where mass loss due to stellar evolution is not properly tracked by the hydrodynamical code.

- *Concentration and Sérsic index*: as these quantities are only observational, we can not compare with any value derived directly from the simulations. We found that the effects of the projection (face-on/edge-on) is significant and can change the position of the galaxies in the SDSS magnitude-concentration and magnitude-Sérsic index diagrams.

- *Stellar age and stellar metallicity*: for the mean ages and metallicities the different calculations strongly affect the results. The SDSS observational techniques give younger and more metal-poor galaxies compared to the values extracted directly from the simulations by $\sim 2 - 4$ Gyr in ages and $0.2 - 0.6$ dex in metallicities, however depending on the age and metallicity of the galaxies.

- *Gas oxygen abundance*: recovering the gas metallicities applying the SDSS emission line calibration to the synthetic spectra sampled in the region covered by the fiber of the SDSS spectrograph, the gas metallicities are higher than the ones given by the simulations snapshots by $0.1 - 0.8$ dex.

- *Star formation rate*: the values of SFR extracted directly from the simulations can be meaningfully compared with the ones derived observationally, at least for normal star-forming galaxies.

Comparing our simulations with observations mimicking SDSS gives the following results:

- our simulations lie in the blue sequence/green valley regions of the colour-magnitude and colour-mass diagrams in SDSS,
- the concentrations and Sérsic indices are in most of the cases in the range of the SDSS galaxies, even though the values depend strongly on the galaxy formation code adopted and the orientation,
- they are in general older compared to most of the SDSS spirals,
- the set has mean stellar metallicities similar to the metal-poor spirals in the SDSS dataset,
- the gas oxygen abundances are in relatively good agreement with observations, even though the simulations are more metal-poor and below the Tremonti et al. (2004) mass-metallicity relation,
- most of the galaxies have SFRs in the region of the SDSS green valley/blue sequence galaxies, although there are objects with SFRs in different regions of the SFR-mass diagram.

In addition to comparing our galaxies with SDSS in an unbiased manner, for each physical property that we studied we provide fitted linear relations, in general with high correlation, that can be easily used to convert the values extracted from other simulations in order to more reliably compare them with SDSS data, at least in the stellar mass range studied here.

To summarize, in this work we have shown that reliably testing the physics implemented in galaxy formation codes against given observations requires the conversion of the direct results of simulations into synthetic observations that mimic the observational dataset, and then the physical quantities in the simulations must be extracted using the same algorithms applied in the derivation of the properties in the observational dataset.

6.3 Synthetic observations mimicking the CALIFA survey

We present in Chapter 5 a different way to connect observations and simulations, exploiting the new possibility offered by the recent Integral Field Spectroscopy (IFS) technique, that is able to recover the spectra of a galaxy sampling different regions at the same time, therefore allowing to study its spatially-resolved properties.

We have developed a pipeline to create synthetic observations that mimic the Calar Alto Legacy Integral Field Area (CALIFA) survey from hydrodynamical simulations of galaxies in cosmological context. To this end, we have applied the radiative transfer technique to 9 hydrodynamical simulations (3 galaxies with three different orientations) to obtain their spatially-resolved spectra including stellar and nebular emission, the kinematic broadening of the absorption and emission lines, and dust extinction assuming a constant dust-to-metals ratio.

We have converted the simulated spectra into the CALIFA data format, namely 3-dimensional data (datacubes) in the two following setups:

– *V500*: the CALIFA low-resolution setup, that has spectral range $3749 - 7501$ Å and sampling every 2 Å.

– *V1200*: the blue mid-resolution V1200 setup, that covers the wavelength range 3650 – 4840 Å with sampling 0.7 Å.

We have mimicked in these datacubes the most important observational effects of the PMAS instrument used in the CALIFA survey, namely:

- physical size of the Field of View ($\sim$ 2 half-light radii R_{50}) sampled by an hexagonal FoV of $77'' \times 73''$,
- Point Spread Functions in the spatial dimensions (a 2-D Gaussian function with 2.5" FWHM) and in the spectral dimension (1-D Gaussian with a FWHM of 6 Å and 2.3 Å for the V500 and V1200 setups respectively),
- after characterizing the properties of the noise for a sample of 20 objects observed by the CALIFA telescope fitting a simple parametric formula, we have added a similar noise to the mock datacubes.

In addition to the synthetic datacubes, we provide the resolved maps of the physical properties of our simulations, calculated directly from the snapshots or extracted from the synthetic datacubes before adding the detector noise.

The analysis of IFS data is often a more complex task compared to the analysis of traditional spectroscopy datasets, for instance Signal-to-Noise (S/N) quickly decreases in the outskirt of the target galaxies to potentially unacceptable levels, which often requires complex algorithms to carry out a spatial segmentation (binning) of the data. Our synthetic datacubes are then a useful tool to test this new class of observational algorithms applied to IFS data, in particular the ones developed for CALIFA data analysis.

6.4 Future prospects

The approaches outlined in this dissertation can be extended in several directions; here we list some of the possible applications and extensions of the methods illustrated in this work:

- *Increasing the number of galaxies to obtain a better statistics*

  The galaxy samples used in this work are relatively small and in particular cover a limited range in masses; to achieve a better statistical significance of the results, in particular to asses the general validity of the fitting formulae given in Chapter 4, it is necessary to increase the sample of galaxies analyzed including other morphological types, and to test the methods on galaxies simulated using other galaxy formation codes.

- *Test simulations and observations at different redshifts*

  In this work we used simulations at redshift zero to generate mock observations, since at low redshift many large galaxy datasets (e.g. SDSS, 2dFGRS, 6dFGS) are available, and the data analysis is more reliable and often based on spectroscopic information. However, in order to judge the ability of a galaxy formation code in reproducing the evolution of real galaxies over cosmic time, several tests at different redshifts are needed. On the other side, observations at high redshift often suffer from many uncertainties (mainly related to low S/N) and the observational algorithms may be usefully tested against mock observations at high redshifts.

- *The mass-metallicity relation over cosmic time*

  The mass-metallicity relation is a benchmark of each implementation of chemical and energy feedback in galaxy formation codes, as it encodes complex and highly non-linear physical processes such as gas outflows, accretion rates, metal enrichment, and star formation efficiency. It has been proven that the gas metallicity determination suffers from strong biases depending on the method and calibration applied in the derivation, and the agreement/disagreement between simulations and observations needs to be tested in an unbiased manner at different redshifts in order to be able to asses the validity of a chemical enrichment model.

- *Generating mock datacubes for other IFS surveys*

  The procedure developed in Chapter 5 to generate synthetic datacubes from hydrodynamical simulations can be extended to create mock data that mimic other ongoing (or future) IFS surveys such as MaNGA or MUSE, helping in the interpretation of the data. On the other side, simulations can be usefully compared with IFS data to test the ability of a galaxy formation code in reproducing realistic spatially-resolved properties of galaxies, such as the observed gradients in the ages and metallicities of the stellar populations.

- *Study the spatially-resolved BPT diagram*

  Simulated IFS observations can be applied to the study of the resolved Baldwin, Phillips & Terlevich (BPT) diagram (that plots the emission line intensities [NII]/Hα vs [OIII]/Hβ); as shown by Sánchez et al. (2015), the spatially-resolved version of the BPT diagram have strong correlations that are still not fully understood. Synthetic datacubes generated from hydrodynamical simulations can help in principle to understand the dependence of the line ratios on the excitation mechanism, and on the physical properties of the gas such as oxygen abundance, ionization parameter and electron density.

