Università degli Studi di Trieste
Department of Physics

PhD Thesis

On the Role of Cold Gas in Galaxy Evolution

PhD Student:

Anna Zoldan

Supervisor:

Dr. Gabriella De Lucia

Contents

Introduction **2**

1 Thesis context **3**
1.1 The role of gas in galaxy evolution . . . 3
1.2 Observational framework . . . 6
1.3 Theoretical framework . . . 9
1.4 Thesis objectives . . . 12

2 Background on cosmology and simulations **15**
2.1 Cosmology and Large Scale Structure . . . 15
2.1.1 Cosmology and astrophysics: different views of the same Universe . . . 15
2.1.2 Cosmological principle and cosmological measurements . . . 16
2.1.3 The Big Bang and the baryons' origin . . . 18
2.1.4 Perturbations and growth of structures . . . 18
2.1.5 Statistical description of perturbations . . . 20
2.2 Cosmology and astrophysics in simulations . . . 23
2.2.1 Dark Matter only N-body cosmological simulations . . . 24
2.2.2 Halo and subhalos identification . . . 25
2.2.3 Halo Occupation Distribution . . . 26
2.2.4 Hydrodynamical simulations . . . 27
2.2.5 Semi-Analytic Models: state of the art . . . 28

3 Semi-Analytic Models **33**
3.1 The Millennium Simulation . . . 33
3.2 SUBFIND and the merger trees . . . 34
3.3 The Semi-Analytic Model . . . 35
3.3.1 Re-ionization . . . 38
3.3.2 Cooling . . . 38
3.3.3 Star Formation . . . 39
3.3.4 Stellar feedback . . . 41
3.3.5 Metal enrichment . . . 42
3.3.6 AGN feedback . . . 43
3.3.7 Mergers and disk instabilities . . . 44

4 HI-selected galaxies **47**
4.1 Introduction . . . 47
4.2 Simulations and galaxy formation models . . . 49
4.2.1 Galaxy formation models . . . 49
4.2.2 Dark matter simulation and merger trees . . . 51
4.2.3 Light-cones algorithm . . . 51
4.3 Neutral hydrogen distribution and scaling relations . . . 54
4.3.1 Neutral hydrogen mass estimates . . . 54
4.3.2 Stellar and HI Mass Functions . . . 55
4.3.3 Scaling relations . . . 57
4.4 Two-point correlation function . . . 59
4.4.1 The projected correlation function . . . 60
4.4.2 Model predictions . . . 60
4.4.3 Halo occupation distribution . . . 62
4.5 The role of satellite galaxies . . . 66
4.6 Relations with the dark matter halo . . . 69
4.6.1 HI galaxy content and maximum halo mass . . . 69
4.6.2 HI galaxy content and halo spin . . . 71
4.7 Conclusions . . . 72

5 Sizes and specific angular momenta **77**
5.1 Introduction . . . 77
5.2 The model . . . 79
5.2.1 The cosmological simulation and the merger tree . . . 80
5.2.2 The fiducial semi-analytic model . . . 80
5.3 The size-mass relation . . . 84
5.3.1 The size of galactic components . . . 89
5.3.2 Early Type Central and Satellite Galaxies . . . 91
5.3.3 The size-mass relation for X17 modifications . . . 92
5.4 The specific angular momentum . . . 94
5.4.1 Specific angular momentum determination . . . 94
5.4.2 Comparison with observations . . . 100
5.4.3 Specific angular momentum in X17CA3 and X17G11 . . . 102
5.5 Evolution of galaxies dynamics . . . 104
5.5.1 Morphological dependencies of the size-mass relation . . . 104
5.5.2 Disk instability in central and satellite galaxies . . . 107
5.5.3 The role of cold gas in the dynamical evolution . . . 109
5.5.4 Evolution in X17CA3 and X17G11 . . . 112
5.6 Conclusions . . . 113

6 Conclusions and future perspectives **117**
6.1 Ongoing work: mock catalogs . . . 118
6.1.1 Mock catalogs for VANDELS . . . 119
6.1.2 Mock catalogs for SKA . . . 122
6.2 Future plans . . . 124
6.2.1 Dedicated cosmological simulations . . . 124
6.2.2 A consistent treatment for disk instabilities . . . 126

7 Bibliography **131**

Chapter 1

Thesis context

Galaxy evolution results from a complex interplay of physical processes, that act on a wide range of scales and times. One crucial piece of this complex puzzle is cold gas. Indeed, all physical processes occurring within galaxies, as well as those related to interactions with the external environment, directly influence (or are regulated by) the dynamics and the content of cold gas.

In this work, I take advantage of state-of-the-art models of galaxy formation and evolution embedded in a cosmological context, and use them to explore the role of cold gas in the physical processes driving galaxy evolution. This approach allows a controlled analysis of individual galaxy histories, and of the contributions of the specific processes involved. I can, in this way, follow the origin of the observed relations, providing a physical interpretation for existing data and predictions for higher redshift.

This Thesis work provides results useful for the interpretation of data collected by ongoing and upcoming surveys, such as those planned for ASKAP and MeerKAT, precursors of the Square Kilometre Array.

In this Chapter I provide an overview of the Thesis context, describing the physical processes that rule the cold gas content evolution, and the current status of cold gas observations and related theoretical studies. In the final part of this chapter, I describe in detail the objectives of this study and the methods adopted.

1.1 The role of gas in galaxy evolution

In the current standard picture of structure formation, galaxies are believed to form from gas condensation at the center of dark matter halos. The gas, trapped within the potential wells of collapsing structures, is shock heated to high temperatures. It can then cool and be accreted on the central regions, where it settles in a rotating disk. The accretion of gas on central galaxies takes place very efficiently in the case of accretion along filamentary, anisotropic structures (the so-called "cold mode"), or less efficiently when the accretion is isotropic and undergoes shocks when interacting with the medium (the so-called "hot mode"; Silk, 1977; Rees and Ostriker, 1977; Binney, 1977; White and Rees, 1978). The difference between the two accretion modes is set by the time scale at which the accreted gas cools, compared to the dynamical

time of the halo. The latter is the typical time scale for a particle to cross the halo. In the case of hot accretion, typical of massive halos ($\gtrsim 10^{12} M_{\odot}$) at late times, the cooling time is usually longer than the dynamical time, and the shocked gas forms a quasi-static hot atmosphere that extends out to the virial radius. In cold accretion, instead, typical of small halos at early times, the cooling time is short compared to the dynamical one. In this case, time is not enough to form a quasi-static hot atmosphere, and the shocks occur very near to the halo center, followed by a very rapid cooling of the gas. This two-modes description of gas accretion was introduced in early semi-analytic models (White and Frenk, 1991a), and is included in all recently published models (e.g. Benson et al., 2001b; Croton et al., 2006). The validity of this cooling model has been confirmed by dedicated hydrodynamical simulations (Yoshida et al., 2002; Helly et al., 2003; Saro et al., 2010; Lu et al., 2011; Monaco et al., 2014, and references therein). Observationally, hot gas halos are observed through their X-ray emission around individual galaxies, groups and clusters of galaxies (Crain et al., 2010; Anderson and Bregman, 2011). The existence of cold inflows (gas accreting in cold mode) has not been confirmed observationally yet. Indirect evidences are inferred from observations of atomic hydrogen absorbers in proximity of high redshift galaxies (Giavalisco et al., 2011), or from high-velocity clouds around the Milky Way (Sancisi et al., 2008), whose interpretation is, however, not unequivocal.

Due to conservation of angular momentum, cold gas is expected to settle in a rotating disk. There is general agreement that the hot gas halos statistically follow the same distribution of spin of DM halos (van den Bosch et al., 2002, 2003), and that the angular momenta of the hot gas and of the DM halo are strictly correlated (Sharma and Steinmetz, 2005). Hydrodynamical simulations show that the retention of the hot gas angular momentum in the cold gas disk depends on the assembly history of galaxies (Zavala et al., 2008). In the case of smooth accretion histories, typical of disk dominated galaxies, the angular momentum of the parent halo is transferred to the gaseous disk and preserved. In the case of bulge dominated galaxies, large fractions of the angular momentum are lost during the merger events that lead to the formation of the central bulge component. The dynamical state of the cold gas disk is important, because it determines its typical size and stability.

The cold gas disk is the place where stars form. Proto-stars are observed to form in giant molecular clouds (GMC), that are cold (T$\sim$ 10 K), over-dense gas regions, of size 20-200 pc. GMCs are the ideal place for the condensation of molecules, and, due to their high density, for star formation. The mechanisms driving GMCs formation and their collapse are still not well understood. The most accepted scenario is the fragmentation of the cold gas disk, and the collapse of over-densities onto the GMCs, in a top-down scenario (Elmergreen, 1975; Elmegreen et al., 1995). Other scenarios have been proposed, and this matter is still debated (see McKee and Ostriker, 2007, for a review of the GMC theoretical and observational framework).

First studies on modeling of star formation were based on the direct correlation observed between the surface density of star formation and that of the cold gas disk, in a power law called Kennicutt-Schmidt relation (Schmidt, 1959; Kennicutt, 1998). More recently, high resolution maps of the spatial distribution of star formation, atomic and molecular cold gas have become available for relatively large galaxy samples. These new observations have allowed a deeper

understanding of the distribution of the cold gas main components (atomic and molecular hydrogen, HI and H_2), and of their relation with star formation. Data show a strong correlation between regions of high star formation and those of high H_2 surface density, while atomic hydrogen exhibits a poor correlation with star formation (Wong and Blitz, 2002; Kennicutt et al., 2007; Leroy et al., 2008; Bolatto et al., 2011; Krumholz, 2013). These studies have led to new investigations of the processes that regulate the transition from atomic to molecular gas, and of the mechanisms driving star formation in GMC (see Krumholz, 2014, for a review). Different studies have proposed relations to parametrize the molecular to atomic hydrogen mass ratio. Some are based on empirical relations, e.g. with the mid-plane pressure of the disk (Blitz and Rosolowsky, 2006); some on analytic solutions for the photo-dissociation in molecular clouds (Krumholz et al., 2008, 2009), and others on direct hydrodynamical simulations (Gnedin et al., 2009).

During their late evolutionary stages, stars eject gas, metals and energy, that affect significantly the surrounding gas. During these events, referred to as "stellar feedback", metals, produced by nuclear reactions inside stars, are ejected through winds or Super Novae explosions, and mix with the surrounding medium. The energy released has the double effect of re-heating part of the cold gas, and, carrying momentum, originating galactic outflows. The fate of these outflows depends on their velocity: high velocity winds can carry material out of the potential well of the galaxy, excluding it from the baryon cycle (at least temporarily). If the velocity is, instead, small, the outflow can be re-accreted on a short time-scale, contributing to refuel and enrich the cold gas in the galaxy disk.

Another important feedback source is the energy released by black holes at the center of galaxies. This so-called "AGN feedback" (where AGN stands for Active Galactic Nuclei) is necessary to solve the over-cooling problem: in a scenario with only stellar feedback, the most massive halos would accrete too much cold gas at late times, forming very massive galaxies actively forming stars, which is in disagreement with observations (Benson et al., 2003). One manifestation of AGN feedback is the "quasar mode', associated to the inflow of large amounts of cold gas towards the central black hole. It is believed to occur during mergers, and contributes to most of the mass growth of black holes. Observationally, it is identified by large obscured clouds of accreting material, strong winds of some thousands of $\mathrm{km\,s^{-1}}$ (Reynolds et al., 1997; Crenshaw et al., 2003; Pounds et al., 2003; Reeves et al., 2009; Tombesi et al., 2010), and outflows of tens to thousands of solar masses per year, at some hundred $\mathrm{km\,s^{-1}}$ (Heckman et al., 2000; Veilleux et al., 2005; Strickland and Heckman, 2009; Weiner et al., 2009).

Another manifestation of AGN feedback is the "radio mode", associated with lower accretion rates of hot gas. This mode is typical of central galaxies in groups and clusters, for which X-rays observations predict very high cooling rates, but there is no evidence of a correlated enhanced star formation (Fabian et al., 1991). Radio mode AGN feedback reheats the surrounding material, preventing massive star formation (Pedlar et al., 1990; Baum and O'Dea, 1991; Tabor and Binney, 1993; Tucker and David, 1997). This feedback manifests in the form of powerful jets that inflate bubbles of relativistic plasma (Churazov et al., 2000; McNamara et al., 2000; Dunn and Fabian, 2006, 2008). Because of its balancing effect between cooling and heating, this mode is often referred

to as "maintenance feedback" (Pizzolato and Soker, 2010; Narayan and Fabian, 2011).

Radio AGN feedback was introduced in semi-analytic models (Croton et al., 2006; Bower et al., 2006; Somerville et al., 2008), and in hydrodynamical simulations (Brüggen et al., 2007; Cattaneo and Teyssier, 2007; Dubois et al., 2010; Gaspari et al., 2011, 2012), to counteract cooling flows. The quasar mode is usually implemented as a growing mode for the central black hole during mergers, but its feedback is not always modeled explicitly.

Finally, cold gas is sensitive to the environment, through interactions with other galaxies and with the hot gas halo. When a galaxy falls onto a more massive halo, its hot gas, due to the interaction with the gravitational potential of the cluster, might be stripped from the galaxy, in a process called strangulation (Larson et al., 1980). A similar outcome results when a galaxy moves with high speed through the intra-cluster medium: the pressure exerted by the medium can strip loosely bound gas off the galaxy. This effect, first explained and modeled in Gunn and Gott (1972), is usually referred to as "ram-pressure" stripping. Clusters also host repeated fast galaxy-galaxy encounters: these interactions can disturb the morphologies of galaxies, favoring the compression of gas and, therefore, star formation or star bursts. This process is called harassment (Moore et al., 1996). All the above processes are effective in removing gas in galaxies residing in dense environments, and therefore in suppressing star formation. Their relative importance and dependence on galaxy mass, redshift and mass scale remains subject of debate (Vollmer et al., 2001; Wetzel et al., 2013; Peng et al., 2015; Westmeier et al., 2015; Davies et al., 2016).

1.2 Observational framework

For its importance in all the physical processes described above, cold gas can be considered as a record of star formation and assembly history of galaxies. Therefore, its abundance and dynamics can provide important information on physical processes regulating galaxy evolution. Cold gas can be observed directly through its main components: molecular and atomic hydrogen. Molecular hydrogen is not measurable directly, and its amount is usually inferred indirectly from the emission associated with rotational lines of the CO molecule, the second most abundant molecule in giant molecular clouds. The conversion factor from CO to H_2 is uncertain, and likely depends on the physical and thermodynamical state of the inter-stellar medium (Glover and Mac Low, 2011).

Atomic hydrogen is, in contrast, directly measurable from the 21 cm line, a forbidden line associated with the spin-flip transition of the electron. It was first detected in the Milky Way and Magellanic Clouds (Kerr and Hindman, 1953), and in extragalactic systems later on (Roberts, 1975). These first observations were made using single dish telescopes. Later on, spectral line aperture synthesis provided new powerful instrumentation for radio measurements. Nowadays, the HI galaxy content has been measured for thousands of galaxies up to $z \sim 0.1$, providing the necessary statistics for quantifying its correlation with other galaxy properties.

The relation between HI and stellar mass in a statistical sample of galaxies was first characterized using data from the GASS survey (Catinella et al., 2010, 2012). This observational project provided the HI content (or upper limits in

the case of non-detections) for $\sim$ 1000 galaxies selected from the SDSS survey (York et al., 2000), with mass $M_* > 10^{10}M_\odot$ and redshift $0.025< z <0.05$. The HI gas mass decreases linearly with increasing galaxy stellar mass, with a large scatter. This work was successively extended to the H_2 content, with the COLDGASS survey (Saintonge et al., 2011), that provided an insight on the relation between H_2, HI and stellar mass for a large sample of mass-selected galaxies.

The cold gas content of a galaxy depends strongly on its morphology. Both HI and H_2 are detected in spirals, in the form of rotating disks (Haynes and Giovanelli, 1984; Thronson et al., 1989). Ellipticals are typically gas poor and dispersion supported, but some measurements of cold gas were already carried out in the eighties (Sanders, 1980; Knapp et al., 1985; Wiklind and Henkel, 1989). Since this gas was observed to be offset with respect to the optical image of the galaxy, and to have a different kinematics, it was believed to have an external origin. In more recent years, the survey ATLAS^{3D} (Cappellari et al., 2011) has analyzed the stellar and gaseous kinematics of 260 nearby Early Type galaxies (Davis et al., 2011; Serra et al., 2012, 2014). They have found that ET galaxies can be divided in two groups, according to the dominance of rotation against dispersion: fast and slow rotators. Outside the cluster environment, fast and slow rotators are characterized by the presence of gaseous disks: these are co-rotating with the stars in most of the slow rotators, while often counter-rotating in fast rotators. The authors argue that the properties of these gas disks are compatible with recent accretion, and, for this reason, they are not strongly correlated with the stellar dynamics. An analysis of the dynamics of stars, HI and H_2, and of their relation with stellar mass and galaxy morphology in disk galaxies, was performed by Obreschkow and Glazebrook (2014). These authors used 16 nearby spiral galaxies from THINGS (Walter et al., 2008), a high spectral and spatial resolution survey of HI emission in 34 nearby galaxies. They have found that the specific angular momenta of the stellar and of the stellar+gaseous components of the galaxies in their sample lie on a plane in the $j - M - B/T$ space, where B/T is the ratio between bulge and total stellar mass. The gas is found to have an angular momentum slightly larger than that of the stars.

The HI mass function, i.e. the number density of galaxies with different HI mass, has been measured by Zwaan et al. (2005) and Martin et al. (2010), for the blind HI surveys HIPASS (Meyer et al., 2004) and ALFALFA (Giovanelli et al., 2005), limited to redshifts $z < 0.04$ and $z < 0.06$, respectively. These blind surveys were used also to estimate the clustering of HI-selected galaxies (Basilakos et al., 2007; Papastergis et al., 2013). Results show that HI-selected galaxies have a 2-point correlation function that depends weakly on the HI mass, and that, in general, is lower than that measured for luminosity-selected galaxies. Huang et al. (2012) matched the ALFALFA and SDSS samples, and showed that HI-dominated galaxies are usually star-forming, but have stellar populations less evolved than those of a non HI-selected sample of the same stellar mass. Huang et al. (2012) also argued that HI rich galaxies preferentially reside in high spin DM halos. In this work, the spin parameter was evaluated from the rotational velocity of the gaseous disk (estimated from the 21 cm line width) and the disk scale radius (following Hernandez et al., 2007).

An estimate of the H_2 galaxy mass function has been obtained by Keres et al. (2003), from the FCRAO Extragalactic CO Survey (Young et al., 1995).

Galaxies residing in dense environments were early found to be "cold gas deficient" (see Haynes and Giovanelli, 1984, for a review). The deficiency is usually quantified in terms of the difference between the observed HI mass and that expected for a "normal" galaxy (i.e. residing in a region with average density) of similar linear size and morphology. Recent studies, as those based on the already cited ATLAS^{3D} and GASS surveys, have confirmed these results for a wide range of morphologies and stellar masses (Davis et al., 2011; Brown et al., 2017).

All relations described above for HI are measured in the nearby Universe. A few very recent programs are focusing on measuring the HI content of galaxies at redshifts higher than $z \sim 0.1$. These are either based on targeted surveys, that need long integration times (Catinella and Cortese, 2015), or relying on stacking techniques or intensity mapping in the redshift range $z \sim 0.1 - 0.8$ (e.g. Lah et al., 2007; Chang et al., 2010; Geréb et al., 2013). At intermediate redshift ($0.16 < z < 0.22$), BUDHIES is conducting a deep HI survey of galaxies of two clusters and of the large scale structure around them (Verheijen et al., 2007; Jaffé et al., 2012). The COSMOS HI Large Extragalactic Survey is ongoing at the VLA (Fernández et al., 2016), probing the HI content of galaxies in the redshift range $0 < z < 0.5$.

Different is the case for H_2 (see Solomon and Vanden Bout, 2005; Omont, 2007; Carilli and Walter, 2013, for a review), because the excitation lines of CO are visible in a large range of bands, and are easily detectable at high redshift. The main issue, in this case, is the target selection, and the uncertainty related to possible variations of the CO-to-H_2 conversion factor.

Observations of CO from low to high redshift have recently received a boost with ALMA[1], an interferometer operating at wavelengths from 0.32 to 3.6 mm, allowing observations with unprecedented spectral and spatial resolution. Its location is on the Chilean Andes, where the atmosphere is thin and the sub-millimeter radiation is poorly absorbed by atmospheric water vapor. It is composed of 66 high-precision antennas, that can be arranged in different interferometric configurations, setting the sensitivity and resolution according to the scientific goal of the observation. ASPECS is a recent blind CO survey carried out with ALMA, in the same region of the Hubble Ultra Deep Field (Walter et al., 2016). Results from this survey have been used to constrain the CO luminosity function and the evolution of cosmic molecular density up to $z \sim 4.5$ (Decarli et al., 2016). Although uncertainties are still large, preliminary results suggest that the cosmic molecular gas density drops by a factor 3-10 from $z \sim 2$ to $z \sim 0$, with a possible decline at $z > 3$.

The same improvements in resolution and redshift range for HI observations will be possible with the new generation of radio instruments, as the Square Kilometre Array (SKA, Schilizzi et al., 2008) and its precursors. SKA is a radio interferometer of unprecedented sensitivity and resolution, currently under design and construction[2]. The construction is divided in two phases, between 2018 and late 2020s: Phase 1 includes the construction of 500 stations, each containing 250 individual antennas, in Australia, and 200 dishes, including the 64 of MeerKAT, in South Africa. Phase 2 will complete the arrays to 2000 high and mid frequencies dishes and aperture arrays, and a million low-frequency

[1] More information can be found at the official web page: almascience.nrao.edu/sitemap

[2] More information on SKA configurations and building timescales can be found at the official web page: skatelescope.org

antennas.

Different pathfinders will pave the way for SKA, among them MeerKAT, a 64 dishes system now under construction in South Africa, and ASKAP, in Australia, that counts 36 dishes. Some surveys have been already planned. For example, WALLABY and DINGO will be conducted with ASKAP (Johnston et al., 2008). The former is a wide HI blind survey that should cover $\sim 75\%$ of the sky; the latter is a blind HI survey that should provide information on the evolution of HI since $z \sim 0.5$. In the case of MeerKAT, LADUMA (Holwerda et al., 2012) will observe the 21 cm line in a deep field to $z \sim 1$, covering the same field of surveys in other bands.

1.3 Theoretical framework

Until recently, galaxy evolution theoretical studies accounted for cold gas as a single phase component, with temperatures T$< 10^4$ K. Star formation was typically related to the cold gas total surface density, using a Kennicutt-Schmidt law (Schmidt, 1959; Kennicutt, 1998).

Driven by recent observational results, more attention has been devoted to the inclusion of an HI-to-H_2 partition of the modeled cold gas. For example, Blitz and Rosolowsky (2006) proposed a relation to calculate the molecular-to-atomic gas ratio in disks, assuming it depends on the mid-plane hydrostatic pressure. They tested their model against a sample of nearby galaxies, that cover a wide range of magnitudes and metallicities. Its validity has not yet been confirmed for galaxies at higher redshift. Another approach is that of Krumholz et al. (2008, 2009), who calculated the balance between photo-dissociation, molecular self-shielding, and formation of molecules on the surface of dust grains in the central region of a molecular cloud. They evaluated the HI and H_2 masses over several galactic properties, and validated their results against observational data from nearby galaxies. Other studies, such as Gnedin et al. (2009), studied strategies to include the atomic-to-molecular transition in cosmological hydrodynamical simulations. They used hydrodynamical simulations including non equilibrium chemistry, radiative transfer, and self-shielding from photo-dissociation with a phenomenological treatment. Other simulations focused on the HI-H_2 partition in single galaxies or molecular clouds (see, for example, Pelupessy et al., 2006; Robertson and Kravtsov, 2008; Glover and Clark, 2012; Christensen et al., 2012; Walch et al., 2015).

The implementation of a self-consistent treatment of HI and H_2 in large volume cosmological hydrodynamical simulations is difficult. In this case, H_2 formation can either be modeled using sub-grid prescriptions, or the fraction of HI and H_2 can be evaluated in post-processing. This is the case for recent large simulations as OWLS (Schaye et al., 2010), EAGLE (Schaye et al., 2015) and Illustris (Vogelsberger et al., 2014b).

Duffy et al. (2012) extrapolated the HI and H_2 content for the OWLS simulations adopting the relation proposed by Blitz and Rosolowsky (2006) in post-processing, and found realistic gas contents at low redshift, but too many H_2-rich galaxies.

In EAGLE, Lagos et al. (2015a) assigned the H_2 galaxy content in post-processing, using prescriptions proposed by Gnedin and Kravtsov (2011) and Krumholz (2013). The simulation reproduces well the H_2 mass function and scaling rela-

tions, and predicts that molecular gas resides preferentially in high mass, star forming galaxies. For the same simulation, the HI content has been inferred from the neutral hydrogen amount (evaluated using a model for photo-ionization), and using the Blitz and Rosolowsky (2006) relation. The simulation reproduces nicely the cosmic HI mass density evolution (Rahmati et al., 2015), the radial distribution and morphology of HI observed for the GASS galaxies (Bahé et al., 2016), and the dependence on the environment (Marasco et al., 2016). Crain et al. (2017) evaluated the dependence of the HI content on the star formation efficiency, the feedback prescription and the resolution of the simulation, finding that at fixed stellar mass, a stronger feedback would lead to a larger HI content. In the Illustris simulations, the HI is self-consistently evolved in the chemical scheme of the simulation (as described in Ferland et al., 1998), while H_2 is assumed to be proportional to the amount of stars formed. The obtained scaling relations and their evolution show a good agreement with observations (Vogelsberger et al., 2014a; Genel et al., 2014).

In semi-analytic models, the partition of cold gas into HI and H_2 is modeled employing relations similar to those used in hydrodynamical simulations. The first analyses were performed by partitioning the cold gas in post-processing. For example, Obreschkow and Rawlings (2009) partitioned the cold gas in the De Lucia and Blaizot (2007) semi-analytic model, using the relation by Blitz and Rosolowsky (2006). This work was used to make predictions for the HI and H_2 content at high redshift, and to build mock catalogs of the 21 cm line for the Square Kilometre Array. In this Thesis work, I will show that the model described in De Lucia and Blaizot (2007) is not able to reproduce a realistic evolution of the HI galaxy content. Therefore, available mock catalogs for SKA need to be improved by the adoption of a self-consistent model, able to provide a more realistic description of the HI content of galaxies and its evolution.

Other studies, based on a post-processing partition of the cold gas, are those by Power et al. (2010) and Kim et al. (2011), who analyzed the HI mass function and the spatial distribution of HI-selected galaxies predicted by two independent semi-analytic models. Both found a general good agreement with observations, and little evolution with the HI content with redshift.

In several recent studies based on semi-analytic models, the transition from HI to H_2 is modeled self-consistently: at each time step, the cold gas is partitioned in HI and H_2 using an empirical or theoretical prescription (Fu et al., 2010; Lagos et al., 2011b,a; Popping et al., 2014; Xie et al., 2017). These models predict scaling relations in agreement with observations. Interestingly, the resulting stellar properties of model galaxies (mass, star formation rate and metallicity) are not dependent on the specific prescription used to partition the cold gas (Lagos et al., 2011b; Somerville et al., 2015; Xie et al., 2017). This is due to the self-regulation between star formation and stellar feedback, that regulates the stellar production depending on the available quantities of gas. The differences among different prescriptions can be more easily identified at the high redshift or in low mass, recently formed galaxies, where the self-regulation mechanism has not yet become effective. The model by Lagos et al. (2011a) was also used to evaluate the HI mass function and the two-point correlation function of low mass HI-selected galaxies, finding a good agreement with the observed relations for masses $M_{HI} > 10^8 M_\odot$ (Kim et al., 2015, 2017).

The recent availability of integral field spectroscopy for large surveys of galaxies allowed a deeper analysis of the internal dynamics of galaxies. As

explained above, cold gas links the dynamics of the halo and that of the stars. It cools, preserving the angular momentum of the hot gas, coupled with that of the DM halo; the stars in the disk carry the angular momentum of the cold gas they formed from. As a consequence, the size and the dynamics of a galaxy are determined by the history of gas accretion and star formation.

First hydrodynamical simulations suffered of excessive loss of angular momentum, resulting in too compact galaxies, with disks too small compared to observations (Steinmetz and Navarro, 1999; Navarro and Steinmetz, 2000). Several studies suggested that excessive cooling and star formation during the early stages of galaxy formation prevent the formation of extended disks. Suppression of this early cooling produced more realistic disks (Weil et al., 1998; Eke et al., 2000; Abadi et al., 2003; Governato et al., 2004), suggesting that feedback plays an important role in the galaxy dynamics. Thanks to the improvement of numerical resolution and stellar feedback schemes included, many groups have succeeded to reproduce realistic thin disks supported by rotation (Scannapieco et al., 2005, 2006, 2008; Zavala et al., 2008; Governato et al., 2010; Guedes et al., 2011; Danovich et al., 2015).

The distribution of sizes and angular momenta of galaxies were studied using the cosmological simulations cited above. In EAGLE, the feedback scheme is calibrated to reproduce the stellar mass function and galaxy sizes at redshift zero, but the simulation reproduces well the size-mass relation observed at higher redshifts (Furlong et al., 2015). Lagos et al. (2017) have shown that this simulation produces a realistic distribution of specific angular momentum with stellar mass, and that it strongly correlates with the cold gas content.
Illustris does not reproduce a realistic distribution of galactic morphologies, having too many disk-dominated galaxies. This influences the observed size-luminosity relation, that does not agree with the observed trends (Bottrell et al., 2017). The specific angular momentum-stellar mass relation is, instead, well reproduced. This relation was analyzed against different morphological selections and wind schemes (Genel et al., 2015; DeFelippis et al., 2017), finding that strong winds are able to increase the angular momentum, and to prevent angular momentum loss. Stellar feedback is thus confirmed as a fundamental ingredient in regulating the angular momentum evolution, as already found in other hydrodynamical simulations (Übler et al., 2014). Similar investigations were performed using other simulations, finding similar results (Teklu et al., 2015; Pedrosa and Tissera, 2015; Zavala et al., 2016)

Recently, size and angular momentum have been studied in the context of semi-analytic models. The size-mass relation was studied in several works, highlighting the necessity for a treatment for gas dissipation during mergers to have realistic sizes for early type galaxies (Shankar et al., 2013; Tonini et al., 2016). Some recent models include a specific treatment for the exchange of angular momentum among the disk components of the galaxy and the hot gas halo, and a consistent estimation of the disk scale radius (e.g. Lagos et al., 2009; Guo et al., 2011; Benson, 2012; Padilla et al., 2014; Tonini et al., 2016; Stevens et al., 2016).
Lagos et al. (2015b) studied the misalignment between the stellar and the gas angular momenta in ET galaxies, finding a good agreement with observations. For this purpose, they assigned in post-processing the specific angular momenta to model galaxies, following the statistical analysis of the spin-flips in mergers performed by Padilla et al. (2014).

A detailed analysis on the relation between specific angular momentum and stellar mass of galaxies, in the framework of semi-analytic model, has not yet been performed.

1.4 Thesis objectives

This thesis is focused on the role of cold gas in driving the evolution of the physical and structural properties of galaxies in the local Universe. The primary tools used in this PhD work are semi-analytic models of galaxy formation. These methods offer the possibility to access a large range of masses and environments, at a limited computational costs.

I first analyze the agreement between available HI observations in the local Universe, and results from state-of-the-art semi-analytic models. This comparison allows us to identify the prescriptions that mostly influence the cold gas content of galaxies, and to characterize the dependence of the results on the specific model used. The models I consider in my study include different treatments for star formation, stellar feedback, AGN feedback, chemical enrichment, and gas stripping in satellites, which makes it possible to quantify the relative importance of different processes. In particular, I take advantage of the GAEA model (Hirschmann et al., 2016), and its more recent updates described in Xie et al. (2017). The model includes a non-instantaneous recycling approximation and an improved feedback scheme, which leads to a better agreement with a number of observational results. These include the observed correlation between gas fraction and galaxy mass at different cosmic epochs. Furthermore, the model presented in Xie et al. (2017) includes an explicit treatment for the partition of cold gas into HI and H_2, and a star formation law based on the H_2 surface density. For the other models, HI and H_2 masses are extracted in post-processing, using the same prescriptions described in Obreschkow and Rawlings (2009). All the models I consider are based on the same cosmological DM only simulation. While previous studies have focused on results from specific models (e.g. Power et al., 2010; Kim et al., 2011; Lagos et al., 2011a), this is the first time such a systematic comparison is carried out. This allows us to identify robust successes and failures of existing models, as well as to highlight possible directions for future improvements. In addition, this analysis is required to assess the ability of current models to make specific predictions for future observational programs.

In the second part of the Thesis, I focus on the ability of one particular model (that presented in Xie et al., 2017) to reproduce the observed trends for the size-mass relation (Shen et al., 2003; van der Wel et al., 2014; Lange et al., 2015) and the specific angular momentum-mass relation (Romanowsky and Fall, 2012; Cortese et al., 2016). Observations have shown that the scatter in these relations depends on the morphology of galaxies: disky galaxies are larger and have larger specific angular momenta with respect to spheroidal ones. I characterize the role of cold gas in these relations and their scatter. The model I use is ideal for this study, because it includes a specific treatment for the angular momentum of the gaseous and stellar disks. In addition, the scale radii of the disks are computed self-consistently from the specific angular momentum. I include a treatment for dissipation of gas during mergers, following results from hydrodynamical simulations. While the size-mass relation has

been already subject of investigation in previous studies based on semi-analytic models (Shankar et al., 2013; Tonini et al., 2016), a systematic analysis of the specific angular momentum has not yet been performed, and current work is limited to post-processing analyses (Lagos et al., 2015b) or to dedicated models (Stevens et al., 2016). Therefore, my study provides additional insight on cold gas dynamical evolution and its influence on the stellar content.

The success of the galaxy formation model I use, combined with the flexibility of the semi-analytic approach, represent ideal requirements to generate mock light-cones that can be used to interpret ongoing observations, or to guide future observational programs. For the analysis carried out in the first part of the Thesis, I developed a software that creates mock catalogs starting from the model outputs. This software is able to add galaxies in the mock light-cone at the evolutionary stage corresponding to their distance from the observer, and to avoid repetitions of structures along the line of sight. With this software, I created mock catalogs for the ongoing deep field survey VANDELS (McLure et al., 2017). These cones are currently used by the VANDELS group to analyze the environmental dependences of high redshift galaxies. I am currently modifying this software to create realistic mock catalogs of 21 cm lines. This work will be very useful in the scientific planning for the Square Kilometre Array and its precursors. The mock catalogs currently used for this purpose are based on the output of an out-of-date semi-analytic model, that, as I demonstrate in the first part of this work, is not able to reproduce a realistic evolution of the galactic cold gas content.

Chapter 2

Background on cosmology and simulations

In this Chapter, I introduce the essential background of this Thesis work. Specifically, in the following sections I will discuss the basics of cosmology and structure formation, and notions of simulations, with a focus on N-body dark matter simulations and semi-analytic models of galaxy formation and evolution.

2.1 Cosmology and Large Scale Structure

2.1.1 Cosmology and astrophysics: different views of the same Universe

Cosmology and astrophysics are two disciplines of modern physics, both dealing with the evolution of our Universe. They focus, however, on different specific questions, as well as on different phenomena and physical scales.

Cosmology deals with the origin and the evolution of the Universe from a distribution of fluctuations of matter and energy in a relativistic background. The typical scales span from that of the dark matter halos ($\sim$2 Mpc) to that typical of the baryon acoustic oscillations ($\sim$100 Mpc). Astrophysics deals with the physics of the astra, from the smallest objects, like atoms and grains of dust, to collapsed systems, like stars, galaxies and clusters of galaxies.

The relative fields of interest are not well divided, and in several cases the two disciplines overlap, especially at the scales where galaxies grow and evolve. This overlap is the natural consequence of the evolution of structures in a cosmological context, and the formation of galaxies in these structures.

The specific cosmology driving the evolution of our Universe can be described by a set of "cosmological parameters". The values of these parameters uniquely identify a cosmological model, and include the relative densities of the different components of the Universe. Several observational studies are devoted to the measurement of these parameters, and recent results suggest we live in an expanding Universe, where baryons are only a small fraction of the total matter. The main component is dark energy, a repulsive force that drives the Universe expansion (Λ), and accounts for $\sim 68\%$ of the total energy-matter budget. The rest of the Universe is composed of matter, with baryons accounting

for almost $\sim 5\%$, and Dark Matter (DM) for the remaining part ($\sim 27\%$). The nature of DM is still uncertain, but the most accepted theory is that it is composed of weakly interacting massive particles (WIMPs) which do not interact electromagnetically.

According to the standard model of structure formation, structures in the Universe form from the growth of small perturbations in the initial density field. These perturbations grow and separate from the Universe expansion due to their self-gravitation, leading to the collapse of DM halos. The baryons fall into the potential wells formed by DM, preserving its dynamical state.

In this scenario, galaxies and galaxy clusters settle at the center of the DM halos and, for this reason, can be used as proxies for the large scale matter distribution. Measurements of the large scale structure and statistical distribution of matter can then be used to estimate the cosmological parameters. Nevertheless, the transition from baryonic to matter distribution is not straightforward, and requires a good knowledge of the physics that regulates galaxy observable properties and their distribution inside the halos. This bias evaluation is thus the result of a complex interplay between cosmology and astrophysics.

2.1.2 Cosmological principle and cosmological measurements

Modern cosmology is based on the so-called Cosmological Principle, which states that, at sufficiently large scales, the Universe is spatially homogeneous and isotropic. This was confirmed by the first measurements of the Cosmic Microwave Background (CMB, Penzias and Wilson, 1965): this background is a relic radiation from the early Universe, characterized by very small anisotropies. In an homogeneous and isotropic Universe, the only motion allowed to preserve the Cosmological Principle is a pure expansion (or a pure contraction). The expansion of the Universe was first demonstrated by Hubble (1929). He discovered the existence of a linear relation between the distances (x) and the radial velocities (v) of 46 galaxies in the Local Universe: $v = H_0 \cdot x$, where H_0 is the Hubble constant, that Hubble estimated to be equal to $H_0 = 500$ km s^{-1}Mpc^{-1}. Recent measurements of the cosmological parameters set the value of the Hubble constant to $H_0 = 100 \cdot h = 67.3 \pm 1.2$ km s^{-1}Mpc^{-1} (Planck Collaboration et al., 2014).

The expansion paradigm implies an "initial time", when the Universe is concentrated in an extremely hot and dense singularity, a scenario called Big Bang. The time describing the evolution since the Big Bang is called *cosmic time*, and we can define a set of spatial coordinates that satisfy the symmetry of the Cosmic Principle. An observer located in these coordinates would move along the overall cosmic expansion, and, for this reason, they are called *co-moving coordinates*. We can relate the co-moving coordinates and the cosmic time geometrically, using the Robertson-Walker metric:

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - Kr^2} + r^2 (d\vartheta^2 + \sin^2 \vartheta d\varphi^2) \right] \tag{2.1}$$

where ds is the distance in space-time, c is the speed of light, t is the cosmic time, and (r, ϑ, φ) are the spatial coordinates. K is a constant, whose value can be -1, 0 or 1, according to the geometry of the Universe, and describes its curvature.

Data from various independent experiments confirm its value to be 0, i.e we live in a flat Universe, with a very small error (see for example measurements from the CMB power spectrum, Spergel et al., 2003; Planck Collaboration et al., 2014). The scale factor $a(t)$ is a time dependent factor which relates the comoving coordinates to true physical distances. The expansion of the Universe can be described applying general relativity to the cosmological principle, and assuming an appropriate equation of state. Assuming a perfect fluid, the result is given by the Friedmann equations (Friedmann, 1922). These express the scale factor in terms of density and pressure of the Universe:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right). \tag{2.2}$$

In this equation, $\ddot{a}$ is the second derivative of $a(t)$, G is the gravitational constant, $\rho(t)$ is the mass density of the Universe, and P is the pressure. Immediately after the Big Bang, the Universe is radiation dominated, and the scale factor scales as $a(t) \sim t^{1/2}$. While the Universe expands, radiation becomes less important, and matter becomes the dominant component (from $t \sim 5 \cdot 10^4$yr to $t \sim 9.8 \cdot 10^9$ yr). In this case, the scale factor scales with $a(t) \sim t^{2/3}$. After this period, the matter density and the radiation density continue to drop, because of the expansion. The pressure term that drives the expansion dominates, leading into a Dark Energy-dominated era, that lasts to the present day. In this era, the scale factor scales as $a(t) \sim \exp(\sqrt{8\pi G\rho_{cr}/3}\, t)$, where ρ_{cr} is the critical density of the Universe, the mean density the Universe should have to halt its expansion at infinite time.