- *Better description of the ISM medium in radiative transfer algorithms*

  A possible improvement in the radiative transfer simulations is the additional modelling of the radiation emitted by the Diffuse Ionized Gas (DIG, Haffner et al. 2009), which in turn will provide more realistic synthetic SEDs in terms of the properties of gas emission, that may be used to study emission lines in galaxies at a higher level of detail.

6.5 Closing remarks

In conclusion, the possibility to combine observations and simulations in the study of galaxy formation is still far from being fully explored. Connecting observations and simulations will be a key aspect in the development of the new generations of hydrodynamical codes in the **near** future, as well as in the interpretation of the data coming from the many ongoing or planned galaxy surveys.

The work done in this thesis is a first step in this direction. We have shown how important is to consider the observational uncertainties and biases when connecting observations and simulations, as neglecting them in the comparison between the two gives inconsistent results. We have also shown some of the possibilities offered by generating synthetic datasets from simulations, that can be used to test the observational algorithms or to interpret the

observational results. Hence, future studies of galaxy formation in cosmological context should exploit all the advantages given by reliably connecting the models and the data.

Bibliography

Abadi M. G., Moore B., Bower R. G., 1999, MNRAS, 308, 947

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 591, 499

Abazajian K., Adelman-McCarthy J. K., Agüeros M. A., et al., 2003, ApJ, 126, 2081

Abazajian K. N. et al., 2009, ApJS, 182, 543

Abraham R. G., Valdes F., Yee H. K. C., van den Bergh S., 1994, ApJ, 432, 75

Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., Anderson K. S. J., Anderson S. F., et al., 2006, ApJS, 162, 38

Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25

Agertz O. et al., 2007, MNRAS, 380, 963

Aguerri J. A. L., González-García A. C., 2009, A&A, 494, 891

Allen J. T., Croom S. M., Konstantopoulos I. S., Bryant J. J., Sharp R., et al., 2015, MNRAS, 446, 1567

Allington-Smith J. R., 2007, in Revista Mexicana de Astronomia y Astrofisica Conference Series, Vol. 28, Revista Mexicana de Astronomia y Astrofisica Conference Series, Kurtz S., ed., pp. 17–23

Amendola L., Gannouji R., Polarski D., Tsujikawa S., 2007, Phys. Rev. D, 75, 083504

Andrews B. H., Martini P., 2013, ApJ, 765, 140

Anglés-Alcázar D., Davé R., Özel F., Oppenheimer B. D., 2014, ApJ, 782, 84

Antonuccio-Delogu V., Dobrotka A., Becciani U., Cielo S., Giocoli C., Macciò A. V., Romeo-Veloná A., 2010, MNRAS, 407, 1338

Arnett W. D., 1978, ApJ, 219, 1008

Arons J., Silk J., 1968, MNRAS, 140, 331

Asari N. V., Cid Fernandes R., Stasińska G., Torres-Papaqui J. P., Mateus A., Sodré L., Schoenell W., Gomes J. M., 2007, MNRAS, 381, 263

Aumer M., White S. D. M., Naab T., Scannapieco C., 2013, MNRAS, 434, 3142

Bacon R., Bauer S.-M., Bower R., Cabrit S., Cappellari M., Carollo M., et al., 2004, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 5492, Ground-based Instrumentation for Astronomy, Moorwood A. F. M., Iye M., eds., pp. 1145–1149

Balcells M., Quinn P. J., 1990, Merger origin for counterrotating cores in elliptical galaxies., Wielen R., ed., pp. 210–211

Baldry I. K. et al., 2012, MNRAS, 421, 621

Baldry I. K., Glazebrook K., 2003, ApJ, 593, 258

Baldry I. K., Glazebrook K., Baugh C. M., et al., 2002, ApJ, 569, 582

Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, ApJ, 527, 54

Balsara D. S., 1995, Journal of Computational Physics, 121, 357

Barnes J. E., 1992, ApJ, 393, 484

Barnes J. E., 2012, MNRAS, 425, 1104

Barnes J. E., Hernquist L., 1996, ApJ, 471, 115

Bastian N., Covey K. R., Meyer M. R., 2010, ARA&A, 48, 339

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Baumann D., 2009, ArXiv e-prints

Beckwith S. V. W., Stiavelli M., Koekemoer A. M., et al., 2006, AJ, 132, 1729

Bekki K., Couch W. J., 2003, ApJ, 596, L13

Bekki K., Couch W. J., Shioya Y., 2002, ApJ, 577, 651

Bell E. F., de Jong R. S., 2000, MNRAS, 312, 497

Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289

Bellovary J. M., Holley-Bockelmann K., Gültekin K., Christensen C. R., Governato F., Brooks A. M., Loebman S., Munshi F., 2014, MNRAS, 445, 2667

Belokurov V. et al., 2006, ApJ, 642, L137

Bennett C. L. et al., 2013, ApJS, 208, 20

Benson A. J., 2010, Phys. Rep., 495, 33

Benson A. J., Bower R., 2010, ArXiv e-prints

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2002, MNRAS, 333, 156

Benson A. J., Madau P., 2003, MNRAS, 344, 835

Bertone G., 2010, Particle Dark Matter : Observations, Models and Searches. Cambridge University Press

Bertschinger E., 1998, ARA&A, 36, 599

Bertschinger E., 2001, ApJS, 137, 1

Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, MNRAS, 376, 215

Birnboim Y., Dekel A., 2003, MNRAS, 345, 349

Blanton M. R., Dalcanton J., Eisenstein D., Loveday, 2001, AJ, 121, 2358

Blanton M. R. et al., 2005, AJ, 129, 2562

Blumenthal G. R., Pagels H., Primack J. R., 1982, Nature, 299, 37

Bode P., Ostriker J. P., Turok N., 2001, ApJ, 556, 93

Boggess N. W. et al., 1992, ApJ, 397, 420

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bond J. R., Szalay A. S., Turner M. S., 1982, Physical Review Letters, 48, 1636

Booth C. M., Schaye J., 2009, in American Institute of Physics Conference Series, Vol. 1201, American Institute of Physics Conference Series, Heinz S., Wilcots E., eds., pp. 21–24

Boselli A., Gavazzi G., 2006, PASP, 118, 517

Bouchet P., Lequeux J., Maurice E., Prevot L., Prevot-Burnichon M. L., 1985, A&A, 149, 330

Bournaud F., Jog C. J., Combes F., 2005, A&A, 437, 69

Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Bresolin F., 2008, in The Metal-Rich Universe, Israelian G., Meynet G., eds., p. 155

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Brodwin M. et al., 2013, ApJ, 779, 138

Brook C. B., Kawata D., Gibson B. K., Flynn C., 2004, MNRAS, 349, 52

Brooks A. M., Governato F., Quinn T., Brook C. B., Wadsley J., 2009, ApJ, 694, 396

Brooks A. M., Kuhlen M., Zolotov A., Hooper D., 2013, ApJ, 765, 22

Brüggen M., Scannapieco E., 2009, MNRAS, 398, 548

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517

Bundy K., Bershady M. A., Law D. R., Yan R., Drory N., MacDonald N., et al., 2015, APJ, 798, 7

Burkert A. M., D'Onghia E., 2004, in Astrophysics and Space Science Library, Vol. 319, Penetrating Bars Through Masks of Cosmic Dust, Block D. L., Puerari I., Freeman K. C., Groess R., Block E. K., eds., p. 341

Butcher H., Oemler, Jr. A., 1978, ApJ, 226, 559

Butcher H., Oemler, Jr. A., 1984, ApJ, 285, 426

Byrd G., Valtonen M., 1990, ApJ, 350, 89

Calzetti D., 2008, in Astronomical Society of the Pacific Conference Series, Vol. 390, Pathways Through an Eclectic Universe, Knapen J. H., Mahoney T. J., Vazdekis A., eds., p. 121

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, ApJ, 429, 582

Calzetti D., Sheth K., Churchwell E., Jackson J., 2009, in The Evolving ISM in the Milky Way and Nearby Galaxies, p. 8

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Casado J., Ascasibar Y., Gavilán M., Terlevich R., Terlevich E., Hoyos C., Díaz A. I., 2015, MNRAS, 451, 888

Catelan P., Theuns T., 1996a, MNRAS, 282, 436

Catelan P., Theuns T., 1996b, MNRAS, 282, 455

Chabrier G., 2003, PASP, 115, 763

Charlot S., Fall S. M., 2000, ApJ, 539, 718

Charlot S., Longhetti M., 2001, MNRAS, 323, 887

Chen X. Y., Liang Y. C., Hammer F., Prugniel P., Zhong G. H., Rodrigues M., Zhao Y. H., Flores H., 2010, A&A, 515, A101

Chieffi A., Limongi M., 2004, ApJ, 608, 405

Chiueh T., Lee J., Lin L., 2002, ApJ, 581, 794

Chomiuk L., Povich M. S., 2011, AJ, 142, 197

Christensen C. R., Brooks A. M., Fisher D. B., Governato F., McCleary J., Quinn T. R., Shen S., Wadsley J., 2014, MNRAS, 440, L51

Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, MNRAS, 358, 363

Cid Fernandes R. et al., 2013, A&A, 557, A86

Cid Fernandes R. et al., 2009, in Revista Mexicana de Astronomia y Astrofisica Conference Series, Vol. 35, Revista Mexicana de Astronomia y Astrofisica Conference Series, pp. 127–132

Ciotti L., Ostriker J. P., Proga D., 2009, ApJ, 699, 89

Clowes R. G., Harris K. A., Raghunathan S., Campusano L. E., Söchting I. K., Graham M. J., 2013, MNRAS, 429, 2910

Coil A. L., 2013, The Large-Scale Structure of the Universe, Oswalt T. D., Keel W. C., eds., p. 387

Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781

Coles P., Lucchin F., 1995, Cosmology. The origin and evolution of cosmic structure

Colless M., 1999, Royal Society of London Philosophical Transactions Series A, 357, 105

Conroy C., 2013, ARA&A, 51, 393

Conroy C., Gunn J. E., 2010, ApJ, 712, 833

Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486

Couch W. J., Barger A. J., Smail I., Ellis R. S., Sharples R. M., 1998, ApJ, 497, 188

Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53

Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, ApJ, 452, 797

Cox T. J., Dutta S. N., Di Matteo T., Hernquist L., Hopkins P. F., Robertson B., Springel V., 2006, ApJ, 650, 791

Crain R. A. et al., 2015, MNRAS, 450, 1937

Crain R. A. et al., 2009, MNRAS, 399, 1773

Creasey P., Scannapieco C., Nuza S. E., Yepes G., Gottlöber S., Steinmetz M., 2015, ApJ, 800, L4

Crill B. P., Ade P. A. R., Artusa D. R., Bhatia R. S., Bock J. J., Boscaleri A., et al., 2003, ApJS, 148, 527

Croton D. J. et al., 2006, MNRAS, 365, 11

Cuesta A. J., Prada F., Klypin A., Moles M., 2008, MNRAS, 389, 385

Czerny B., 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, Ho L. C., Wang J.-W., eds., p. 586

da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595

Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431

Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140

Dalton G., Trager S., Abrams D. C., Bonifacio P., López Aguerri J. A., et al., 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 0

David L. P., Nulsen P. E. J., McNamara B. R., Forman W., Jones C., Ponman T., Robertson

B., Wise M., 2001, ApJ, 557, 546
Davis A. J., Natarajan P., 2009, MNRAS, 393, 1498
Davis M., Huchra J., Latham D. W., Tonry J., 1982, ApJ, 253, 423
de Blok W. J. G., 2010, Advances in Astronomy, 2010, 789293
de Felice A., Tsujikawa S., 2010, Living Reviews in Relativity, 13
Dekel A., Silk J., 1986, ApJ, 303, 39
Del Popolo A., 2006, A&A, 454, 17
Denicoló G., Terlevich R., Terlevich E., 2002, MNRAS, 330, 69
Di Matteo P., Lehnert M. D., Qu Y., van Driel W., 2011, A&A, 525, L3
Di Matteo T., Colberg J., Springel V., Hernquist L., Sijacki D., 2008, ApJ, 676, 33
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Dolag K., Borgani S., Schindler S., Diaferio A., Bykov A. M., 2008, Space Sci. Rev., 134, 229
Dolag K., Jubelgas M., Springel V., Borgani S., Rasia E., 2004, ApJ, 606, L97
Dolag K., Stasyszyn F., 2009, MNRAS, 398, 1678
Donato F. et al., 2009, MNRAS, 397, 1169
Doroshkevich A., Tucker D. L., Allam S., Way M. J., 2004, A&A, 418, 7
Doroshkevich A. G., Zel'dovich Y. B., Novikov I. D., 1967, AZh, 44, 295
Draine B. T., 2003, ApJ, 598, 1017
Dressler A., 1980, ApJ, 236, 351
Dressler A. et al., 1997, ApJ, 490, 577
Drory N., Bender R., Hopp U., 2004, ApJ, 616, L103
Dubinski J., Carlberg R. G., 1991, ApJ, 378, 496
Dubois Y., Teyssier R., 2008, A&A, 477, 79
Durrer R., Neronov A., 2013, A&A Rev., 21, 62
Dwek E., 1998, ApJ, 501, 643
Efstathiou G., 1992, MNRAS, 256, 43P
Efstathiou G., 2000, MNRAS, 317, 697
Efstathiou G., Davis M., White S. D. M., Frenk C. S., 1985, ApJS, 57, 241
Einasto J., Haud U., 1989, A&A, 223, 89
Elbert O. D., Bullock J. S., Garrison-Kimmel S., Rocha M., Oñorbe J., Peter A. H. G., 2015, MNRAS, 453, 29
Ellis R. S., 2008, Observations of the High Redshift Universe, Loeb, A. and Ferrara, A. and Ellis, R. S., pp. 259–364
Ellison S. L., Mendel J. T., Patton D. R., Scudder J. M., 2013, MNRAS, 435, 3627
Elmegreen B. G., 1997, in Revista Mexicana de Astronomia y Astrofisica Conference Series, Vol. 6, Revista Mexicana de Astronomia y Astrofisica Conference Series, Franco J., Terlevich R., Serrano A., eds., p. 165
Elmegreen B. G., Scalo J., 2004, ARA&A, 42, 211
Emsellem E. et al., 2007, MNRAS, 379, 401

Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813

Fabian A. C., 2012, ARA&A, 50, 455

Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994a, A&AS, 104, 365

Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994b, A&AS, 105, 29

Fall S. M., 1979, Nature, 281, 200

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415

Fardal M. A., Katz N., Gardner J. P., Hernquist L., Weinberg D. H., Davé R., 2001, ApJ, 562, 605

Farouki R. T., Shapiro S. L., 1982, ApJ, 259, 103

Fathi K., Allen M., Boch T., Hatziminaoglou E., Peletier R. F., 2010, MNRAS, 406, 1595

Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, ApJ, 725, 633

Ferland G. J., 2003, ARA&A, 41, 517

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Ferrari C., Govoni F., Schindler S., Bykov A. M., Rephaeli Y., 2008, Space Sci. Rev., 134, 93

Ferraro R., 2012, in American Institute of Physics Conference Series, Vol. 1471, American Institute of Physics Conference Series, Alcaniz J., Carneiro S., Chimento L. P., Del Campo S., Fabris J. C., Lima J. A. S., Zimdahl W., eds., pp. 103–110

Fioc M., Rocca-Volmerange B., 1997, A&A, 326, 950

Fioc M., Rocca-Volmerange B., 1999, arXiv:astro-ph/9912179

Fischera J., Dopita M., 2005, ApJ, 619, 340

Fitzpatrick E. L., 1986, AJ, 92, 1068

Fraser C. W., 1972, The Observatory, 92, 51

Frenk C. S., Evrard A. E., White S. D. M., Summers F. J., 1996, ApJ, 472, 460

Frenk C. S., White S. D. M., 2012, Annalen der Physik, 524, 507

Frenk C. S., White S. D. M., Efstathiou G., Davis M., 1985, Nature, 317, 595

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385

Fryxell B. et al., 2000, ApJS, 131, 273

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Gajda G., Lokas E. L., 2015, ArXiv e-prints

Gallazzi A., Charlot S., Brinchmann J., White S. D. M., 2006, MNRAS, 370, 1106

Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41

Gao L., White S. D. M., 2007, MNRAS, 377, L5

García-Benito R. et al., 2015a, AAP, 576, A135

García-Benito R. et al., 2015b, A&A, 576, A135

Gebhardt K. et al., 2000, ApJ, 539, L13

Geller M. J., Huchra J. P., 1989, Science, 246, 897

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Gnedin N. Y., Hollon N., 2012, ApJS, 202, 13

Gnedin N. Y., Hui L., 1998, MNRAS, 296, 44

Gnedin O. Y., 2003, ApJ, 589, 752

Gómez P. L. et al., 2003, ApJ, 584, 210

González Delgado R. M. et al., 2015, A&A, 581, A103

González Delgado R. M. et al., 2014, A&A, 562, A47

González-García A. C., Oñorbe J., Domínguez-Tenreiro R., Gómez-Flechoso M. Á., 2009, A&A, 497, 35

González-García A. C., van Albada T. S., 2003, MNRAS, 342, L36

Goriely S., Siess L., 2005, in IAU Symposium, Vol. 228, From Lithium to Uranium: Elemental Tracers of Early Cosmic Evolution, Hill V., Francois P., Primas F., eds., pp. 451–460

Gottloeber S., Klypin A., 2008, ArXiv e-prints

Governato F. et al., 2010, Nature, 463, 203

Governato F. et al., 2009, MNRAS, 398, 312

Governato F. et al., 2004, ApJ, 607, 688

Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479

Govoni F., Feretti L., 2004, International Journal of Modern Physics D, 13, 1549

Graham A. W., Driver S. P., Petrosian V., Conselice C. J., Bershady M. A., Crawford S. M., Goto T., 2005, AJ, 130, 1535

Greggio L., Renzini A., 1983, A&A, 118, 217

Greif T. H., Glover S. C. O., Bromm V., Klessen R. S., 2009, MNRAS, 392, 1381

Greisen E. W., Calabretta M. R., 2002, A&A, 395, 1061

Groves B., Dopita M. A., Sutherland R. S., Kewley L. J., Fischera J., Leitherer C., Brandl B., van Breugel W., 2008, ApJS, 176, 438

Groves B. A., Dopita M. A., Sutherland R. S., 2004, ApJS, 153, 9

Guidi G., Scannapieco C., Walcher C. J., 2015, MNRAS, 454, 2381

Guidi G., Scannapieco C., Walcher J., Gallazzi A., 2016, MNRAS, 462, 2046

Gunn J. E., 1977, ApJ, 218, 592

Gunn J. E. et al., 1998, AJ, 116, 3040

Gunn J. E., Gott, III J. R., 1972, ApJ, 176, 1

Gunn J. E. et al., 2006, AJ, 131, 2332

Guth A. H., 1981, Phys. Rev. D, 23, 347

Haffner L. M. et al., 2009, Reviews of Modern Physics, 81, 969

Hales C. A., 2013, ArXiv e-prints

Hambrick D. C., Ostriker J. P., Naab T., Johansson P. H., 2011, ApJ, 738, 16

Häring N., Rix H.-W., 2004, ApJ, 604, L89

Hayward C. C., Kereš D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743, 159

Hayward C. C. et al., 2014, MNRAS, 445, 1598

Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, MNRAS, 428, 2529

Hayward C. C., Smith D. J. B., 2015, MNRAS, 446, 1512

Heckman T. M., Lehnert M. D., 2000, ApJ, 537, 690

Heitmann K. et al., 2008, Computational Science and Discovery, 1, 015003

Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, MNRAS, 396, 535

Henyey L. G., Greenstein J. L., 1941, ApJ, 93, 70

Hernandez-Jimenez J. A., Pastoriza M. G., Rodrigues I., Krabbe A. C., Winge C., Bonatto C., 2013, MNRAS, 435, 3342

Hodge J. A., Karim A., Smail I., Swinbank A. M., et al., 2013, ApJ, 768, 91

Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, ApJ, 624, 54

Hopkins P. F., 2013, MNRAS, 428, 2840

Hopkins P. F. et al., 2005, in Bulletin of the American Astronomical Society, Vol. 37, American Astronomical Society Meeting Abstracts, p. 1354

Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, MNRAS, 427, 968

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950

Hoversten E. A., Glazebrook K., 2008, ApJ, 675, 163

Howell D. A., 2011, Nature Communications, 2, 350

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Husemann B., Jahnke K., Sánchez, S. F. e. a., 2013, A&A, 549, A87

Huynh M., Lazio J., 2013, ArXiv e-prints

Ibata R. A. et al., 2013, Nature, 493, 62

Irwin M. J. et al., 2007, ApJ, 656, L13

Izotov Y. I., Stasińska G., Meynet G., Guseva N. G., Thuan T. X., 2006, A&A, 448, 955

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jimenez R., Padoan P., Matteucci F., Heavens A. F., 1998, MNRAS, 299, 123

Jonsson P., 2006, MNRAS, 372, 2

Jonsson P., Groves B. A., Cox T. J., 2010, MNRAS, 403, 17

Jubelgas M., Springel V., Dolag K., 2004, MNRAS, 351, 423

Jungman G., Kamionkowski M., Griest K., 1996, Phys. Rep., 267, 195

Kamionkowski M., Liddle A. R., 2000, Physical Review Letters, 84, 4525

Kaplinghat M., Knox L., Turner M. S., 2000, Physical Review Letters, 85, 3335

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Kauffmann G., Heckman T. M., White S. D. M., Charlot, 2003, MNRAS, 341, 33

Kaufmann T., Bullock J. S., Maller A. H., Fang T., Wadsley J., 2009, MNRAS, 396, 191