The scale factor regulates the expansion velocity and the distances among objects in the expanding Universe. For this reason, the Hubble constant in the Hubble equation is a time dependent parameter, that can be expressed in terms of variations of the scale factor in time: $H(t) = \dot{a}(t)/a(t)$. This means that the Hubble relation, introduced above, scales linearly only in the Local Universe, while at larger distances it depends on the scale factor.

The expansion produces effects similar to the Doppler effect on the light emitted by astrophysical systems. Analyzing the propagation of light in a Robertson-Walker metric, for photons emitted at time t_e with wavelength λ_e and received at time t_0 with wavelength λ_0, one obtains:

$$\frac{\lambda_0}{\lambda_e} = \frac{a(t_0)}{a(t_e)} = 1 + z. \tag{2.3}$$

z is the redshift, and provides a measure of the distance of galaxies.

The effect of the expansion is measured also as a variation of the measured magnitude of a source. The magnitude is a logarithmic measure of the brightness of an object in a specific band. The absolute magnitude is defined as that measured at a distance of 10 pc from the source. The absolute (M) and apparent (m) magnitudes are related through the distance (d) from the observed source: $m - M = 5(\log_{10} d - 1)$. One definition of distance often used in cosmology is the luminosity distance:

$$d_L = \left(\frac{L}{4\pi F}\right)^{1/2} = a_0 \cdot r \cdot (1 + z) \tag{2.4}$$

where L is the intrinsic luminosity of the source, F its flux, a_0 is the scale factor at redshfit zero, and r is the comoving distance of the source.

2.1.3 The Big Bang and the baryons' origin

As explained above, the current paradigm of an expanding Universe assumes that, during its earliest evolutionary stages, the Universe was extremely hot and dense. Its thermal history can be described by applying the laws of thermodynamics to a homogeneous and isotropic plasma.

At the very beginning, the hot plasma was made of photons, electrons, positrons and neutrinos. Until the temperature was above $T \sim 10^{13}$ K, the physics regulating the interactions is not understood.

It is believed that during this epoch, two important processes took place: baryogenesis, that created an asymmetry in favor of baryons (and against antibaryons), and inflation, an extreme expansion period. The inflation paradigm was introduced to solve some problems raised by the standard Big Bang scenario. One is the horizon problem: the CMB shows very small, casually independent fluctuations. This can be naturally explained if the Universe originated from a very small region where all matter and energy were in close and uniform contact. Another problem is the observed flatness of the Universe: without inflation, it can be obtained only with a fine tuned set of initial values for the matter and energy density. Small deviations from these values would have had extreme effects on the currently observed Universe. The last problem is the monopole problem: if the early Universe was very hot, a large number of heavy magnetic monopoles would have been produced, and should have persisted to present day. Observationally, we do not see any evidence of their existence. An inflationary period, just after their formation, would have diluted them significantly.

As the temperature of the Universe decreases to $T \sim 5 \cdot 10^9$ K, quarks and hadrons form in the hot plasma, pions and muons decay, and the number of neutrons and protons starts to differentiate. At this point, electrons and neutrinos ν_e and $\bar{\nu}_e$ decouple from the plasma, and the fraction of neutrons/protons freezes at a value $\sim 1/10$. At $T \sim 10^9$ K, nucleosynthesis starts, and some heavier elements are produced from the neutron-proton plasma, as deuterium, helium and lithium. The measure of the pristine abundance of these elements, namely the metallicity of material not processed in stellar nuclear reactions, provides information on the initial neutron/proton and photon/baryon ratios. When the temperature reaches $T \sim 3000$ K (t$\sim 2 \cdot 10^5$ yrs), neutral atoms become dominant (until this moment all matter was ionized), and the Universe becomes transparent to radiation. The photons of the CMB originate from this epoch, called recombination.

2.1.4 Perturbations and growth of structures

While at large scales the Cosmological Principle appears a reasonable approximation, structures of variable dimensions and distributions are observed at smaller scales.

Observationally, baryonic structures can be measured directly, from their light emission/absorption. The presence of DM can be inferred indirectly from the dynamics of the baryonic systems. For example, the rotational velocity profile of a disk dominated galaxy requires a major dark mass component in the galaxy, and the velocity distribution of galaxies inside a galaxy cluster requires the existence of an invisible halo that dominates the gravitational potential. The potential of DM structures can be measured also through gravitational lensing:

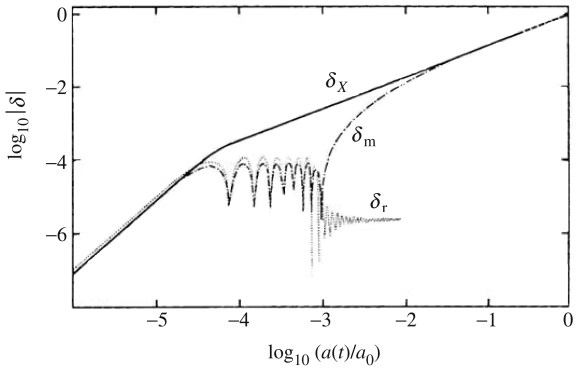

Figure 2.1: The evolution of a perturbation on a scale of M$\sim 10^{15} M_{\odot}$. δ represents the density contrast, a normalization of the density field, for DM (solid), baryons (dash-dotted) and photons (dotted line). The time is expressed in terms of the growing factor. Illustration from Coles and Lucchin (2002).

general relativity predicts that a massive object (a lens) produces a deformation of the nearby space-time, that bends the light traveling from a distant source. The lens potential can be evaluated from the distortion it produces on the lensed objects.

Cosmic structures originate from fluctuations in the initial matter density field, that have grown under their own self-gravity. In the case of baryons, the growth of structures is prevented by the initial high temperature. DM perturbations growth, instead, is prevented only by the Universe expansion, that tends to dilute the over-densities. Structure formation occurs when a perturbation reaches a critical density, and decouples from the expansion of the Universe, collapsing due to its own gravity. The collapse process can be described with several theoretical treatments. The simplest solution (and the only analytical one) is obtained assuming a spherical symmetry and no shell crossing (when an external shell of the structure collapses faster than an internal one). The solution is calculated assuming that, when the over-density leaves the expansion, it can be treated as a closed Universe, using the Friedmann equations. In this way, one obtains an analytic description of the radius and the time describing the collapse, and finds that spherical perturbations can separate from the Universe expansion when the initial over-density is ~ 180 times the average density of the Universe. Additional complexities, such as a non spherical shape, shell crossing, or rotational torques in the collapsing over-density, require a numerical treatment. Numerical simulations confirm that, on average, structures separate from the cosmological background when their density is around ~ 100 times the mean density of the Universe. When their internal kinematics support their gravitational potential, structures become "virialized". We define a virialized DM structure a "DM halo". This halo can grow further through accretion of material from the external regions, or through mergers with other halos. When a small halo is merged or accreted into a larger one, without being destroyed by its tidal field, it becomes a self-bound substructure, and we call it a subhalo.

In this DM structure formation scenario, baryons are accreted onto the col-

lapsed DM halo potentials. At the beginning of the baryonic collapse, radiation pressure is still dominant, and baryons are spread out of the perturbations. When photons are completely decoupled from baryonic matter, radiation pressure no longer prevents the collapse, and baryons fall in the DM perturbation, that has grown during the phase of baryonic oscillations. At decoupling, baryons have a distribution around the DM potential that is determined by their sound velocity at that time. For this reason, the typical scale length of this distribution is called the *Baryonic Acoustic Oscillations scale*. This scale is imprinted in the total matter distribution, and can be used to measure cosmological parameters, because it is a standard ruler. We show a schematic illustration of the growth of structures in Fig. 2.1, where different lines represent the normalized density field of each component evolving with the scale factor. In particular, DM (solid) perturbation grows all over the time, while baryons (dot-dashed line) oscillate with photons until decoupling, when they finally fall into the DM potential.

During structure growth, the asymmetric distribution of matter in the neighborhood of the over-density can impress a tidal torque on the collapsing material. The tidal torque theory describes this effect in the linear regime (before the structure viralizes). Usually, the dynamical state of a collapsed halo is described through a dimensionless spin parameter:

$$\lambda = \frac{|E|^{\frac{1}{2}}|\vec{L}|}{GM^{\frac{5}{2}}} \tag{2.5}$$

where $E = \frac{M\sigma^2}{2}$ is the kinetic energy of the halo, M and σ are the mass and velocity dispersion, $\vec{L} = MRV_{rot}$ is the angular momentum, R is the virial radius and V_{rot} the rotational velocity. One expects that this spin is transferred to the baryons falling onto the halo potential, but observations show that galaxies are characterized by a λ much lower than expectations from DM. This difference must be attributed to the energy that, during baryonic collapse, is partly dissipated through cooling.

2.1.5 Statistical description of perturbations

Observationally, the distribution of cosmic structures provides information on the cosmological parameters. For this reason, large effort is spent in surveys that study large volumes of the sky to measure the statistical properties of galaxies and/or galaxy over-densities.

In the analysis of the matter (or galaxy) density field $\rho(\vec{x})$, an often used quantity is the density contrast $\delta(\vec{x}) = (\rho(\vec{x}) - \bar{\rho}_m)/\bar{\rho}_m$, namely the difference between the density field at the position $\vec{x}$, and its average value $\bar{\rho}$. The density contrast easily characterizes over-densities and under-densities with positive and negative values. It can be used to select, at each cosmic time, the over-densities that will collapse in a DM halo, by setting a minimum threshold, for example, the critical density predicted by the spherical collapse model. The variance of the density contrast distribution is defined as $\sigma^2 =< \delta^2 >= 1/V \int_V \delta^2(\vec{x}) d^3\vec{x}$, where V is a representative volume of Universe.

When looking at the over-density distribution, it is useful to smooth the density contrast on a specific scale length. This is made convolving it with a window function $W(\vec{x}, R)$ (or $\tilde{W}(\vec{k}, R)$ in Fourier space): $\delta_R(\vec{x}) = \int \delta(\vec{x}) W(\vec{x} - \vec{x'}, R) d^3\vec{x}$. The most used window function is the top hat, which assigns to $\delta_R(\vec{x})$

the mean density contrast inside R:

$$W(\vec{x}, R) = \begin{cases} \frac{3}{4\pi R^3} \text{ if } |\vec{x}| \leq R \\ 0 \text{ if } |\vec{x}| > R \end{cases}$$

We can define the variance also for the smoothed density contrast field:

$$\sigma^2(R) = < \delta_R^2(\vec{x}) > .$$

The cosmological parameter σ_8 is defined as the variance of the density field smoothed with a top-hat filter of size $R = 8h^{-1}$ Mpc. Its measured value is $\sigma_8 = 0.8 \pm 0.01$, and characterizes the typical scale of structures. A larger value corresponds to larger fluctuations in the density field, and structure formation starting earlier.

The smoothing length can be translated in a smoothing mass $M = 4/3\pi\bar{\rho}R^3$, and thus $\delta_M(\vec{x}) = \delta_R(\vec{x})$. At a time t, we can select the peaks of the smoothed density contrast, using a threshold that identifies the regions that will collapse in a halo of mass M. Assuming a Gaussian random field, Press and Schechter (1974) postulated that the probability that, at time t, δ_M is larger than a critical threshold δ_c, is the same as the mass fraction that at the time t is contained in halos with mass greater than M:

$$n(M,t)dM = \frac{\bar{\rho}}{M}\frac{\partial f(>M)}{\partial M}dM = \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{M^2}\frac{\delta_c}{\sigma}e^{\frac{-\delta_c^2}{2\sigma^2}}\left|\frac{d\ln\sigma}{d\ln M}\right|dM \tag{2.6}$$

This formalism is used to predict the halo mass function. It can be extended to identify the progenitors of a collapsed spherical region of mass M_2 at time t_2. At a time $t_1 < t_2$, the fraction of mass collapsed in halos of mass M_1 can be determined from eq. 2.6, and we can use it to determine the mean number of progenitors of mass M_1 of a structure of mass M_2, using the expression:

$$n(M_1,t_1|M_2,t_2)dM_1 = \frac{1}{\sqrt{2\pi}}\frac{M_2}{M_1}\frac{\delta_c(t_1)-\delta_c(t_2)}{(\sigma_{M_1}-\sigma_{M_2})^{\frac{3}{2}}}e^{\frac{(\delta_c(t_1)-\delta_c(t_2))^2}{2(\sigma_{M_1}-\sigma_{M_2})^2}}\left|\frac{d\sigma_{M_1}}{dM_1}\right|dM_1 \tag{2.7}$$

In the above equation, $\delta_c(t)$ is the critical density at the epoch t, and σ_M is the variance for the mass M. This formula can be used to generate analytic merger trees, i.e. a description of the assembly history of dark matter halos (see e.g. Bower, 1991; White and Frenk, 1991b; Lacey and Cole, 1993). Each halo is assigned a probability to have progenitors of a certain mass at the previous time-step, and the merger tree is built using a random selection on this probability distribution. We show a classical example of a merger tree in Fig. 2.2: the width of the branches represents the mass of the progenitor, while the height is time. This formalism has several limitations, that were highlighted by results from numerical simulations (Sheth and Tormen, 1999; Jenkins et al., 2001). Several modifications were proposed to account for a more complex treatment, for example analytic solutions accounting for elliptical collapse, or direct fits of numerical simulations (Bond et al., 1991; Jenkins et al., 2001; Sheth et al., 2001; Warren et al., 2006).

Another important statistics is provided by the spatial distribution of the halos. The most used measure of spatial distribution is the two-point correlation function $\xi(r)$, which is the expectation value of the product of the density

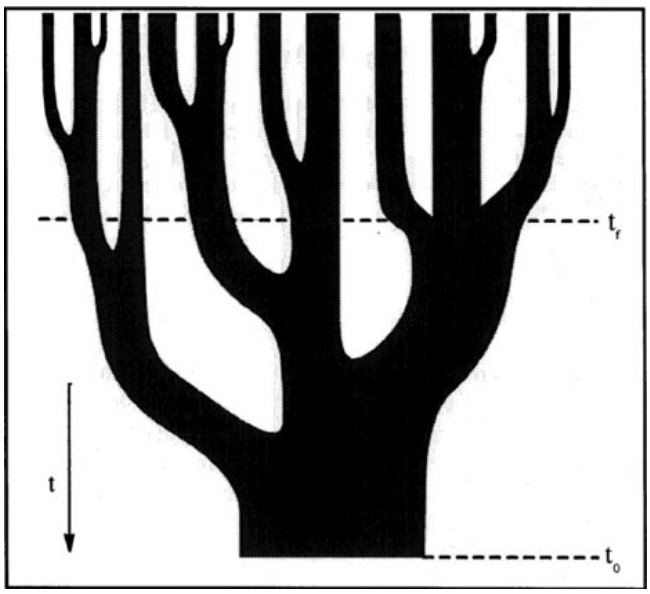

Figure 2.2: A schematic illustration of a merger tree. From Lacey and Cole (1993).

contrast at two points at distance r:

$$\xi(r) = \xi(\vec{r}) = \xi(|\vec{x}_1 - \vec{x}_2|) = \langle \delta_{\vec{x}_1} \delta_{\vec{x}_2} \rangle \tag{2.8}$$

It quantifies the excess probability dP_{12} of finding two objects at a distance r, with respect to a Poisson distribution of δ: $dP_{12}(r) = \bar{n}^2(1 + \xi(r))dV_1 \cdot dV_2$, with $\bar{n}$ the average number density of the objects, and dV_1 and dV_2 the volumes occupied by the two objects. In this way, the correlation function will take positive values at the typical scales of structures. Observationally, the correlation function is estimated directly from the number of couples of galaxies at a certain distance ($DD(r)$), compared to the number of couples one would obtain from a random distribution ($RR(r)$). The simplest estimator is $\xi(r) = DD(r)/RR(r) - 1$, but several others were proposed to account for observational and computational limits, as the limited number of data, the limited volume analyzed, or the small random sample generated for the $RR(r)$ evaluation. For example, the Landy and Szalay (1993) estimator is less sensitive to the size of the random catalog, and is able to handle well edge corrections: $\xi(r) = (DD(r) - 2RD(r) + RR(r))/RR(r)$, with $RD(r)$ the number of galaxy-random couples at distance r.

When the distribution of matter is analyzed in the Fourier space, the analog of the 2-point correlation function is the power spectrum, which describes the density contrast as a function of the mode k:

$$P(k_1)\delta^D(\delta_{\vec{k}_1} + \delta_{\vec{k}_2}) = \langle \delta_{\vec{k}_1} \delta_{\vec{k}_2} \rangle, \tag{2.9}$$

where δ^D is the Dirac delta function, and $\delta(\vec{k})$ is the Fourier transform the density contrast. The power spectrum and the correlation function are one the Fourier transform of the other.

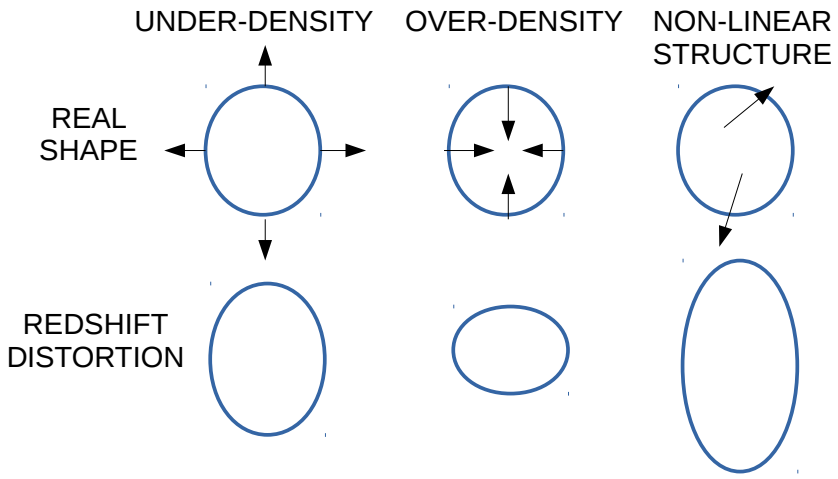

Figure 2.3: A schematic illustration of the effect of redshift distortions.

When the analysis is performed on visible matter, e.g. galaxies, used as proxies of the total matter distribution, we should consider the bias that links their density contrasts. If we assume a direct proportionality through a factor b, we obtain: $P_{gal}(k) = b^2 P_m(k)$ (Kaiser, 1984; Bardeen et al., 1986), with P_{gal} and P_m the power spectra of galaxies and total matter. Theoretical studies demonstrated that this model applies only at large scales, because the b parameter is scale dependent in the non-linear regime (Mo and White, 1996; Mann et al., 1998).

When the clustering functions are estimated from observational data, the real distribution is distorted by peculiar velocities, an effect called "redshift-space distortion". The measured velocity along the line of sight is the sum of the redshift due to Hubble expansion, and of the redshift due to the peculiar velocity of the galaxy in the system (see the schematic representation in Fig. 2.3). These peculiar velocities depend on the gravitational state of the system the galaxy resides in. If galaxies are in a collapsing over-density, they will be falling to the center of the system, and the system will appear flattened. If galaxies are in a virialized system, they will have almost random motions inside it, and the system will appear elongated. Redshift distortions damp the power spectrum at small scales, because of virialized structures, and enhance it on large scales, because of the collapsing coherent flows. These distortions are present also in the 2-point correlation function. For this reason, when using these statistical measurements for cosmological purposes, one should account for the redshift distortions, and correct the obtained clustering measure.

2.2 Cosmology and astrophysics in simulations

As described above, the formation of structures in a cosmological context is driven by the gravitational collapse of DM over-densities. Although analytic treatments are possible (e.g. the spherical collapse model illustrated earlier) the process can be described in its full geometrical complexity only resorting to

numerical simulations. In this framework, DM particles interact as particles of a non-collisional fluid, driven by their gravitational interactions. Once the initial conditions are set, the subsequent evolution can be uniquely and precisely computed, as the physics driving the Universe expansion and the particle-particle gravitational interactions is well known.

When the scientific aim is the study of galaxies and baryons, the treatments adopted in simulations become more complicated. Baryon particles interact through gravity and a number of other physical processes that cover a wide range of time and physical scales, and for which we do not have a satisfactory understanding. These processes are typically included using "sub-grid" models or "prescriptions" that are built from available observations or theoretical models.

Several approaches have been adopted to include baryons in cosmological simulations. The explicit treatment for hydrodynamics has an important cost in terms of computational time, and, for this reason, the simulated volumes are usually small. The evolution of the baryonic components is modeled using subgrid prescriptions, as mentioned above. Alternatively, semi-analytic models can access a large dynamic range in mass and spatial resolution at a limited computational cost. They lack, however, an explicit treatment for hydrodynamics. Finally, the Halo Occupation Distribution (HOD) method populates DM halos using a purely statistical approach, bypassing an explicit treatment for baryon physics.

In the next sections, I describe briefly the basics of cosmological simulations and of the techniques adopted to treat baryons and their evolution in a cosmological context, discussing the limits and advantages of each.

2.2.1 Dark Matter only N-body cosmological simulations

In cosmological DM-only N-body simulations, the density field is represented through a set of mass particles interacting only through gravity. At the beginning of the simulation, they are distributed on a regular grid that fills the simulated volume. This represents a homogeneous and isotropic density field. The perturbations are generated imposing that the power spectrum of the perturbed particles reproduces the initial cosmic power spectrum predicted by inflation, that scales as $P(k) = Ak^n$, with $n = 1$ (Harrison, 1970; Zeldovich, 1972). In addition, the simulation assumes periodic boundary conditions, to avoid issues due to the limited simulated volume, and to allow the solution of Poisson's equation in the Fourier space.

Once the initial conditions are set, the simulation is run by considering the gravitational interaction among the particles. The simulation is divided in time-steps, a discrete set of timings at which the simulated quantities are updated. At each time-step the position of a particle is calculated according to the total force acting on it, using a leapfrog integration. With this method, positions and velocities are updated at interleaved time steps, arranged in a leapfrog. In this way, the smaller is the time step, the more precise is the motion. Of course, more time steps mean a larger computational time.

The evaluation of the total gravitational force acting on each particle is the heaviest part of the calculation. The simple sum of all gravitational interactions results in computational times that scale with $\sim N^2$, where N is the number of particles. This simple approach limits the maximum affordable number of

particles, and, consequently, the mass resolution. Several techniques were developed to reduce the computational time, allowing for simulations with more particles, and, as a consequence, with higher resolution.

Tree methods (Barnes and Hut, 1986) adopt a direct summation method for the closest particles, while more distant groups of particles are aggregated in a single large particle located at the center of mass of the group. The aggregation of particles in groups is obtained by building an octree, namely a recursive division of the simulation volume in octants, until a certain number of particles per octant is reached. In this way the interactions to be calculated for each time step scale with $\sim N \log N$.

In the particle mesh (PM) approach, the simulation volume is discretized on a mesh, and a particle mass is distributed among nearby mesh cells. With this method, the discrete distribution of particles is smoothed onto a density distribution ρ. This, Fourier transformed in ρ_k, returns the gravitational potential in the Fourier space: $\Phi_k = -4\pi G\rho_k/k^2$. The inverse Fourier Transform of Φ_k gives the real space gravitational field. Further developments of this method are the P^3M code (Particle-Particle-Particle Mesh), where small scales are resolved by direct force summation, and APM (Adaptive Particle Mesh), where the mesh size is adaptable to the density of the volume considered. PM and tree codes can be combined to further improve the computational efficiency.

The size of the simulated box and the number/mass of the particles are fixed according to the scientific goal of the simulation.

The box size defines how much the simulation will be able to reproduce the average Universe, e.g. a too small box can be biased to an over-dense (or under-dense) region, while a large volume better samples the average Universe and its dispersion ($\gtrsim 100$ Mpc).

The total number of particles is chosen according to the mass resolution required by the scientific purposes. For example, if the study is focused on the spatial distribution of galaxies (or halos), the mass of the particles should be smaller than the typical mass of a galaxy (or halo), in order to easily identify them as groups of particles.

Another important parameter of a simulation, is the softening scale length. This length defines the distance between particles below which the force between them is "softened". This artifice is necessary because of the nature of particles in simulation: each of them represents many real particles, and it should be considered a fluid element rather than a point-mass. This length is chosen to be approximately the size of the fluid element, and sets the scale at which the forces are halted, to realistically treat the merger of the fluid elements.

2.2.2 Halo and subhalos identification

The final output of a N-body simulation is a set of snapshots, which are a subsample of the time-steps of the simulation. The 3D positions and velocities of the particles are recorded at each snapshot.

The next step of the analysis is the identification of "halos" and of the particles they are composed of. This task is performed by a dedicated algorithm: a halo finder. Several algorithms are used for this purpose, based on different definitions of halo.

Some methods, called density peak locators, find peaks in the density matter field. Each peak identifies a halo, that collects all the particles from the neigh-

borhood that satisfy a specific criterion (for example the spherical over-density method, Press and Schechter, 1974).

Other methods, called particle collectors, are based on the distances between particles, in 3D or 6D phase space. The particles at a distance below a given length are linked together in the same halo. A typical particle collector is the Friends-of-Friends (FoF) method (Davis et al., 1985), which links together particles with distance in the 3D space below a fraction of the mean particle-particle distance. As this method is purely geometric, it does not account for the actual gravitational bound among the FoF particles. For this reason, collector methods can identify as a single halo gravitationally separated structures.

Many algorithms have been developed to identify DM subhalos. As explained in the previous sections, they represent the cores of halos that have been accreted at earlier times. Some subhalo identification methods are recursive applications of the above described algorithms, iteratively looking for smaller structures inside halos (Gill et al., 2004; Knollmann and Knebe, 2009; Behroozi et al., 2013). Others are dedicated, e.g. the SUBFIND algorithm developed by Springel et al. (2005).

Therefore, for each snapshot, a series of halos and/or subhalos can be identified, and their properties, such as mass, position, size, spin, etc. measured. Finally, these halos (or subhalos) can be connected from one snapshot to the previous and to the subsequent one, in order to reconstruct the halo merger tree.

2.2.3 Halo Occupation Distribution

That of the Halo Occupation Distribution (HOD) is a method to populate DM halos with galaxies using a probabilistic approach. DM halos are assigned a probability of hosting a certain number of galaxies that depends on their mass. This probability is usually assumed to be a power law, and is typically constrained using clustering measurements.

The HOD approach can be split in several steps: the first one is the evaluation of the probability that a halo of mass M_h contains N galaxies, with respect the average number $< N >$ of galaxies in a halo of that mass; subsequently one needs to assign spatial and velocity distribution to the galaxies in the halo. This kind of analysis can be applied also to a subsample of the total galaxy population, for example to the red/blue galaxies, to central/satellites, to Late/Early Type Galaxies. In this case, the HOD parameters are set against the clustering of the individual population. For example, central galaxies are usually more massive and older than satellites, that, on the other hand, are more numerous. Thus, old and young galaxies have a different clustering estimation, and lead to a different HOD parametrization (Zheng et al., 2005).

These methods provide a very efficient tool to create large mock-catalogs of galaxies at a limited computational cost. On the other hand, they are not ideal to study the physical processes regulating galaxy evolution, as they are based on a purely statistical approach that does not treat explicitly baryon physics.

One strong assumption adopted in this method is that the galaxy content of the halos depends only on halo mass. Numerical simulations have shown that halos of the same mass, characterized by different assembly histories, lead to different clustering, spin distributions and fraction of substructures (Gao and White, 2007; Li et al., 2008), an effect called “assembly bias”. This introduces

systematic errors in clustering studies, and must be taken into account when HOD is used for cosmological studies (Zentner et al., 2014).

Some variations of the classical HOD approach are not affected by assembly bias. In particular, the SubHalo Abundance Matching methods consider, instead of the present mass of the subhalo, its mass (or circular velocity) at the accretion time, and assign galaxies to subhalos by matching the mass function of the former to the mass (velocity) function of the latter, assuming the most massive galaxy lives in the most massive subhalo (Vale and Ostriker, 2004; Conroy et al., 2006; Shankar et al., 2006). This method resolves on some extent the assembly bias problem, letting galaxies hosted in subhalos of the same mass to have a realistic scatter in their properties (Moster et al., 2010; Wetzel et al., 2009). On the other hand, the SHAM approach requires many more information than a HOD, and in particular needs simulations at higher resolution to identify subhalos. In addition, SHAM does not fully account for the environment of the subhalos.

2.2.4 Hydrodynamical simulations

Hydrodynamical simulations include an explicit description of hydrodynamics in N-body simulations. Different techniques were developed to evolve the baryonic and DM components consistently, and here I will briefly review the most used ones: Smoothed-Particle Hydrodynamics and Adaptive Mesh Refinement.

Smooth-particle Hydrodynamics (SPH, first developed by Gingold and Monaghan, 1977; Lucy, 1977) includes baryons in the N-body simulation as mass particles, that move according to the gravitational and hydrodynamical interactions. SPH is a Lagrangian method, namely the coordinate system is assumed to move with the particles. In setting the initial conditions, baryon particles are coupled with DM particles. When the simulation starts, baryons and DM evolve interacting only through gravity. The baryon-baryon interaction is based on the Lagrangian implementation of the Euler equations: a set of equations describing the fluid physics, linking together density, velocity, gravitation and temperature of the fluid. Particle properties are smoothed with a kernel, in order to correctly apply the Euler equations.

The Lagrangian approach has several advantages: it is naturally adaptive, namely the over and under-densities are identified by a larger or smaller number of particles, and it preserves mass, energy and angular momentum, because it follows the individual particle motions.

On the other hand, the Lagrangian method is not able to correctly treat discontinuities and shocks, namely the interactions between nearby particles in a different thermal state. This problem is partially solved with the adoption of a more sophisticated smoothing kernel, and an artificial viscosity term, in some cases activated only during shocks (Price, 2008; Cha et al., 2010; Read et al., 2010). The artificial viscosity prevents particles in colliding streams to pass through each other, decelerating them.

A different approach to hydrodynamics implementation in simulations is that of the Adaptive Mesh Refinement (AMR, first developed by Berger and Colella, 1989). It is based on an Eulerian set of coordinates for the baryonic fluid: mass particles are distributed onto a grid, and the fluid physics is solved in each cell of the grid. In this case, the main limit is given by the finite dimensions of the grid cell. A cartesian set of coordinates is not the best choice in an astrophysical

framework, because some zones of the simulation are typically over resolved (as those in voids) while others (for example discontinuities) are not well resolved. For this reason, AMR typically involves adaptive grids: volumes that need a higher resolution correspond to a denser grid (Anninos et al., 1994; Gnedin, 1995). Eulerian methods resolve more precisely and easily the dynamical interactions among baryons, with respect to Lagrangian methods. In particular, they are better suited to describe discontinuities. However, they have problems in the conservation of the angular momentum of gas particles. This is due to the grid structure, that favors the formation of artificial diffusion along its main axis, and rapidly degrades the orbit of gas particles.

Some attempts have been made to solve the limits of both the SPH and AMR approaches. For example, Springel (2010) proposed a new code, called AREPO, based on a moving unstructured mesh, which tiles a set of discrete points. The laws of hydrodynamics are solved on the mesh, that moves with the velocity of the local flow. This model is able to resolve shocks with the precision of the Eulerian schemes, and to resolve the particle motions as in a Lagrangian approach.

Hydrodynamical models, with the limits described above, are able to simulate the dynamical interactions of gas, providing a description of the spatial and velocity distributions of the gas particles. This is not enough to reproduce realistically baryon evolution in an astrophysical context. In fact, the physical processes involved cover a wide range of scales (of mass, size and time), often below the resolution limit of the simulation. In addition, the mechanisms driving these processes are almost always not well understood. For this reason, they are called "sub-grid processes", and are modeled using prescriptions based either on observational measurements or on higher resolution and more sophisticated (but lower scale) simulations. Examples of sub-grid processes in cosmological simulations are star formation, stellar and AGN feedback, and black hole growth.

Despite the simplified implementation of the sub-grid physics, hydrodynamical interactions are computationally costly. Therefore, this kind of simulations are usually performed on small cosmic volumes. Recent simulations, as EAGLE (Schaye et al., 2015, based on SPH) or Illustris (Vogelsberger et al., 2014a, based on AREPO) have simulated volumes boxes of $\sim$ 100 Mpc on a side. In Fig. 2.4, I show a slice of the EAGLE simulation, color coded according to the gas temperature. In the zoom-in box, one can see a simulated galaxy, color coded according to its stellar density.

2.2.5 Semi-Analytic Models: state of the art

Semi Analytic Models (SAM) combine an analytic treatment of the physical processes driving galactic evolution, with halo histories obtained from a cosmological simulation. Unlike HOD, the SAM approach attempts a direct treatment of the physical processes driving baryonic evolution.

The analytic treatment consists of a collection of prescriptions regulating the evolution and the interactions of the baryonic components. These prescriptions are extracted, as in the case of sub-grid processes in hydrodynamic simulations, from observational relations or from dedicated high-resolution hydrodynamic simulations.

The halo merger trees provide the skeleton for the SAM analytic treatment,

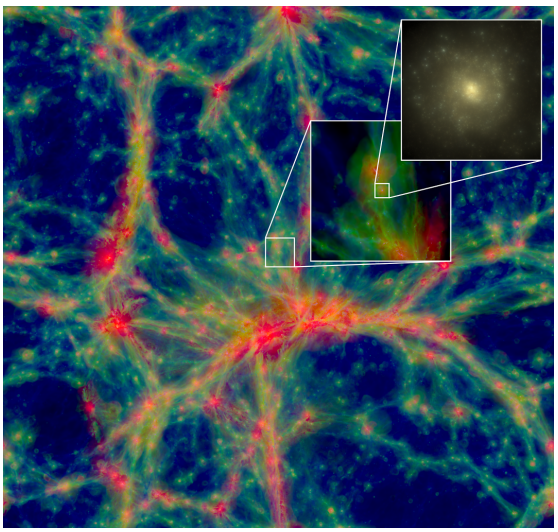

Figure 2.4: Example of a hydrodynamical simulation, with different colors representing gas at different temperatures, and sub-panels showing zoom-in of the largest box. The last zoom-in panel shows the stellar component of a galaxy similar to the Milky Way. Figure from the web page of the EAGLE project, http://icc.dur.ac.uk/Eagle/.

allowing galaxy evolution to be followed in a cosmological context. Merger trees can be constructed from a cosmological N-body simulation (see Sec. 2.2.2), or using Monte-Carlo methods based on the Press-Schechter formalism, or its modifications. A hybrid approach is also possible. Methods based on analytic merger trees offer the advantage of high resolution at small computational cost. However, the direct comparison with results of a full N-body simulation highlights the limits of this approach, and several corrections have been proposed to overcome these limits (see for example Parkinson et al., 2008; Neistein and Dekel, 2008).

The first SAMs were introduced by White and Frenk (1991b); Cole (1991); Lacey and Silk (1991). These models included prescriptions for gas cooling, star formation and stellar feedback. They were applied to halo distributions at various redshifts, calculated using the Press-Schechter formalism, or using other analytic models of the collapse of the density fluctuation field. SAMs were first applied to merger trees describing the evolution of individual halos only some years later, in Kauffmann et al. (1993); Cole et al. (1994), using merger trees evaluated with Monte Carlo techniques applied to the Press-Schechter formalism. Roukema et al. (1997); Kauffmann et al. (1999) later coupled the SAM method to merger trees extracted from cosmological N-body simulations. Since then, SAMs have been updated and extended to investigate more in detail the various aspects of galaxy formation and evolution (see for reviews Baugh, 2006; Benson, 2010).

SAMs are the ideal tool to study galaxy evolution in large cosmic volumes, because of their limited computational cost and the consequent relatively fast analysis of different prescriptions for the same physical process. Their flexibility

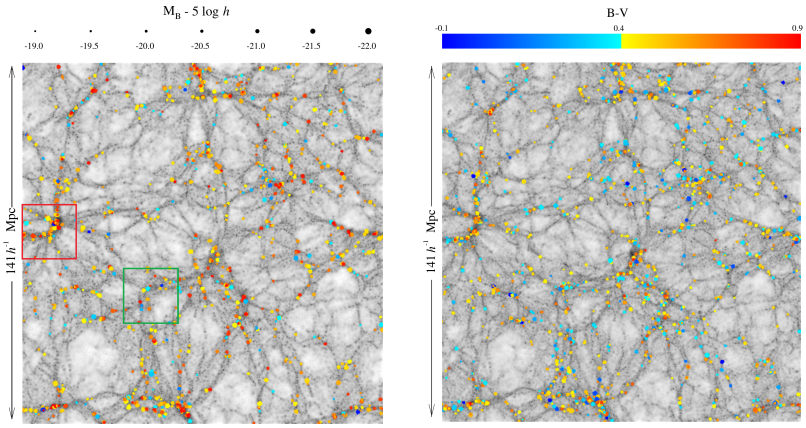

Figure 2.5: Illustration of a DM simulation populated by galaxies through a semi-analytic model. DM structures are in gray, while galaxies are color coded by their B-V color. The two panel represent the same simulation at redshift 0 (left) and redshift 0.5 (right). Figure from Benson et al. (2001a).

is, indeed, an advantage in large statistical studies, especially when compared to hydrodynamical simulations.

Different models have been published in the last years, and different groups have released their model outputs through public databases (see for example the Millennium database, Lemson and Virgo Consortium, 2006). The public diffusion of these results favored the use of SAMs for the interpretation of observational data and as a tool for making predictions at high redshift. The access to several independent models allowed a direct comparison among their results, identifying their respective abilities and weaknesses.

State-of-the-art models include a treatment for all the main processes driving galaxy evolution: gas cooling, star formation, stellar feedback, chemical enrichment, black hole growth, AGN feedback, mergers, and disk instabilities. Models are continuously updated to introduce higher degrees of complexity and to investigate new specific aspects, often driven by new observational data. For example, recent attention has been given to the implementation of a star formation law based on the molecular surface density. Lagos et al. (2011a,b) first included an explicit partition of cold gas into atomic and molecular gas in the GALFORM SAM, followed by other groups (see, for more details, Sec. 1.3).

Another strength of SAMs is their applicability to very large cosmic volumes. For this reason, SAMs are recommended both for statistical studies of galaxies and their evolution in large cosmic volumes, and for the construction of large mock catalogs. Large mocks can be built also using HOD techniques, but, in this case, the limit is the lack of an explicit treatment for baryonic evolution.

As discussed above, the strongest weakness of SAMs is the lack of an explicit treatment of gas dynamics. Recent models have suffered from systematic (common to different, independent models) discrepancies with observational data. For example, SAMs tend to predict too many small and passive galaxies. This problem has been partially solved with the adoption of better prescriptions for star formation and feedback scheme (Lagos et al., 2013; Henriques et al., 2013;

White et al., 2015; Lu et al., 2015; Hirschmann et al., 2016). Some problems remain, e.g. the color distributions of galaxies, that does not reproduce the bimodal distribution of observed galaxies (Henriques et al., 2013; Hirschmann et al., 2016), and their metallicity evolution (White et al., 2015).

Chapter 3

The GAlaxy Evolution and Assembly (GAEA) semi-analytic model

In this chapter, I provide a detailed description of the semi-analytic model used in this Thesis work, and of the cosmological simulations used to generate the input merger trees. In the following, I describe the main physical processes driving galaxy evolution, and how they are implemented in our model.