Kawata D., Gibson B. K., 2003, MNRAS, 346, 135

Kay S. T., Pearce F. R., Frenk C. S., Jenkins A., 2002, MNRAS, 330, 113

Keller B. W., Wadsley J., Benincasa S. M., Couchman H. M. P., 2014, MNRAS, 442, 3013
Keller B. W., Wadsley J., Couchman H. M. P., 2015, MNRAS, 453, 3499
Kelz A. et al., 2006, PASP, 118, 129
Kennicutt, Jr. R. C., 1998a, ARA&A, 36, 189
Kennicutt, Jr. R. C., 1998b, ApJ, 498, 541
Kennicutt, Jr. R. C., Bresolin F., Garnett D. R., 2003, ApJ, 591, 801
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Geller M. J., Jansen R. A., 2004, AJ, 127, 2002
Kewley L. J., Geller M. J., Jansen R. A., Dopita M. A., 2002, AJ, 124, 3135
King A., 2009, ApJ, 695, L107
Klypin A., Kravtsov A. V., Bullock J. S., Primack J. R., 2001, ApJ, 554, 903
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Klypin A. A., Shandarin S. F., 1983, MNRAS, 204, 891
Knebe A., Power C., 2008, ApJ, 678, 621
Kobulnicky H. A., Kewley L. J., 2004, ApJ, 617, 240
Kochanek C. S., Pahre M. A., Falco E. E., 2000, arXiv:astro-ph/0011458
Kofman L. A., Gnedin N. Y., Bahcall N. A., 1993, ApJ, 413, 1
Kormendy J., Kennicutt, Jr. R. C., 2004, ARA&A, 42, 603
Kravtsov A., 2010, Advances in Astronomy, 2010, 281913
Kravtsov A. V., 1999, PhD thesis, NEW MEXICO STATE UNIVERSITY
Kravtsov A. V., Gnedin O. Y., Klypin A. A., 2004, ApJ, 609, 482
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kron R. G., 1980, ApJS, 43, 305
Kroupa P., 2002, Science, 295, 82
Krumholz M. R., Thompson T. A., 2013, MNRAS, 434, 2329
Lanz L., Hayward C. C., Zezas A., Smith H. A., Ashby M. L. N., Brassington N., Fazio G. G., Hernquist L., 2014, ApJ, 785, 39
Larson R. B., 1974, MNRAS, 169, 229
Larson R. B., 1998, MNRAS, 301, 569
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692
Lee S.-K., Ferguson H. C., Somerville R. S., Wiklind T., Giavalisco M., 2010, ApJ, 725, 1644
Leitch E. M. et al., 2002, ApJ, 568, 28
Leitherer C. et al., 1999, ApJS, 123, 3
Li Y. et al., 2007, ApJ, 665, 187
Li Z., 2013, arXiv:1302.0100
Lia C., Portinari L., Carraro G., 2002, MNRAS, 335, 864
Liang Y. C., Yin S. Y., Hammer F., Deng L. C., Flores H., Zhang B., 2006, ApJ, 652, 257
Libeskind N. I., Hoffman Y., Tully R. B., Courtois H. M., Pomarède D., Gottlöber S., Steinmetz M., 2015, MNRAS, 452, 1052
Liddle A. R., Lyth D. H., 2000, Cosmological Inflation and Large-Scale Structure. p. 414

Lintott C. et al., 2011, MNRAS, 410, 166

Lintott C. J. et al., 2008, MNRAS, 389, 1179

Longair M. S., 2008, Galaxy Formation

Lovell M. R. et al., 2012, MNRAS, 420, 2318

Lucy L. B., 1977, AJ, 82, 1013

Lupton R., Blanton M. R., Fekete G., Hogg D. W., O'Mullane W., Szalay A., Wherry N., 2004, PASP, 116, 133

Lykins M. L., Ferland G. J., Porter R. L., van Hoof P. A. M., Williams R. J. R., Gnat O., 2013, MNRAS, 429, 3133

Mac Low M.-M., Ferrara A., 1999, ApJ, 513, 142

Mac Low M.-M., Klessen R. S., 2004, Reviews of Modern Physics, 76, 125

MacArthur L. A., Courteau S., Bell E. F., Holtzman J. A., 2004, in Bulletin of the American Astronomical Society, Vol. 36, American Astronomical Society Meeting Abstracts, p. 1444

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940

Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55

Madau P., 1998, in Astronomical Society of the Pacific Conference Series, Vol. 146, The Young Universe: Galaxy Formation and Evolution at Intermediate and High Redshift, D'Odorico S., Fontana A., Giallongo E., eds., p. 289

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Madau P., Diemand J., Kuhlen M., 2008, ApJ, 679, 1260

Maller A. H., Berlind A. A., Blanton M. R., Hogg D. W., 2009, ApJ, 691, 394

Maller A. H., Bullock J. S., 2004, MNRAS, 355, 694

Maoz D., Mannucci F., 2012, PASA, 29, 447

Maraston C., 2005, MNRAS, 362, 799

Marconi A., Hunt L. K., 2003, ApJ, 589, L21

Marigo P., 2001, A&A, 370, 194

Martin C. L., Shapley A. E., Coil A. L., Kornei K. A., Bundy K., Weiner B. J., Noeske K. G., Schiminovich D., 2012, ApJ, 760, 127

Martin D. C., Small T., Schiminovich D., Wyder T. K., et al., 2007, ApJS, 173, 415

Martínez-Serrano F. J., Serna A., Domínguez-Tenreiro R., Mollá M., 2008, MNRAS, 388, 39

Martins L. P., González Delgado R. M., Leitherer C., Cerviño M., Hauschildt P., 2005, MNRAS, 358, 49

Martizzi D., Mohammed I., Teyssier R., Moore B., 2014, MNRAS, 440, 2290

Mastropietro C., Moore B., Mayer L., Debattista V. P., Piffaretti R., Stadel J., 2005, MNRAS, 364, 607

Mathews W. G., Baker J. C., 1971, ApJ, 170, 241

Mattsson L., De Cia A., Andersen A. C., Zafar T., 2014, MNRAS, 440, 1562

McGaugh S. S., 1991, ApJ, 380, 140

McGaugh S. S., de Blok W. J. G., Schombert J. M., Kuzio de Naray R., Kim J. H., 2007, ApJ, 659, 149

McKee C. F., Ostriker E. C., 2007, ARA&A, 45, 565

McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148
Merritt D., 1984, ApJ, 276, 26
Mestel L., Spitzer, Jr. L., 1956, MNRAS, 116, 503
Mezcua M., Prieto M. A., 2014, ApJ, 787, 62
Michałowski M. J., Hayward C. C., Dunlop J. S., Bruce V. A., Cirasuolo M., Cullen F., Hernquist L., 2014, A&A, 571, A75
Miedema P. G., 2011, ArXiv e-prints
Mihos J. C., Hernquist L., 1994, ApJ, 425, L13
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Milgrom M., 1983a, ApJ, 270, 371
Milgrom M., 1983b, ApJ, 270, 365
Miller G. E., Scalo J. M., 1979, ApJS, 41, 513
Miller R. H., 1986, A&A, 167, 41
Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2013, MNRAS, 435, 87
Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Monaghan J. J., 1992, ARA&A, 30, 543
Monaghan J. J., 1997, Journal of Computational Physics, 136, 298
Monaghan J. J., Gingold R. A., 1983, Journal of Computational Physics, 52, 374
Moore B., 1994, Nature, 370, 629
Moore B., Diemand J., Madau P., Zemp M., Stadel J., 2006, MNRAS, 368, 563
Moore B., Katz N., Lake G., 1996, ApJ, 457, 455
Moore B., Lake G., Quinn T., Stadel J., 1999, MNRAS, 304, 465
Morris J. P., Monaghan J. J., 1997, Journal of Computational Physics, 136, 41
Mosconi M. B., Tissera P. B., Lambas D. G., Cora S. A., 2001, MNRAS, 325, 34
Mouschovias T. C., 1976, ApJ, 207, 141
Mukhanov V., 2005, Physical Foundations of Cosmology. p. 442
Munshi F. et al., 2013, ApJ, 766, 56
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
Naab T., Burkert A., 2003, ApJ, 597, 893
Naab T., Johansson P. H., Ostriker J. P., 2009, ApJ, 699, L178
Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007, ApJ, 658, 710
Naab T., Khochfar S., Burkert A., 2006, ApJ, 636, L81
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Narayan R., Medvedev M. V., 2001, ApJ, 562, L129
Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Navarro J. F. et al., 2004, MNRAS, 349, 1039
Navarro J. F., Steinmetz M., 2000, ApJ, 538, 477
Navarro J. F., White S. D. M., 1993, MNRAS, 265, 271