The semi-analytic model I employ was developed by the "Trieste group", and is described in detail in Xie et al. (2017). It represents a further development of the GAlaxy Evolution and Assembly (GAEA) model, featuring a sophisticated chemical enrichment scheme that accounts for the non-instantaneous recycling of metals, gas and energy (De Lucia et al., 2014b), and a new stellar feedback scheme partly based on results from numerical simulations (Hirschmann et al., 2016). Xie et al. (2017) have further improved the model by introducing a treatment for partitioning the cold gas into its molecular and atomic components, a star formation law based on the molecular surface density, and a treatment for tracing the specific angular momentum of the gaseous and stellar disks.

3.1 The Millennium Simulation

All the semi-analytic models used in this Thesis are run on the Millennium Simulation (Springel et al., 2005). This cosmological simulation was completed in 2004 at the Max Planck Society's supercomputer center in Garching, and was one of the projects of the Virgo Consortium, an international team developing state-of-the-art computer simulations of galaxy formation. Simulation results were made publicly available in 2005. At the time, the Millennium Simulation was the largest simulation of structure formation in a ΛCDM cosmology of such high resolution. Despite being now more than 15 years old, it still provides an excellent compromise between large volume and high resolution.

The Millennium Simulation follows the evolution of $N = 2160^3$ particles of mass $8.6 \times 10^8\, h^{-1}\mathrm{M}_\odot$ within a cubic region of comoving size of $500\, h^{-1}\mathrm{Mpc}$, from z=127 to z=0. The cosmology assumed is consistent with WMAP1 cos-

mological parameters (Spergel et al., 2003): $\Omega_b = 0.045$, $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $H_0 = 100h\,\mathrm{Mpc}^{-1}\,\mathrm{km\,s}^{-1}$, $h = 0.73$, $\sigma_8 = 0.9$ and $n_s = 1$. These cosmological parameters are nowadays out of date, and more precise measurements were obtained e.g. by the PLANCK mission (Planck Collaboration et al., 2014). A parameter that heavily influences the clustering of structures is σ_8, and its value changed drastically with the PLANCK estimate: $\sigma_8 = 0.829$. This may have an effect on galaxy evolution, but previous work (Wang et al., 2008; Guo et al., 2013) has shown that, after a retuning of the physical parameters of the model, results are qualitatively similar when run on a simulation with a lower σ_8.

The code used to run the Millennium Simulation was a modified version of the publicly available code GADGET-2. Simulation data were stored at 64 output redshifts, each corresponding to a snapshot number n (with n varying from 0 to 63) through the following formula: $\log(1+z_n) = n(n+35)/4200$. For each snapshot, halos and subhalos were identified and used to build substructure based merger trees[1], that are used as input for the semi-analytic model used in this Thesis (see Sec. 3.2 for more details).

In 2008, a second simulation was completed with the same cosmology, output structure and number of particles, but with a 5 times smaller box, and a 125 times better resolution: the Millennium-II Simulation (Boylan-Kolchin et al., 2009). The combined Millennium-I and Millennium-II provide an unprecedented range of scales, from halos similar to the Local Group dwarf spheroidal galaxies, to halos corresponding to the richest galaxy clusters. In this Thesis, we will use the Millennium-II to improve the resolution of the studies performed on the Millennium-I, where necessary.

3.2 SUBFIND and the merger trees

In this section, I describe how halos are identified from the particles of the Millennium simulation, and how they are linked in merger trees.

Dark matter halos are identified using a standard Friends-of-Friends algorithm (FoF) with a linking length of 0.2 in units of the mean particle separation. FoF halos are then decomposed using the algorithm SUBFIND (Springel et al., 2001), which identifies gravitationally bound substructures and estimates their properties (i.e. mass, radius, spin). Only structures with at least 20 bound particles are retained as genuine substructures, a limit that corresponds to a halo mass of $M_h = 1.7 \times 10^{10}\ \mathrm{M}_\odot h^{-1}$.

Merger trees are constructed for all identified subhalos at each snapshot. For each subhalo, a unique descendant subhalo is identified, usually (>99.9% of the cases) in the subsequent snapshot in the future. In the other cases, usually at the limit of resolution, the direct descendant drops below the resolution limit in the next snapshot, and is recognized again in the second next snapshot, when it raises again above the minimum resolution. The descendant is chosen by tracing the majority of the most bound particles of the subhalo from a given snapshot to the following one. A representation of a merger tree is given in Figure 3.1, where the descendant of each subhalo is indicated with a black arrow.

The identification of a unique descendant for each subhalo allows the construction of a uniquely defined merger tree. Merger trees from the entire Mil-

[1] These merger trees are publicly available through a relational database that can be accessed at: http://wwwmpa.mpa-garching.mpg.de/millennium/

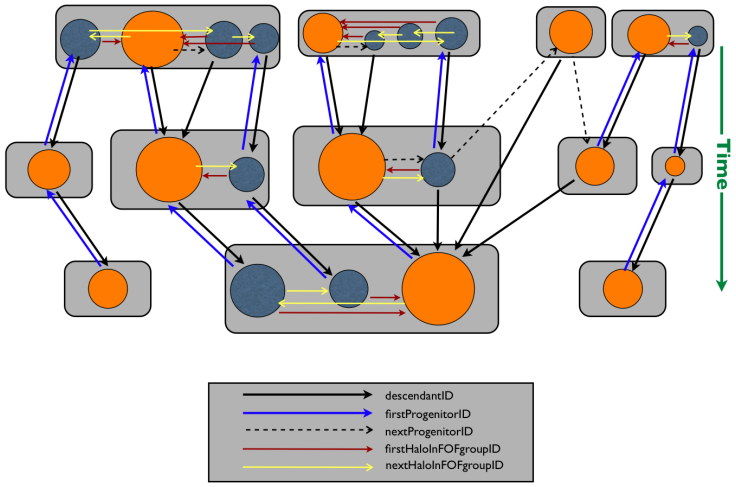

Figure 3.1: Schematic representation of a merger tree from the Millennium Simulation. Each row represents a snapshot of the simulation, FoF halos are represented as gray rectangles, subhalos are orange (centrals) and blue (satellites) circles. Subhalos of subsequent snapshots are linked together in a descendant/first progenitor relation through black/blue arrows.

lennium Simulation are stored in 512 separate files, and include all necessary information to be used as unique input for the galaxy formation model I describe below.

The distinction between FoF halos (rectangles in Fig, 3.1) and subhalos (circles) is important in the semi-analytic model, because it is used to distinguish between central and satellite galaxies. The former are typically identified as the central galaxies of the most massive subhalo, that coincides with the bound part of the FoF group (orange circles in the figure). The latter are associated with the other gravitationally bound subhalos of the FoF (dark blue circles in the figure).

For each subhalo, a main progenitor is defined (blue arrows in the figure) as the most massive among the progenitors of the subhalo at the previous snapshot. The branch of the tree built by recursively following back in time the main progenitor of a subhalo is called “main branch”.

3.3 The Semi-Analytic Model

In this section, we describe in detail the prescriptions adopted in the GAEA model, as updated in Xie et al. (2017).

Galaxies are attached to dark matter substructures. A galaxy residing at the center of a FoF halo is a central galaxy, or type 0 in the jargon of our model, while a galaxy that resides in a subhalo is a satellite, or type 1. Our model also considers “orphan satellite” galaxies, or type 2, whose parent dark matter substructure has been stripped below the resolution limit of the simulation. In our model, the position and velocity of these galaxies are traced by following the position and velocity of the most bound particle of the parent substructure

at the last time it was identified.
A hot gas reservoir is assigned to each FoF halo, assuming it follows an isothermal distribution:

$$\rho_{hot}(r) = \frac{M_{hot}}{4\pi R_{200} r^2} \tag{3.1}$$

where r is the distance from the center of the halo, M_{hot} is the mass of the hot gas and R_{200} is the virial radius of the halo, namely the radius enclosing an overdensity corresponding to 200 times the critical density of the Universe.

The hot gas halo cools onto the central galaxy as described more in detail in Sec. 3.3.2, originating a disk of cold gas. This disk is assumed to follow an exponential profile, and to be rotationally supported. Its specific angular momentum is acquired during gas cooling, from the specific angular momentum of the hot halo, assumed to be equal to that of the DM halo:

$$\vec{j}^{f}_{cold} = \frac{\vec{j}^{0}_{cold} M^{0}_{cold} + \vec{j}_{DM} M_{cooling}}{M^{0}_{cold} + M_{cooling}} \tag{3.2}$$

In the above equation, j_{DM} and j^{0}_{cold} are the specific angular momenta of the DM halo and of the cold gas disk before cooling, and j^{f}_{cold} is the specific angular momentum of the cold gas after gas cooling. M^{0}_{cold} and $M_{cooling}$ are the mass of the cold gas before cooling and the cooled mass. The scale radius of the disk is evaluated from the angular momentum of the cold gas disk:

$$R_{cold} = \frac{|\vec{j}_{cold}|}{2V_{max}}, \tag{3.3}$$

where $\vec{j}_{cold} = \vec{J}_{cold}/M_{cold}$ is the specific angular momentum of the cold gas and V_{max} is its rotational velocity, assumed to be equal to that of the parent halo.

Star formation (see details in Sec. 3.3.3) converts part of the gas into stars, forming a stellar disk. Both the gaseous and stellar disks are assumed to be characterized by an exponential surface density profile:

$$\Sigma_{disk}(r) = \Sigma_0 e^{-\frac{r}{R_{disk}}}, \tag{3.4}$$

where $\Sigma_0 = M_{disk}/(2\pi R^2_{disk})$ is the central surface density of the disk, M_{disk} is the stellar/cold gas disk mass, and R_{disk} is the scale radius of the stellar/cold gas disk. The stellar disk inherits the specific angular momentum of the cold gas as follows:

$$\vec{j}^{f}_{*,disk} = \frac{\vec{j}^{0}_{*,disk} M^{0}_{*,disk} + \vec{j}_{cold} M_{*,new}}{M^{0}_{*,disk} + M_{*,new}}, \tag{3.5}$$

where $\vec{j}^{0}_{*,disk}$ and $\vec{j}^{f}_{*,disk}$ represent the specific angular momentum of the stellar disk before and after star formation, $M_{*,disk}$ and $M_{*,new}$ are the disk stellar mass before star formation and the mass of stars formed. The scale radius is evaluated through:

$$R_{*,disk} = \frac{|\vec{j}_{*,disk}|}{2V_{max}}. \tag{3.6}$$

During the latest evolutionary stages of massive stars, large amounts of energy and momentum are injected into the interstellar medium (stellar feedback,

see Sec. 3.3.4). This feedback is due to winds from massive stars, or supernovae explosions. As a consequence of stellar feedback, part of the cold gas can be reheated, and part ejected from the galactic potential well. In our model, the reheated gas and the ejected gas are assigned to the hot gas and to the "ejected" gas reservoirs of the central galaxy, respectively. This means that gas reheated/ejected from satellite galaxies is assigned to the central, and will never return to the satellite. The ejected gas can then be reaccreted onto the hot reservoir at later times. Another source of gas reheating is the central black hole, that grows during mergers and prevents cooling through the so-called AGN feedback (see Sec. 3.3.6).

The disk composed of cold gas and stars can undergo disk instabilities or the galaxy can merge with another one (see Sec. 3.3.7). These events lead to the formation of another stellar component: the central stellar bulge. The bulge is a spheroidal component, dispersion supported, assumed to have a Jaffe profile:

$$\rho(r) = \frac{\rho_0}{4\pi}\left(\frac{r}{r_B}\right)^{-2}\left(1+\frac{r}{r_B}\right)^{-2}, \tag{3.7}$$

where ρ_0 is the central density and r_B is the scale length of the bulge. The latter is calculated from energy conservation. During a disk instability event, the stars moved from the disk to the bulge are assumed to form a spheroid, whose scale length is equal to the radius enclosing the moved mass. Then the spheroid is merged with the pre-existing bulge, and the final size is calculated with assumptions similar to those adopted for mergers. Assuming no energy dissipation, one only needs to consider the gravitational internal energies and the energy associated with the gravitational interaction. The energy before the merger can be expressed as:

$$E_i = CG\left[\frac{M_*^{p\,2}}{R_p} + \frac{M_*^{s\,2}}{R_s}\right] + \alpha G\frac{M_*^p M_*^s}{R_p + R_s}. \tag{3.8}$$

And the energy after the merger:

$$E_f = CG\frac{M_*^{f\,2}}{R^f}, \tag{3.9}$$

where G is the gravitational constant, M_*^x is the total stellar mass of the $x = p$ (primary) or $x = s$ (secondary) galaxy (or bulge, in the case of disk instabilities). R_x is an estimate of the half mass radius of the galaxy, obtained by averaging, and weighting by their respective mass, the size of the disk and of the bulge (or the scale radius of the bulge for disk instabilities). $C = 0.5$ and $\alpha = 0.5$ are parameters depending on the shape of the halo potential and on the kind of interaction between spheroids (Cole et al., 2000). In the last equation, M_*^f is the final stellar mass of the spheroid, and R_f is the bulge size. During disk instabilities, we assume that the angular momentum of the disk is preserved, and, thus, the specific angular momentum varies as:

$$j^f_{*,disk} = \frac{j^0_{*,disk} M^0_{*,disk}}{M^0_{*,disk} - \delta M} \tag{3.10}$$

with δM representing the mass moved from the disk to the bulge to restore stability.

3.3.1 Re-ionization

As described in the previous chapter, after recombination, the Universe becomes transparent to radiation. At this epoch, baryons are in their neutral state, and their collapse in the potential wells of dark matter halos is not stopped by radiation pressure, leading to the formation of the first population of baryonic structures. The first stars (and/or AGNs) form and emit radiation, ionizing again the gas. Knowledge of the emissivity and the optical depth of the Universe at the time of re-ionization are required for a consistent treatment of the re-ionizing background, and of its effect on the cooling rate. In the model, dark matter halos are populated with a fraction of baryons corrected for this effect. The adopted prescription is based on results by Gnedin (2000). This work used simulations to quantify the characteristic mass scale (M_F), below which the baryon fraction is reduced compared to the Universal value. Gnedin (2000) found that, below this mass scale, the fraction of baryons that condenses in a halo of mass M_{200} can be parametrized as:

$$f_b(z, M_{200}) = \frac{f_b^{cosmo}}{(1 + 0.26 M_F(z)/M_{200})^3} \tag{3.11}$$

where $f_b^{cosmo} = 0.17$ is the Universal baryon fraction (Spergel et al., 2003). The filtering mass is a function of redshift. In the GAEA model, $M_F(z)$ is fixed to $4 \cdot 10^9$ $M_\odot$ at the re-ionization epoch, and reaches $3 \cdot 10^{10}$ $M_\odot$ at redshift 0. The model also assumes that re-ionization starts at redshift $z = 15$, and lasts for about 0.12 Gyr. Okamoto et al. (2008) have argued that the filtering mass should be significantly lower than that found in Gnedin (2000). We do not expect this to have a significant impact on the results of the model discussed in this Thesis.

3.3.2 Cooling

In the GAEA model, the hot gas is assumed to have an isothermal profile with density $\rho_{hot}(r)$. A cooling time can be defined as the ratio between the specific thermal energy of the gas and the cooling rate per unit volume:

$$t_{cool} = \frac{3 \mu m_H \kappa_b T_{vir}}{2 \rho_{hot}(r) \Lambda(T, Z)} \tag{3.12}$$

where μm_H is the mean particle mass (m_H is the hydrogen particle mass), κ_b is the Boltzmann's constant, and T_{vir} is the virial temperature of the hot gas. The cooling function $\Lambda(T, Z)$ depends on the metallicity of the hot gas, Z, and on the virial temperature $T = 35.9(V_{vir}/[\mathrm{km\ s^{-1}}])^2$ K of the host halo. In our model, this function is modeled using the collisional ionization cooling curves of Sutherland and Dopita (1993).

Cooling occurs through two possible channels: the "rapid cooling regime", or cold accretion, and the "slow cooling regime", or hot accretion. The former is typical of small halos at early epochs, when the cooling time is short and the gas is accreted directly onto the central regions of the galaxy. The latter, instead, is typical of large halos, where the accretion is slowed by shocks and occurs on longer time-scales (Silk, 1977; Rees and Ostriker, 1977; Binney, 1977; White and Rees, 1978, see also Sec. 1.1).

In the GAEA model, the rapid and slow regimes are implemented comparing a characteristic cooling radius to the virial radius R_{200} of the halo. The cooling radius r_{cool} is defined as the radius at which the cooling time is equal to some characteristic time (different definitions are employed in different models, see De Lucia et al., 2010). In the case of GAEA, r_{cool} is where $t_{cool} = t_{dyn}$, with t_{dyn} given by the dynamical time of the halo:

$$r_{cool} = \left[\frac{t_{dyn} M_{hot} \Lambda(T, Z)}{6\pi\mu m_H \kappa_b T_{vir} R_{200}} \right]^{\frac{1}{2}} \tag{3.13}$$

When $r_{cool} > R_{200}$, we assume that we are in the rapid accretion regime, and that all the hot gas available is cooled onto the cold gas disk in a code time-step. The slow accretion regime occurs when $r_{cool} < R_{200}$. In this case, only the fraction of the hot mass that is enclosed within the cooling radius can cool, and cooled mass can be written as:

$$M_{cool} = \frac{M_{hot}}{R_{200}} \frac{R_{cool}}{t_{dyn}} dt \tag{3.14}$$

where dt is the time step of the model.

3.3.3 Star Formation

Stars form from the collapse of dense clouds of cold gas. These clouds are characterized by low temperatures and high fractions of molecular gas. The physics and the chemistry driving the collapse of molecular clouds are complex, and still not well understood. On the scale of the galactic disk, star formation can be modeled by resorting to empirical relations between the surface density of the star formation and that of the cold gas. First implementations were based on the Kennicutt-Schmidt law (Schmidt, 1959; Kennicutt, 1998), that relates the star formation in the disk to the total cold gas surface density. Recent observations have highlighted that the star formation surface density strongly correlates with the molecular gas surface density, while the correlation with the atomic gas is weak (Wong and Blitz, 2002; Kennicutt et al., 2007; Leroy et al., 2008). The GAEA model used in Hirschmann et al. (2016) adopts prescriptions depending on the total cold gas in the disk, while Xie et al. (2017) introduces an explicit dependence of star formation on the molecular content.

The Kennicutt-Schmidt star formation law adopted in Hirschmann et al. (2016) can be expressed as:

$$\dot{M}_* = \alpha_{SF} \times M_{SF} / \tau_{dyn} \tag{3.15}$$

where $\alpha_{SF} = 0.03$ is a parameter of the model, kept fixed as a function of cosmic time and any other physical property of the galaxy or of the inter-stellar medium. $\tau_{dyn} = r_{cold,disk}/V_{vir}$ is the dynamical time of the disk, and $r_{cold,disk} = 3 \cdot R_{cold}$ is the radius of the star forming region. M_{SF} is the cold gas available for star formation, calculated integrating the surface density of the cold gas disk, out to the radius r_{crit} where the gas surface density drops below a critical value (Kauffmann, 1996):

$$\Sigma_{crit}[\mathrm{M_\odot pc^{-2}}] = \frac{0.59 \times V_{vir}[\mathrm{km/s}]}{r[\mathrm{kpc}]} \tag{3.16}$$

This can be converted to a critical gas mass:

$$M_{crit} = 3.8 \cdot 10^9 \left(\frac{V_{vir}}{200[\mathrm{km/s}]}\right) \left(\frac{r_{cold,disk}}{10[\mathrm{kpc}]}\right) \mathrm{M}_\odot \tag{3.17}$$

Therefore, $M_{SF} = M_{cold} - M_{crit}$.

In Xie et al. (2017), we have implemented an explicit treatment for the partition of cold gas into its atomic (HI) and molecular components (H_2), and a star formation law based on the molecular abundance. In our updated model, the disk is divided in 21 concentric annuli, and the star formation rate is estimated in each of them. The number of rings used is a compromise between computational time and good convergence of model results (see Xie et al., 2017, for details). In each annulus, the star formation rate density is evaluated as:

$$\dot{\Sigma}_{i,SF} = \nu^i_{SF} \Sigma^i_{H_2}, \tag{3.18}$$

where ν^i_{SF} is the star formation efficiency and $\Sigma^i_{H_2}$ is the molecular hydrogen surface density. The total star formation rate of the disk is the sum over all the annuli: $\dot{M}_* = \sum_{i \in annuli} \dot{\Sigma}_{i,SF} S_i$, with S_i representing the area of the i-th annulus.

The partition in HI and H_2 is evaluated at each time-step (just before star formation takes place), and in each annulus. Xie et al. (2017) explored a set of prescriptions proposed in literature (Blitz and Rosolowsky, 2006; Krumholz et al., 2009; Gnedin and Kravtsov, 2011; Krumholz, 2013), but in this Thesis work we focus on the model that provides the best agreement with observational data. This is the model based on the empirical relation found by Blitz and Rosolowsky (2006) between the molecular fraction $R_{mol} = H_2/HI$ and the hydrostatic mid-plane pressure of the disk (P_{ext}). Note that this relation was measured in the local Universe, but in our model we assume it is valid at all cosmic epochs and in all environments. Specifically, the molecular fraction can be written as:

$$R_{mol} = \frac{\Sigma_{H_2}}{\Sigma_{HI}} = \left(\frac{P_{ext}}{P_0}\right)^\alpha \tag{3.19}$$

where P_0 is the external pressure of molecular clumps. Blitz and Rosolowsky (2006) find $P_0 \in [0.4;\ 7.1] \times 10^4 \mathrm{cm}^{-3}\mathrm{K}$, and $\alpha \in [0.58;\ 1.64]$ for a sample of galaxies in the Local Universe, that includes dwarfs, HI-rich and H_2-rich galaxies, as well as the Milky Way. In GAEA, we use the average value of the observational measurements: $\log_{10}(P_0/\kappa_B[\mathrm{cm}^{-3}\mathrm{K}]) = 4.54$ and $\alpha = 0.92$. The hydrostatic pressure at the mid-plane is evaluated as follows (Elmegreen, 1989):

$$P_{ext} = \frac{\pi}{2} G \Sigma_{cold} [\Sigma_{cold} + f(\sigma)\Sigma_*], \tag{3.20}$$

where Σ_* is the surface density of the stellar disk, and $f(\sigma) = \sigma_{cold}/\sigma_*$ is the ratio between the vertical velocity dispersion of the cold gas and that of the stellar component. We assume $\sigma_{cold} = 10$ km s^{-1} (Leroy et al., 2008), $\sigma_* = \sqrt{\pi G h_* \Sigma_*}$ (as in Lagos et al., 2011b), and $h_* = r_*/7.3$ (Kregel et al., 2002) for the height of the disk. For the star formation efficiency we assume:

$$\nu_{SF} = \nu_0 \left[1 + \left(\frac{\Sigma_{cold}}{\Sigma_0}\right)^{0.4}\right], \tag{3.21}$$

where $\Sigma_0 = 200$ M$_\odot$pc^{-2} is the critical density where the slope of the relation between Σ_{SF} and Σ_{H_2} steepens (Bigiel et al., 2008). Finally, we treat ν_0 as a free parameter, and find that $\nu_0 = 0.4$ Gyr^{-1} reproduces well the observed HI and stellar mass function at redshift z=0.

3.3.4 Stellar feedback

At the end of their lives, stars are subject to significant mass ejections through winds, or (depending on their mass) to Super Novae explosions. These events cause the heating and the ejection of part of the gas, and are responsible for what is usually referred to as "stellar feedback".

Hirschmann et al. (2016) evaluated the impact of different stellar feedback schemes within the GAEA model. The scheme that better reproduces the observational relations examined in that work is a feedback model where outflows are parametrized following the results from the FIRE hydrodynamical simulations (Hopkins et al., 2014; Muratov et al., 2015).

The energy released by supernovae and stellar winds is assumed to reheat the cold gas of the disk, and to eject a part of it, that is driven out of the parent halo through large-scale galactic outflows.

The reheating rate of the cold gas is assumed to depend on the star formation rate and to scale with both redshift and the potential well of the galaxy, following the Muratov et al. (2015) parametrization:

$$\dot{M}_{reheat} = \epsilon_{reheat}(1+z)^{1.25}\left(\frac{V_{max}}{60 \text{ km s}^{-1}}\right)^{\alpha} \times \dot{M}_*. \tag{3.22}$$

V_{max} is the maximum circular velocity of the parent halo, and the α index is:

$$\alpha = \begin{cases} -3.2 & \text{if } V_{max} < 60 \text{ km s}^{-1} \\ -1.0 & \text{if } V_{max} > 60 \text{ km s}^{-1} \end{cases}$$

The reheating efficiency is $\epsilon_{reheat} = 0.3$. We have used the same parametrization for the total energy released by massive stars:

$$\dot{E} = \epsilon_{eject}(1+z)^{1.25}\left(\frac{V_{max}}{60 \text{ km s}^{-1}}\right)^{\alpha} \times 0.5 \cdot \dot{M}_* \cdot V_{SN}^2, \tag{3.23}$$

where $0.5V_{SN}^2$ is the mean kinetic energy of SN ejecta per unit mass of stars formed, and the ejection efficiency is $\epsilon_{eject} = 0.1$.

The ejection rate is evaluated following energy conservation arguments for the escaping material:

$$\dot{M}_{ejected} = \frac{\dot{E} - 0.5\dot{M}_{reheat}V_{vir}^2}{0.5V_{vir}^2}. \tag{3.24}$$

Finally, the ejected gas is assumed to be re-accreted on a time-scale that depends on the virial mass of the halo: $\dot{M}_{reinc} = \gamma M_{ejected}/t_{reinc}$, where $t_{reinc} = \frac{10^{10} \text{ M}_\odot}{M_{vir}}$ and $\gamma = 1$. Hirschmann et al. (2016) found that this dependence, first proposed by Henriques et al. (2013), is necessary to reproduce the observed stellar mass function.

3.3.5 Metal enrichment

During stellar feedback, stars release into the surrounding gas the metals produced by their internal nuclear reactions. The amount of metals produced depends on the initial mass of the star and on its metallicity: the stellar mass determines the gravitational potential that makes possible the nuclear reactions, favoring the formation of heavier elements in more massive stars; the initial metallicity of the star influences the efficiency of the nuclear reactions. The amount of metals that is returned to the interstellar medium as newly produced elements is called stellar yield.

Theoretically, knowing the initial mass distribution of the newly formed stars, the metallicity of the stellar population, and the yields for each element, one can calculate the amount of metals returned by the stellar population. In practice, several approximations need to be made to include chemical enrichment in a galaxy formation model.

The initial mass distribution of a stellar population, called Initial Mass Function (IMF), may depend on the physical properties of the inter-stellar medium (Weidner and Kroupa, 2005; Conroy and van Dokkum, 2012; Cappellari et al., 2012). The GAEA model assumes a universal IMF, in particular that described in Chabrier (2002). The effect of a variable IMF in the GAEA framework was analyzed in Fontanot et al. (2017a).

The yield of a certain element is evaluated theoretically studying the network of nuclear reactions in a star of a certain mass and metallicity. The abundances predicted using the theoretical yields in chemical evolution models are compared to the observed abundances. Uncertainties in these estimates come both from theoretical predictions, that can e.g. miss some nuclear process, and from the observational measurements, because some elements are difficult to measure. The yields adopted in GAEA are those of Karakas (2010) for low and intermediate-mass stars ($M < 8$ M$_\odot$), and those of Chieffi and Limongi (2004) for high mass stars SNeII ($M > 8$ M$_\odot$). The model includes also the enrichment due to SNeIa; the SNeIa rate follows the analytical formulation of Greggio (2005), and the yields are those of Thielemann et al. (2003).

Only some published semi-analytic models account for the finite stellar lifetimes (see for example Nagashima et al., 2005; Arrigoni et al., 2010; Yates et al., 2013; Gargiulo et al., 2015). Most models published assume an Instantaneous Recycling Approximation (IRA), that means that the elements produced by the stars are recycled at the moment of star formation. The main advantage of the IRA approximation is computational: relaxing this approximation requires an algorithm that distributes the metals at the correct time-step. While the IRA approximation is acceptable for massive stars that explode as SNeII, it is a strong one for AGBs and SNIa, that explode on significantly longer timescales and are the main producers of Iron peak elements.

The GAEA model accounts for a complex treatment of metal recycling, employing a non-instantaneous recycling approximation. The model was introduced in De Lucia et al. (2014b): at each time-step the model calculates the metals produced by the newly formed stellar population. These metals are stored in an array that is used, after a time equal to the appropriate stellar lifetime has elapsed, to incorporate metals in the galaxy inter-stellar medium. The stellar life times assumed are those calculated by Padovani and Matteucci (1993). This scheme allows the evolution of the α-elements, Oxygen and Iron

to be followed (on the correct time-scales).

3.3.6 AGN feedback

A large fraction of galaxies host a bright nucleus, with strong emission in X-rays, radio and ultraviolet radiation, and luminosity varying on short time scales (hours or days). Often these galaxies are characterized by radio synchrotron emission, collimated in jets, that can reach several Mpc from the galactic center. These galaxies are called Active Galactic Nuclei (AGN), and their different observational properties are explained through the so-called "unified model". In this model, AGNs host a central super-massive black hole, surrounded by a gaseous accretion disk of few pc diameter. Around the central black hole, at distances of some hundreds of parsecs, gaseous clouds rotate at large velocities around it. Just outside, a dusty torus of colder gas, optically thick, prevents direct observations of the central region when the AGN is observed edge-on. At kpc scales, low density clouds move at slow velocities, while at Mpc scales only the collimated jets, ejected perpendicularly to the plane of accretion, are visible.

In this scenario, AGNs release energy during the growth of their central supermassive black hole, by accretion of material from the accretion disk. The presence of an AGN influences the surrounding environment and the host galaxy properties. Typically, a distinction is made between two modes of feedback. The quasar mode occurs when the black hole accretes material at high rates, for example during mergers. This is the main channel of black hole growth, and is the responsible for the large scale outflows observed for these systems. Radio mode feedback is believed to be associated with low accretion rates of hot material, and manifests itself through radio bubbles observed around the central regions of massive galaxy clusters. The released energy is believed to reheat the surrounding material, stopping cooling, in a sort of "maintenance feedback".

The GAEA model includes both the radio and the quasar AGN feedback mode as detailed below.

Quasar mode

During galaxy mergers, the black holes of the progenitor galaxies are assumed to merge, and a fraction of the combined cold gas is accreted onto the remnant black hole. The prescriptions adopted in the GAEA model are the same of Croton et al. (2006), inspired to the model proposed by Kauffmann and Haehnelt (2000). Specifically, the amount of gas accretion on black holes during mergers is proportional to the total cold gas available and to the mass ratio of the merging systems:

$$\Delta M_{BH} = M_{BH,s} + f_{BH} \left(\frac{M_s}{M_p}\right) \left(\frac{M_{cold}}{1 + (280\ [\mathrm{km\ s^{-1}}]/V_{vir})^2}\right) \tag{3.25}$$

$M_{BH,s}$ is the black hole mass of the secondary, $f_{BH} = 0.03$ is a free parameter, tuned to reproduce the observed $M_{BH} - M_{bulge}$ relation in the local Universe (see Croton et al., 2006, and references therein), M_s/M_p is the baryonic mass ratio of the merging galaxies, and M_{cold} is the total cold gas in the remnant system.

The feedback due to this mode is not explicitly accounted for in our model, but galaxy mergers trigger a star burst in the cold gas remnant disk (described

in Sec. 3.3.7), and the energy released by stellar feedback can be considered as a manifestation of the quasar mode.

Radio mode

In radio mode feedback, the black hole is assumed to grow material from a static hot halo formed around the galaxy, in a continuous and quiescent accretion, described by a simple phenomenological model (following Croton et al., 2006):

$$\dot{M}_{BH,radio} = \kappa_{AGN} \left(\frac{M_{BH}}{10^8 [\mathrm{M}_\odot \mathrm{h}^{-1}]} \right) \left(\frac{f_{hot}}{0.1} \right) \left(\frac{V_{vir}}{200\ [\mathrm{km\ s}^{-1}]} \right)^3 \mathrm{M}_\odot \mathrm{yr}^{-1}. \tag{3.26}$$

M_{BH} is the black hole mass, f_{hot} the ratio between the hot gas mass and the dark matter halo mass, and $V_{vir} \sim T_{vir}^{1/2}$ the virial velocity of the halo. $\kappa_{AGN} = 7.5 \cdot 10^{-6} \mathrm{M}_\odot\ \mathrm{yr}^{-1}$ is a free parameter that sets the efficiency of the accretion, and whose value is chosen to reproduce the turn-over at the bright-end of the galaxy luminosity/mass function.

The energy released in this mode is supposed to reduce, or stop, gas cooling. Specifically, we assume that the accretion of gas onto the black holes generates a mechanical energy:

$$\dot{E}_{radio} = \eta \dot{M}_{BH} c^2, \tag{3.27}$$

with c the speed of light, and $\eta = 0.1$ the efficiency of the energy coupling. The energy injected in the medium has the effect of reducing the cooling rate:

$$\dot{M}_{cool,new} = \dot{M}_{cool} - 2 \frac{\dot{E}_{radio}}{V_{vir}^2}. \tag{3.28}$$

3.3.7 Mergers and disk instabilities

In the GAEA model, the central bulge component forms through two channels: mergers and disk instabilities. These two processes are thus important to understand the morphological mix of galaxies and its evolution.

Mergers

In a ΛCDM Universe, structures grow hierarchically, through a bottom-up scenario: the smallest halos/subhalos merge together, forming larger halos. The galaxies residing in the merging halos are also subject to mergers, and these events can have relevant effects on their internal structure.

The influence of galaxy mergers on morphology has been studied in the framework of dedicated hydrodynamical simulations (see for example Hernquist, 1992, 1993; Weil and Hernquist, 1994; Heyl et al., 1996; Barnes, 2002; Hopkins et al., 2009a), finding that the remnant properties are largely determined by the initial merger mass ratio, orbital parameters, pre-merger cold gas fraction, and initial morphology. The modeling used in GAEA is based on results from these simulations.

Mergers between galaxies are determined by mergers between their respective halos. In the initial stages of a merger, the smaller subhalo (the secondary) orbits around the larger one (the primary), losing mass through tidal stripping. The secondary subhalo is stripped until its mass falls below the resolution of

the simulation. At this point, the galaxy hosted by the secondary subhalo is assumed to survive for a certain time, becoming a type 2, orphan satellite. Its position in the model is assumed to be that of the most bound particle of the subhalo it resided in. The orphan satellite survives for a time given by dynamical friction, following the formula of Binney and Tremaine (2008):

$$t_{merge} = f_{fudge} \frac{1.17 \cdot D^2 M_p}{\ln \Lambda \cdot R_{vir,p}^2 M_s} t_{dyn,p}. \tag{3.29}$$

$f_{fudge} = 2$ is a factor accounting for results from recent numerical work (Boylan-Kolchin et al., 2008; Jiang et al., 2008), D is the distance between the secondary and the primary centers, $R_{vir,p}$ and M_p are the virial radius and the DM mass of the primary subhalo, M_s is the total mass (DM + baryons) of the secondary, $t_{dyn} = R_{vir,p}/V_{vir,p}$ is the dynamical time of the primary subhalo, and $\Lambda = 1 + M_p/M_s$ is the Coulomb logarithm. When the residual merger time has elapsed, the primary and the secondary galaxies are merged together. In the GAEA model, mergers are allowed only between centrals and type 2 satellites, or between type 1 and type 2 galaxies, but not between two type 2 satellites.

As an outcome of the merger, the cold gas of the secondary is added to the cold gas disk of the primary. The model assumes that this gas is stripped from the secondary, settling into the primary halo. Here the stripped gas takes the specific angular momentum of the primary DM halo, and, when it is finally accreted onto the primary cold gas disk, it transfers the acquired specific angular momentum to the disk, as done for cooling gas (Eq. 5.1). The total stellar mass of the secondary galaxy is added to the bulge of the primary galaxy. In this process, the specific angular momentum of the stars of the secondary is lost, because the bulge is considered to be dispersion dominated.

After the merger, a star burst is triggered in the cold gas disk of the primary, following prescriptions from controlled merger simulations, as in Somerville et al. (2001), with revised coefficients by Cox et al. (2008). Specifically, the fraction of gas in the remnant converted in stars can be expressed as:

$$e_{SB} = e_{1:1} \cdot \left(\frac{M_s}{M_p} \right)^{0.7}, \tag{3.30}$$

where $e_{1:1} = 0.55$ is the fraction of cold gas forming stars in a merger between equal mass galaxies.

The masses of the galaxies involved in the merger can influence the properties of the merger remnant: a very small secondary only marginally affects the central structure, while a secondary with a mass comparable to the primary would strongly influence it, with possibly destructive effects. In our model this is accounted for by dividing the mergers in two categories, minor and major mergers. The threshold depends on the baryonic mass ratio:

$$R_{merge} = \frac{M_{s,baryon}}{M_{p,baryon}} = 0.3, \tag{3.31}$$

where $M_{s,baryon}$ and $M_{p,baryon}$ are the total baryonic masses of the secondary and of the primary. In the case of major mergers ($R \geq R_{merge}$), all the stars in the merger remnant are moved to its bulge.

Disk instability

The galactic disk is a rotating structure, composed of cold gas and stars. Its structure is maintained by the equilibrium between the mass distribution and its rotation. Any perturbation of this equilibrium may drive to an instability in the disk, driving fragmentations and collapse of material towards the central regions. A disk instability can trigger the formation of dense gas clouds and star-bursts, while the stars can acquire random velocities, which move them out of the disk plane.

In the GAEA model, the adopted instability criterion is based on results from old N-body simulations by Efstathiou et al. (1982). An instability occurs when:

$$\frac{V_{max}^2 R_{*,disk}}{G \cdot M_{*,disk}} \leq 1 \tag{3.32}$$

where V_{max} is the maximum rotational velocity of the stellar disk, assumed equal to that of the hosting dark matter halo, $R_{*,disk}$ is the scale length of the stellar disk, and $M_{*,disk}$ is the stellar mass of the disk. When the instability criterion is met, a fraction of the stellar disk, δM, is subtracted from the central disk, and is added to the bulge. The quantity δM is calculated as the minimum mass necessary to restore stability.

This treatment is rather simplistic, and its performance and limits were discussed in several previous studies (Athanassoula, 2008; Benson and Devereux, 2010; De Lucia et al., 2012a). As I will show below, these approximations affect significantly the results found in the second part of this Thesis work, where we analyze how the dynamical properties of model galaxies are influenced by their morphology.

Chapter 4

HI-selected galaxies in models of galaxy formation and evolution

This chapter is based, with limited modifications, on the published paper Zoldan et al. (2017).

4.1 Introduction

Cold gas has a central role in galaxy evolution as it is involved in virtually all processes at play: cooling, star formation, feedback, mergers and processes related to the 'environment' such as ram-pressure and strangulation. The main component of cold gas is neutral hydrogen, in its atomic and molecular forms, each distributed in a 'disk' with a specific density profile (Leroy et al., 2008). Observationally, HI can be observed directly through the 21 cm line, while H_2 must be inferred from the ^{12}CO content through a factor, α_{CO}, whose value and dependence on other physical properties (in particular the gas metallicity) are not well known (Magdis et al., 2011; Bolatto et al., 2011; Leroy et al., 2011; Narayanan et al., 2012; Hunt et al., 2015; Amorín et al., 2016).

Surveys such as the HI Parkes All Sky Survey (HIPASS, Zwaan et al., 2005) or the Arecibo Legacy Fast ALFA survey (ALFALFA, Martin et al., 2010), have recently collected large amounts of data, but since the HI signal is relatively faint, these data are generally limited to low redshift and relatively HI rich galaxies. These observations have provided detailed information on the basic statistical properties of HI selected galaxies in the local Universe, in particular the HI mass function (Zwaan et al., 2005; Martin et al., 2010) and the 2-point correlation function (Meyer et al., 2007; Martin et al., 2012; Papastergis et al., 2013). We live in a very exciting era for HI observations, as new radio facilities such as the Square Kilometre Array (SKA) and its precursors/pathfinders will soon extend significantly the redshift range and dynamical range in HI mass probed (Obreschkow and Rawlings, 2009; Kim et al., 2011), providing a new crucial testbed for our models of galaxy formation and evolution. It is therefore important to develop our theoretical tools, and assess how their basic predictions

compare with the available data.