Navarro J. F., White S. D. M., 1994, MNRAS, 267, 401

Nehlig F., Vollmer B., Braine J., 2016, ArXiv e-prints

Nelson D. et al., 2015, arXiv:1504.00362

Neto A. F. et al., 2007, MNRAS, 381, 1450

Newman J. A., Cooper M. C., Davis M., Faber S. M., et al., 2013, ApJS, 208, 5

Nomoto K., Thielemann F.-K., Yokoi K., 1984, ApJ, 286, 644

Norberg P., Cole S., Baugh C. M., Frenk C. S., et al., 2002, MNRAS, 336, 907

Nulsen P. E. J., 1982, MNRAS, 198, 1007

Nuza S. E., Parisi F., Scannapieco C., Richter P., Gottlöber S., Steinmetz M., 2014, MNRAS, 441, 2593

Oñorbe J., Garrison-Kimmel S., Maller A. H., Bullock J. S., Rocha M., Hahn O., 2014, MNRAS, 437, 1894

Oh S.-H., Brook C., Governato F., Brinks E., Mayer L., de Blok W. J. G., Brooks A., Walter F., 2011, AJ, 142, 24

Oh S.-H., de Blok W. J. G., Walter F., Brinks E., Kennicutt, Jr. R. C., 2008, AJ, 136, 2761

Oh S.-H. et al., 2015, AJ, 149, 180

Okamoto T., 2008, Astrophysics and Space Science Proceedings, 4, 367

Okamoto T., Gao L., Theuns T., 2008, MNRAS, 390, 920

Oke J. B., 1965, ARA&A, 3, 23

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Oliva-Altamirano P., Brough S., Jimmy, Kim-Vy T., Couch W. J., McDermid R. M., Lidman C., von der Linden A., Sharp R., 2015, MNRAS, 449, 3347

Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265

Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577

O'Shea B. W., Bryan G., Bordner J., Norman M. L., Abel T., Harkness R., Kritsuk A., 2004, ArXiv Astrophysics e-prints

Osterbrock D. E., Ferland G. J., 2006, Astrophysics of gaseous nebulae and active galactic nuclei

Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600

Ott J., Walter F., Brinks E., 2005, MNRAS, 358, 1453

Owen J. M., Villumsen J. V., Shapiro P. R., Martel H., 1998, ApJS, 116, 155

Padmanabhan T., 1993, Structure Formation in the Universe. p. 499

Pastrav B. A., Popescu C. C., Tuffs R. J., Sansom A. E., 2012, in IAU Symposium, Vol. 284, The Spectral Energy Distribution of Galaxies - SED 2011, Tuffs R. J., Popescu C. C., eds., pp. 306–308

Patton D. R., Torrey P., Ellison S. L., Mendel J. T., Scudder J. M., 2013, MNRAS, 433, L59

Peacock J. A., 1999, Cosmological Physics. p. 704

Peacock J. A., Heavens A. F., 1990, MNRAS, 243, 133

Peebles P. J. E., 1969, ApJ, 155, 393

Peebles P. J. E., 1970, AJ, 75, 13

Peebles P. J. E., 1971, A&A, 11, 377

Peebles P. J. E., 1982, ApJ, 263, L1

Peebles P. J. E., 1983, ApJ, 274, 1

Peebles P. J. E., 1993, Principles of Physical Cosmology

Peebles P. J. E., 2000, ApJ, 534, L127

Pence W. D., Chiappetti L., Page C. G., Shaw R. A., Stobie E., 2010, A&A, 524, A42

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097

Perlmutter S. et al., 1999, ApJ, 517, 565

Petrosian V., 1976, ApJ, 209, L1

Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59

Pettini M., Shapley A. E., Steidel C. C., Cuby J.-G., Dickinson M., Moorwood A. F. M., Adelberger K. L., Giavalisco M., 2001, ApJ, 554, 981

Pforr J., Maraston C., Tonini C., 2012, MNRAS, 422, 3285

Pforr J., Maraston C., Tonini C., 2013, MNRAS, 435, 1389

Pilbratt G. L. et al., 2010, A&A, 518, L1

Pilkington K. et al., 2012, A&A, 540, A56

Pilyugin L. S., 2001, A&A, 369, 594

Pilyugin L. S., Vílchez J. M., Thuan T. X., 2010, The Astrophysical Journal, 720, 1738

Piontek F., Steinmetz M., 2011, MNRAS, 410, 2625

Planck Collaboration et al., 2015a, ArXiv e-prints

Planck Collaboration et al., 2015b, ArXiv e-prints

Platen E., van de Weygaert R., Jones B. J. T., Vegter G., Calvo M. A. A., 2011, MNRAS, 416, 2494

Pollack J. B., Hollenbach D., Beckwith S., Simonelli D. P., Roush T., Fong W., 1994, ApJ, 421, 615

Pontzen A., Governato F., 2012, MNRAS, 421, 3464

Pontzen A. et al., 2008, MNRAS, 390, 1349

Portinari L., Chiosi C., Bressan A., 1998, A&A, 334, 505

Powell L. C., Slyz A., Devriendt J., 2011, MNRAS, 414, 3671

Power C., Knebe A., 2006, MNRAS, 370, 691

Prada F., Klypin A. A., Simonneau E., Betancort-Rijo J., Patiri S., Gottlöber S., Sanchez-Conde M. A., 2006, ApJ, 645, 1001

Prantzos N., 2008, in EAS Publications Series, Vol. 32, EAS Publications Series, Charbonnel C., Zahn J.-P., eds., pp. 311–356

Press W. H., Schechter P., 1974, ApJ, 187, 425

Prieto M. A., Reunanen J., Tristram K. R. W., Neumayer N., Fernandez-Ontiveros J. A., Orienti M., Meisenheimer K., 2010, MNRAS, 402, 724

Putman M. E., Bland-Hawthorn J., Veilleux S., Gibson B. K., Freeman K. C., Maloney P. R., 2003, ApJ, 597, 948

Putman M. E., Peek J. E. G., Joung M. R., 2012, ARA&A, 50, 491

Quinn P. J., Salmon J. K., Zurek W. H., 1986, Nature, 322, 329

Rabii B. et al., 2006, Review of Scientific Instruments, 77, 071101

Raiteri C. M., Villata M., Navarro J. F., 1996, A&A, 315, 105

Ratra B., Peebles P. J. E., 1988, Phys. Rev. D, 37, 3406

Read J. I., Wilkinson M. I., Evans N. W., Gilmore G., Kleyna J. T., 2006, MNRAS, 366, 429

Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2009, MNRAS, 394, 624

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Renzini A., Buzzoni A., 1986, in Astrophysics and Space Science Library, Vol. 122, Spectral Evolution of Galaxies, Chiosi C., Renzini A., eds., pp. 195–231

Ricotti M., Gnedin N. Y., 2005, ApJ, 629, 259

Rosa-González D., Terlevich E., Terlevich R., 2002, MNRAS, 332, 283

Rosswog S., 2009, New A Rev., 53, 78

Roth M. M. et al., 2005, PASP, 117, 620

Ruiz-Lara T. et al., 2016, MNRAS, 456, L35

Ruszkowski M., Brüggen M., Begelman M. C., 2004, ApJ, 611, 158

Ruszkowski M., Enßlin T. A., Brüggen M., Begelman M. C., Churazov E., 2008, MNRAS, 383, 1359

Salim S., 2014, Serbian Astronomical Journal, 189, 1

Salim S., Rich R. M., Charlot S., Brinchmann J., Johnson, 2007, ApJS, 173, 267

Salpeter E. E., 1955, ApJ, 121, 161

Sánchez S. F., Cardiel N., Verheijen M. A. W., Martín-Gordón D., Vilchez J. M., Alves J., 2007, A&A, 465, 207

Sánchez S. F. et al., 2016a, ArXiv e-prints

Sánchez S. F. et al., 2012, A&A, 538, A8

Sánchez S. F., Pérez E., Rosales-Ortega F. F., et al., 2015, A&A, 574, A47

Sánchez S. F. et al., 2016b, Rev. Mexicana Astron. Astrofis., 52, 171

Sánchez S. F. et al., 2016c, Rev. Mexicana Astron. Astrofis., 52, 21

Sánchez-Blázquez P., Gorgas J., Cardiel N., González J. J., 2006, A&A, 457, 787

Sawala T., Scannapieco C., Maio U., White S., 2010, MNRAS, 402, 1599

Scalo J., 1998, in Astronomical Society of the Pacific Conference Series, Vol. 142, The Stellar Initial Mass Function (38th Herstmonceux Conference), Gilmore G., Howell D., eds., p. 201

Scalo J. M., 1986, Fund. Cosmic Phys., 11, 1

Scannapieco C., Creasey P., Nuza S. E., Yepes G., Gottlöber S., Steinmetz M., 2015, A&A, 577, A3

Scannapieco C., Gadotti D. A., Jonsson P., White S. D. M., 2010, MNRAS, 407, L41

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2005, MNRAS, 364, 552

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2006, MNRAS, 371, 1125

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2008, MNRAS, 389, 1137

Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, MNRAS, 396, 696 (S09)

Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2011, MNRAS, 417, 154

Scannapieco C., et al., 2012, MNRAS, 423, 1726

Schaye J., Crain R. A., Bower R. G., Furlong M., et al., 2015, MNRAS, 446, 521

Schmidt M., 1959, ApJ, 129, 243

Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573

Scudder J. M., Ellison S. L., Torrey P., Patton D. R., Mendel J. T., 2012, MNRAS, 426, 549

Seaton M. J., 1979, MNRAS, 187, 73P

Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978

Shen S., Wadsley J., Stinson G., 2010, MNRAS, 407, 1581

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Shimasaku K. et al., 2001, AJ, 122, 1238

Shu F. H., 1983, ApJ, 273, 202

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Silk J., 1977, ApJ, 211, 638

Silk J., 1997, ApJ, 481, 703

Silva L., Granato G. L., Bressan A., Danese L., 1998, arXiv:astro-ph/9806314

Skrutskie M. F., Cutri R. M., Stiening R., Weinberg M. D., et al., 2006, AJ, 131, 1163

Smee S. A. et al., 2013, AJ, 146, 32

Smith B., Sigurdsson S., Abel T., 2008, MNRAS, 385, 1443

Smith D. J. B., Hayward C. C., 2015, MNRAS, 453, 1597

Snyder G. F., Cox T. J., Hayward C. C., Hernquist L., Jonsson P., 2011, ApJ, 741, 77

Snyder G. F., Hayward C. C., Sajina A., Jonsson P., Cox T. J., Hernquist L., Hopkins P. F., Yan L., 2013, ApJ, 768, 168

Snyder G. F., Lotz J., Moody C., Peth M., Freeman P., Ceverino D., Primack J., Dekel A., 2015, MNRAS, 451, 4290

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Sommer-Larsen J., Dolgov A., 2001, ApJ, 551, 608

Sommer-Larsen J., Götz M., Portinari L., 2003, ApJ, 596, 47

Song M. et al., 2015, ArXiv e-prints

Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C.-A., Muratov A. L., Kereš D., Hopkins P. F., 2015, ArXiv e-prints

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2012, Astronomische Nachrichten, 333, 515

Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V. et al., 2008, MNRAS, 391, 1685

Springel V. et al., 2005, Nature, 435, 629

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Stadel J., Potter D., Moore B., Diemand J., Madau P., Zemp M., Kuhlen M., Quilis V., 2009, MNRAS, 398, L21

Starobinsky A. A., 1980, Physics Letters B, 91, 99

Stasińska G., 2002, arXiv:astro-ph/0207500

Stasińska G., 2005, A&A, 434, 507

Stasińska G., 2007, arXiv:0704.0348

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Steinmetz M., 1996, MNRAS, 278, 1005

Steinmetz M., Bartelmann M., 1995, MNRAS, 272, 570

Steinmetz M., Mueller E., 1994, A&A, 281, L97

Steinmetz M., Muller E., 1995, MNRAS, 276, 549

Stevens A. R. H., Martig M., Croton D. J., Feng Y., 2014, MNRAS, 445, 239

Stierwalt S., Besla G., Patton D., Johnson K., Kallivayalil N., Putman M., Privon G., Ross G., 2015, ApJ, 805, 2

Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074

Stinson G. S., Bailin J., Couchman H., Wadsley J., Shen S., Nickerson S., Brook C., Quinn T., 2010, MNRAS, 408, 812

Stinson G. S. et al., 2013a, MNRAS, 436, 625

Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013b, MNRAS, 428, 129

Stinson G. S. et al., 2012, MNRAS, 425, 1270

Stone J. M., Gardiner T. A., Teuben P., Hawley J. F., Simon J. B., 2008, ApJS, 178, 137

Strateva I., Ivezić Ž., Knapp G. R., Narayanan V. K., et al., 2001, AJ, 122, 1861

Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253

Tapia M. T., Balcells M., Eliche-Moral M. C., 2010, in American Institute of Physics Conference Series, Vol. 1240, American Institute of Physics Conference Series, Debattista V. P., Popescu C. C., eds., pp. 423–424

Tempel E., Stoica R. S., Martínez V. J., Liivamägi L. J., Castellan G., Saar E., 2014, MNRAS, 438, 3465

Teyssier R., 2002, A&A, 385, 337

Thacker R. J., Couchman H. M. P., 2000a, ArXiv Astrophysics e-prints

Thacker R. J., Couchman H. M. P., 2000b, ApJ, 545, 728

Thacker R. J., Scannapieco E., Couchman H. M. P., 2006, ApJ, 653, 86

Thielemann F.-K. et al., 2003, Nuclear Physics A, 718, 139

Thielemann F.-K., Nomoto K., Hashimoto M., 1993, in Origin and Evolution of the Elements, Prantzos N., Vangioni-Flam E., Casse M., eds., pp. 297–309

Thielemann F.-K., Nomoto K., Yokoi K., 1986, A&A, 158, 17

Thomas D., 1999, MNRAS, 306, 655

Thoul A. A., Weinberg D. H., 1995, ApJ, 442, 480

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tissera P. B., White S. D. M., Scannapieco C., 2012, MNRAS, 420, 255