Until recently, the inter-stellar medium (ISM) of galaxies, in both semi-analytic models of galaxy formation and hydrodynamical simulations, was treated as a single phase constituted of gas with temperature below 10^4 K. Stars were assumed to form from this gaseous phase adopting different parametrizations, generally based on some variation of the Kennicutt-Schmidt law (Kennicutt, 1989). In the last decade, our understanding of how star formation depends on the local conditions of the ISM has improved significantly thanks to new high-resolution observations of the multiphase gas in samples of nearby galaxies. These data have shown that the density of star formation correlates strongly with the molecular gas density, while there is almost no correlation with the total gas density (e.g. Blitz and Rosolowsky, 2006; Leroy et al., 2008, and references therein). These results have triggered significant activity in the theoretical community aimed at including an explicit treatment for the transition from atomic to molecular hydrogen, and molecular hydrogen based star formation laws (Fu et al., 2010; Lagos et al., 2011a; Christensen et al., 2012; Kuhlen et al., 2013). Recent studies have focused on the impact of different physical processes and of cosmology on the HI distribution and properties of HI selected galaxies (e.g. Popping et al., 2009; Obreschkow and Rawlings, 2009; Kim et al., 2011; Davé et al., 2013; Rafieferantsoa et al., 2015; Kim et al., 2017; Crain et al., 2017).

In this chapter, we focus on results from semi-analytic models of galaxy formation. While these methods do not include an explicit treatment of the gas dynamics and prevent (at least in their standard implementations) studies of spatially resolved properties of galaxies, their computational cost is typically significantly lower than that of high-resolution hydrodynamical simulations. They therefore offer a flexible and efficient tool to analyze the dependence of model predictions on specific assumptions and parameterizations. We focus here on the basic statistical properties of HI selected galaxies (i.e. their distributions in mass, their halo occupation distribution, and their 2-point clustering function). To analyze the dependence on the different modeling adopted, we use four different models whose results are publicly available on the Virgo-Millennium Database (Lemson and Virgo Consortium, 2006), and other two models recently developed by our group (Hirschmann et al., 2016; Xie et al., 2017). The latter model, in particular, includes an explicit treatment for partitioning the cold gas associated with model galaxies in HI and H_2. For all other models, as we will discuss below, the HI content of model galaxies has been derived post-processing model outputs.

The layout of the chapter is as follows. In Sec. 4.2, we briefly describe the models used in our study, their backbone simulations, and the algorithm we use to build mock light-cones from model outputs. In Sec. 4.3, we explain how we assign HI masses to model galaxies where this is not an explicit model output, and compare the distribution of HI masses and basic scaling relations predicted by the models with available observational data. In Sec. 4.4, we analyze the predicted 2-point correlation function of HI-selected galaxies, and compare model predictions with recent measurements by Papastergis et al. (2013). We discuss the results on the basis of the predicted halo occupation distribution. In Sec. 4.5, we discuss the role of satellite galaxies and study the predicted evolution of their HI content. In Sec. 4.6, we study the relation between the HI in galaxies and the physical properties of the hosting dark matter haloes. Finally, in Sec. 4.7, we summarize our results and give our conclusions.

4.2 Simulations and galaxy formation models

In this work, we take advantage of 6 different semi-analytic models of galaxy formation and evolution, that originate from the work of three independent research groups. All models are run on the same N-body cosmological simulation, i.e. the Millennium Simulation (Sec. 4.2.2), and all follow the principles outlined in White and Frenk (1991a) including specific (different) modeling for cooling, star formation, stellar feedback, mergers and starbursts, disk instabilities, chemical enrichment and feedback from Active Galactic Nuclei (AGN). Using models that employ different prescriptions for processes involving cold gas, it is possible to quantify the relative importance in determining the observed relations. At the same time, systematic (i.e. common to all models) disagreements can be used to identify specific aspects of the models that need to be improved. Below, we briefly outline the main differences between the models considered, the main characteristics of the simulations used, and describe the software built to construct mock light-cones from the available galaxy catalogues.

4.2.1 Galaxy formation models

The Bower et al. (2006, B06 hereafter) model has been developed by the 'Durham group' and is an extension of the GALFORM model published in Cole et al. (2000) and Benson et al. (2003). The model published in De Lucia and Blaizot (2007, hereafter DLB07) has been developed by the 'Munich group' and represents an extension of the model described earlier in Springel et al. (2001), De Lucia et al. (2004), and Croton et al. (2006). These two models have been the first whose results were made publicly available through a relational database[1] that we have heavily used for our study.

It was early realized that both these models tend to over-predict the number densities of galaxies smaller than the Milky Way, and the overall fraction of passive galaxies - problems that have turned to be of difficult solution and shared by all hierarchical models published in the last years, including hydrodynamical cosmological simulations (see e.g. De Lucia et al., 2014a, and references therein). The other models we use in this study represent two independent branches, both based on subsequent upgrades of the DLB07 model, that provide significant improvements on these problems.

The model described in Guo et al. (2011, G11) differs from the DLB07 model primarily for a more efficient feedback and a non-instantaneous stripping of the hot reservoir associated with galaxies at the time of infall on larger systems. The modified treatment of satellite galaxies improves the agreement with observational data as for the fraction of active galaxies, while the more efficient feedback brings the predicted mass function in agreement with data in the local Universe. The model, however, still suffers of an excess of intermediate to low-mass galaxies at higher redshift. Henriques et al. (2013) adopt the same physical model as in G11, but include a variation with cosmic time and halo mass of the efficiency with which galactic wind ejecta are re-accreted. They then use Monte Carlo Markov Chain methods to identify the parameter space that allows the measured evolution of the galaxy mass (and luminosity) function from z=0 to z=3 to be reproduced. Henriques et al. (2015, H15) have later extended this

[1] http://wwwmpa.mpa-garching.mpg.de/millennium/

work to the Planck first-year cosmology, by "re-scaling" the Millennium Simulation merger trees, as explained in Sec. 4.2.2. In this work, we use model outputs based on the re-scaled simulation.

The other independent branch we use in our study has been developed by the 'Trieste group', and is provided by the recently published GAlaxy Evolution and Assembly (GAEA) model (Hirschmann et al., 2016), and by an extension of this model (Xie et al., 2017, XBR16). The GAEA model differs from DLB07 primarily for the inclusion of a sophisticated chemical enrichment scheme that accounts for the finite lifetimes of stars and independent yields from massive stars and both SNII and SNIa, and for an updated stellar feedback scheme, based on the results obtained in the framework of the Feedback In Realistic Environments (FIRE) simulations (Hopkins et al., 2014). The GAEA model includes modifications of the re-incorporation rate of ejected gas similar to those suggested by Henriques et al. (2015). We refer to the original paper for more details. It is worth stressing that GAEA adopts an instantaneous stripping of the hot gas reservoir associated with infalling galaxies. Yet, Hirschmann et al. (2016) demonstrate that this model is successful in reproducing simultaneously the evolution of the galaxy stellar mass function, the fraction of passive galaxies[2] as a function of stellar mass observed in the local Universe, and the observed evolution of the relation between galaxy stellar mass and metallicity content of the gaseous phase.

All models introduced so far treat the gas as a single phase component, and therefore need to be *post-processed* to infer the HI content associated with each model galaxy, as described in Sec. 4.3.1. The XBR16 model represents an update of the GAEA model including a treatment for the partition of the gaseous phase in atomic and molecular gas, based on the empirical relation by Blitz and Rosolowsky (2006). Therefore, among all models considered in our study, this is the only one including a self-consistent treatment of the star formation rate, which depends on the molecular hydrogen content of galaxies. In this model the disk is divided in 21 concentric annuli (in all other models star formation is modelled assuming a single value for the entire disk), and the star formation rate is estimated in each of the annuli assuming it is proportional to the surface density of molecular gas. A partition of the cold gas into its molecular and atomic components is performed at each time-step (just before star formation takes place) of the simulation, in each annulus. This allows star formation to continue for longer times than in the other models in the inner parts of the disk. This more sophisticated treatment of the star formation process leads to relevant differences with respect to the GAEA model, as we will show in the next sections.

Among the models used in this chapter, B06, DLB07, G11 and H15 are publicly available at the relational database mentioned above, developed as part of the activities of the German Astrophysical Virtual Observatory (GAVO). The GAEA and XBR16 models are not yet public, but will soon be released on the same database. Additional models, including more recent versions of the GALFORM model featuring a treatment for the partition of cold gas in atomic and molecular hydrogen, are accessible form an alternative database server at

[2]The quiescent fraction measured for the lowest mass galaxies is still larger than observational estimates, but improved significantly with respect to predictions from the DLB07 model.

the Institute for Computational Cosmology[3].

4.2.2 Dark matter simulation and merger trees

All models presented in the previous section have been run on the Millennium Simulation (Springel et al., 2005), based on WMAP1 cosmology, and described in detail in Sec. 3.1 and Sec. 3.2.

As mentioned above, H15 is based on a different cosmology: the authors still use the Millennium Simulation but 'rescale' it to the PLANCK cosmology (Planck Collaboration et al., 2014) using the technique discussed in Angulo and White (2010), as updated by Angulo and Hilbert (2015). Specifically, the adopted cosmological parameters are: $\Omega_b = 0.0487$ ($f_b = 0.155$), $\Omega_m = 0.315$, $\Omega_\Lambda = 0.685$, $h = 0.673$, $\sigma_8 = 0.829$ and $n_s = 0.96$. We note that as found in previous studies (e.g. Wang et al., 2008; Guo et al., 2013), relatively small variations of the cosmological parameters are found to have little influence on the overall model predictions once model parameters are (slightly) modified to recover the same model normalization.

To control the resolution limits of our models, we have also taken advantage of the Millennium II simulation (Boylan-Kolchin et al., 2009, MS-II). The size of this simulation is one fifth that of the Millennium Simulation (i.e. 100 Mpc h^{-1} on a side), but the particle mass is 125 times lower, resulting in a minimum resolved halo mass of $M_h = 1.4 \times 10^8$ $\mathrm{M_\odot h^{-1}}$. Below, we will use G11, H15 and XBR16 outputs based on the MS-II in order to check the reliability of some relations, in particular those involving dark matter halo masses in the $10^{11} - 10^{12}$ $\mathrm{M_\odot h^{-1}}$ range.

4.2.3 Light-cones algorithm

In Sec. 4.4, we analyze the 2-Point Correlation Function of galaxies selected according to their HI content. We use mock light-cones and the projected correlation in order to carry out a straightforward comparison with observations. In addition, the construction of many light-cones (some dozens) allows us to give an estimate of the error on the projected 2PCF due to the cosmic variance, as well as of the random noise expected for the real 2PCF.

We have developed an algorithm that creates mock catalogues from the outputs of galaxy formation models. We use these catalogues in the calculation of the 2PCF (Sec. 4.4). The approach is very similar to that adopted in the Mock Map Facility developed by Blaizot et al. (2005, MoMaF), and we briefly summarize it below. We refer to the original paper for more details.

As explained above, the outputs of the galaxy formation models used in this study are given by galaxy catalogues stored at a finite number of snapshots, each corresponding to a different redshift z_i. Galaxy catalogues contain a number of physical properties (i.e. masses, metallicities, luminosities, etc.) and consistent redshift and spatial information, i.e. the position of each model galaxy within the simulated box, as well its velocity components. The first problem that arises when constructing mock observations from these types of outputs is that redshift varies continuously along the past light cone while outputs are stored at a finite number of redshifts (in our particular case spaced at approximately 300

[3] http://galaxy-catalogue.dur.ac.uk:8080/Millennium/

Myr intervals out to $z = 1$). The other problem arises because of the need to 'replicate' the same simulated region of the Universe (in our case corresponding to a box of $500\,h^{-1}$ Mpc comoving on a side) several times to fill the light-cone.

We try to minimize these problems by building the light-cone as follows (we refer to Fig. 4.1 for a schematic description of our method).

1. We build a 3-D grid made of cubes of the same size l_{box} of the simulation box (the red grid in the figure).
2. The cone is placed inside this grid horizontally and moved so as to overlap with the smallest number of grid boxes (the blue cone in the figure).
3. The cone is divided into different regions whose edges are given by the average redshift of two subsequent snapshots, namely $z_i^{lc} = (z_i + z_{i+1})/2$ (black arrows in the figure).
4. Model galaxies are then placed inside the cone according to the following rules:
   - Each redshift region z_i^{lc} of the cone is filled with galaxies from the box corresponding to the closest redshift z_i (the correspondence is colour-coded in the figure).
   - The positions (and velocities) of model galaxies are randomly transformed, in order to avoid replications of structures along the line of sight. A specific random transformation is assigned to each 3-D grid box. This is a combination of a random shift, a random rotation of 0, $\pi/2$, π or $3\pi/2$ around a random axis or an inversion of coordinates along a random axis. In the figure, a specific transformation is indicated with a big arrow.

Referring to the specific example shown in Fig. 4.1, the lightcone element corresponding to the second grid box (from the left) is divided in two redshift regions by the edge $z_2^{lc} = (z_1 + z_2)/2$. The positions (and velocities) of the galaxies in the snapshots corresponding to z_1 and z_2 are transformed using the same specific transformation T_2. The transformed snapshots are then used to fill the grid element considered. Galaxies below z_2^{lc} are extracted from the transformed snapshot z_1, while those above z_2^{lc} are taken from the transformed snapshot z_2.

Adopting this approach, large scale structures are affected at the edges of the 3D grid boxes, but not at the redshift separation between subsequent snapshots. This allows accurate measurements of the 2-point correlation function (2PCF).

Blaizot et al. (2005) carried out a detailed analysis of replication effects on the clustering signal. In particular, they quantified the effect due to (i) random tiling and (ii) finite volume. The random tiling approach introduces a negative bias in the mock catalogues because it decorrelates pairs of galaxies when reshuffling them to suppress replication effects. Blaizot et al. (2005) gave a theoretical estimate of the relative error in the number of pairs due to shifting of boxes. For the spatial correlation function, this is found to be less than 10 per cent on scales ranging from 1 to $\sim 10\,h^{-1}$ Mpc. A numerical estimate was also computed by subdividing the volume of the simulation used in their study (a cube of $100\,\mathrm{h}_{100}^{-1}$ Mpc on a side) in 8^3 sub-volumes to which translations, rotations and inversions were applied as described above. We have repeated the same

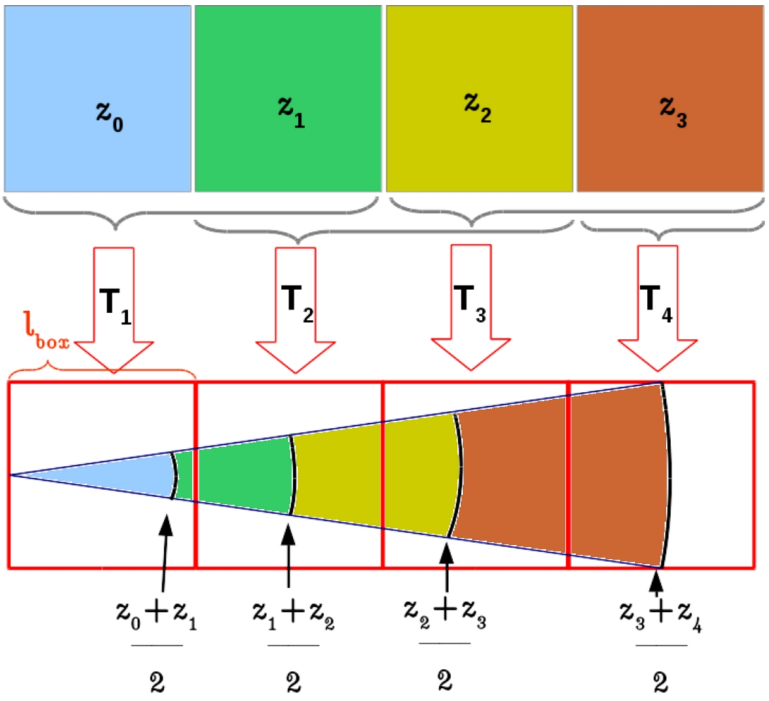

Figure 4.1: A schematic illustration of our light-cone construction algorithm: a grid of cubic boxes with size length l_{box} (red) is built over the cone, and galaxies inside each box of the grid are randomly shifted/rotated/inverted depending on the specific box they end up in. Galaxies are extracted from the simulation output with redshift closest to the co-moving distance to the observer.

exercise by taking advantage of the larger volume of the Millennium Simulation and considering sub-boxes of different sizes ($L_{box} = 250,\ 125$ Mpc h^{-1}). For the analysis presented in the following, light-cones are built starting from the entire volume of the MS ($L_{box} = 500$ Mpc h^{-1}), and we have verified that the random tiling approach is not expected to significantly affect our measurements.

The finite volume of the simulation prevents us from studying fluctuations larger than $\sim V^{1/3}$. Since in this study we focus on clustering on relatively small scales (~ 20 Mpc h^{-1}), we do not expect the finite volume of our simulation to affect significantly our measurements.

4.3 Neutral hydrogen distribution and scaling relations

In this section, we study the distribution of HI masses and basic scaling relations predicted by our models, and compare them with recent observational determinations. Given the resolution mass limit of the Millennium Simulation, all the following analysis is limited to galaxies with $M_* > 10^9\,M_\odot$. Below this limit the sample is incomplete (this is evaluated comparing the galaxy stellar mass function for models based on the MS and on the MS-II). Using the HI mass function, Xie et al. (2017) have determined the completeness limit in HI mass, that is assessed to be $M_{HI} \sim 10^{8.8}\ M_\odot$. In the following analysis, we study relations involving HI masses slightly below this limit. We checked the influence of incompleteness using the MS-II for the G11, H15 and XBR16 models, and found that, qualitatively, our results do not change.

The following analysis is performed on a sub-volume of the MS of one sixth, but we have verified that results do not change significantly when considering a larger volume.

4.3.1 Neutral hydrogen mass estimates

XBR16 is the only model among those considered in our study that accounts self-consistently for a direct dependence of star formation rate on the molecular hydrogen density (Blitz and Rosolowsky, 2006; Leroy et al., 2008). This is done by introducing a partition of the cold gas phase in atomic and molecular hydrogen, based on the empirical relation by Blitz and Rosolowsky (2006), at the time of each star formation episode. Therefore, the HI and H_2 content of model galaxies are direct outputs of this model, even if they are not explicitly followed as two separated components. For all the other models, star formation is assumed to depend on the total cold gas content through a classical Kennicutt (1989) relation. In this case, the HI content of galaxies can be estimated using two different methods:

- Combining the observed gas density profile (Leroy et al., 2008) and the dependence of $R_{mol} = M_{H_2}/M_{HI}$ on the disk mid-plane pressure (Blitz and Rosolowsky, 2006), Obreschkow and Rawlings (2009) find the following relation:
$$R_{mol} = [3.44 R_c^{-0.506} + 4.82 R_c^{-1.054}]^{-1}.$$
In this equation, $R_c \sim [r_{disk}^{-4} M_{cold}(M_{cold} + 0.4 M_*)]^{0.8}$ is the central R_{mol}, r_{disk} is the disk scale length, M_{cold} the cold gas mass in the disk, and M_*

the stellar mass of the disk.

- Following Baugh et al. (2004) and Power et al. (2010), we can assume a direct proportionality between the cold gas content and HI. Helium amounts to about 24 per cent of the total cold gas, while metals represent a negligible fraction. Assuming the HII content is negligible, hydrogen can be divided in $\sim$ 71 per cent of HI and $\sim$ 29 per cent of H_2, namely $R_{mol} = M_{H_2}/M_{HI} = 0.4$ (Keres et al., 2003; Zwaan et al., 2005).

Power et al. (2010) analyzed the impact of these two approaches on different semi-analytic models (including some used in our study) and found little differences between them at low redshift (differences are larger at high redshift). We confirm that there are only small differences between the HI content of model galaxies estimated using the two methods described above, for all models used in our study. Therefore, unless otherwise stated, we will always compute HI using the Obreschkow and Rawlings (2009) prescription.

4.3.2 Stellar and HI Mass Functions

In Fig. 4.2, we compare the galaxy stellar mass function predicted by the models with observational measurements by Moustakas et al. (2013) and Bernardi et al. (2013). Both determinations are based on SDSS data, but assume different estimates of stellar mass. In particular, the latter is based on a Sérsic-exponential fit to the surface brightness profiles of galaxies, and translates in significantly higher number densities of massive galaxies. The two measurements highlight that, while uncertainties are relatively small for the intermediate and low-mass regime, they are much larger above the knee of the mass function. Furthermore, in this mass regime also statistical uncertainties are larger because of the small number of galaxies.

All the models reproduce well the observational measurements around the knee, with a slight underestimation for the B06 and H15 models. The low-mass end of the mass function is well reproduced only by the H15 model, while the largest over-prediction is found for the DLB07 and B06 models. As mentioned above, the good agreement found for the H15 model is obtained by construction as these authors tune their model parameters to reproduce the evolution of the galaxy mass function from z=0 to z=3. At high masses, models tend to be closer to the Moustakas et al. (2013) determination up to stellar masses $\sim 10^{11}\,M_\odot$, while they basically cover all the range between the two measurements shown for larger masses. For the most massive galaxies, the lowest number densities are obtained by the B06 and H15 models, the largest by GAEA and XBR16.

Fig. 4.3 shows the HI mass function predicted by the models used in our study, obtained using the methods discussed in Sec. 4.3.1 (with the exception of XBR16, that returns a direct estimate of the HI mass). Model predictions are compared to observational data from Zwaan et al. (2005), based on HIPASS, and from Martin et al. (2010), based on the ALFALFA survey.

The two different HI estimates (solid and dashed lines in the figure) return quite similar results, with the Power et al. (2010) prescription slightly under-predicting the HI high mass end with respect to the method suggested by Obreschkow and Rawlings (2009).

B06 is the model that deviates most from the data. In particular, it over-predicts the number densities of galaxies with high HI masses ($M_{HI} \geq 10^9\,M_\odot h^{-2}$),

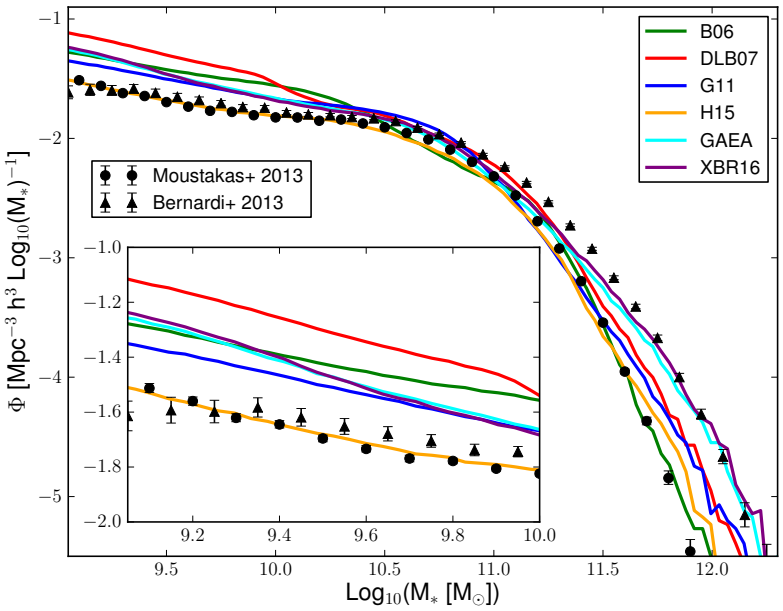

Figure 4.2: The galaxy stellar mass function predicted by the six models used in our study, compared to observational measurements by Moustakas et al. (2013, circles) and Bernardi et al. (2013, triangles), both based on SDSS. The inset shows a zoom on the low-mass end.

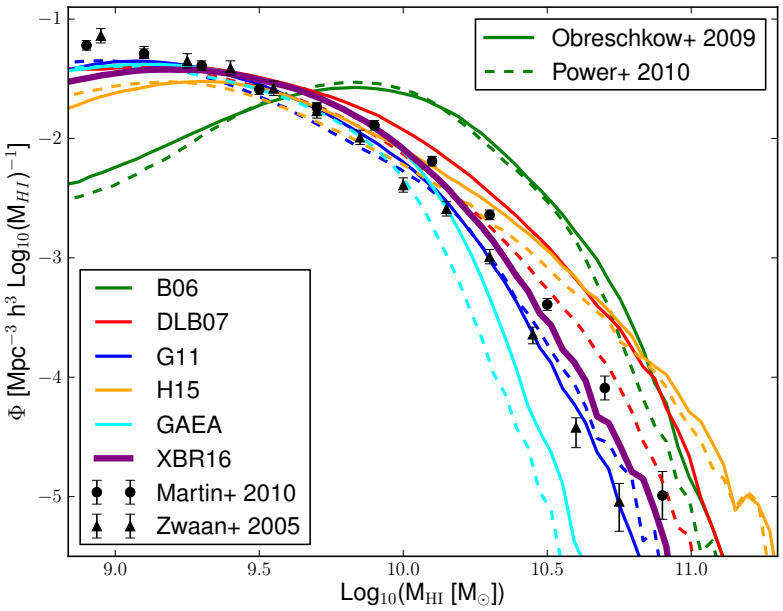

Figure 4.3: The HI mass function predicted by the models, compared to observational data by Martin et al. (2010, circles) and Zwaan et al. (2005, triangles). Solid lines correspond to the Obreschkow and Rawlings (2009) method to assign HI masses, while dashed lines correspond to the assumption of a constant proportionality between total cold gas and HI mass as in Power et al. (2010).

and under-predicts it for lower masses. Kim et al. (2011) already discussed this limitation of B06 and proposed a new version of this model tuned to reproduce the HIMF of Zwaan et al. (2005). This modified model also resulted in slightly better agreement with the observed 2PCF based on HIPASS. G11 and XBR16 reproduce quite well the observations, while the GAEA model underestimates the high mass end. We remind that GAEA and XBR16 are based on the same physical model but the latter includes an explicit treatment for the partition of cold gas in atomic and molecular hydrogen in disk annuli, and a star formation law that depends on H_2. The division in annuli translates in the XBR16 model having less H_2 at fixed star formation rate, and this leads to a larger amount of HI (at fixed cold gas). This effect is enhanced for gas rich galaxies. Finally, we note that the XBR16 model is tuned to reproduce the HIMF.

The slight underestimation of the HIMF at the low mass end in all the models is due to resolution (Xie et al., 2017).

4.3.3 Scaling relations

Fig. 4.4 shows the predicted relation between the HI mass and the galaxy stellar mass for all models used in our study. Solid lines show the median of the distributions, while dashed lines (and the shaded region that corresponds to the black line) mark the 1σ spread. Lines of different colours correspond to centrals (blue) and satellite galaxies (orange). Symbols correspond to observational data from the GALEX Arecibo SDSS Survey (Catinella et al., 2013), divided in detected HI (green pentagons) and non detections, and thus upper limits (purple triangles).

Given the large spread of the observational data and the detection limits, it is hard to use these data to put strong constraints on the models. All models cover the region sampled by the observational data but for the B06 model, that predicts no HI in galaxies more massive than $\sim 10^{10}\,M_\odot$ likely because of too efficient AGN (radio-mode) feedback. We accounted for the same detection limit of observations (for $M_* > 10^{10}$ $M_\odot$, only considered galaxies with M_{HI} larger than the maximum no detection in the observed sample). Results are shown in this case as black diamonds with 1σ error-bars: this narrows the distributions of model galaxies, particularly for the G11, GAEA, and XBR16 models. Overall, the agreement between model predictions and observations is good, but some models (e.g. G11, GAEA and XBR16) exhibit a deficit of HI-rich galaxies at any given stellar mass. This remains statistically significant even when applying a detection threshold similar to that of the survey used as a comparison.

As expected, central galaxies always have larger HI masses than satellite galaxies. The difference is largest for the B06 and DLB07 models that adopt an instantaneous stripping of the hot gas reservoir associated with infalling satellites. The G11 and H15 models relax this assumption, which brings the median HI mass of satellite galaxies closer to that of central galaxies. Interestingly, the difference between these medians in the GAEA and XBR16 models, that also assume an instantaneous stripping of the hot gas, is only slightly larger than that found in G11 and H15. As discussed in Hirschmann et al. (2016), this is driven by significantly larger amounts of cold gas at higher redshift and a reduced reheating rate from stellar feedback. Interestingly, at the most massive end, the HI content of satellite galaxies is comparable to that of centrals in all models but B06 and DLB07.

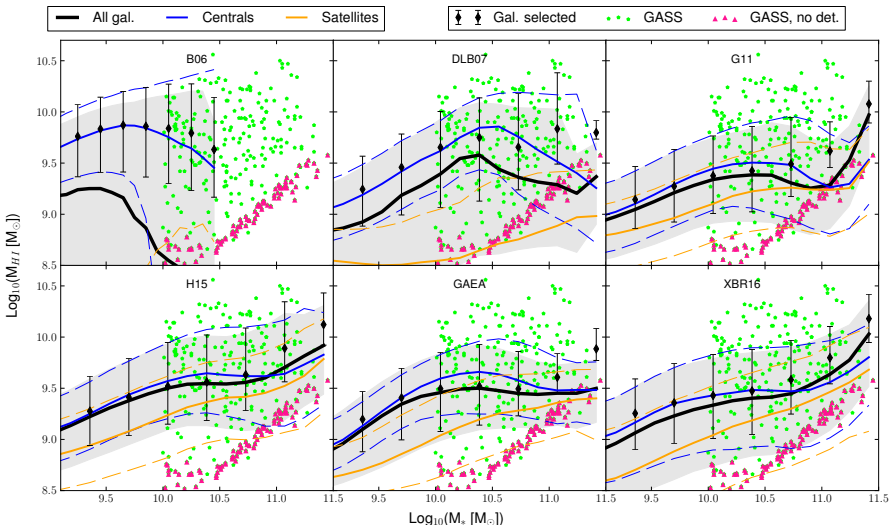

Figure 4.4: The predicted relation between HI mass and galaxy stellar mass, compared to results from the GALEX Arecibo SDSS Survey (symbols). Lines of different colours are used for central galaxies (blue) and satellites (orange), while the black lines and shaded regions correspond to the entire galaxy population. Solid lines correspond to the median of the distributions, while dashed lines (and the shaded region) mark the 1σ distribution. The black diamonds with error-bars show the distribution of all model galaxies when using the same detection threshold of the data. Symbols of different colours are used for detections (green pentagons) and non detections (upper limits, purple triangles).

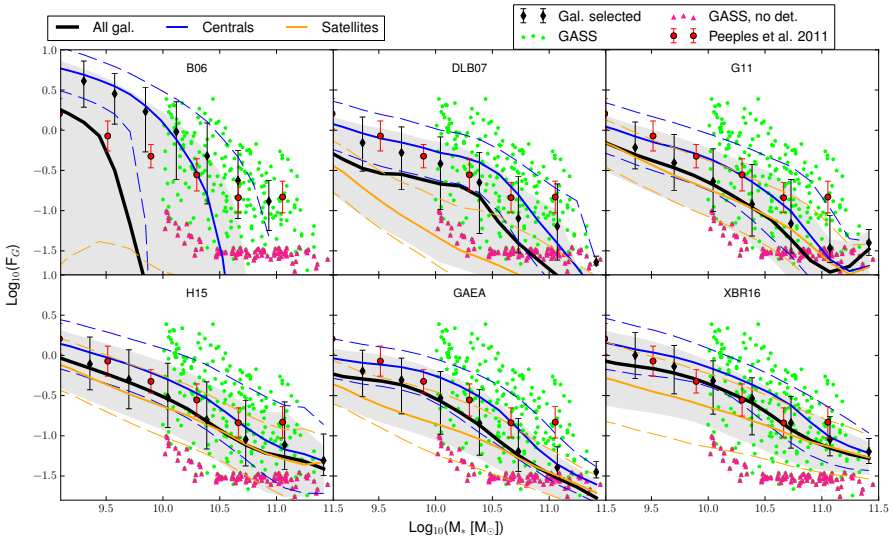

Figure 4.5: The gas fraction $F_G = M_{cold}/M_*$ of model galaxies as a function of their stellar mass, compared to observational measurements by Catinella et al. (2013, green pentagons and pink triangles), and Peeples and Shankar (2011, red circles with error bars). Lines (solid and dashed), shaded region and diamonds with error-bars have the same meaning as in Fig. 4.4.

In Fig. 4.5, we show the cold gas fraction $F_g = M_{cold}/M_*$ as a function of the galaxy stellar mass for all models considered in our study. Symbols show observational measurements by Catinella et al. (2013, GASS survey, pentagons and triangles) and Peeples and Shankar (2011, red points with error bars). The latter are a collection of previous HI measurements from McGaugh (2005), West et al. (2009, 2010) and Leroy et al. (2008). To convert the measurements in cold gas fraction, we have assumed HI represents a constant fraction ($\sim$71 per cent) of the total cold gas available.

As discussed above, the B06 model clearly under-predicts the gaseous content of galaxies of intermediate to high mass. In addition, virtually all gas in this model is associated with central galaxies. Comparing Fig. 4.5 and Fig. 4.4 with the predicted HI mass function for this model (Fig. 4.3), we can see that the excess of HI rich galaxies is driven by the high gas fractions of galaxies with low stellar masses. As noted above, even accounting for the same selection limits of the observations (black diamonds and error-bars), only some of the models predict gas fractions as large as those observed (DLB07, H15, XBR16). Finally, both the DLB07 and the GAEA models under-predict the gas fractions estimated for galaxies with the largest stellar mass. When applying the same selection of the data, however, model predictions appear consistent with observations. It should be noted that observational measurements are sparse and likely incomplete in this mass regime, thus further measurements are necessary to constrain this relation at the highest stellar masses.

4.4 Two point projected correlation function of HI selected galaxies

In this section we will analyze the clustering properties of model galaxies, selecting them in HI bins. We use, as observational reference, the work by Papastergis et al. (2013, P13 hereafter). They used 6,123 HI-selected galaxies from the ALFALFA survey, covering a contiguous rectangular sky region of $\sim 1,700$ deg^2 in the redshift range $z \sim 0.0023 - 0.05$. They also used an optical sample of 18,516 galaxies in the same volume using the SDSS DR7 (Abazajian et al., 2009), and applying a magnitude cut of $M_r < -17$. As expected and shown in previous studies (e.g. Catinella et al., 2010; Huang et al., 2012), the most luminous galaxies tend to be HI poor, while HI rich galaxies tend to be the dominant population among galaxies that populate the blue cloud.

Below, we will compare our model predictions with the estimated projected correlation functions for the following HI mass bins: $\log_{10}(M_{HI}[\mathrm{M}_\odot]) \in [8.5;\ 9.5]$, $[9.5;\ 10.0]$ and $[10.0;\ 10.5]$. Lower mass bins were considered in P13, but these fall below the resolution limit of our models. The lowest bin considered here is already partially below the completeness limit. We will account for this in the following.

P13 found little evidence of dependence of the clustering signal on the HI content, with some uncertainties for the lowest HI mass bin. For galaxies in this bin, a lower clustering signal is measured with respect to HI richer galaxies, but P13 argue this is not statistically significant due to the lower volume sampled by galaxies in this HI bin.

4.4.1 The projected correlation function

We compute the two-point correlation function (2PCF) for all models used in our study taking advantage of the mock light-cones described above in Sec. 4.2.3. To compute the predicted 2PCF, we use the Landy and Szalay (1993) estimator:

$$\xi(r) = \frac{DD(r) + RR(r) - 2DR(r)}{RR(r)}$$

where $DD(r)$, $RR(r)$ and $DR(r)$ represent the galaxy-galaxy, random-random, and galaxy-random number of galaxy pairs separated by a distance r. In observations, the physical separation r is not directly available, and the observables are the position on the sky and the recessional velocity. Using our mock light-cones (Sec. 4.2.3) we mimic the data and carry out all calculations in redshift space. We measure the separation among two galaxies as $s = \sqrt{(v_1^2 + v_2^2 - 2v_1v_2\cos\theta)}/H_0$, where v_1 and v_2 are the recessional velocities of the galaxies (km s^{-1}), θ is the angle between them in the sky, and H_0 is the Hubble constant. Hence we calculate the correlation function using the separation along the line of sight ($\pi = |v_1 - v_2|/H_0$) and on the sky plane ($\sigma = \sqrt{\pi^2 - s^2}$) to obtain $\xi(\pi, \sigma)$. The projected correlation function used in the following discussion corresponds to

$$w(\sigma) = 2\int_0^\infty \xi(\pi, \sigma)d\pi.$$

P13 had a non-uniform radial selection function for their sample and accounted for it in their random sample. We do not attempt to mimic the selection function of the data, and simply use for the random catalogue the same smooth redshift distribution of the selected model galaxies.

4.4.2 Model predictions

Fig. 4.6 shows the 2PCF for all models considered in this study. P13 found that the middle and highest HI mass bins ($M_{HI} \in [10^{9.5};\ 10^{10}]$ and $[10^{10};\ 10^{10.5}]$ M$_\odot$) have almost the same 2PCF, while they considered the measurements obtained for the lowest HI bin ($M_{HI} \in [10^{8.5};\ 10^{9.5}]$ M$_\odot$) not reliable due to the small sampling volume. To better show the differences between model predictions and observational measurements, we show in the bottom sub-panels model results in each HI mass bin divided by the observational measurement corresponding to the bin $M_{HI} \in [10^{9.5};\ 10^{10}]$ M$_\odot$.

For the lowest HI mass bin ($[10^{8.5};\ 10^{9.5}]$ M$_\odot$), a problem is immediately evident: P13 found these galaxies to be less clustered than their HI richer counterparts, but argued that this might be due to finite volume effects. All models considered in this study, in contrast, predict a higher clustering signal for galaxies in this bin. The B06 model is the closest to the observational measurements, but still standing a little above them. All other models largely over-predict the measured clustering signal, particularly at small scales, with G11 and H15 deviating most from the data. As we will see in the following, this HI mass bin is sensitive to various physical prescriptions and to numerical resolution. In addition, it is dominated by satellite galaxies and therefore strongly dependent on the adopted treatment for satellite evolution. We verified how the incompleteness of the model sample influences the final 2PCF of this bin for the G11,

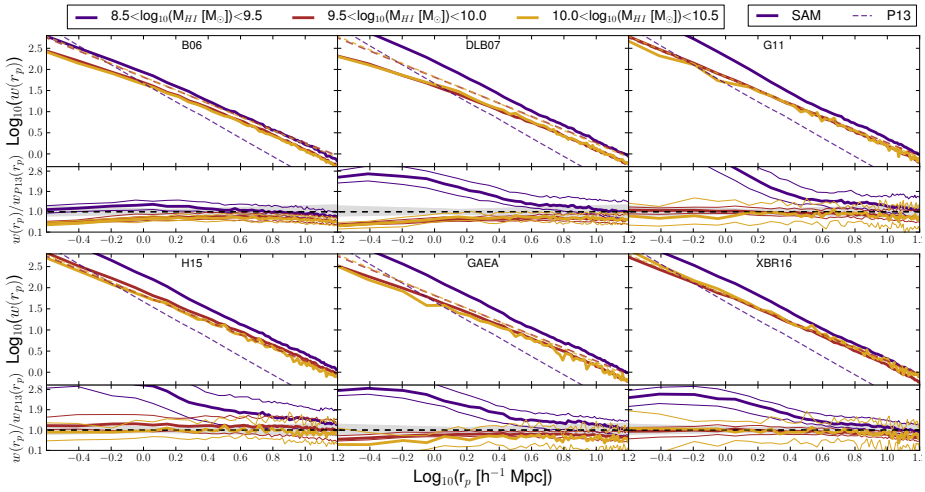

Figure 4.6: The two-point projected correlation function for different bins in HI mass (solid thick lines of different colours). Each panel corresponds to a different model, and the lower sub-panels show the ratio between the model predictions and the corresponding observational estimates (dashed lines) by P13 for the medium HI mass bin. As discussed in the text, this is equal to the one of the highest HI mass bin and more reliable than the measurements corresponding to the lowest HI mass bin. The 1σ scatter is shown in the lower sub-panels using thin lines for models and a shaded region for the data.

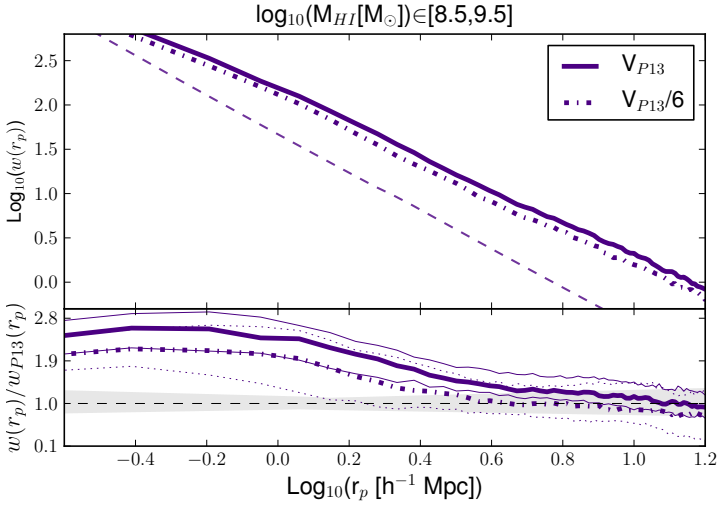

Figure 4.7: The 2PCF for the lowest HI mass bin for the XBR16 model. The solid line corresponds to model predictions for the full volume considered (i.e. this line corresponds to that shown in Fig. 4.6). The dot-dashed line corresponds to the results obtained when considering only one sixth of the volume, as in the observational sample by P13. In the lower sub-panel, we show the ratio between model predictions and the observational estimate of the medium HI bin. Shaded regions and thin lines show the 1σ scatter.

H15 and XBR16 models using the MS-II: we find that the predicted clustering signal is lower than that found for the MS, but still a factor ~ 2 larger than that measured.

For HI richer galaxies ($[10^{9.5};\ 10^{10}]$ and $[10^{10};\ 10^{10.5}]$ $M_{\odot}$), B06, DLB07 and GAEA systematically under-predict the clustering signal, while G11, H15 and XBR16 are in good agreement with observational measurements in the corresponding mass bins.

The relatively noisy behaviour of the 2PCF in the highest HI bin is due to the small number of galaxies with such large HI masses.

As noted above, P13 argue that for the lowest HI mass bin, results are affected by smaller sampling volume. We can test the influence of a small sampling volume using our mocks. We show results of this test in Fig. 4.7. In this case, we use only the XBR16 model (results are similar for the other models) and show the 2PCF only for the lowest HI mass bin. The corresponding model predictions obtained by using only one sixth of the volume are shown as a dotted-dashed line, and are not statistically different from those obtained using the entire volume (solid line).

4.4.3 Halo occupation distribution

The results in the previous section can be understood by considering the number occupation of haloes of different mass by galaxies with different HI content, i.e. what is typically referred to as halo occupation distribution (HOD). The HOD gives information about the distribution of galaxies in dark matter haloes of

different masses, with the possibility to distinguish between centrals and satellites. It can be used to interpret the 2PCF at the scales of the halo dimensions. For this analysis, we use the data from the $z = 0$ snapshots to have a larger statistics.

Results from all models used in this study are shown in Fig. 4.8. Each panel shows the average number of galaxies with stellar mass larger than $M_* > 10^9 \mathrm{M}_\odot$ (approximately corresponding to the resolution limit of the MS) and with different HI mass (different columns) that reside in FoF haloes of mass M_{200}^{FoF}. Blue and orange lines correspond to central and satellite galaxies respectively, while black solid lines show the total. We checked the convergence of the results obtained for low mass FoF haloes by considering the G11, H15 and the XBR16 model on the MS-II (results are shown as dashed lines in the corresponding rows). The HOD of the central galaxies has a Gaussian shape in B06 (as already noted in Kim et al., 2011), with basically no central galaxies for halo masses larger than $\sim 10^{12.5}$ $\mathrm{M}_\odot \mathrm{h}^{-1}$. In the other models the distributions corresponding to central galaxies are generally broader and extend to significantly larger halo masses than in B06. As noted above, HI in central galaxies of massive haloes is likely suppressed in B06 by efficient radio-mode feedback.

In Fig. 4.9 we show the mass function of FoF haloes hosting N galaxies with HI mass in different bins (different columns), namely the volume density (per M_{200}^{FoF} bin) of haloes hosting at least N centrals/satellites with a selected HI content. Different rows correspond to N$>$ 0, N$>$ 1 and N$>$ 10. Galaxies are divided in centrals (solid lines) and satellites (dashed lines). Note that the FoF mass function for the central galaxies is shown only in the N$>$ 0 row, as each FoF contains only one central galaxy by construction. Only a few haloes have large number of satellites containing large amounts of HI. The GAEA and XBR16 models in particular, have significantly larger number of satellite galaxies with modest to significant HI content with respect to the other models. For all other models, the distributions peak at lower halo masses. It is interesting that GAEA and XBR16 predict a significant contribution from the HI gas rich satellite population albeit assuming an instantaneous stripping of the hot gas associated with infalling galaxies. As noted above and in Hirschmann et al. (2016), this is driven by the modified stellar feedback scheme. For the intermediate and high HI mass bins considered ($[10^{9.5}; 10^{10.5}]$ $\mathrm{M}_\odot$), there are less haloes hosting large numbers of galaxies in the GAEA model than in XBR16. We remind that these two models are based on the same physical parametrizations and differ only for the explicit modelling of the partition of cold gas in atomic and molecular gas, the division of the disk in annuli, and on the explicit dependence of star formation on the molecular gas content in XBR16.

In the top row of Fig. 4.9 , haloes hosting only one central galaxy dominate the low M_{200}^{FoF} end, and outnumber the FoF with HI-selected satellites up to $M_{200}^{FoF} \sim 10^{12}$ $\mathrm{M}_\odot \mathrm{h}^{-1}$. Above this limit, HI rich satellites become the main contributors.

In Fig. 4.8, in the lowest HI mass bin ($[10^{8.5}; 10^{9.5}]$ $\mathrm{M}_\odot$), satellites outnumber centrals at $M_{200}^{FoF} \sim 10^{12}$ $\mathrm{M}_\odot \mathrm{h}^{-1}$ in the B06 and H15 models. In DLB07, G11, GAEA and XBR16, the number of satellites becomes larger than the corresponding number of centrals at $M_{200}^{FoF} \sim 10^{11.5}$ $\mathrm{M}_\odot \mathrm{h}^{-1}$. All models have more than 10 galaxies per FoF only above $M_{200}^{FoF} \sim 10^{13}$ $\mathrm{M}_\odot \mathrm{h}^{-1}$ (B06 only at $\sim 10^{14}$). The volume density of such haloes is significant only for the GAEA and XBR16 models. B06 has the lowest number of haloes with more than one

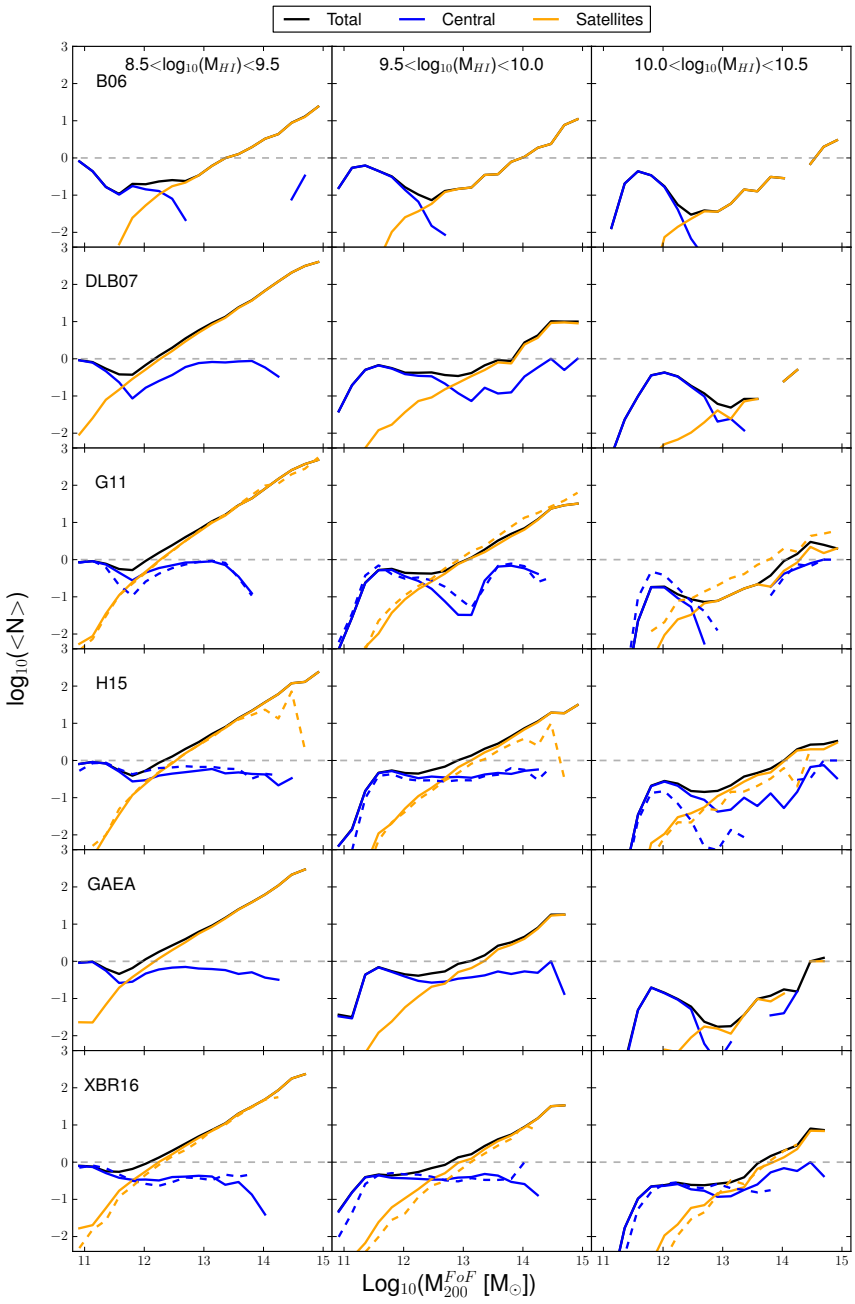

Figure 4.8: Average number of galaxies with stellar mass larger than $M_* > 10^9 \mathrm{M}_\odot$ and in different HI mass bins (different columns) for haloes of mass M_{200}^{FoF}. Different rows correspond to the different models considered in our study. The colour coding is the same as in Fig. 4.4: black for all model galaxies, blue for centrals and orange for satellites. In the rows corresponding to the G11, H15 and the XBR16 models, the dashed lines correspond to results based on the MS-II simulation.

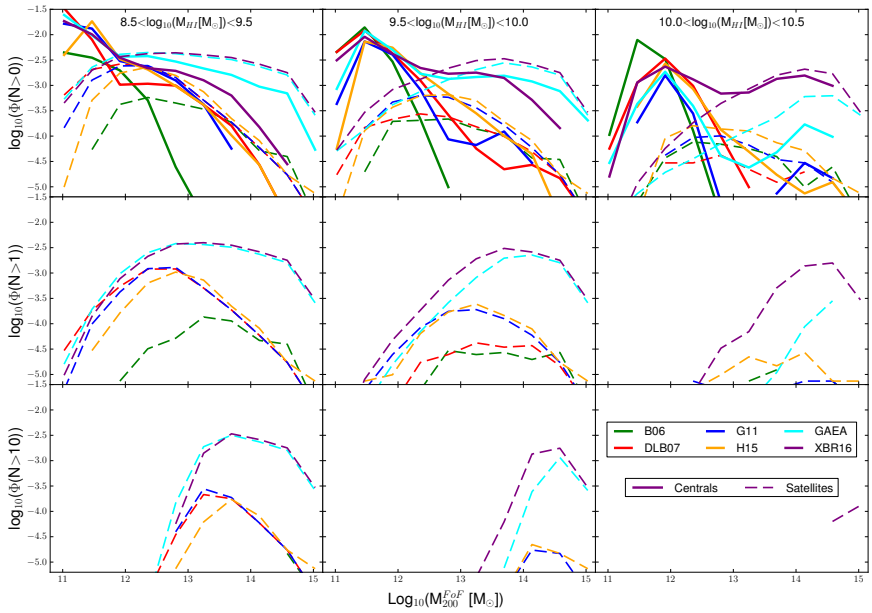

Figure 4.9: Mass function of FoF haloes hosting at least one (top panel), more than one (middle panel), and more than ten (bottom panel) galaxies with HI mass in different bins (different columns). Solid lines are used for central galaxies while long-dashed lines correspond to satellite galaxies.

satellite in all the HI mass bins, but the difference in the lowest bin is more relevant, in particular compared to the DLB07 model, which instead is aligned with G11 and H15 both in the HOD and in the mass function. This results in different 2PCFs for B06 and the other models.

For the middle HI mass bin ($[10^{9.5};\ 10^{10}]$ $M_\odot$), we always find less than 10 galaxies per halo in the B06 and DLB07 models. In all other models, haloes with mass larger than $M_{200} \sim 10^{14}$ $M_\odot h^{-1}$ host more than 10 satellite galaxies. However, as for the lowest HI mass bin, the number of these haloes is large only for the GAEA and XBR16 models. In general B06 and DLB07 have less satellite galaxies than the other models. This can explain the underestimation of the 2PCF: a lower number of satellite galaxies lowers the correlation signal. We remind that these models are characterized by a simple treatment for satellites and gas stripping (see Sec. 4.5 for details on satellite evolution), that leads to the well known problem of too many passive galaxies (Weinmann et al., 2006; Wang et al., 2007; Fontanot et al., 2009). This simplified treatment is, however, assumed also in the GAEA and XBR16 models, in which the effect of the instantaneous stripping of gas is mitigated by a different treatment of stellar feedback.

There are generally very few galaxies in the highest HI mass bin ($[10^{10};\ 10^{10.5}]$ $M_\odot$) considered. In particular, the B06, DLB07 and GAEA models always have only one galaxy per halo below halo masses $M_{200} \sim 10^{14.5}$ $M_\odot h^{-1}$. For the other models, the number of galaxies becomes larger than one at halo masses larger than $M_{200} \sim 10^{13.7}$ $M_\odot h^{-1}$. The only model with numerous satellites is XBR16.

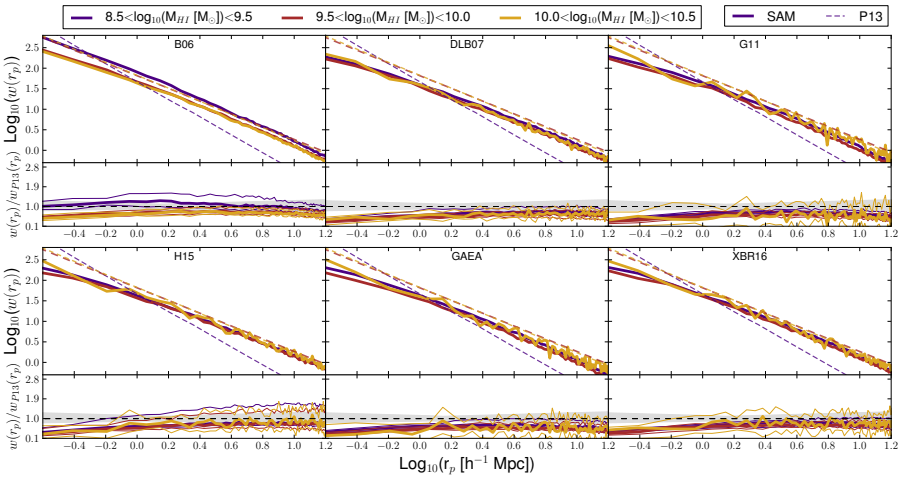

Figure 4.10: As in Fig. 4.6, but considering only central galaxies for all models.

The difference in number of satellite rich haloes does not produce an appreciable difference between the XBR16, G11 and H15 2PCFs. The smaller numbers of satellite rich halos in the GAEA model result in an underestimated 2PCF. The difference with respect to the XBR16 model can again be ascribed to the approach adopted, based on dividing the star forming disk in annuli. As commented above, this results in a lower molecular fraction to total cold gas with respect to the GAEA model, and keeps the star formation ongoing in the central regions of the disk for longer times. This implies that satellite galaxies in XBR16 have more HI left than their counterparts in the GAEA model.

4.5 The role of satellite galaxies

The results discussed above suggest that satellite galaxies play an important role in the disagreement found between model predictions and observational measurements.

As discussed in Section 4.2, all models used in our study are built using subhalo based merger trees extracted from the Millennium Simulation. Dark matter haloes are subject to significant stripping after being accreted on larger systems (e.g. De Lucia et al., 2004; Gao et al., 2004). At the resolution of the MS, substructures fall below the resolution limit when at distances from halo centre that are typically significantly larger than the separation from which galaxy mergers are expected to occur. All models used in our study assume that galaxies in disrupted subhalos survive as 'orphan' galaxies. While the specific treatment depends on the model (e.g. G11 and H15 include an explicit treatment for tidal disruption), they all assign to these galaxies a residual merger time that generally depends on the initial orbit and on the mass-ratio between the infalling system and the accreting one.

When the infalling process begins satellites undergo tidal processes that strip away a part or all their hot gas halo. In B06, DLB07, GAEA and XBR16 this

Table 4.1: Fraction of central and satellite galaxies in different HI mass bins, for all models considered in this study ($M_* > 10^9 M_\odot$).

Model	% of centrals/ satellites	% of central/satellites with $M_{HI}[M_\odot] \in$			
		$[0;\ 10^{8.5}]$	$[10^{8.5};\ 10^{9.5}]$	$[10^{9.5};\ 10^{10}]$	$[10^{10};\ 10^{10.5}]$
B06	54.3 - 45.7	7.8 - 39.5	5.6 - 3.1	20.4 - 1.4	16.8 - 0.6
DLB07	54.4 - 45.6	1.5 - 19.8	33.0 - 24.6	15.6 - 1.2	4.0 - 0.1
G11	53.3 - 46.7	1.9 - 3.7	34.0 - 37.6	14.1 - 4.8	3.2 - 0.6
H15	61.4 - 38.6	0.9 - 2.3	33.7 - 28.8	20.2 - 6.3	5.7 - 1.0
GAEA	59.3 - 40.7	1.0 - 7.2	39.7 - 30.9	16.8 - 2.6	1.8 - 0.1
XBR16	57.0 - 43.0	4.7 - 12.7	35.0 - 25.8	15.0 - 4.1	2.3 - 0.4

process is instantaneous, and all the hot gas is stripped away (in B06, all the gas outside the dynamical radius of the halo) at the infall time (when the galaxy becomes a satellite). In G11 and H15, the stripping is gradual, and the hot gas that remains associated with satellite galaxies can cool providing fresh material for star formation.

The effect of satellite galaxies on the predicted correlation function of HI selected galaxies is shown in Fig. 4.10. In this case, we are considering only central galaxies, i.e. we are excluding from galaxy catalogues both 'orphan' galaxies and satellite galaxies associated with distinct dark matter substructures. In all models, the clustering signal becomes weaker, with a shift dependent on the number of the satellites in each HI mass bin. For the B06 model, the 2PCF remains the same as if satellites are included. This is expected as Fig. 4.4 shows that satellite galaxies in this model are typically HI poor. Also for the DLB07 model, small differences are found between the clustering signal predicted including and excluding satellite galaxies in the medium-high HI bins ($[10^{9.5};\ 10^{10.5}]\ M_\odot$). Again, this is expected because satellite galaxies are HI poor in this model, although the effect is not as strong as for the B06 model.

In Tab. 4.1 the fraction of satellites in each HI mass bin is listed for each model considered. Satellites account for $39-46\%$ of the total, depending on the model. The majority are in the lowest HI bin ($[10^{8.5};\ 10^{9.5}]\ M_\odot$) considered in this analysis, with the exception of the B06 model where satellites have typically lower HI masses.

Taking advantage of model results, we can also quantify the evolution of the HI content in satellite galaxies. To this aim, we have selected all satellite galaxies at $z = 0$ and followed back in time their main progenitors until they become central galaxies. In Fig. 4.11, we show the HI mass that satellite galaxies have at the last time they are centrals, as a function of lookback time. Galaxies are split in different bins according to the HI mass at present. We plotted as vertical dotted lines the median times of accretion for each HI bin. The figure shows that, for the two bins corresponding to the largest HI masses, accretion times tend to be lower than for galaxies with lower HI mass. In other words, the HI richest galaxies tend to be accreted later. This is particularly significant for the B06 and DLB07 models that are characterized by the most rapid consumption of the cold gas in satellites. The figure also shows that for some models (e.g. B06, DLB7, GAEA) the slope of the lines tend to be steeper than for the other

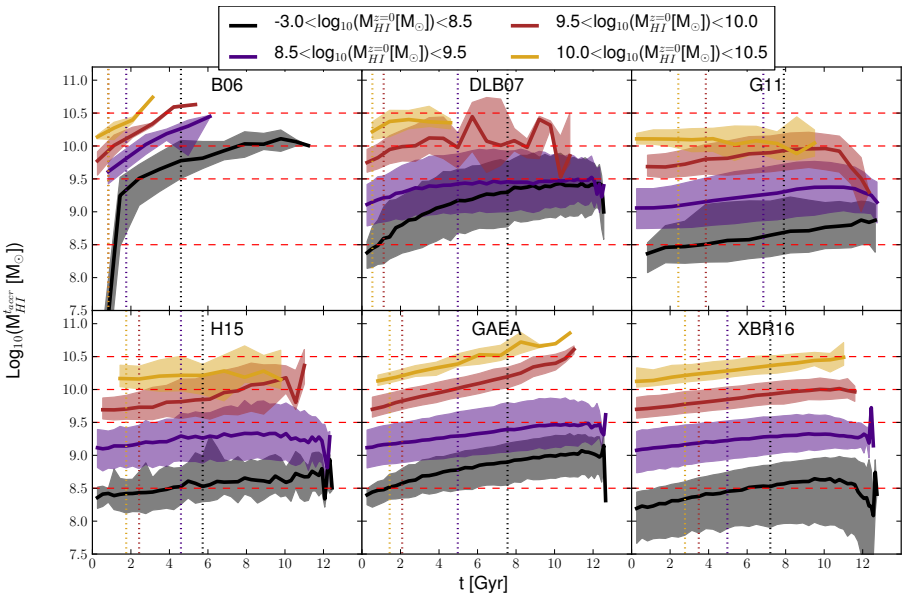

Figure 4.11: The HI content of satellites at the time of accretion vs the look-back time of accretion. Galaxies are divided according to their final HI content, with different colours representing different HI mass bins. The solid lines correspond to the median of the distributions, while the shaded areas show the one sigma scatter. The vertical dotted lines are the median times of accretion of model galaxies in each HI bin (colour coded as above). The red horizontal dashed lines correspond to the limits of the HI mass bins considered, and are plotted as a reference.

models, indicating a more rapid depletion of the HI content of model galaxies.

A more direct way to quantify gas depletion in satellite galaxies is to choose a specific redshift of accretion and consider the average evolution of their HI content down to present time. Results of this analysis are shown in Fig. 4.12 for satellites accreted at $z \sim 1$. Different colours correspond to different HI masses at the time of accretion. The B06 model is characterized by the most rapid depletion rate. Also satellites in the DLB07 model consume their gas rapidly but they tend to flatten when the average HI content in satellites reaches a value $M_{HI} < 10^{8.7}$ $M_\odot$. The figure also clearly shows the different satellite treatment in G11 and H15 compared to GAEA and XBR16: the latter are characterized by a faster depletion rate soon after accretion. In DLB07, G11, H15, and GAEA the lines tend to flatten after reaching some value. This is due to fact that these models assume a critical surface density of gas for star formation. In the XBR16 model, the flattening is less evident. This model does not assume an explicit threshold for star formation but, as discussed above, evaluates the star formation rates in different annuli after estimating the amount of molecular gas available using the (Blitz and Rosolowsky, 2006) empirical relation.

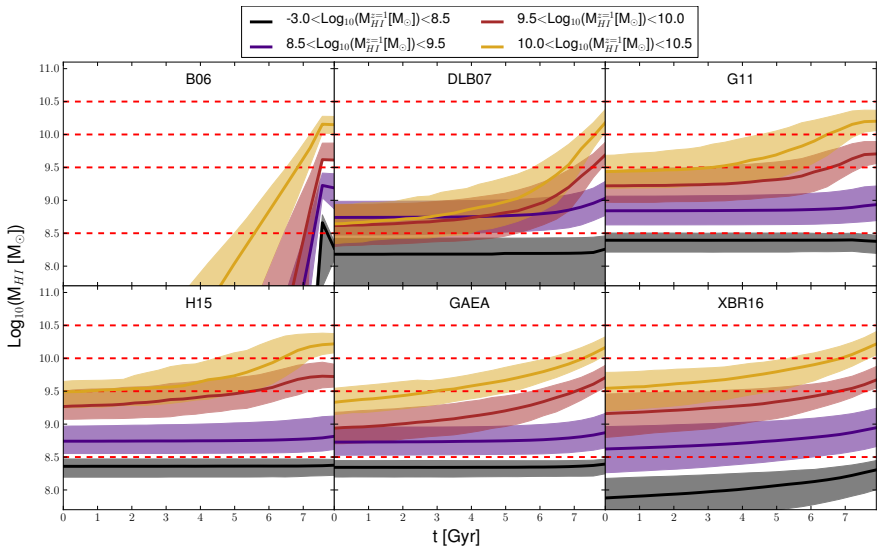

Figure 4.12: The HI content of satellites accreted at redshift 1 vs the look-back time. Galaxies are divided according to their initial HI content, with different colours representing different HI mass bins at z=1. The solid lines correspond to the median of the distributions, while the shaded areas show the one sigma scatter. The red dashed lines correspond to the limits of the HI mass bins considered, and are plotted as a reference.

4.6 Relations with the dark matter halo

In this section we analyze the relation between HI-selected galaxies and the hosting dark matter haloes. We take advantage of our knowledge of the halo mass to characterize the HI-M_{halo} relation in Sec. 4.6.1, and we analyze the dependence of HI content on the dark matter halo spin in Sec. 4.6.2.

In the following, for satellite galaxies, we will consider the mass and spin of the parent dark matter halo corresponding to the last time the galaxy was central.

4.6.1 HI galaxy content and maximum halo mass

P13 used the measured 2PCF for the HI-selected galaxies to estimate the shape and the scatter of the HI-M_{halo} relation. They took advantage of the Bolshoi dark matter only simulation (Klypin et al., 2011), based on a WMAP7 cosmology, and populated haloes and subhalos with HI through abundance matching. In their analysis, they linked the current (z=0) HI content of each subhalo, $M_{HI}^{z=0}$, to the maximum value of the subhalo mass during its past evolution, M_{200}^{max} (this corresponds with good approximation to the time just before a halo is accreted on a larger structure). The implied assumption is that the HI attached to subhalos at their maximum mass does not change too much down to redshift 0. They note, but do not discuss further, that while this assumption can be valid for stellar masses, it is generally not a good one for the HI masses, because of ram-pressure stripping and gas consumption due to quiescent star

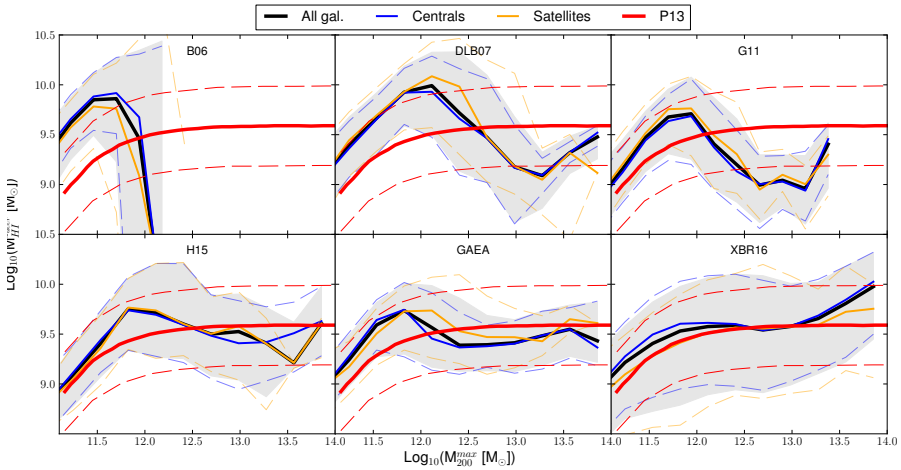

Figure 4.13: The HI-halo mass relation for the models considered in this study, for central (blue), satellite (orange) and all (black) galaxies. The relation inferred by P13 is shown for comparison (red). The HI mass shown on the y axis is that measured at the time the parent halo mass peaks (i.e. just before infall). Solid lines show the median values, shaded region corresponds to the 1σ scatter for all galaxies, while dashed lines correspond to the 1σ scatter for centrals and satellites.

formation in satellite galaxies.

In our work, we take advantage of our knowledge of the HI content of model galaxies both at the time the halo mass was maximum (we assume this corresponds to the last time the galaxy was central), and at redshift 0. We can thus verify the influence of the different satellite treatments on the predicted relation between HI mass and halo mass.

In Fig. 4.13 we show the relation between the HI mass and the halo mass at the last time the galaxy was central. This figure represents what we would find if we suppose the HI content of galaxies does not change after accretion, as done in P13. Solid lines show the median relation for all galaxies (black), centrals (blue) and today satellites (orange). The 1σ spread is shown by the shaded area (for all galaxies) and the dashed lines (for central and satellites). The relation obtained in P13 is over-plotted in red (solid line is the median and dashed lines show the 1σ spread) for a direct comparison. In this figure the relation obtained for satellite galaxies is very close to that obtained for centrals, with some small differences due to redshift evolution of the HI content of galaxies at the time of accretion (see Fig. 4.11).

As already noted for the HI-M_* relation (see Sec. 4.3.3), B06 fails to reproduce the HI content in medium to high mass haloes because satellites are generally too gas poor and central galaxies are depleted of their cold gas by efficient radio mode feedback. The other models exhibit a wide scatter in the relation that is almost independent of $M_{200}^{\rm max}$, with values of $\sigma_{\rm HI} \sim 0.3 - 0.5$ dex depending on the model. The shape is qualitatively similar to the one inferred in P13 if we consider H15, GAEA and XBR16, while the other models predict more HI for $M_{200}^{\rm max} \lesssim 10^{12}$ M$_\odot$, and less HI than that inferred by P13 in the

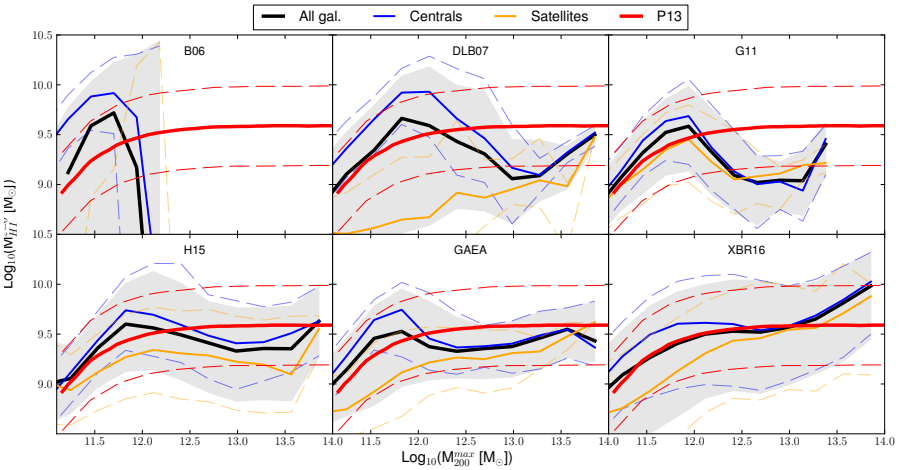

Figure 4.14: As in Fig. 4.13, but the HI mass shown on the y axis is that measured from the models at redshift 0.

most massive haloes.

In Fig. 4.14 we take into account the evolution of HI in satellites, that is we used for all galaxies their HI content at redshift 0, while $M_{200}^{\rm max}$ corresponds again to the parent halo mass at the last time the galaxy was central. In this figure we can appreciate a down-shift of 0.1-0.4 dex of the relation for $M_{200}^{\rm max} < 10^{12.5-13}\ M_{\odot}$. The shift is driven by the satellites with lowest masses (both halo and HI), whose HI is depleted after accretion because consumed through star formation. The gas depletion increases the scatter of the total relation, with the effect being enhanced for B06, DLB07 and GAEA.

As noted in Sec. 4.5, the lowest HI bin considered ($[10^{8.5};\ 10^{9.5}]$ $M_{\odot}$) is dominated by satellite galaxies, and therefore the one most affected by a different treatment for the evolution of these galaxies.

4.6.2 HI galaxy content and halo spin

A relatively tight correlation is expected between the values of the halo spin and the gas content of galaxies. Huang et al. (2012) studied this relation for the ALFALFA HI-selected galaxies. The spin of the dark matter halo was calculated using the λ estimator proposed by Hernandez et al. (2007), assuming a dark matter isothermal density profile, an exponential surface density profile for the stellar disk, and a flat disk rotation curve. Using these assumptions: $\lambda \propto r_{disk}/V_{rot}$, thus the spin depends on the disk scale radius and the rotational velocity. Huang et al. (2012) found that ALFALFA HI rich galaxies favour high-λ values. Kim and Lee (2013) simulated the evolution of dwarf-size haloes with varying halo-spin parameters and initial baryon fractions, and found a correlation between disk radius (and therefore gas mass) and λ.

P13 used their measurements of the 2PCF to infer a relation between halo spin and gas content. They divided the haloes of the Bolshoi simulation into three conveniently chosen spin bins: low spin with $\lambda \in [0.002; 0.02]$, medium spin with $\lambda \in [0.02; 0.05]$ and high spin, with $\lambda \in [0.05; 0.20]$. Assigning an HI

content to each halo according to the relation presented in Sec. 4.6.1, they found that the high and medium spin bins have almost the same 2PCFs measured for the HI selected galaxies, while low spin haloes have larger clustering signal. Based on these results, they argued that halo spin is the main driver of the HI content of galaxies. We have verified that selecting halos in the same range of spin values used by P13, we find results consistent with theirs.

Since the models we have used in our study are coupled with a high-resolution cosmological simulation (MS), we can explicitly analyze the correlation between the HI content of model galaxies and the spin of their parent haloes. For the B06 model, information about the spin is not available, so we exclude this model from the following analysis. For central galaxies, we just use the spin of the parent halo. For satellite galaxies, we consider the spin measured for the parent halo at the last time they were central galaxies.

We show the histograms of the spin for some chosen $M_{200}^{\rm max}$ (different columns) and for the usual ranges in HI content in Fig. 4.15. Dashed lines correspond to the distributions obtained for each HI mass, independently of $M_{200}^{\rm max}$, in order the easily compare each $M_{200}^{\rm max}$ bin to the total. The red vertical lines correspond to the bins used in P13.

The majority of the galaxies with the lowest HI mass ($[10^{8.5};\ 10^{9.5}]$ $M_{\odot}$) are found in haloes with low spin values ($\lambda \in [0.002; 0.02]$), while the majority of galaxies in the two HI richer bins ($[10^{9.5};\ 10^{10.5}]$ $M_{\odot}$) correspond to mid-spin values ($\lambda \in [0.02; 0.05]$). Interestingly, this distribution is very similar for all models, i.e. the HI-spin correlation/distribution is not dependent on the specific prescriptions of each model. This result can be understood as follows: the halo spin parameter is used to set the initial value of the disk scale-length, and this quantity is then used to compute a density threshold for star formation. In the case of a low spin halo, the initial radius will be small, and the surface density will be large. The star formation rate is directly dependent on the surface density and thus, for a fixed value of HI, a smaller radius will correspond to a faster HI depletion. This explains why HI rich galaxies are rare in haloes with low spin values, and why high spin (large disk radius) haloes host galaxies with a broad range of HI masses.

The figure shows that, in all models considered in our study, there is no tight correlation between the HI content of galaxies and the spin of their parent haloes: low spin haloes are more likely populated by galaxies with low HI mass, and HI rich galaxies are most likely hosted by haloes with large spin values, but high spin haloes are populated by galaxies in a wide range of HI mass.

4.7 Conclusions

In this work, we study the basic statistical properties of HI selected galaxies extracted from semi-analytic models of galaxy formation, and compare theoretical predictions with available data. In particular, we use four models whose galaxy catalogues are publicly available (Bower et al., 2006; De Lucia and Blaizot, 2007; Guo et al., 2011; Henriques et al., 2015), and two models recently developed by our group (Hirschmann et al., 2016; Xie et al., 2017). Only one of these models (that described in Xie et al., 2017) includes an explicit modelling for the partition of cold gas into atomic and molecular components, and a molecular hydrogen based star formation law. All models are run on the same cosmological

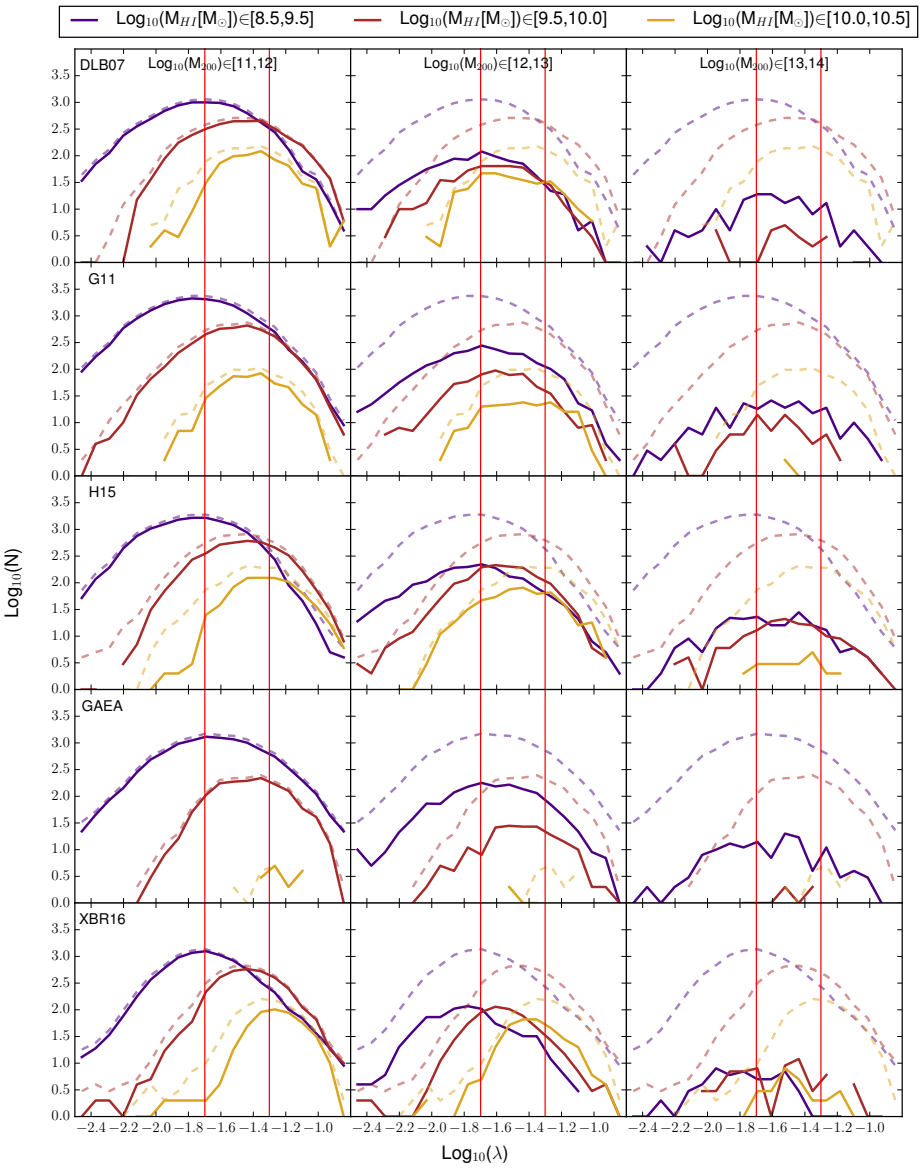

Figure 4.15: Distribution of halo spin values for galaxies selected in haloes with different $M_{200}^{\max}$ (columns) and HI mass (same colours of Fig. 4.6), for all models used in our study (different rows). The solid lines represent the distribution for each selection, while dashed lines are the total distribution over all the $M_{200}^{\max}$ range, reproduced in every panel as reference. The vertical red lines show the spin division considered by P13, and are plotted as reference.

simulation, the Millennium Simulation (Springel et al., 2005).