Tojeiro R., Heavens A. F., Jimenez R., Panter B., 2007, MNRAS, 381, 1252

Tojeiro R., Wilkins S., Heavens A. F., Panter B., Jimenez R., 2009, ApJS, 185, 1

Tolstoy E., Hill V., Tosi M., 2009, ARA&A, 47, 371

Toomre A., 1976, in BAAS, Vol. 8, Bulletin of the American Astronomical Society, p. 354

Toomre A., Toomre J., 1972, ApJ, 178, 623

Tormen G., 1996, in Dark Matter in Cosmology Quantam Measurements Experimental Gravitation, Ansari R., Giraud-Heraud Y., Tran Thanh Van J., eds., p. 207

Tormen G., Bouchet F. R., White S. D. M., 1997, MNRAS, 286, 865

Torrey P. et al., 2015, MNRAS, 447, 2753

Tosi M., 2007, in Astronomical Society of the Pacific Conference Series, Vol. 378, Why Galaxies Care About AGB Stars: Their Importance as Actors and Probes, Kerschbaum F., Charbonnel C., Wing R. F., eds., p. 353

Trager S. C., Faber S. M., Worthey G., González J. J., 2000a, AJ, 120, 165

Trager S. C., Faber S. M., Worthey G., González J. J., 2000b, AJ, 119, 1645

Trager S. C., Worthey G., Faber S. M., Burstein D., González J. J., 1998, ApJS, 116, 1

Tremonti C. A. et al., 2004, ApJ, 613, 898

Treu T., Ellis R. S., Kneib J.-P., Dressler A., Smail I., Czoske O., Oemler A., Natarajan P., 2003, ApJ, 591, 53

van Albada G. B., 1961, AJ, 66, 590

van den Bosch F. C., Abel T., Croft R. A. C., Hernquist L., White S. D. M., 2002, ApJ, 576, 21

van Haarlem M. P. et al., 2013, A&A, 556, A2

Vázquez G. A., Leitherer C., 2005, ApJ, 621, 695

Vazquez-Semadeni E., 2015, IAU General Assembly, 22, 2250878

Veilleux S., Rupke D. S., 2005, in Astronomical Society of the Pacific Conference Series, Vol. 331, Extra-Planar Gas, Braun R., ed., p. 313

Velazquez H., White S. D. M., 1999, MNRAS, 304, 254

Verheijen M. A. W., Bershady M. A., Andersen D. R., Swaters R. A., Westfall K., Kelz A., Roth M. M., 2004, Astronomische Nachrichten, 325, 151

Villalobos Á., Kazantzidis S., Helmi A., 2010, ApJ, 718, 314

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M. et al., 2014, MNRAS, 444, 1518

Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 425, 3024

von Hippel T., Gilmore G., Tanvir N., Robinson D., Jones D. H. P., 1996, AJ, 112, 192

Wadsley J. W., Stadel J., Quinn T., 2004, New A, 9, 137

Walcher C. J., Lamareille F., Vergani D., Arnouts S., Buat V., et al., 2008, A&A, 491, 713

Walcher C. J., Wisotzki L., Bekeraité, S. e. a., 2014, A&A, 569, A1

Walcher J., Groves B., Budavári T., Dale D., 2011, Ap&SS, 331, 1

Walker I. R., Mihos J. C., Hernquist L., 1996, ApJ, 460, 121

Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83

Wang Y., Ferland G. J., Lykins M. L., Porter R. L., van Hoof P. A. M., Williams R. J. R., 2014, MNRAS, 440, 3100

Weinberg D. H., Davé R., Katz N., Hernquist L., 2004, ApJ, 601, 1

Weinberg S., 1972, Gravitation and Cosmology: Principles and Applications of the General

Theory of Relativity. p. 688
Welikala N., Kneib J.-P., 2012, arXiv:1202.0494
Werner M. W. et al., 2004, ApJS, 154, 1
Wetzstein M., Nelson A. F., Naab T., Burkert A., 2009, ApJS, 184, 298
White S. D. M., 1976, MNRAS, 177, 717
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Whitmore B. C., Gilmore D. M., Jones C., 1993, ApJ, 407, 489
Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Wild V. et al., 2014, A&A, 567, A132
Willman B. et al., 2005, AJ, 129, 2692
Wood P. R., Olivier E. A., Kawaler S. D., 2004, ApJ, 604, 800
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181
Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, ApJS, 94, 687
Worthey G., Ottaviani D. L., 1997, ApJS, 111, 377
Wright E. L., 2006, PASP, 118, 1711
Wuyts S., Franx M., Cox T. J., Hernquist L., Hopkins P. F., Robertson B. E., van Dokkum P. G., 2009, ApJ, 696, 348
Wyse R. F. G., Gilmore G., Franx M., 1997, ARA&A, 35, 637
Yadav J. K., Bagla J. S., Khandai N., 2010, MNRAS, 405, 2009
Yano T., Nagashima M., Gouda N., 1996, ApJ, 466, 1
Yasuda N., Fukugita M., Narayanan V. K., Lupton R. H., Strateva I., et al., 2001, AJ, 122, 1104
York D. G. et al., 2000, AJ, 120, 1579
Yoshida N., Hosokawa T., Omukai K., 2012, Progress of Theoretical and Experimental Physics, 2012, 01A305
Zafar T., Watson D., 2013, A&A, 560, A26
Zaritsky D., Kennicutt, Jr. R. C., Huchra J. P., 1994, ApJ, 420, 87
Zaroubi S., 2013, in Astrophysics and Space Science Library, Vol. 396, The First Galaxies, Wiklind T., Mobasher B., Bromm V., eds., p. 45
Zentner A. R., Bullock J. S., 2003, ApJ, 598, 49
Zuckerman B., Evans, II N. J., 1974, ApJ, 192, L149

Acknowledgments

I want to thank Cecilia Scannapieco and Matthias Steinmetz for giving me the possibility to work at the AIP and write there my dissertation. I thank Cecilia Scannapieco for supervising and providing help and advices whenever I needed. I acknowledge Cecilia Scannapieco and Michael Aumer for granting me access to their suites of simulations, and Pierre-Antoine Poulhazan and Peter Creasey for sharing their new chemical enrichment code. I also thank Yago Ascasibar for sharing is passion in scientific resarch, as well as ideas and helpful advices. I acknowlege Jakob Walcher for fruitful scientific discussions and for providing his photometry fitting code, and Anna Gallazzi for processing our mock observations with the SDSS pipeline. I thank all my colleagues at the AIP and UAM for their friendship and help during these years, in particular Javier Casado, Umberto Maio, Metin Ata, Pierre-Antoine Poulhazan, Ugur Ural, Peter Creasey and Tilmann Piffl. I would like to thank also Volker Müller and Lutz Wisotzki for giving me the possibility to teach at the Universität Potsdam and Humboldt-Universität zu Berlin.

I acknowledge support from the Leibniz Gemeinschaft through SAW-Project SAW-2012-AIP-5 129, from the High Performance Computer in Bavaria (SuperMUC) through the Project pr94zo, and from the DAAD through the Spain-Germany Collaboration programm (Project ID 57050803).

The SDSS dataset is founded by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. The SDSS Web Site is http://www.sdss.org/.

The major partners of the CALIFA survey are the Institute of Astrophysics of Andalusia, the Max-Planck-Institut für Astronomie, and the Calar Alto observatory. The web site of the CALIFA collaboration is http://califa.caha.es.

This document was typeset by LaTeX.