For those model in which the cold gas is treated as a single star forming phase, we estimate the HI content of model galaxies *a posteriori* assuming (as in Obreschkow and Rawlings, 2009) that the cold gas is distributed in an exponential disk and that the ratio between molecular and atomic hydrogen is determined by the kinetic gas pressure. All models used in our study include different specific modelling for the various physical processes considered and, in particular, for the evolution of satellite galaxies. We find this to have relevant consequences on model predictions for HI selected galaxies.

All models considered are in relatively good agreement with the observed local HI mass function, with the exception of the model by Bower et al. (2006) that predict too many galaxies with intermediate to large HI mass and too few galaxies with small HI content ($M_{HI} < 10^{9.6}\mathrm{M}_{\odot}$). We find this is due to excessive HI masses in low stellar mass galaxies. This particular model also fails to reproduce the observed scaling relations as it predicts very little HI associated with satellites (likely because of instantaneous stripping of hot gas and efficient stellar feedback), and central galaxies with stellar masses larger than $\sim 3\times10^{10}\,\mathrm{M}_{\odot}$ (because of too efficient radio mode feedback). The observed scaling relations are relatively well reproduced by all other models, but some of them (Guo et al., 2011; Hirschmann et al., 2016; Xie et al., 2017) exhibit a deficit of HI rich galaxies ($M_{HI} > 10^{10}\mathrm{M}_{\odot}$) at any stellar mass.

The HI content of satellite galaxies varies significantly among the models considered, due to a different treatment for the hot gas associated with infalling galaxies and stellar feedback. As expected, models that assume a non instantaneous stripping of this hot gas reservoir tend to predict larger HI masses for satellite galaxies. The most massive satellites, in particular, tend to have an average HI content that is very close to that of central galaxies of the same stellar mass. This is due to the fact that these galaxies were accreted relatively recently and evolved as central galaxies for most of their lifetime (De Lucia et al., 2012b). Interestingly, assuming an instantaneous stripping of the hot gas at the time of accretion, as in GAEA, does not necessarily imply low HI content for satellite galaxies. As already noted in Hirschmann et al. (2016), this model is characterized by significantly lower fractions of passive (and therefore larger fractions of gas-rich, star forming) satellites with respect to e.g. the model presented in De Lucia and Blaizot (2007). This is a consequence of suppressed and delayed star formation at early times, and leads to larger cold gas fractions at the time of accretion.

Using galaxy catalogues from each model, we have built mock light-cones that we have used to analyze how the clustering of HI selected galaxies compares to recent measurements by Papastergis et al. (2013). In particular, we have considered three HI mass bins: low ($M_{\mathrm{HI}} \in [10^{8.5};\ 10^{9.5}]\ \mathrm{M}_{\odot}$), intermediate ($M_{\mathrm{HI}} \in [10^{9.5};\ 10^{10.}]\ \mathrm{M}_{\odot}$) and high ($M_{\mathrm{HI}} \in [10^{10};\ 10^{10.5}]\ \mathrm{M}_{\odot}$). The lowest HI mass bin is likely affected by limited sampling volume in the observations so, although the data suggest a lower clustering signal for this particular bin, Papastergis et al. (2013) argue that this is not significant and that the 2-point clustering function measured for these three bins are not statistically different. In contrast, we find that all models predict for galaxies in the lowest HI bin a clustering signal *higher* than for the HI richer galaxies. For the other two bins ($[10^{9.5};\ 10^{10.5}]\ \mathrm{M}_{\odot}$), half of the models are in relatively good agreement with data while the other half tend to under-predict slightly the measured clustering.

Interestingly, the model by Bower et al. (2006) that has the worst performance for the HI mass distribution and scaling relations, exhibits the best agreement with data by Papastergis et al. (2013) in the lowest HI bin, with a clustering signal only slightly stronger than that for HI richer galaxies. We show that the lowest HI bin is strongly affected by the adopted treatment for satellite galaxies as this bin is dominated by this galaxy population.

The relation between the HI mass (at the accretion time for satellite galaxies) and halo mass (at its maximum) predicted by all models considered is in quite good agreement with that inferred by Papastergis et al. (2013). Again, the exception is the model by Bower et al. (2006) that predicts negligible HI in haloes more massive than $\sim 10^{12}\,\mathrm{M}_{\odot}$. This is interesting because, as noted above, the clustering signal predicted by this particular model is the one that is closest to observational measurements. Thus, taken at face value, these results suggest that the clustering of HI selected galaxies does not provide enough information to constrain the relation between halo mass and HI mass, and that a crucial element is represented by the evolution of the HI content of satellite galaxies. The scatter of the predicted relation increases in case one considers the HI associated with galaxies at present time, because of gas depletion in satellite galaxies. Specifically, we find that the Bower et al. (2006) model exhibits the shortest gas consumption times: a galaxy accreted at $z \sim 1$ with HI mass $\sim 10^{10}\,\mathrm{M}_{\odot}$ conserves only about 10 per cent of this gas after 2 Gyr (1 per cent after 3 Gyr). This is likely due to an efficient stellar feedback, coupled with an instantaneous stripping of the hot gas associated with infalling galaxies. The gas consumption timescale is typically longer in the other models, and the assumption of a gas density threshold for star formation implies that the gas associated with satellite galaxies never falls below such limit. The gas consumption timescales are longest in the Xie et al. (2017) model where star formation is evaluated in radial annuli, considering the local physical conditions of the inter-stellar medium.

Finally, we have examined the relation between the HI content of galaxies and the spin of the parent dark matter halo (we have considered the value at the time of accretion for satellite galaxies). We find that low spin haloes ($\lambda \in [0.002; 0.02]$) are more likely populated by HI poor galaxies ($[10^{8.5};\ 10^{9.5}]\,\mathrm{M}_{\odot}$), and HI richer galaxies tend to reside in haloes with large spin ($\lambda \in [0.02; 0.2]$). The scatter, however, is relatively large and haloes with intermediate and large spin values tend to host galaxies with a large dynamic range in HI mass. Interestingly, the distributions are very similar in all models considered, i.e. they are not significantly affected by the specific modelling of the various physical processes affecting the HI content of galaxies. This is somewhat surprising as the spin enters the calculation of the disk radius and this, in turn, affects the star formation rate (and therefore the gas content). In most models, however, the halo spin is only used to determine the initial radius of the gaseous component and does not affect significantly the subsequent evolution. The dependence of the gas initial disk radius on halo spin explains the trends found: an initial small radius and a large gas fraction translates into a high surface density and therefore into a large star formation rate, which consumes the gas rapidly. This explains why haloes with small spin values tend to be associated with gas poor galaxies.

Our analysis shows that different models lead to very similar results for galaxies with intermediate to large HI mass ($[10^{9.5};\ 10^{10.5}]\,\mathrm{M}_{\odot}$), while signifi-

cant differences can be found for relatively HI poor galaxies ($[10^{8.5};\ 10^{9.5}]$ $M_{\odot}$). As discussed above, this bin is dominated by satellite galaxies and therefore mostly affected by the different treatment for this particular galaxy population. More detailed data in this particular HI mass range are needed to put stronger constraints on galaxy formation models. Dedicated controlled simulations would be useful to quantify the effect of stripping processes in satellites (see e.g. Tonnesen and Bryan, 2009). Some attempts in this direction were made using semi-analytic models based on Monte Carlo merger trees (Lagos et al., 2011a; Kim et al., 2015) or on the Millennium Simulation II (Kim et al., 2017), and have shown that these models can be used to constrain the physics of low mass satellites.

At the same time, we need larger statistical samples and the possibility to estimate the galaxy 'hierarchy' (i.e. being a satellite or a central) in observations. New radio instruments, such as SKA[4] and its precursors (Johnston et al., 2008; Booth et al., 2009, for ASKAP and MeerKAT, respectively), will provide valuable data for these analyses.

[4]https://www.skatelescope.org/project/

Chapter 5

Sizes and Specific Angular Momenta: Dynamical Properties in Semi-Analytic Models

This chapter is based on a paper in preparation, soon to be submitted.

5.1 Introduction

The history of a galaxy (and of its components) is determined by a network of physical processes, that drive a complex exchange of mass, energy, metals and dynamics. A detailed description of these processes is provided in Chap. 1. When a galaxy forms from the collapse of baryons in the parent Dark Matter halo potential, its hot gas acquires a specific angular momentum that is strictly correlated to that of the halo (van den Bosch et al., 2002, 2003; Sharma and Steinmetz, 2005). When the hot gas cools, its specific angular momentum is transferred to the cold gas, which leads to the formation of a rotationally supported gas disk. The dynamical state of the cold gas is transferred, in turn, to the stars formed therein, creating a rotationally supported stellar disk. Disk rotation makes it an extended structure, with a size proportional to its specific angular momentum. Bulges are typically supported by dispersion velocity, but, in some cases, rotation is measured. In fact, observationally, bulges can be divided in two different categories: classical spheroidal bulges, that are believed to originate from mergers, and pseudo-bulges, that tend to be thicker, disky and rotation supported, and that are believed to originate from internal secular processes. The two bulge categories have different dynamical properties and sizes (Fisher and Drory, 2008; Gadotti, 2009).

As explained above, the sizes and angular momenta of galaxies are determined by a combination of different physical processes. The specific assembly history of a galaxy determines the conservation (or consumption) of the angular momentum acquired from secular evolution (see, for example, hydrodynamical simulations by Zavala et al., 2008). And, in turn, the size and the specific

angular momentum of the galaxy represent a record of the galaxy formation and assembly history.

The processes driving the dynamical evolution of galaxies have long been studied (Fall and Efstathiou, 1980; Dalcanton et al., 1997; Mo et al., 1998), and the progress in observational techniques has recently allowed a systematic study of size and dynamics for large samples of galaxies. Good resolution imaging in several photometric bands became available for large galaxy samples from projects like SDSS or GAMA (York et al., 2000; Driver et al., 2011) in the local Universe, and, more recently, up to high redshifts with e.g. CANDLES (Grogin et al., 2011). The advent of Integral Field Spectroscopic surveys allowed the measurement of resolved properties for thousands of galaxies, starting with the pioneering work done with SAURON (Bacon et al., 2001) culminating in ongoing projects like CALIFA, SAMI, and MaNGA (Sánchez et al., 2012; Bryant et al., 2015; Bundy et al., 2015).

The size-mass relation was first analyzed for a statistical sample of galaxies in Shen et al. (2003), using SDSS data. This relation presents a large scatter, that depends on the morphology of the galaxies. Late Type (LT) galaxies have, on average, larger radii than Early Types (ETs). This result was confirmed by subsequent studies, based on different methods to estimate the size and different ET/LT selections (e.g. GAMA, Lange et al., 2015), in different environments (e.g. Weinmann et al., 2009; Huertas-Company et al., 2013, both based on SDSS), and for galaxies at higher redshifts (Ichikawa et al., 2012; van der Wel et al., 2014). The size of the galaxies decreases with increasing redshift, for both LT and ET galaxies. Finally, the median size-mass relation is only slightly dependent on the method used to select LT/ET galaxies, as demonstrated in Lange et al. (2015).

The first analysis of the relation between the specific angular momentum (j_*) and galaxy stellar mass (M_*) was performed in Fall (1983), for a small sample of galaxies of various morphological types. This study was extended in the work of Romanowsky and Fall (2012), using a larger sample, with spatially resolved dynamical information available to large aperture radii. With integral field spectroscopic surveys, such as ATLAS^{3D} (Cappellari et al., 2011) or SAMI (Bryant et al., 2015), the velocity maps of hundreds of galaxies became available. In ATLAS^{3D}, ET galaxies are classified, according to their rotational component, as fast and slow rotators, demonstrating that ET galaxies exhibit a complex internal dynamics, manifestation of different galaxy assembly histories. Using data from SAMI, Cortese et al. (2016) have provided an estimate of the j_*-M$_*$ relation both for LT and ET galaxies, obtaining results similar to those of Romanowsky and Fall (2012). A similar analysis, performed on a limited sample of spiral, gas rich, galaxies (THINGS, Walter et al., 2008), was performed by Obreschkow and Glazebrook (2014). They confirmed previous results, and showed that their galaxies lie on a plane in the 3D space described by specific angular momentum, mass, and bulge over total mass ratio. Further analysis of the dynamical properties of large samples of galaxies will be possible with data from MaNGA, an integral field spectroscopic survey of SDSS galaxies (Bundy et al., 2015).

The observational studies carried out in recent years have shown that the scatter in the relation between the specific angular momentum and the mass depends on galaxy morphology, as in the case of the size-mass relation. Furthermore, the slope of the relation is consistent with that predicted from theory

(Fall and Efstathiou, 1980; Dalcanton et al., 1997; Mo et al., 1998).

Recent studies have shown that hydrodynamical simulations are able to reproduce the measured correlation between specific angular momenta and galaxy masses. This relation was analyzed against different morphological selections and wind schemes (Übler et al., 2014; Teklu et al., 2015; Genel et al., 2015; Zavala et al., 2016; DeFelippis et al., 2017), and simulations have shown that strong winds reflect in an increased angular momentum, and prevent angular momentum loss. High resolution simulations of merging galaxies have also demonstrated the importance of cold gas dissipation in determining the final size of the merger remnants. In particular, these simulations show that wet mergers tend to form smaller bulges (Hopkins et al., 2009b, 2014; Porter et al., 2014). This dependence has been implemented in the framework of semi-analytic models by Shankar et al. (2013), who have studied the ability of a previous published model to reproduce the size-mass relation measured for ET galaxies.

Several semi-analytic models include a treatment for the exchange of angular momentum among galactic components (e.g. Lagos et al., 2009; Guo et al., 2011; Benson, 2012; Padilla et al., 2014; Tonini et al., 2016), and this implementation is used to infer other properties, such as the disk size. Some studies have focused on the dynamical properties of model galaxies, but they either were based on post-processing implementations (Lagos et al., 2015b), or on dedicated models specifically aimed at reproducing these properties (Stevens et al., 2016).

In this paper, we will take advantage of the state-of-the-art model described in Xie et al. (2017), to perform a systematic analysis of the size and specific angular momentum distributions of model galaxies. This model includes a treatment for specific angular momentum evolution of the gaseous and stellar disks; disk radii are estimated from disk dynamics; and bulge sizes are calculated from energy conservation. We include a treatment for gas dissipation during mergers, and analyze the size and total stellar specific angular momentum versus mass relations, and their scatter. We analyze the origin of these relations, and their dependence on the total amount of cold gas and on galaxy morphology. In addition, we also study how the sizes and angular momenta of model galaxies, as well as the predicted scaling relations, are affected by a modified treatment of gas cooling during the rapid cooling regime, and stellar feedback.

In Section 5.2, we describe the semi-analytic model we use in this work, the dark matter simulations used to build the halo merger tree, and the main prescriptions of the models, with particular emphasis on those influencing sizes and angular momenta. In Section 5.3, we describe the size-mass relation obtained from the model and its variants, and we compare it to observations both for ET and LT galaxies. We analyze the stellar specific angular momentum in Section 5.4, where we analyze its relation with stellar mass and the dependence on galaxy morphology and cold gas content. In Section 5.5, we analyze in detail the evolution of galaxies, and the dependence on morphology, cold gas content and the effect of a different implementation for gas cooling and stellar feedback. In Section 5.6, we discuss the main results of this work.

5.2 The model

In this work, we take advantage of an updated version of the GAlaxy Evolution and Assembly (GAEA) model, described in Hirschmann et al. (2016), as updated

in Xie et al. (2017). A detailed description of all the prescriptions included in this model is provided in Chap. 3.

In the following we provide a brief description of the Xie et al. (2017) model, focusing on those prescriptions that are more relevant for this work. Although many details have been provided earlier in this Thesis, we find useful to re-iterate on these aspects that are most relevant for this analysis. For a detailed description of the model, we refer to the original papers by Hirschmann et al. (2016) and Xie et al. (2017). In the following, we refer to the Xie et al. (2017) model, based on the Blitz and Rosolowsky (2006) prescription for the HI-H_2 partition of cold gas, as X17.

5.2.1 The cosmological simulation and the merger tree

The merger trees used in this work are based on the Millennium Simulation (MRI, Springel et al., 2005), and on its higher resolution companion, the Millennium II Simulation (MRII, Boylan-Kolchin et al., 2009). Both are described in detail in Sec. 3.1. The resolution limits of these simulations translate in stellar mass limits for the X17 model of about $M_{*,lim}^{MRI} \sim 10^9$ $M_\odot$ for the MRI, and $M_{*,lim}^{MRII} \sim 10^8$ $M_\odot$ for the MRII.

5.2.2 The fiducial semi-analytic model

Halo merger trees from dark matter only simulations are used as input for the semi-analytic model. This assigns baryonic components to the simulated halos, accounting for their merger histories and for the different physical processes at play. In this section, we briefly describe the processes driving the evolution of sizes and angular momenta, the main subjects of this work.

Cooling

When a halo collapses, we assign to it a hot gas component, whose mass is $M_{hot} = f_b M_{200}$ (f_b is the Universal baryon fraction, and M_{200} is defined as the mass enclosing an over-density of 200 times the critical density of the Universe). We assume that the hot gas follows an isothermal distribution. In our model, this hot gas can cool only onto central galaxies. The process is modeled as described in detail in De Lucia et al. (2010), and following the original prescriptions suggested in White and Frenk (1991b): a cooling radius is defined as the radius at which the local cooling time is equal to the halo dynamical time. Two different gas cooling regimes are considered, depending on how the cooling radius compares to the virial radius R_{200} (the radius corresponding to M_{200}). At high redshift and for small halos, the formal cooling radius is much larger than the virial radius. In this case, the infalling gas is not expected to reach hydrostatic equilibrium. Gas accretion is anisotropic (filamentary) and limited by the infall rate. In this "rapid cooling regime" or "cold mode", we assume that all the hot gas available cools in one time step. When the cooling radius is smaller than the halo virial radius, the hot gas is assumed to reach hydrostatic equilibrium and to cool quasi-statically. In this "slow cooling regime" or "hot mode", the cooling rate is modeled by a simple inflow equation.

In both regimes, the hot gas transfers angular momentum to the cold gas disk, proportionally to the cooled mass, $M_{cooling}$. As in previously published

models (Guo et al., 2011), the hot halo is assumed to have the same specific angular momentum of the dark matter halo $\vec{j}_{DM}$, so that the specific angular momentum of the cold gas, after cooling, can be written as:

$$\vec{j}^{f}_{cold} = \frac{\vec{j}^{0}_{cold} M^{0}_{cold} + \alpha_A \vec{j}_{DM} M_{cooling}}{M^{0}_{cold} + M_{cooling}}. \tag{5.1}$$

$\vec{j}^{f}_{cold}$ and $\vec{j}^{0}_{cold}$ are the specific angular momenta of the cold gas after and before gas cooling, M^{0}_{cold} is the mass of the cold gas disk, and α_A is assumed to be 1 in our fiducial model. Several recent studies based on hydrodynamical simulations (Stewart et al., 2011; Pichon et al., 2011; Danovich et al., 2015) have shown, however, that gas accreted through cold mode carries an angular momentum from 2 to 4 times the one of the DM halo. To quantify the influence of this effect on our model results, we have carried out a test run assuming for cold accretion $\alpha_A = 3$. We will refer to this run as X17CA3 in the following.

Star formation and stellar feedback

Xie et al. (2017) introduce modified prescriptions for the star formation law, accounting for recent observational results on the dependence of star formation on the molecular gas content (Wong and Blitz, 2002; Kennicutt et al., 2007; Leroy et al., 2008). In the updated model, the total cold gas reservoir associated with each galaxy is partitioned into its molecular and atomic components, and the star formation rate depends on the molecular density in the galaxy disk. In X17, we adopt prescriptions based on the Blitz and Rosolowsky (2006) empirical relation, as this model appears to be in better agreement with the observational measurements considered in Xie et al. (2017).

We assume that the newly formed stars, $M_{*,new}$, carry the angular momentum of the cold gas they originated from. The specific angular momentum of the stellar disk, after a star formation episode, $\vec{j}^{f}_{*,disk}$, can be written as:

$$\vec{j}^{f}_{*,disk} = \frac{\vec{j}^{0}_{*,disk} M^{0}_{*,disk} + \vec{j}_{cold} M_{*,new}}{M^{0}_{*,disk} + M_{*,new}}, \tag{5.2}$$

with $\vec{j}^{0}_{*,disk}$ and $M^{0}_{*,disk}$ the initial specific angular momentum and stellar mass of the disk before star formation, and $\vec{j}_{cold}$ the specific angular momentum of the cold gas disk.

We assume that reheating and/or ejection does not affect the specific angular momentum of the cold gas, and that of the hot gas (that is always equal to that of the parent dark matter halo). The ejected gas is stored in a reservoir, from where it can be re-accreted onto the hot gas associated with the parent halo, on a time-scale that depends on the virial mass of the halo.

Gas recycled from stars is directly returned to the cold gas, carrying the specific angular momentum of the stellar disk, or, in the case of gas originating from bulge stars, a zero specific angular momentum. The stellar feedback model adopted in our reference X17 model is described in detail in Hirschmann et al. (2016), and corresponds to that including the outflow parametrization based on the FIRE hydrodynamical simulations (Hopkins et al., 2014; Muratov et al., 2015).

It is well known that stellar feedback is a key ingredient for simulating thin disks with sizes and rotation properties comparable to those of our own Milky Way. Recent work has focused on angular momentum (Übler et al., 2014; Genel et al., 2015), and evidenced that gas ejected through winds during stellar feedback is accelerated so that a larger angular momentum is transferred to the disk when the gas is re-accreted. The treatment of feedback has, thus, a non-negligible effect on the final stellar specific angular momentum of the galaxy.

To quantify how much our results depend on the feedback scheme adopted, we also consider a run based on the feedback scheme used in Guo et al. (2011), assuming the same parameters used in Hirschmann et al. (2016). Note that this particular run does not exhibit the same level of agreement with observational data as our fiducial model, and, in particular, does not reproduce the observed galaxy mass function at high redshift. Hirschmann et al. (2016) demonstrated that the feedback scheme adopted in X17 translates in re-accretion times for the ejected gas that are delayed with respect to those expected from Guo et al. (2011). This different re-accretion history is expected to affect significantly the star formation history of model galaxies, and therefore also their sizes and angular momenta. In the following, we refer to results from this modified run as X17G11.

Bulge formation

Mergers and disk instabilities are the two possible channels that in our model produce a bulge, that we assume to be a spheroidal with zero angular momentum, supported by velocity dispersion.

Following previous work, we distinguish between two types of mergers, based on the baryonic (stars+cold gas) ratio between the secondary (less massive) and primary (more massive) galaxy. In the case this ratio is larger than 0.3, we have a "major merger" event, for which we assume that the stars of both galaxies are added to the bulge of the primary. In the case of a minor merger (ratio less than 0.3), the stellar disk of the primary is unperturbed, and the stars of the secondary are added to the primary bulge. In both cases, the cold gas of the secondary is added to the cold gas disk of the primary. We assume that the cold gas is first stripped from the satellite, and acquires the same specific angular momentum of the primary dark matter halo, obtaining an equation similar to that used for cooling (Eq. 5.1), but with the secondary cold gas mass instead of $M_{cooling}$.

All mergers trigger a star burst in the cold gas disk, which is modeled following the "collisional starburst" prescription introduced in Somerville et al. (2001), with revised coefficients from Cox et al. (2008). The amount of new stars added to the stellar disk of the central, $M_{*,SB}$, is a fraction of the cold gas of the progenitors, proportional to the merger mass ratio.

In the X17 model, disk instability is modeled as described in detail in Croton et al. (2006, see also De Lucia et al. 2011). The instability criterion is based on results from old simulations by Efstathiou et al. (1982). When a disk becomes unstable, a fraction of stars δM, necessary to restore the stability, is subtracted from the center of the disk, and is added to the bulge. During a disk instability episode, we assume that the angular momentum of the stellar disk is preserved,

and thus the specific angular momentum of the disk becomes:

$$j_*^f = \frac{j_*^0 M_{*,disk}^0}{M_{*,disk}^0 - \delta M} \tag{5.3}$$

with $M_{*,disk}^0$ the initial mass of the disk.

As discussed in previous work (see Athanassoula, 2008; Benson and Devereux, 2010; De Lucia et al., 2011), our modeling of disk instability is rather uncertain: this treatment does not take into account the gas component of the disk, which could be unstable too, collapse, and trigger a star-burst. In addition, the instability criterion adopted has been criticized by e.g.Athanassoula (2008), because it cannot distinguish bar stable from bar unstable disks. Improving the modeling adopted for this physical process is highly needed, but goes beyond the aims of this work.

The disk radius and bulge size

In X17, the radii of the cold gas and the stellar disks are based on their specific angular momentum and rotational velocity. Specifically, following Guo et al. (2011):

$$R_x = \frac{j_x M_x}{2V_{max}}, \tag{5.4}$$

where R_x, j_x and M_x are the radius, the specific angular momentum and the mass, x stands for cold gas or stellar disk, and V_{max} is the rotational velocity of the halo.

The bulge is supposed to be a dispersion dominated spheroid, and its size is estimated from energy conservation considerations. During mergers of spheroids, the energies involved are those due to their gravitation and gravitational interaction. Assuming no energy dissipation, we obtain the energy before:

$$\begin{aligned} E_i = CG \left[\frac{(M_*^p + M_{*,SB}^p)^2}{R_p} + \frac{(M_*^s + M_{*,SB}^s)^2}{R_s} \right] \\ + \alpha G \frac{(M_*^p + M_{*,SB}^p)(M_*^s + M_{*,SB}^s)}{R_p + R_s} \end{aligned} \tag{5.5}$$

and after the merger:

$$E_f = CG \frac{{M_*^f}^2}{R^f}. \tag{5.6}$$

In the first equation, G is the gravitational constant, $M_*^x + M_{*,SB}^x$ is the total stellar mass of the $x = p$ (primary) or $x = s$ (secondary) galaxy, comprehensive of the stars formed in the starburst. R_x is an approximation of the half mass radius of the total galaxy (cold gas included), obtained from the average of the different components sizes weighted for their masses. $C = 0.5$ is the form factor, that accounts for the self-binding energy of each galaxy, and $\alpha = 0.5$ is a parameter accounting for the mutual orbital energy between the spheroids (Cole et al., 2000). In the last equation, M_*^f is the final stellar mass of the spheroid, and R_f is the final bulge size.

In the case of disk instability, we calculate the radius enclosing the stellar mass removed from the disk to restore stability, assuming a disk with exponential

surface density, and we assume that this radius is the scale radius of the newly formed spheroidal component. If a bulge is already present, we merge the newly formed spheroid with the pre-existing bulge, assuming energy conservation as detailed above.

Different studies have highlighted how this simple treatment leads to unrealistic sizes of galaxies, especially at the low mass end (Hopkins et al., 2009b; Covington et al., 2011; Porter et al., 2014). This problem arises from the fact that the model outlined above ignores gas dissipation in bulge formation through mergers. Using high-resolution hydro-simulations, Hopkins et al. (2009b) proposed a simple formula to account for gas dissipation without modifying the energy conservation equation, but simply the final radius:

$$r_f = \frac{r_{no\,diss}}{1 + f_{gas}/f_0}. \tag{5.7}$$

In the above equation, $r_{no\,diss}$ is the size when dissipation is not considered, f_{gas} is the ratio between the gas involved in the merger and the stellar mass involved (including stars formed in star burst), and f_0 is a parameter varying between 0.25 and 0.30 (we assume it equal to 0.275). This formula was calculated from a set of controlled simulations of binary mergers with mass-ratio larger than 1:6, and the strongest influence was found in the case of disk-disk mergers (for example Porter et al., 2014, their Table 1). Therefore, it is not straightforward to apply this correction to all mergers. In the following, we will show results of a run where we have accounted for gas dissipation during all mergers (model X17allM) and also a run where we have applied this correction only during major mergers (model X17MM). We have not distinguished between mergers between disks or spheroids (we expect that mergers between spheroids involve only a small fraction of gas). We also tested an alternative implementation for gas dissipation, that includes a dissipation term in the energy conservation during mergers, as proposed in Covington et al. (2011), finding similar results.

Corrections for gas dissipation during mergers were already considered in the framework of semi-analytic models. For example, Shankar et al. (2013) used the formula proposed in Hopkins et al. (2009b) to correct the size of Early Type galaxies, in the framework of the model by Guo et al. (2011). He demonstrated that this correction is important in order to obtain realistic sizes for stellar masses $M_* < 10^{10.5}\mathrm{M}_\odot$. Tonini et al. (2016) included dissipation directly in the energy conservation calculation, finding similar results.

5.3 The size-mass relation

In previous work, we have shown that our model reproduces reasonably well several observed properties of galaxies, such as the Stellar, HI and H_2 Mass Functions, the HI-M_* relation, the mass-metallicity relation, the sSFR-M_* relation and the size-M_* relation for star forming, disky, gas-rich galaxies (De Lucia et al., 2014b; Hirschmann et al., 2016; Xie et al., 2017).

In this section, we study the size-mass relation for galaxies divided in Late Type (LT) and Early Type (ET), and compare it to available observational measurements. The results of this comparison are shown in Fig. 5.1, for different LT/ET selections (different panels). Different colors correspond to different galaxy types (red for ET and blue for LT galaxies), while different line styles

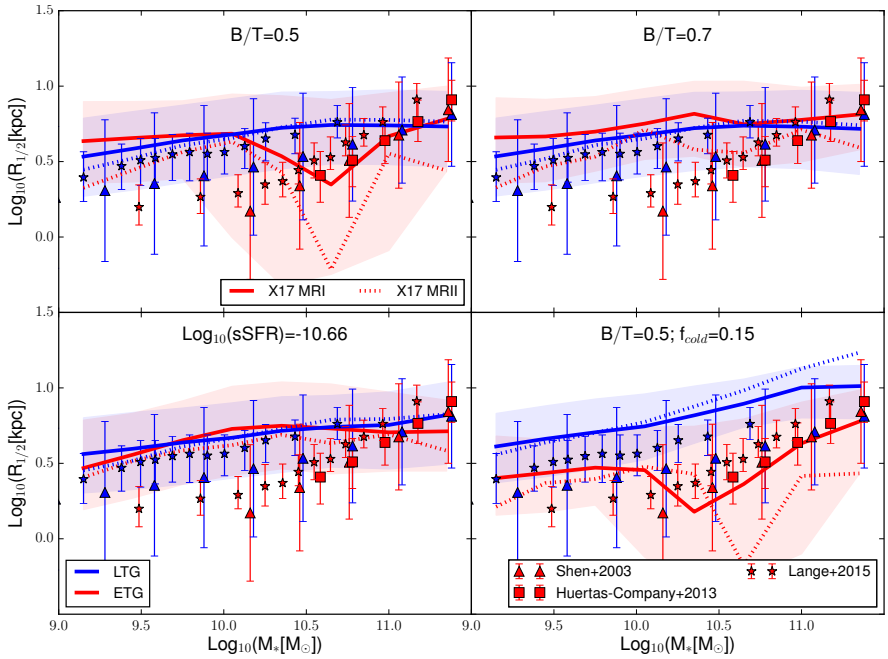

Figure 5.1: The $R_{1/2}$-M_* relation for LT and ET galaxies (blue and red lines) for our reference model X17. Solid lines correspond to results from the MRI and dashed lines are used for the MRII. Different panels show different selections for LT and ET galaxies: $B/T = 0.5$ (top left panel), $B/T = 0.7$ (top right panel), $\log_{10}(sSFR[\mathrm{yr}^{-1}]) = -10.66$ (bottom left panel) and a combination of $B/T = 0.5$ and $f_{cold} = 0.15$ (bottom right panel). Shaded areas show the 16^{th}-84^{th} percentiles of the MRI distribution; we find a similar scatter for the MSII. Different symbols correspond to observational estimates, as indicated in the legend.

are used for model results based on the MRI (solid) and on the MRII (dotted lines). The shaded areas represent the 16^{th}-84^{th} percentiles of the distribution obtained for the MRI, but a similar scatter is found for MRII.

The sizes shown in the figure correspond to the half-mass radii of model galaxies, namely the radii enclosing half of the total stellar mass. We calculated them using the stellar mass profiles assumed in the model. The stellar disk (as well as the cold gas disk) is described by an exponential surface density profile, that can be expressed as a function of the scale radius, R_{disk}, and the average surface density, $\Sigma_0 = M_{*,disk}/(2\pi R_{disk}^2)$:

$$\Sigma_{disk}(r) = \Sigma_0 e^{-\frac{r}{R_{disk}}} \tag{5.8}$$

For the bulge, we assume a stellar distribution that follows a Jaffe profile (Jaffe, 1983):

$$\rho(r) = \frac{M_B}{4\pi R_B^3}\left(\frac{r}{R_B}\right)^{-2}\left(1+\frac{r}{R_B}\right)^{-2} \tag{5.9}$$

where R_B is the scale radius, and M_B is the mass of the bulge. To have a fair comparison with observations, we consider galaxies as they were all face-on, and project their mass profiles on the plane. In this way, $R_{1/2}$ is the radius that encloses half of the projected mass.

Observational estimates are shown in Fig. 5.1 as symbols with different shape and error bars (red and blue are used for ET and LT galaxies). The estimates shown correspond all to half mass radii, but are based on observations at different wavebands, different assumptions about the light distribution, and different selections for ET and LT galaxies.
Shen et al. (2003) (triangles) used SDSS data in the z-band, and circular profiles to estimate galaxies Petrosian half-light radius. LT and ET galaxies were classified according to their concentration, a parameter related to the Sérsic index (Blanton et al., 2003). Specifically, Shen et al. (2003) separated E/S0 galaxies (considered ET) from spirals using as threshold $c = 2.86$ (Nakamura et al., 2003). Huertas-Company et al. (2013) (squares) estimated the half-light radii using a circular double component Sérsic fitting of galaxies from the SDSS DR7 spectroscopic sample. They selected ET galaxies via a morphological classification performed using a machine learning technique. Lange et al. (2015) (stars) estimated the half-light radii along the major axis of GAMA galaxies, using single Sérsic fits. They used elliptical fits, instead of the circular ones used in previous studies, and showed that these give systematically larger radii at fixed stellar mass. They divided LT from ET galaxies using four different methods: a visual morphology classification, a Sérsic index threshold of $n = 2.5$, a color-color (u-r)-(g-i) division, a combination of Sérsic index and (u-r) color. They found that the different methods select different galaxy samples, but the size-mass relations obtained are very similar. In this work we use their estimates based on the Sérsic index.

For our comparison, we are assuming that light distribution perfectly traces mass distribution. We do not attempt to reproduce the selection of LT and ET galaxies adopted in observational studies, because we cannot always mimic Sérsic index, concentration, or color-mass diagram separation, without applying additional assumptions for model galaxies. In Fig. 5.1, we show four different LT/ET selections based on model outputs. In the top panels we consider a simple selection based on the bulge over total stellar mass ratio, B/T: on the

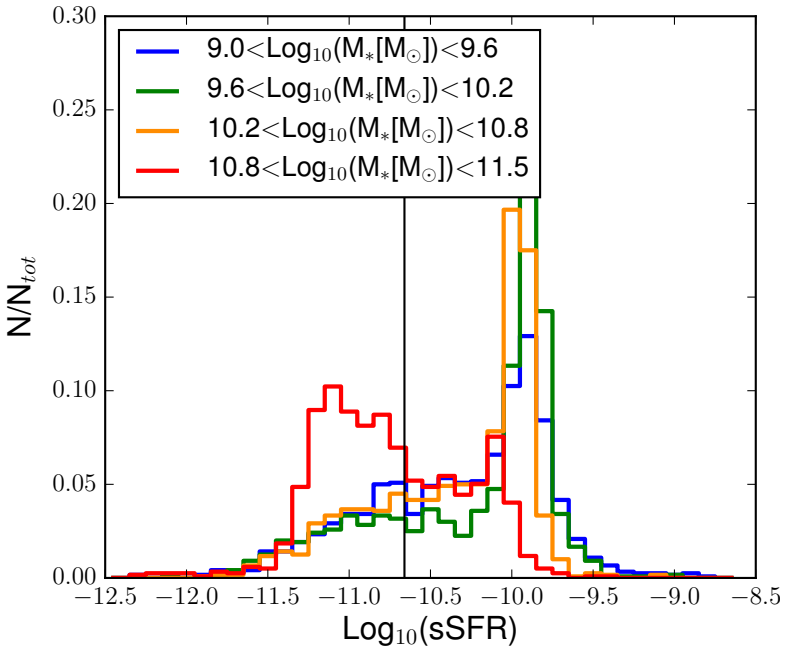

Figure 5.2: The distribution of sSFR of model galaxies with different stellar masses (different colors,as indicated in the legend). The vertical black solid line indicates the threshold adopted to separate LT from ET galaxies.

left, we use a threshold of $B/T = 0.5$, and on the right $B/T = 0.7$ to separate ET from LT galaxies. In the bottom left panel, we show a selection based on the specific Star Formation Rate (sSFR), i.e. the mass of stars formed per unit mass per unit time. This proxy separates star forming galaxies from quiescent ones. The chosen threshold is $Log_{10}(sSFR\ [\mathrm{yr}^{-1}]) = -10.66$, that approximately separates the two peaks of the sSFR bimodal distribution (Fig. 5.2). We warn the reader that the sSFR distribution of our model galaxies does not reproduce well the observed distribution. In particular, massive galaxies are more star forming than measured, as illustrated in detail in Hirschmann et al. (2016, see their Fig. 8). Finally, in the bottom right panel of Fig. 5.1, we show a selection that also considers the cold gas fraction of model galaxies, $f_{cold} = M_{cold}/(M_{cold}+M_*)$. Specifically, we select as LT galaxies those with $f_{cold} > 0.15$ and $B/T < 0.5$, and as ET galaxies those with $f_{cold} < 0.15$ and $B/T > 0.5$.

We now get back to the comparison shown in Fig. 5.1. We expect the MRII to provide a more precise estimate of the relation in the stellar mass range $10^9 - 10^{10}$ $\mathrm{M}_\odot$, where the MRI is close to the resolution limit. The MRII median size for galaxies with stellar mass 10^9 $\mathrm{M}_\odot$ is 0.2 dex lower than that based on the MRI. Predictions based on the two simulations are in good agreement for stellar masses $\sim 10^{10}$ $\mathrm{M}_\odot$. At larger stellar masses, since the MRII simulation volume is small, predictions based on this simulation are more noisy. For this reasons, in the rest of the paper we will show results from MRII up to $M_* = 10^{10}$ $\mathrm{M}_\odot$, and results from MRI above this mass.

The predicted size-mass relation for LT galaxies is in nice agreement with the latest observational estimates (those by Lange et al., 2015), independently of the

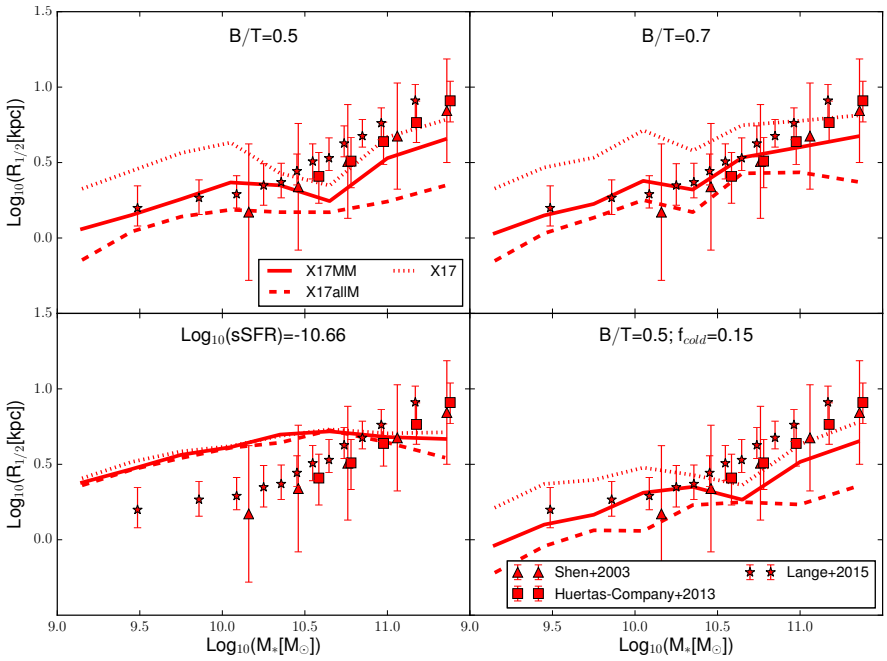

Figure 5.3: The $R_{1/2}$-M_* relation, as in Fig. 5.1, but for models including a treatment for gas dissipation in all mergers and in major mergers (dashed and solid lines). Predictions from our fiducial model are shown as dotted lines, as reference. Only the relation and data for ET galaxies are shown, as LT galaxies are not affected by this treatment.

selection adopted. In contrast, for ET galaxies, the size-mass relation is not well reproduced by any of the selections considered. The two selections that include a cut at $B/T = 0.5$ are in good agreement with data by Shen et al. (2003) and Huertas-Company et al. (2013) for galaxy masses larger than $M_* > 10^{10.5}$ $M_\odot$. For less massive galaxies, and for the other two selections considered, the model significantly over-predicts galaxy sizes. In particular, when selecting galaxies on the basis of the sSFR or using a cut at $B/T = 0.7$, model predictions for ET galaxies are very close to those obtained for LT galaxies.

This problem has been noted earlier (Hopkins et al., 2009b; Covington et al., 2011; Porter et al., 2014), as explained in Sec. 5.2.2, and is due to the fact that our model does not include a treatment for dissipation of energy during gas rich mergers. To account for dissipation, we have added the empirical prescription provided by Hopkins et al. (2009b), and considered two possibilities: dissipation in all mergers (model X17allM) and only in major mergers (model X17MM).

Results are shown in Fig. 5.3, where we show the size-mass relation for X17allM (dashed lines) and for X17MM (solid lines). Predictions from our standard model without dissipation are also shown as reference (dotted lines). LT galaxies are not affected by the inclusion of dissipation, and we do not show their relation for clarity. The relation for ET galaxies is more affected by dissipation for the selections based on B/T, while the sSFR selection is only slightly affected. For the $B/T = 0.5$ and $B/T = 0.7$ selections, the inclusion

of dissipation lowers significantly the relation, slightly in the case of X17MM and more in the case of X17allM. At the largest masses, where the reference model was in good agreement with data, both runs including gas dissipation predict sizes that are smaller than observed. The relation for galaxies selected on the basis of their sSFR is not affected by the introduction of a treatment for gas dissipation. This is due to the fact that many disky galaxies are selected in the passive ET group, making the two relations very similar. The selection including a cut based on f_{cold} results in a relation similar to the division based on $B/T = 0.5$.

Interestingly, dissipation during minor mergers largely affects high mass galaxies ($M_* > 10^{10.2}$ $M_\odot$). As we will see in Sec. 5.5, this is due to the fact that these galaxies have experienced many minor mergers during their life.

5.3.1 The size of galactic components

In this section, we analyze the size-mass relation for the disk and bulge components separately. We plot the half-mass radii versus stellar mass of disks and bulges of the X17MM model galaxies in Fig. 5.4. We also show the observed distribution by Lange et al. (2016), as dashed contours, blue for disks and red for spheroids. The median size-mass relation obtained for disks in our X17MM run nicely follows observations, while the bulge median relation is offset about 0.5 dex below the observational estimates in the stellar mass range $M_* \in [10^{9.8} - 10^{11}]$ $M_\odot$. The scatter is large and there is large overlap between data and model predictions, except at the most massive end, where the model tends to under-predict bulge sizes significantly.

To better understand the behavior of our model, we quantify the relative contribution of mergers and disk instabilities to the mass of each bulge. We then divide model bulges accordingly to the channel that contributed most to their mass: if at least 50% of the bulge mass formed from disk instabilities it is identified as DI bulge, otherwise it is a merger bulge. The predicted sizes for these two different populations are shown in Fig. 5.4 as green and orange solid lines (thick for the median and thin for the 16^{th}-84^{th} percentiles). Merger bulges are systematically larger than DI bulges, and their median size-mass relation is only slightly flatter than the observed distribution, especially at high masses.

The division of bulges into these two categories is motivated by observational evidences (see Kormendy and Kennicutt, 2004, for a review). Classical bulges are spheroidal and dispersion dominated, while pseudo-bulges are characterized by exponential mass distributions (as disks) and rotation. There is still no consensus on their origins, but several studies suggest that classical bulges form through mergers, while pseudo-bulges are formed from secular processes, such as disk instabilities. Recent work has demonstrated that these different bulge families have similar sizes (Gadotti, 2009), or that pseudo-bulges are only slightly larger than classical ones (Lange et al., 2016). In our model, the two channels of bulge formation are included, but we do not explicitly differentiate between a classical and a pseudo component. For the purpose of size estimation, the bulge is considered as a single, dispersion dominated component. Recently, Tonini et al. (2016) included an explicit division between classical and pseudo bulges in a semi-analytic model. In their model, pseudo-bulges have masses similar to those of classical bulges, but most are concentrated in the $M_B \in [10^{10}, 10^{11}]$ $M_\odot$ range. In general, their pseudo-bulges have small half mass radii (up to 5 kpc),

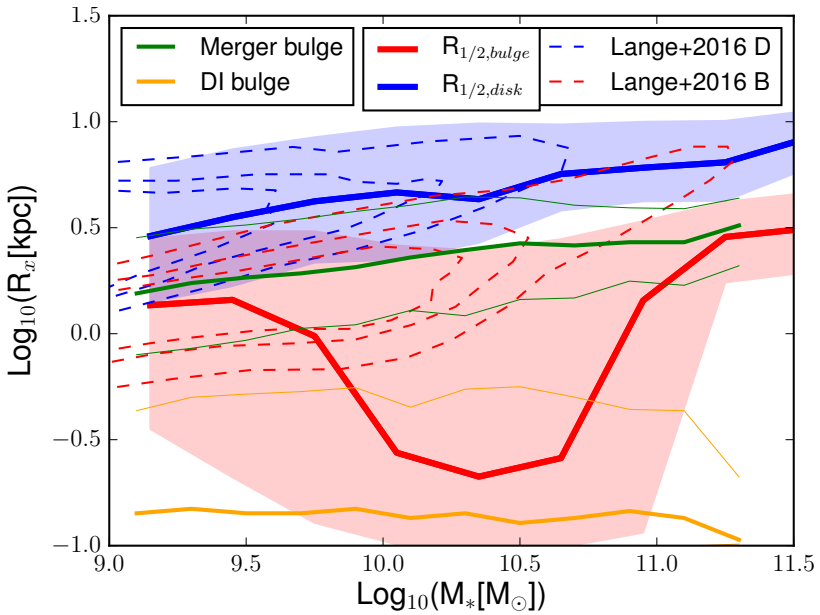

Figure 5.4: The size-mass relation for galactic components. The thick solid blue and red lines show the median sizes of the stellar disks and of the bulges of model galaxies from our X17MM run. The shaded areas correspond to the 16^{th}-84^{th} percentiles of the distributions. The green and orange thick lines show, respectively, the median relations for merger and disk instability dominated bulges. The thin lines of the same colors correspond to 16^{th}-84^{th} percentiles of their distribution. We show as dashed contours the distribution of observational measurements from Lange et al. (2016), for disk components (blue) and spheroids (red).

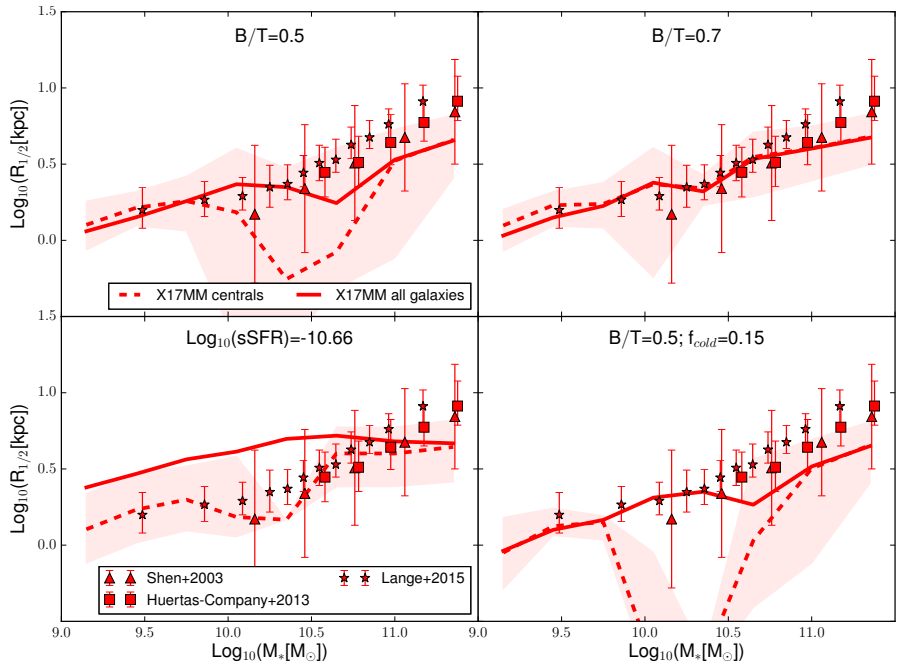

Figure 5.5: The $R_{1/2}$-M_* relation for ET galaxies from the X17MM model. Predictions for all ET galaxies are shown as solid lines, while for central ET galaxies are shown as dashed lines. The shaded areas represent the 16^{th}-84^{th} percentiles of the central galaxies distribution.

while classical bulges are larger (up to 20 kpc). In our model, DI bulges are much smaller than 1 kpc. These results suggest we should revise our treatment of disk instability, and in particular the model assumed to estimate the size of bulges forming through this channel. Assuming that disk instability produces a bulge with an exponential mass profile, would shallow the DI bulge mass distribution, leading to larger sizes. These larger sizes would also affect the classical bulge formed through mergers, because in the energy estimation the DI bulge would contribute with a larger size. In this work, we have not explored the explicit implementation of this treatment, but we have evaluated its effect in post-processing: we assume the DI bulges are disk-like components, and evaluate their half mass radius assuming an exponential distribution for their mass. As a result, the median size-mass relation for DI bulge is shifted upwards by ~ 0.5 dex. We reserve a self-consistent implementation of a two-bulge model for a future work.

5.3.2 Early Type Central and Satellite Galaxies

Observational studies suggest that central and satellite ET galaxies follow the same size-mass relation, at least at the high mass end ($M_* > 10^{10.5}$ $M_\odot$, see for example Huertas-Company et al., 2013). In Fig. 5.5, we show the median size-mass relation for ET galaxies, for all galaxies (solid) and only for centrals (dashed lines).

LT centrals follow the same relation of all galaxies, and we do not show them for clarity. The relations found for ET central galaxies, instead, are different from the relations found for all ET galaxies for all the selections considered, but for $B/T = 0.7$. In the selections assuming a cut at $B/T = 0.5$, the relation for central galaxies presents a strong dip in the stellar mass range $M_* \in [10^{10} - 10^{10.8}]$ $M_\odot$. This dip is not present in the relation for all ET galaxies. The different relations found for ET centrals must be ascribed to the different origin of their bulges. As explained above, model bulges can form through mergers or disk instabilities, and the latter form bulges with a substantially smaller size. When we select ET galaxies using $B/T > 0.7$, we are selecting bulges formed mainly through mergers (from 93% to 100% depending on the stellar mass range considered). Selecting ET galaxies with $0.5 < B/T < 0.7$, we find mainly bulges formed through disk instability (from 24% to 91% of the bulges, depending on the mass range). For this reason, the selection based on a cut at $B/T = 0.5$ is most affected by the small bulges formed in disk instabilities.

The results found in Fig. 5.5 must thus be attributed to a different influence of disk instability for centrals and satellites. We will come back to this in Sec. 5.5.2. In addition, there is also a numerical difference between centrals and satellites in the stellar mass range $M_* \in [10^{10} - 10^{10.8}]$ $M_\odot$. In particular, there are many more centrals in this mass range with $0.5 < B/T < 0.7$ than with $B/T > 0.7$. When considering the entire ET population, the proportions are inverted. As a consequence, in this mass range, central galaxies account for a larger fraction of the small size bulges with respect to the entire ET population. When the $B/T > 0.5$ selection is combined with the $f_{cold} < 0.15$ cut, the dip in the ET central galaxies size-mass relation is much deeper than for the simple $B/T > 0.5$ selection. We will show in Sec. 5.5.3 that gas poor galaxies and disk instabilities are both more likely to occur in halos with low specific angular momentum. For this reason, a selection based on low gas fraction likely correlates with the frequent occurrence of disk instabilities, and thus with small bulges.

Interestingly, the relation for ET central galaxies selected using a sSFR cut is in agreement with observational data. The entire population includes many disky quenched galaxies, but this is not the case for central galaxies. The reason is in the different distributions of sSFR for centrals and satellites. If we re-analyze Fig. 5.2, dividing the sSFR distributions in centrals and satellites, we find that all satellite galaxies are around our threshold, $Log_{10}(sSFR) = -10.66$, while central galaxies are distributed at higher values, with a tail below the sSFR threshold. Only central galaxies with large stellar mass ($M_* \in [10^{10.8} - 10^{11.5}]$ $M_\odot$) are found in the low sSFR region. Satellites are all quenched with respect to our threshold because of the hot gas stripping. Moreover, satellite morphology is generally preserved after accretion, because mergers between satellites are rare, and the only channel for bulge accretion is disk instability. Thus, star formation and morphology in satellites are uncorrelated. For central galaxies, instead, there is still a strong correlation with star formation activity and morphology.

5.3.3 The size-mass relation for X17 modifications

In Sec. 5.2 we described our fiducial model and the relevant physical prescriptions, and highlighted those processes for which we have considered alternative

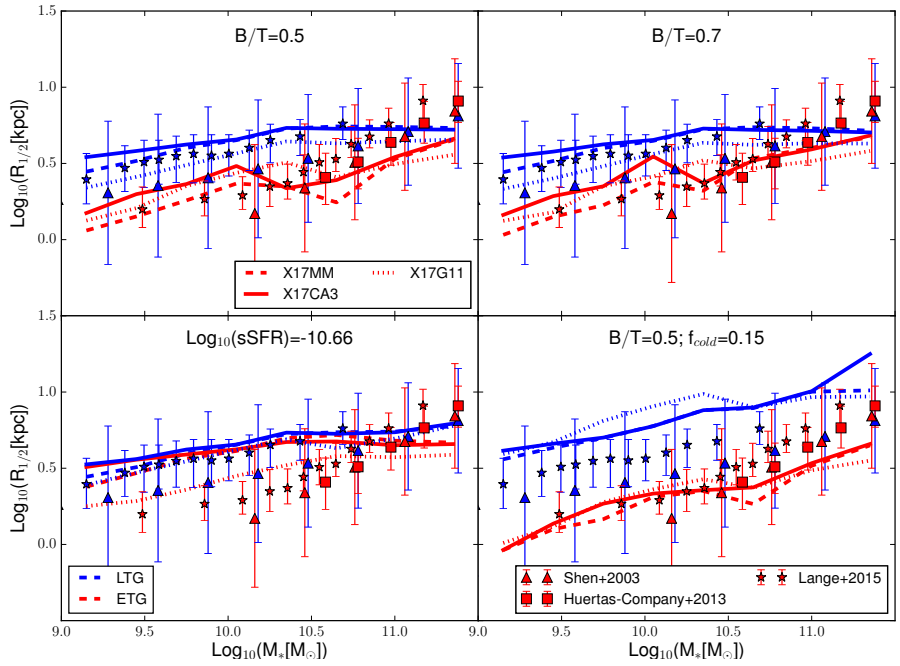

Figure 5.6: The $R_{1/2}$-M_* relation, as in Fig. 5.3, for the X17CA3 (dashed lines) and X17G11 (dotted lines) runs, both with dissipation during major mergers. Predictions from the X17MM model are shown as solid lines, as reference.

treatments. Specifically, we assume that gas accreted through cold accretion has a specific angular momentum higher than that of the dark matter halo (X17CA3), and, alternatively, we adopt the stellar feedback scheme described in Guo et al. (2011) (X17G11). As explained in Sec. 5.2, this latter run is parametrized as described in Hirschmann et al. (2016), who have shown that, in the GAEA framework, this feedback scheme does not reproduce the observed evolution of the galaxy stellar mass function. Both runs include gas dissipation during major mergers.

We show the size-mass relations for these models in Fig. 5.6, with results from X17CA3 shown as dashed lines, and X17G11 as dotted lines. Predictions from the X17MM model are plotted as solid lines for reference.

Results from the X17CA3 run are very similar to those obtained from the X17MM model for stellar masses larger than $M_* > 10^{10.5}$ $M_\odot$. For lower masses, the former model predicts sizes that are offset high with respect to the X17MM run of about 0.1 dex for LT and 0.2 dex for ET galaxies. The larger specific angular momentum in cold accretion affects significantly low mass galaxies. We will see in Sec. 5.5.4 that, in this mass range, the angular momentum of galaxies (and therefore their size) is determined at early times (where cold accretion is important), and is not significantly modified during subsequent evolution.

A different stellar feedback influences the sizes of both ET and LT galaxies over the entire stellar mass range considered. LT galaxies in the X17G11 run have sizes that are systematically below the results from the X17MM model,

by about 0.2 dex. We will show in Sec. 5.5.4, that in X17G11 run most of the stars form earlier than in X17MM. This is due to the different re-accretion times of the ejected material in the two feedback schemes, that allow galaxies in the X17MM run to preserve more cold gas for star formation at late times. The peak of star formation in the X17G11 run occurs earlier than in X17MM for both LT and ET galaxies, so that we expect a lower size-mass relations for both. This is not the case for ET galaxies, that, on average, have sizes larger of ~ 0.2 dex with respect to X17MM, for selections based on a cut in B/T. This is due to the fact that ET galaxies were subject to a similar number of mergers in X17MM and X17G11, but disk instabilities occurred at earlier times in X17G11. The ET galaxies of the X17MM and of the X17G11 runs have thus different origins, and in the X17G11 model those formed during disk instabilities are fewer than those in X17MM.

5.4 The specific angular momentum

The correlation between sizes and masses of our model galaxies can be interpreted in relation with the angular momentum treatment. Indeed, as discussed in Sec. 5.2, disks radii (both of the stellar and of the gaseous component) are calculated from their specific angular momenta.

Below, we first describe how we treat model outputs to have a fair comparison to observational estimates, and then analyze the relation between specific angular momentum and stellar mass, and its dependence on morphology and cold gas fraction.

5.4.1 Specific angular momentum determination

Romanowsky and Fall (2012) has become the observational reference paper in this field. The authors selected a sample of galaxies of different morphology, with spatially resolved rotational velocity and photometric information, available from literature. For a subsample, they estimated the specific angular momentum by direct integration along the major axis of the galaxies. They also compared this estimate against an empirical formula, finding a good agreement with results from the full integration.
In order to estimate the angular momentum of model galaxies in a way that mimics the observations, we need: a resolved rotational velocity field, a resolved light (mass) projected distribution, and a maximum aperture enclosing the observed galaxy. These information are not direct outputs of the X17 model, that, as described in Sec. 5.2, provides an estimate for the total specific angular momentum of the cold gas and stellar disks, evolved according to mass exchanges between the different galactic components. In the following, we describe how we have computed alternative estimates of the galaxy angular momentum, that can be compared to observational estimates. We refer to the cartoons in Fig. 5.7, to help the reader in visualizing the description below.

The first estimate we consider is based on a three dimensional (3D) model of the galaxy (see the top panel of Fig. 5.7). We assume to have information on the mass and velocity 3D distribution for each galaxy, and we calculate the specific angular momentum within the radius R, by integrating the velocity and mass profiles over all radii $r < R$. The stellar mass distribution of the

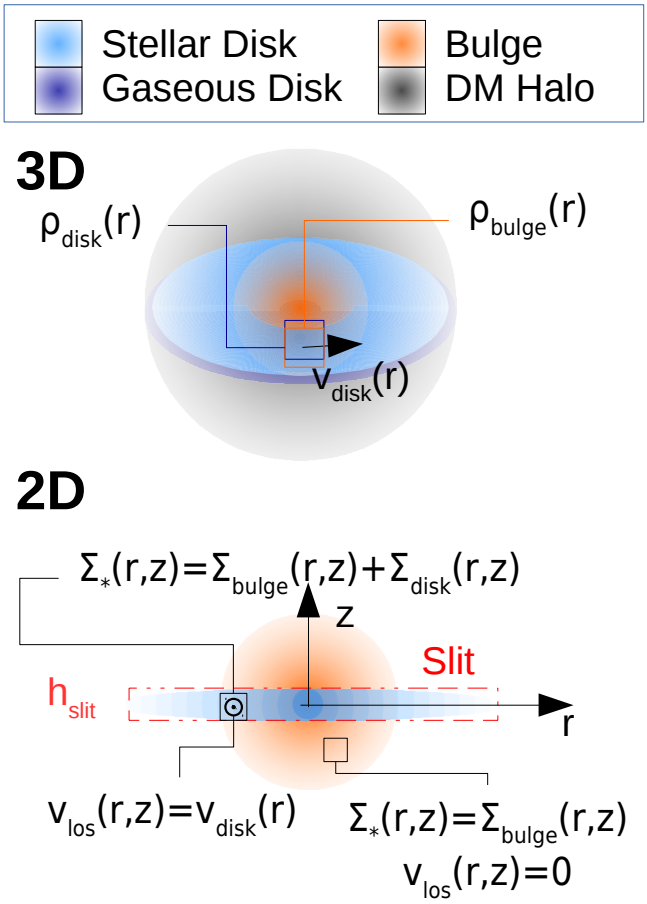

Figure 5.7: A schematic representation of the method adopted to calculate the 3D (top) and 2D (bottom) estimates of the angular momentum for model galaxies. See text for details.

galactic components assumed in the model was discussed in Sec. 5.3: the mass distribution of gas and stars in the disks is described by an exponential profile, while the bulge is assumed to have a Jaffe profile. The model also provides an estimate for the maximum rotational velocity of the disk, assumed to be equal to that of the dark matter halo. The bulge is assumed to be dispersion supported: $v_{bulge}(\vec{r}) = 0$. We estimate the rotational velocity profile for the stellar disk, assuming it is symmetric and supported by rotation:

$$v_{rot}^{disk}(r) = \sqrt{\frac{GM(<r)}{r}}, \tag{5.10}$$

where G is the gravitational constant, and $M(<r)$ is the total mass enclosed within r. It includes the stellar disk and the bulge, the cold gas disk, and the corresponding fraction of the parent DM halo. This last component is described in the model through its mass, M_{200}, corresponding to an over-density of $200\rho_{crit}$ (ρ_{crit} is the critical density of the Universe). This mass is measured directly from the simulations for all halos hosting central galaxies, while for satellites it corresponds to the particle mass times the number of bound particles in the parent subhalo, at the last time this was detected. A size is provided by R_{200}, that can be calculated from M_{200} and the redshift of the halo. These two parameters are sufficient to estimate the DM mass distribution assuming a Navarro-Frenk-White profile (Navarro et al., 1996):

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_{DM}}\left(1+\frac{r}{R_{DM}}\right)^2}, \tag{5.11}$$

where ρ_0 and R_{DM} are a density parameter and the scale radius of the halo. Both can be estimated using the concentration parameter of the halo, that we calculate using the correlation between M_{200} and concentration published in Neto et al. (2007). This relation has a large scatter, but we checked that our results are not significantly affected by this using the extremes of the distribution, instead of the median. The virial radius is directly proportional to the scale radius: $R_{200} = cR_{DM}$ and the density parameter ρ_0 can be obtained integrating Eq. 5.11 to R_{200}, and imposing it to be equal to M_{200}.

We now have all the ingredients to calculate the specific angular momentum of the disk. Integrating on the plane of the disk:

$$j_{disk}^{3D}(R) = \frac{\int_0^R r \cdot v_{rot}^{disk}(r) \cdot \Sigma_{disk}(r) \cdot r \cdot dr}{\int_0^R \Sigma_{disk}(r) \cdot r \cdot dr} \tag{5.12}$$

If we include the bulge, for which we assume $v_{bulge}(r) = 0$, we obtain:

$$j_{tot}^{3D}(R) = \frac{\int_0^R r \cdot v_{rot}^{disk}(r) \cdot \Sigma_{disk}(r) r dr}{\int_0^R \Sigma_{disk}(r) r dr + \int_0^R \rho_{bulge}(r) 4\pi r^2 dr} \tag{5.13}$$

where the integration for the bulge is carried out in 3D space.

In observations, the 3D information is not available: the galaxy is projected on the sky (2D), with a random inclination, and the velocity information is typically available only along the line of sight (los). In addition, it is difficult to separate the contributions from the bulge and the disk. We mimic this

situation assuming all model galaxies are edge-on (see the bottom panel of Fig. 5.7). In this way, the los velocity measured along the disk, at a projected distance r from the center, is exactly the rotational velocity at the 3D distance r: $v_{los}(r) = v_{rot}^{disk}(r)$. We project the stellar profiles of the stellar disk and of the bulge on the edge-on plane, and sum them into a single stellar component: $\Sigma_*(r,z) = \Sigma_{disk}^{edge-on}(r,z) + \Sigma_{bulge}(r,z)$. In this last equation, r is the coordinate along the disk, and z is the distance from the disk plane. We also assume $h_{disk} = R_{disk}/7.3$ (Kregel et al., 2002), and:

$$v_{los}(r,z) = \begin{cases} v_{rot}^{disk}(r) & \text{if } z < h_{disk} \\ 0 & \text{if } z > h_{disk} \end{cases}$$

In this way, the bulge fraction contained in a cylinder of height h_{disk} is assumed to rotate with the disk, while there is no rotation outside the cylinder. The specific angular momentum calculated along a slit of the same height of the disk $h_{slit} = h_{disk}$ is:

$$j_{slit}^{2D}(R) = \frac{\int_0^R \int_0^{h_{slit}} r \cdot v_{los}(r,z)\Sigma_*(r,z) dr dz}{\int_0^R \int_0^{h_{slit}} \Sigma_*(r,z) dr dz} \tag{5.14}$$

Including also the bulge component outside the slit:

$$j_{tot}^{2D}(R) = \frac{\int_0^R \int_0^R r \cdot v_{los}(r,z)\Sigma_*(r,z) dr dz}{\int_0^R \int_0^R \Sigma_*(r,z) dr dz} \tag{5.15}$$

These estimates mimic the integrations performed for observed galaxies, with j_{tot}^{2D} similar to an integration on circular or elliptical concentric annuli, and j_{slit}^{2D} similar to an integration along the major axis, as in Romanowsky and Fall (2012).

As already mentioned earlier, Romanowsky and Fall (2012) found that the 2D estimate is well approximated, after de-projection for inclination, by an empirical formula that depends on the effective radius, the velocity measured at two effective radii, and a factor k_n, that depends on the Sérsic index. In the case of a disk+bulge galaxy, they sum the contributions from the disk and the bulge, weighting them for the corresponding light (mass) fraction, D/T and B/T:

$$j_{D+B} = k_{n_D} v_{disk}(2R_D^e) R_D^e \frac{D}{T} + k_{n_B} v_{bulge}(2R_B^e) R_B^e \frac{B}{T} \tag{5.16}$$

In the above equation, n_x and R_x^e are the Sérsic index and the effective radius of the disk ($x = D$), or the bulge ($x = B$). The disk velocity $v_{disk}(2R_D^e)$ is measured from the ionized gas of the disk, while the bulge rotational velocity $v_{bulge}(2R_B^e)$ is estimated from its relation with the ellipticity ϵ and central velocity dispersion σ_0, through:

$$v_{bulge} = \left(\frac{v}{\sigma}\right)^* \sigma_0 \left(\frac{\epsilon}{1-\epsilon}\right)^{1/2} \tag{5.17}$$

$(v/\sigma)^* \sim 1$ is a parameter describing the relative dynamical importance of rotation and pressure, and its value corresponds to that of an oblate isotropic system viewed edge-on (Kormendy and Illingworth, 1982).

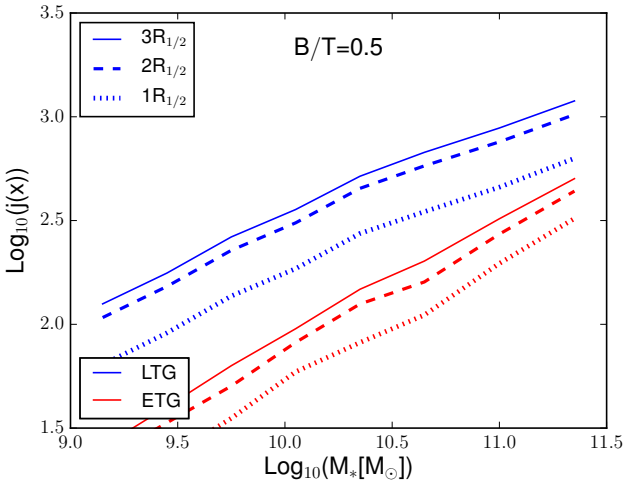

Figure 5.8: The j^{2D}_{slit}-M_* relation for LT and ET galaxies (blue and red), selected by their morphology ($B/T = 0.5$), from the X17MM model. The $j^{2D}_{slit}(r)$ is evaluated considering projected profiles at different radii: $r = 1, 2, 3\ R_{1/2}$.

We assume all galaxies are composed of a disk+bulge, with a disk Sérsic index $n_D = 1$ ($k_1 = 1.19$), $v_{bulge} = 0$, and $\epsilon = 0$. Thus the second part of Eq. 5.16 is always 0:

$$j^{RF} = k_1 v^{disk}_{rot}(2R^e_D)R^e_D\frac{D}{T} \tag{5.18}$$

The assumption of perfectly spherical, dispersion dominated bulges is very strong. We evaluate its impact by assuming, alternatively, that all bulges have a Sérsic index $n_B = 4$ ($k_4 = 2.29$), an ellipticity $\epsilon = 0.2$ (this is the median ellipticity of the elliptical/lenticular SDSS galaxies as found in Hao et al., 2006), and a velocity dispersion evaluated using the virial theorem: $\sigma_0 = \sqrt{GM_B/(2R_B)}$. Eq. 5.16 then writes as:

$$j^{RF}_{\epsilon=0.2} = k_1 v^{disk}_{rot}(2R^e_D)R^e_D\frac{D}{T} + k_4 0.5\sqrt{\frac{GM_B}{2R_B}}R^e_B\frac{B}{T} \tag{5.19}$$

Finally, we use as reference the output from the model, j^{SAM}_{disk}, weighted for the bulge contribution: $j^{SAM}_{tot} = j^{SAM}_{disk}(1 - B/T)$.

Before comparing the various estimates computed, we analyze the dependence of j^{2D} and j^{3D} on the aperture. We plot the median j^{2D}_{slit}-M_* relation evaluated at 1, 2 and 3 $R_{1/2}$ in Fig. 5.8. The difference between $2R_{1/2}$ and $3R_{1/2}$ is much smaller than that between $2R_{1/2}$ and $1R_{1/2}$. In the following, all 2D and 3D measurements shown for the specific angular momentum correspond to $2R_{1/2}$, a distance that provides a good compromise between convergence and observational limit.

We analyze how different estimates reflect on the $j_* - M_*$ relation in Fig. 5.9, where we show results for LT (blue) and ET (red) galaxies from the X17MM model. We only consider a selection assuming a $B/T = 0.5$ cut here, but we have verified that results are qualitatively similar for alternative selections. Different

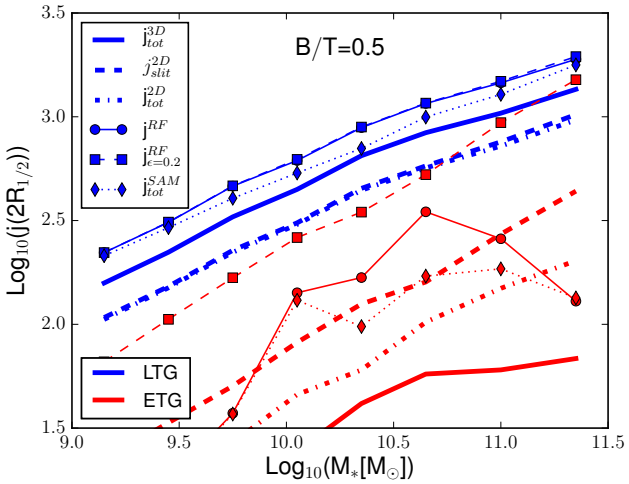

Figure 5.9: The median j-M_* relation for LT and ET galaxies (blue and red) selected by their morphology ($B/T = 0.5$). All j_* estimates are evaluated inside $2R_{1/2}$. Different lines and symbols represent different estimates: j^{3D}_{tot} (solid), j^{2D}_{slit} (dashed), j^{2D}_{tot} (dotted), j^{RF} (solid with circles), $j^{RF}_{\epsilon=0.2}$ (dashed with squares). The relation corresponding to j^{SAM}_{tot} is also shown (as dotted line with diamonds).

line styles correspond to the different estimates introduced earlier, as indicated in the legend.

All relations obtained for LT galaxies, based on different estimates of j_*, are above the ET relations. As we will see in the next section, this is qualitatively in agreement with observations.
The estimates j^{RF} and $j^{RF}_{\epsilon=0.2}$ give almost identical relations for LT galaxies, and correspond to the highest normalization of the j_*-M_* relation for these galaxies. This is expected, because the influence of bulges in eq.s 5.18 and 5.19 for galaxies with $B/T < 0.5$ is negligible. The relation assuming the direct model output, j^{SAM}_{tot}, is only slightly below that based on the empirical formula proposed by Romanowsky and Fall (2012). The relation based on the 3D estimate, j^{3D}_{tot}, is parallel to these but has a ~ 0.2 dex lower normalization. Both the 2D estimates j^{2D}_{slit} and j^{2D}_{tot} lie on the same relation, shifted 0.2 dex below that based on the 3D estimate. This is not surprising, because LT galaxies have a small bulge, and a large fraction of it is contained in the slit. The difference with respect to the 3D estimate is due to the projection of the disk mass: most of the disk mass, residing at the center, has a lower velocity than in the de-projected case. We expect that for inclinations lower than edge-on this relation moves up towards the 3D relation. This argument is valid only for LT galaxies, for which the bulges, whose projected distribution is spherical and does not depend on inclination, do not dominate the central stellar mass distribution. When comparing model predictions for LT galaxies with data below, we will use the area between the 2D and 3D estimates, so as to account for different possible inclinations of the disk mass projection.

While different estimates of j_* for LT galaxies correspond to parallel relations

with relatively small shifts, for ET galaxies they cover a much larger region of the $j_* - M_*$ plane. As for LT galaxies, the highest normalizations are obtained for the empirical estimates by Romanowsky and Fall (2012), $j^{RF}_{\epsilon=0.2}$ and j^{RF}. The $j_* - M_*$ relation based on the estimate that includes rotation in bulges, $j^{RF}_{\epsilon=0.2}$, lies very close to the relation obtained for LT galaxies employing the 2D estimates, with a slightly steeper slope. As expected, the $j^{RF}_{\epsilon=0.2}$ estimate is larger than that obtained assuming $\epsilon = 0$, j^{RF}. The latter does not include a bulge velocity component, and the calculation depends only on the disk size and its velocity. The dependence on the disk size is important, because, observationally, measuring the disk radius in a bulge dominated galaxy is not easy. In fact, using the effective radius of the galaxy in the j^{RF} estimations, instead of the disk effective radius, lowers the relation at intermediate masses to the same position of the relation based on the j^{2D}_{tot} estimate (for $M_* < 10^{10.7}$ M$_\odot$).
The relation obtained considering the direct model output, j^{SAM}_{tot}, is close to that obtained using j^{RF}: the specific angular momentum increases with increasing stellar mass up to $M_* \sim 10^{10.7}$ M$_\odot$; for larger stellar masses, the median value of j^{SAM}_{tot} first flattens and then decreases. The relation assuming the 2D circular estimate, j^{2D}_{tot}, is parallel to that obtained using j^{2D}_{slit}, but is shifted down by 0.3-0.4 dex. We tested the influence of the slit height h_{slit} on the predicted $j_* - M_*$ relation, finding that a smaller h_{slit} would shift j^{2D}_{slit} upwards, and j^{2D}_{tot} downwards. For a slit with height of $\sim 0.1 h_{slit}$, the shift of the relation using j^{2D}_{slit} is of 0.1 dex, while that of the relation based on j^{2D}_{tot} is of 0.2 dex, a quite modest effect.
The relation corresponding to the 3D estimate, j^{3D}_{tot}, is well below the other relations, because it does not mix bulge stars within the rotating disk, as happens in the projected estimates. When comparing model results with observational estimates for ET galaxies, we will show the region enclosed between the relations based on j^{2D}_{slit} and j^{2D}_{tot}. We note that rotating bulges would translate in larger j_* values for both the 3D and the 2D estimates, shifting up the relations similarly to what found for j^{RF} and $j^{RF}_{\epsilon=0.2}$. Therefore, these relations should be considered as lower limits.

5.4.2 Comparison with observations

In this subsection we compare the specific angular momenta of our model galaxies, estimated as described in Sec. 5.4.1, to available data from the literature.

Fig. 5.10 shows the predicted $j_* - M_*$ relation compared with data from Romanowsky and Fall (2012), corrected for a variable light-to-mass ratio as in Fall and Romanowsky (2013). These observational data are represented as circles, and are divided in spirals (S, cyan), ellipticals (E, violet) and lenticulars (S0, light green). The relations from the X17MM model are shown as shaded areas, enclosing the area between j^{3D}_{tot} and j^{2D} for LT galaxies (blue), and between j^{2D}_{slit} and j^{2D}_{tot} for ET galaxies (red). Thin solid lines of the same colors represent the scatter (16^{th}-84^{th} percentiles) of the distributions. We only show the predictions obtained using a $B/T = 0.5$ cut to distinguish between ET and LT galaxies, as different selections based on morphology give similar results.

LT and ET galaxies follow parallel relations, with similar offset for both model and observed galaxies. The slope of the predicted relation is also in nice agreement with the measured one. However, on average, the specific angular momentum of model galaxies is $\sim 0.2 - 0.4$ dex lower than that estimated from

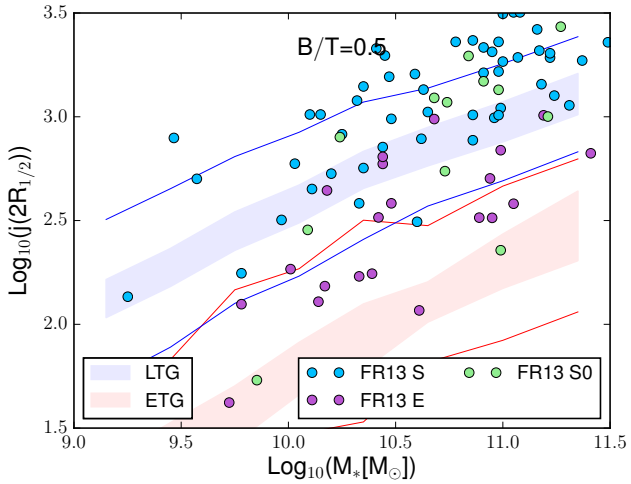

Figure 5.10: The j_*-M_* relation for LT and ET galaxies from the X17MM model (blue and red shaded areas). The thin solid lines correspond to the 16^{th}-84^{th} percentiles of the distribution. The way the shaded regions are determined is explained in detail in Sec. 5.4.1. The LT/ET selection assumes a threshold of $B/T = 0.5$. Observational data (circles) are from Fall and Romanowsky (2013), and are color-coded according to galaxy morphology: spirals (cyan), ellipticals (violet) and lenticulars (light green).

the data for galaxies of similar stellar mass.

In the case of ET galaxies, as explained in Sec. 5.4.1, model estimates do not account for rotating bulges. Comparing the two empirical estimates, j^{RF} and $j^{RF}_{\epsilon=0.2}$, we expect that rotating bulges would raise the median relation of some tenth of a dex. On the other hand, observed ET galaxies were selected to have a measured rotational velocity profile. Therefore, they provide a sample biased towards fast rotators, that have a j_* higher than average elliptical galaxies.

In the case of LT galaxies, as shown in Fig. 5.9, the j^{RF} estimates are 0.2 dex larger than the corresponding j^{3D}_{tot} estimates. This difference is sufficient to move the predicted relation closer to the observational estimates, that are obtained using the empirical formulas proposed. Furthermore, for the observed galaxies, rotational velocity profiles are evaluated from the ionized gas of the disk, selecting gas-rich galaxies. In our model, the cold gas fraction ($f_{cold} = M_{cold}/(M_{cold} + M_*)$) strongly correlates with j_*, as shown in Fig. 5.11. In this figure, we show the distribution of model galaxies in the $j^{2D}_{tot} - M_*$ plane, color-coded by the median f_{cold} in each pixel. For clarity, we show LT and ET galaxies in two distinct panels, dividing them using a cut at $B/T = 0.5$. Gas rich galaxies have high specific angular momenta, and the correlation with stellar mass is somewhat steeper. If we select LT galaxies with $f_{cold} > 0.3$ and $B/T < 0.5$, their median relation shifts up by ~ 0.3 dex. ET galaxies have a similar dependence, although lower values of f_{cold}, as expected.
The correlation between j_* and f_{cold} is driven by galaxy evolution. As a dark matter halo grows, its specific angular momentum increases, and is transferred to the cold gas through gas cooling. When the cold gas fraction is high, the

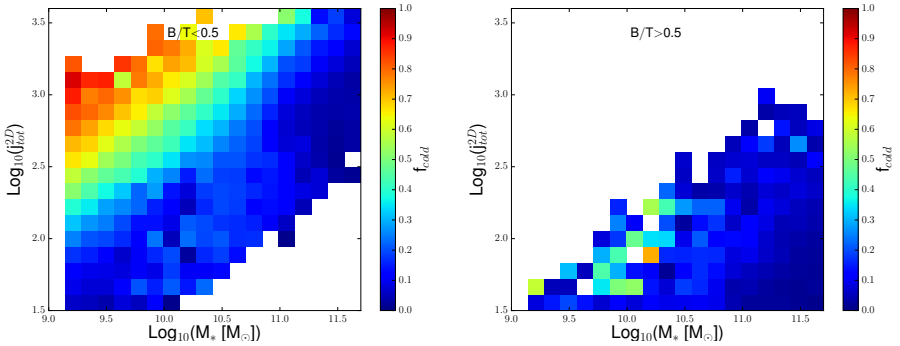

Figure 5.11: The j_{tot}^{2D}-M_* relation for X17MM galaxies, color coded by their median cold gas fraction f_{cold}. The left panel corresponds to LT galaxies ($B/T < 0.5$), the right panel to ET galaxies ($B/T > 0.5$).

galaxy is relatively young, i.e. few stars have formed with relatively high specific angular momentum, resulting in the high $j_* - M_*$ relation we observe. We will discuss further the influence of cold gas and early/late star formation in Sec. 5.5.3.

We show in Fig. 5.12 a comparison with other observational measurements available in literature. Our model predictions are shown as in Fig. 5.10, while the observational estimates are from Obreschkow and Glazebrook (2014) (stars, corresponding to a set of nearby gas-rich spirals), and from Cortese et al. (2016). The latter work is based on galaxies from the SAMI survey (dotted lines), divided according to their morphological type: Sbc (dark blue), S0/Sa-Sb (cyan), E/S0-S0 (pink) and E (red) galaxies. Both these samples are incomplete and somewhat biased. The Obreschkow and Glazebrook (2014) data lie at the high end of our LT galaxies relation. This can be explained, as discussed above, by the fact that the sample is composed of gas rich galaxies. The relations estimated for the SAMI galaxies are steeper than our model predictions. Cortese et al. (2016) discussed the different slope with respect to Fall and Romanowsky (2013), arguing it is due to the aperture used in their measurements, that corresponds to approximately $\sim 1R_{1/2}$. They measured the specific angular momentum up to $\sim 2R_{1/2}$ for a subsample of galaxies, obtaining a better agreement with Fall and Romanowsky (2013). This is not consistent, however, with the different relations obtained using different apertures for our model galaxies. As we noticed in Fig. 5.8, a larger aperture translates in a higher normalization of the j_*-M_* relation, but the slope is largely unaffected.

5.4.3 Specific angular momentum in X17CA3 and X17G11

We show in Fig. 5.13 the $j_* - M_*$ relation for two modified versions of our model: one accounting for a larger angular momentum for gas accreted during the rapid cooling regime (X17CA3, areas with oblique lines), and one with a stellar feedback scheme based on that of Guo et al. (2011) (X17G11, areas with circles). The areas showing model predictions are determined as described in Sec. 5.4.1. Galaxies are classified as LT and ET using a threshold $B/T = 0.5$ (blue and red). We show also results from the X17MM model as a reference

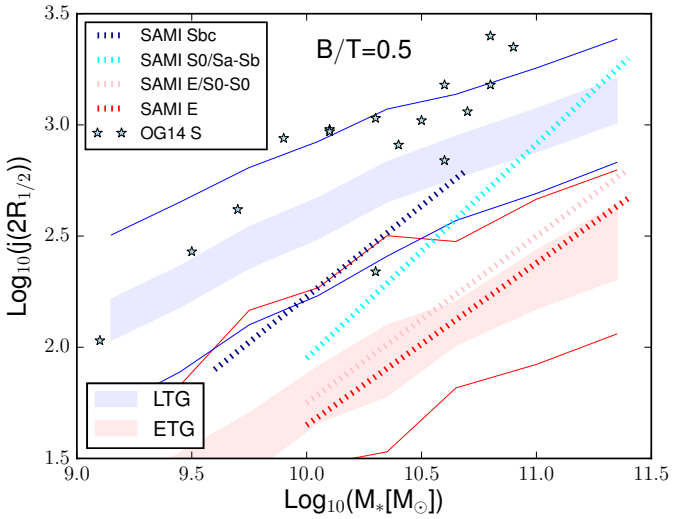

Figure 5.12: The j-M_* relation for LT and ET galaxies (blue and red) from the X17MM model, selected on the basis of their morphology ($B/T = 0.5$). Model predictions are the same as in Fig. 5.10. The observational results shown are: spiral galaxies from Obreschkow and Glazebrook (2014) (stars), and median fits for SAMI galaxies (Cortese et al., 2016, dotted lines). Different colors correspond to different galaxy morphologies, as indicated in the legend.

(shaded areas).

The separation between LT and ET galaxies is evident in all the models. X17CA3 returns predictions that are almost identical to the X17MM run for massive galaxies ($M_* > 10^{10.5}$ M$_\odot$), while at lower masses the former model predicts higher values of j_*. This means that low mass galaxies are more affected by the higher angular momentum acquired during cold accretion, while for high mass galaxies this accretion mode is less important in determining the final value of j_*. We reached similar conclusions when analyzing the $R_{1/2} - M_*$ relation.

X17G11 predicts lower median j_* values for LT galaxies at all masses, with respect to the X17MM run. As mentioned in Sec. 5.3, this feedback scheme causes most of the stars to form earlier than in our the fiducial model. As the angular momentum is lower at earlier cosmic epochs, this leads to a lower normalization of the $j_* - M_*$ relation. For ET galaxies, X17G11 returns predictions consistent with those of X17MM, for $M_* > 10^{10.5}$ M$_\odot$. At lower masses, the angular momentum predicted by the modified feedback model is systematically larger than that obtained using our fiducial X17MM run by ~ 0.2 dex. We have seen in Sec. 5.3.3 that ET galaxies in X17G11 have on average larger sizes that those of the X17MM run. The specific angular momentum integration is influenced by the larger size of the bulges, that translates in larger values of j_*.

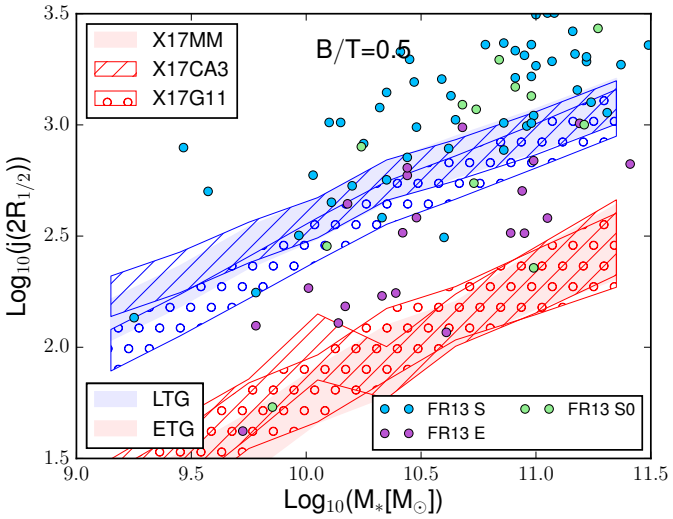

Figure 5.13: The j_*-M_* relation for LT and ET galaxies (blue and red colors), evaluated for galaxies from the X17MM run (shaded area), the X17CA3 run (areas with oblique lines), and the X17G11 run (areas with circles). All models include dissipation during major mergers. LT and ET galaxies were selected using the threshold $B/T = 0.5$. The observational data (circles) are from Fall and Romanowsky (2013).

5.5 Evolution of angular momentum and dependence on other galactic properties

In the previous section, we have demonstrated that the scatter in the j_*-M_* relation correlates both with galaxy morphology and with gas fraction. In this section, we analyze in detail the evolutionary processes that are responsible for this scatter.

5.5.1 Morphological dependencies of the size-mass relation

In this Section, we briefly describe the origin of the LT/ET differentiation for model galaxies. This subject has been discussed, for previous versions of our model, in several studies (De Lucia et al., 2011, 2012a; Wilman et al., 2013). Since the basic results are qualitatively the same, we will focus our analysis on those aspects that are useful to interpret the results presented in Sec. 5.3 and Sec. 5.4.

As explained in Sec. 5.2.2, the model bulge can grow through mergers and disk instabilities. We evaluate the relative importance of these two channels for central galaxies in the X17MM model, in bins of stellar mass and B/T. Table 5.1 lists the fraction of galaxies that have experienced, from z=1 to the present day, no relevant disk instability event (no DI, with relevant we mean an episode characterized by $\delta M_*/M_{*,disk} > 0.1$, where δM_* is the fraction of

Table 5.1: Fraction of galaxies that did not experience a relevant disk instability episode (no DI), a major merger (no MM), or a minor merger (no mM), from $z = 1$ to the present day. Different rows correspond to different stellar mass bins. A further selection is made according to galaxy morphology: $B/T < 0.5$ (first column), $0.5 < B/T < 0.7$ (second column) and $B/T > 0.7$ (third column).

	$B/T < 0.5$			$0.5 < B/T < 0.7$			$B/T > 0.7$		
M_* [$M_\odot$] $\in$	no DI	no MM	no mM	no DI	no MM	no mM	no DI	no MM	no mM
$[10^{9.6};\ 10^{10.2}]$	0.96	0.97	0.79	0.76	0.24	0.83	0.98	0	0.79
$[10^{10.2};\ 10^{10.8}]$	0.93	1.	0.90	0.11	0.89	0.95	0.88	0.07	0.86
$[10^{10.8};\ 10^{11.5}]$	0.91	1.	0.80	0.21	0.93	0.89	0.95	0.03	0.76

stellar disk that is transferred to bulge to restore stability), no major merger (no MM, $M_{sat}/M_{cen} > 0.3$), or no minor merger (no mM, $0.1 < M_{sat}/M_{cen} < 0.3$). Galaxies are binned according to their bulge-to-total ratio: $B/T < 0.5$ (first column), $0.5 < B/T < 0.7$ (second column) and $B/T > 0.7$ (third column), and according to their total stellar mass: low ($M_* \in [10^{9.6} - 10^{10.2}]$ $M_\odot$, first row), intermediate ($M_* \in [10^{10.2} - 10^{10.8}]$ $M_\odot$, second row), and high ($M_* \in [10^{10.8} - 10^{11.5}]$ $M_\odot$, third row).

We show in Fig. 5.14 the evolution of the median of some selected properties as a function of lookback time. In this figure, galaxies are divided according to their final stellar mass, as in the table: low (solid blue lines), intermediate (dashed violet lines), and high (dotted pink lines). Different columns correspond to different B/T bins: $B/T < 0.5$ (left column), $0.5 < B/T < 0.7$ (central column) and $B/T > 0.7$ (right column).

Low stellar mass galaxies exhibit similar evolution for all the B/T selections considered. Their bulges are formed mainly through major mergers, which reflects in an abrupt increase of the B/T value following the merger events. Low-mass LT and ET galaxies reside in halos of similar mass, with halos of ET galaxies only slightly less massive than those hosting LT galaxies ($M_{200}^{LT} \sim 10^{11.5}$ $M_\odot$ and $M_{200}^{ET} \sim 10^{11.3}$ $M_\odot$). On average, halos hosting ET galaxies in this stellar mass bin formed later than those hosting LT galaxies. The former accrete half of their mass 9 Gyrs ago, the latter 10 Gyrs ago. These small differences suggest that ET and LT galaxies in this mass bin belong to the same "halo population", and the differentiation occurs because of the occurrence of major mergers for ET galaxies.

For the median and high stellar mass galaxies, we find more significant differences in their hosting halos. For $B/T < 0.7$, the median halo masses at $z = 0$ are $M_{200}^{med} \sim 10^{11.6} M_\odot$ and $M_{200}^{high} \sim 10^{12.3} M_\odot$ for intermediate and high stellar mass respectively. For $B/T > 0.7$ the numbers are $M_{200}^{med} \sim 10^{12} M_\odot$ and $M_{200}^{high} \sim 10^{12.8} M_\odot$. The $B/T = 0.7$ threshold separates two different galaxy populations, one formed in relatively small halos, less subject to mergers, and the other one formed in more massive halos, that likely experienced more merger events. Mergers between halos trigger mergers between galaxies (after a residual merging time), and galaxies with $B/T > 0.7$ form most of their bulge through mergers. This is confirmed by the fraction of galaxies that did not experience any merger in this B/T range: from 3 to 7%, depending on the stellar mass bin. Galaxies with $B/T < 0.7$, in contrast, form their bulge mainly through disk instability. In this case, the probability of building a relevant bulge depends on

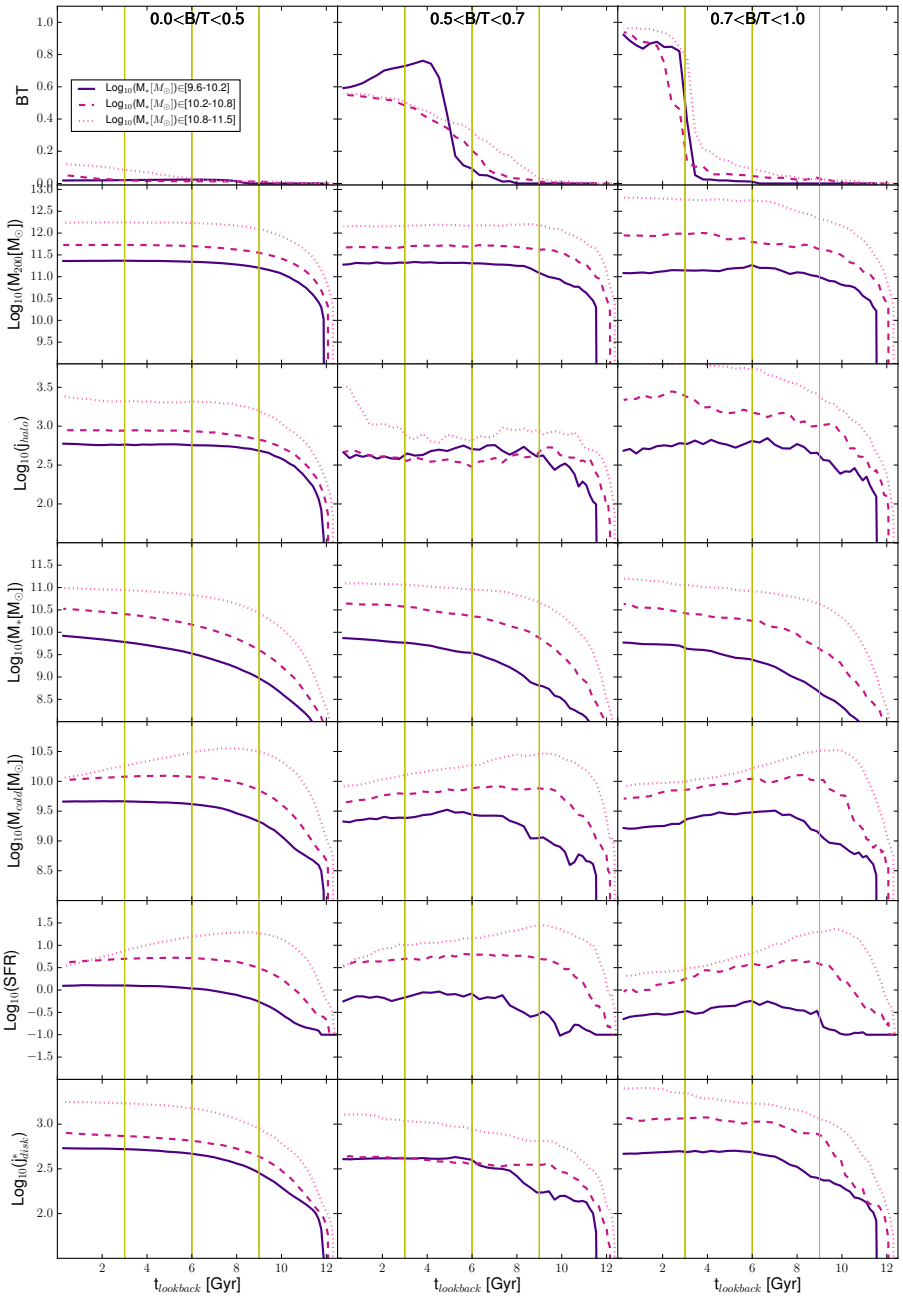

Figure 5.14: Median evolution as a function of look-back time of several galactic properties. From top to bottom: B/T, M_{200}, j_{halo}, M_*, M_{cold}, SFR, and stellar disk specific angular momentum. Model galaxies have been selected according to their stellar mass at redshift 0: $M_* \in [10^{9.6-10.2}]$ (solid blue lines), $[10^{10.2}, 10^{10.8}]$ (dashed violet lines) and $[10^{10.8}, 10^{11.5}]$ $M_\odot$ (dotted pink lines). Different columns correspond to different values of B/T.

the specific history of the galaxy and of its halo. The main difference between galaxies with $B/T < 0.5$ and those with $0.5 < B/T < 0.7$ is in the specific angular momentum of their halos, j_h. Galaxies with a more prominent bulge have a smaller j_h for most of their history. The small j_h is transferred to the cold gas disk through cooling, and then to the stellar disk through star formation. Stellar disks in galaxies with $0.5 < B/T < 0.7$ are thus smaller than those of $B/T < 0.5$ galaxies. This affects the stability of the disk: at fixed stellar mass, halos with smaller j have a higher probability to undergo a disk instability episode. We further discuss the origin and evolution of disk instabilities in the next section.

5.5.2 Disk instability in central and satellite galaxies

In the previous sections, we analyzed the contribution of disk instability to bulge growth, finding it is significant in galaxies with $0.5 < B/T < 0.7$ and $M_* \in [10^{10.2};\ 10^{10.8}]$ $M_\odot$. Fig. 5.14 shows that also the high stellar mass bin is strongly influenced by disk instabilities for $0.5 < B/T < 0.7$, but this fact does not affect the size-mass relation as it does for the intermediate mass bin. This is due to the fact that, in the intermediate mass bin, galaxies with $0.5 < B/T < 0.7$ are more numerous than those with $B/T > 0.7$ ($\sim 1.2/1$), while in the high stellar mass bin the proportions are inverted, with much more galaxies with $B/T > 0.7$ ($\sim 1/2.8$).

In Fig. 5.5, we have shown the size-mass relations obtained considering all ET galaxies, or only the central ones. We find that, in our model central ET galaxies selected using a $B/T = 0.5$ threshold have on average a smaller half-mass radius than the overall population of ET galaxies in the intermediate mass bin. This is in part due to the fact that bulges of central galaxies have a slightly larger contribution form disk instability with respect to those of the overall population. Furthermore, bulges and disks in central galaxies that underwent disk instabilities have smaller sizes than those formed in satellites of the same mass and B/T. We discuss the origin of these differences in the following.

We show in detail the evolution of two representative galaxies that underwent disk instabilities in Fig. 5.15: we show the evolution of the specific angular momentum and mass of the gaseous disk (left column) and of the stellar disk (right column). Each segment of the lines shows a variation of specific angular momentum and stellar mass corresponding to one code time-step, color coded according to the process that determined a change of these properties. Top and bottom panels correspond to a central and a satellite galaxy with $10^{10.2} < M_* < 10^{10.8}$ $M_\odot$ and $0.5 < B/T < 0.7$.

Let us focus first on the evolution of the central galaxy (top panels). At early times, the mass and the specific angular momentum of the gaseous disk grow due to cooling (black lines). The specific angular momentum of the cold gas is unaffected by star formation and stellar feedback (red and magenta lines), and decreases due to recycling (orange lines). When $M_{cold} \sim 0.2 \cdot 10^{10}$ $M_\odot$ and $j_{cold} \sim 300$ kpc/h km/s, the specific angular momentum starts decreasing. This must be due, by construction, to a decrease of the specific angular momentum of the DM parent halo. This decrease is only partially compensated by recycling from older stellar populations, that have higher specific angular momentum.
The evolution of the stellar disk of the central galaxy (right top panel) follows that of the cold gas, through star formation (red lines). The specific angular

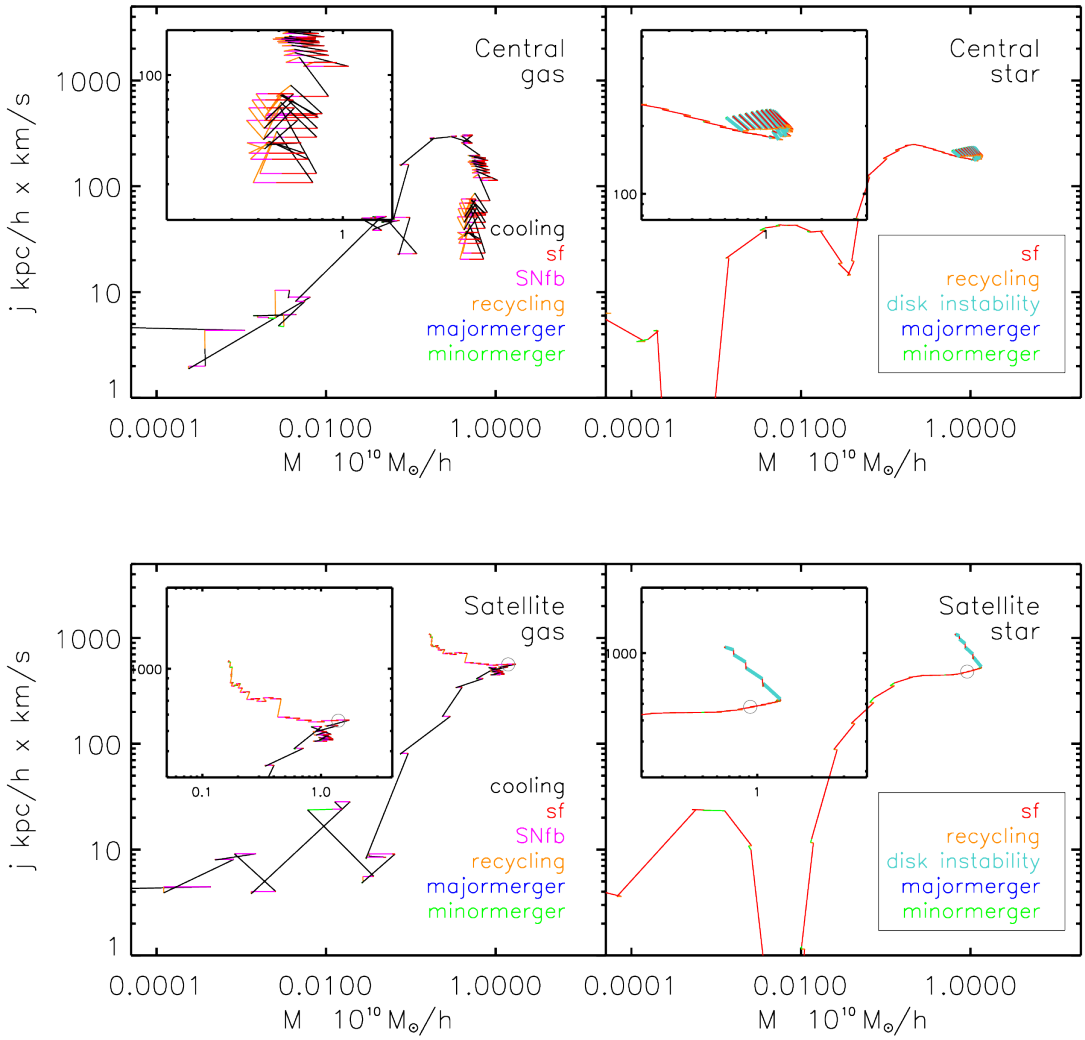

Figure 5.15: Evolution of the specific angular momentum and the mass of two representative galaxies with $10^{10.2} < M_* < 10^{10.8}$ $M_\odot$ and $0.5 < B/T < 0.7$. The top panels show the evolution of a central galaxy, while the bottom panels are for a satellite. The left column shows the evolution of specific angular momentum of the cold gas disk, and the right column that of the stellar disk. The solid lines represent variations of the specific angular momentum and stellar mass of the galaxies considered, at each code time-step, color coded according to the physical process causing them (see legend).

momentum decreases significantly twice, because of merger events. When $M_* \sim 0.2 \cdot 10^{10}$ $M_\odot$ and $j_{*,disk} \sim 200$ kpc/h km/s, the stellar disk starts loosing specific angular momentum because stars are forming from gas with decreased angular momentum (compare right and left panel). This contraction of the stellar disk continues until the disk becomes unstable: the disk instability is triggered when a parameter proportional to $\sim R_{*,disk}/M_{*,disk}$ becomes lower than 1. As described in Sec. 5.2.2, during disk instabilities part of the disk mass is transferred to the bulge to restore the stability, but the disk specific angular momentum (and size) is increased because of angular momentum conservation. These events are visible, in the top right panel, as light blue lines. During subsequent star formation episodes, the stellar disk mass increases again, but the specific angular momentum decreases, following that of the cold gas. Then a new disk instability is triggered. The size of the galaxy oscillates due to a series of consecutive disk instability events.

In the case of the satellite galaxy (bottom panels), the early evolution of the gaseous and stellar disks is similar to that of the central galaxy. This is not surprising, because, at these early stages, the satellite is still a central galaxy, evolving in its own parent DM halo. The accretion event (the last time the galaxy is a central) is marked as a gray circle in the figure. Since in our model we assume that gas cooling occurs only on central galaxies, after accretion the angular momentum of the cold gaseous disk cannot be affected anymore by gas cooling (see inset in bottom left panel). It keeps increasing due to star formation and gas recycling from older stellar populations. As for the central galaxy, the stellar disk follows the evolution of the cold gas due to star formation. After accretion, star formation continues, but at lower rates, because there is no replenishment of the cold gas reservoir via gas cooling. The specific angular momentum of the stellar disk increases slightly. Also in this case, the instability criterion is eventually met, and the satellite undergoes a disk instability event. Similarly to the case of the central galaxy examined above, the satellite enters a recursive cycle of star formation and disk instability, but in this case the specific angular momentum keeps growing. This is due to the increase of the specific angular momentum of the cold gaseous disk, which reflects in an increase of the specific angular momentum of the stellar disk during star formation. In addition, we assume that angular momentum is conserved during disk instability events, which translates in an increase of the specific angular momentum. This sequence of events stops only when the star formation stops, because the cold gas available for star formation is exhausted.

The examples discussed show that disk instabilities lead to a net increase of the size of satellite galaxies, while they do not affect significantly the size of central galaxies.

5.5.3 The role of cold gas in the dynamical evolution

In Fig. 5.11, we have shown the dependence of the specific angular momentum-mass relation on the cold gas fraction of model galaxies. In this section, we investigate in detail the origin of this dependence, by analyzing the evolution of gas-rich and gas-poor galaxies. We show the median evolution of some of their properties in Fig. 5.16, as we did in Fig. 5.14. We divide our model galaxies in the same stellar mass bins at z=0: low ($M_* \in [10^{9.6} - 10^{10.2}]$ $M_\odot$, solid), intermediate ($M_* \in [10^{10.2} - 10^{10.8}]$ $M_\odot$, dashed), and high ($M_* \in$

$[10^{10.8} - 10^{11.5}]$ $M_\odot$, dotted lines). For each stellar mass bin, we select galaxies belonging to three different bins of B/T, to separate LT ($B/T < 0.5$) from ET galaxies, formed mainly through disk instabilities ($0.5 < B/T < 0.7$) and through mergers ($B/T > 0.7$). Different columns correspond to different B/T values. For each of these nine samples, we select galaxies belonging to the extremes of the f_{cold} distribution: i.e. we consider galaxies with f_{cold} smaller than the $10^{\rm th}$ percentile of the distribution as gas-poor, and those with f_{cold} larger than the $90^{\rm th}$ percentile as gas-rich. Lines are color-coded according to the stellar mass bin and the cold gas fraction, as indicated in the legend.

We first focus on the LT galaxies (left column). In all the mass bins, gas-poor galaxies are hosted by halos that form earlier than those of gas-rich galaxies. The halos of gas-poor galaxies grow rapidly in mass, and acquire most of their angular momentum during this phase of rapid accretion. Halo mass, in particular that of low mass galaxies, decreases from 11 Gyrs ago to recent times: this is due to the fact that these model galaxy samples are dominated by satellite galaxies whose parent halo mass decreases (due to stripping) after accretion. We have verified that considering only central galaxies, halo mass remains approximately constant after the peak following the rapid mass accretion phase. The accretion history of halos hosting gas-poor galaxies translates in large amounts of cold gas in these galaxies at early times, which triggers significant early star formation. Most of the stellar mass of gas-poor galaxies is thus formed during the star formation peak, that occurred between 9 and 11 Gyrs ago. At this time, the specific angular momentum of the cold gas is low, as it follows that of the DM halo.

Gas-rich galaxies, in contrast, are hosted by halos that formed more recently than those hosting gas-poor galaxies. These halos accrete their mass more gradually, and their mass increases down to very recent times. As a consequence, star formation occurs on a longer interval of time, and the stellar disk can acquire the high specific angular momentum the halo has at late times. In summary, gas-rich LT galaxies are characterized by a larger specific angular momentum than their gas-poor counterparts, because of a larger contribution from recent star formation. This, in turn, is due to the different accretion histories of their hosting halos.

When considering the other bins in B/T, we find an evolution similar to that found for LT galaxies. A different evolution is found for the j_{halo} of gas-poor low stellar mass ET galaxies: these have, at late times, a j_{halo} larger than that of gas-rich ET galaxies of similar mass. This is due to the significant contribution of satellite galaxies to this gas-poor stellar mass bin. For these satellite galaxies, j_{halo} is poorly constrained. When considering only central gas-poor ET galaxies, the j_{halo} evolution has a behavior similar to that of LT gas-poor galaxies in this mass bin. The main difference between LT and ET galaxies is in the cold gas content of gas-poor galaxies: the gas-poor ET galaxies, at redshift zero, contain less gas than LT galaxies. Another difference is in the bulge formation times of gas-poor and of gas-rich ET galaxies. The bulges of gas-poor galaxies form on average 2-3 Gyrs earlier than those of gas-rich galaxies. This is a selection effect, due to the fact that galaxies with higher star formation rates can regrow a disk more efficiently. As star formation correlates with the gas content, the star formation rates are larger in gas rich galaxies, that can quickly regrow a significant disk component. A certain B/T threshold selects gas-poor ET galaxies that formed their bulge earlier than gas-rich galaxies. This is because

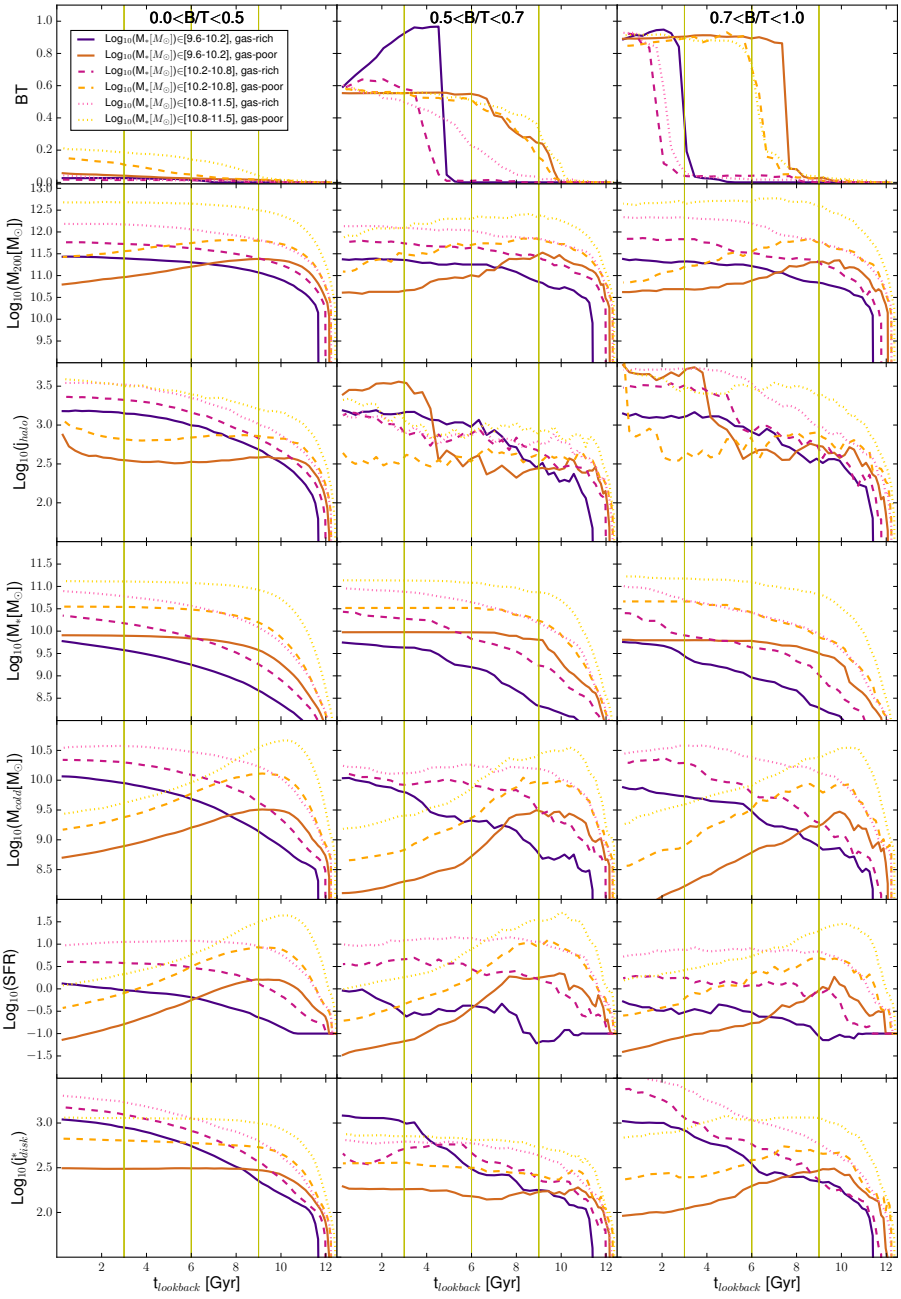

Figure 5.16: Median evolution as a function of look-back time of several galactic properties. From top to bottom: B/T, M_{200}, j_{halo}, M_*, M_{cold}, SFR, and stellar disk specific angular momentum. Model galaxies have been selected according to their stellar mass at redshift 0: $M_* \in [10^{9.6-10.2}]$ (solid lines), $[10^{10.2}, 10^{10.8}]$ (dashed lines) and $[10^{10.8}, 10^{11.5}]\mathrm{M}_\odot$ (dotted lines). A further binning is made as a function of the cold gas fraction f_{cold}, with gas-rich (indigo, purple and pink) and gas-poor galaxies (chocolate, orange and gold) selected as those above or below the upper or lower 10^{th} percentile of the distributions, respectively. Different colors are used for galaxies of different stellar mass, as indicated in the legend.

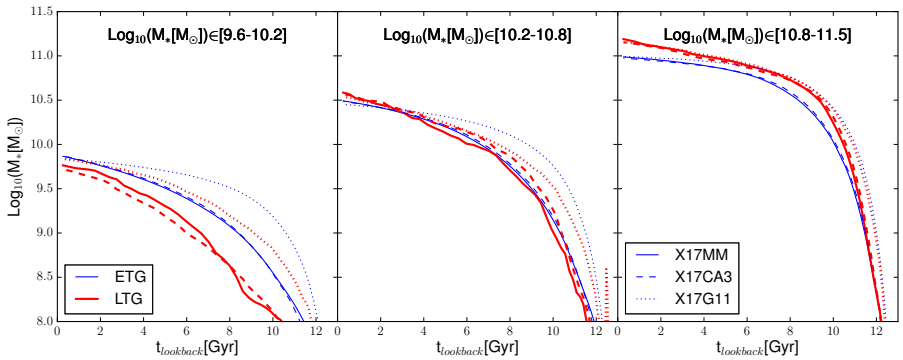

Figure 5.17: Median evolution as a function of look-back time of the total stellar mass, for galaxies selected in stellar mass bins: $M_* \in [10^{9.6-10.2}]$ (left panel), $[10^{10.2}, 10^{10.8}]$ (central panel) and $[10^{10.8}, 10^{11.5}]\mathrm{M}_\odot$ (right panel). Galaxies are classified as LT (blue) and ET (red) according to their B/T, with a threshold $B/T = 0.7$. Different linestyles correspond to different models: solid for the X17MM model, dashed for X17CA3, and dotted for X17G11.

the latter, in the meanwhile, have already regrown their disks.

5.5.4 Evolution in X17CA3 and X17G11

In this section, we study the origin of the different size-mass and specific angular momentum-mass relations for the fiducial model X17MM and its variants X17CA3 and X17G11. In Fig. 5.17 and 5.18, we show the evolution of the total stellar mass and of the stellar disk specific angular momentum for ET (red) and LT (blue) galaxies, divided using a $B/T = 0.7$ threshold. Predictions from different models are shown using different line styles: X17MM with solid, X17CA3 with dashed, and X17G11 with dotted lines.

In the X17CA3 run, where the gas cooled in rapid mode has an angular momentum three times larger than in X17MM, the size-mass relation is slightly shifted upwards at small stellar masses ($M < 10^{10.2}$ $\mathrm{M}_\odot$) with respect to X17MM. The same applies for the relation between specific angular momentum and galaxy mass. When considering the evolution of X17CA3 galaxies (dashed lines in the figures), we see that their specific angular momentum is higher than in the X17MM model during the first 2-3 Gyrs, as expected. This, however, does not imply a higher star formation rate, because a larger j_{cold} translates in a larger disk radius and, as a consequence, a lower surface density. As a result, the stellar mass of galaxies in the X17CA3 run evolves as in X17MM, although the specific angular momentum of the stellar disk is much higher at early epochs (j_* follows j_{cold} through star formation). After 2-3 Gyrs, cold accretion is no more the dominating accretion mode, and the specific angular momentum of the cold gas converges to the same values found in the X17MM run, retaining only a small difference in the final values.

For the X17G11 run, we found a lower final specific angular momentum for LT galaxies, and, as a consequence, a lower normalization of the size-mass relation. In Fig. 5.17, we find that X17G11 galaxies form the bulk of their stars

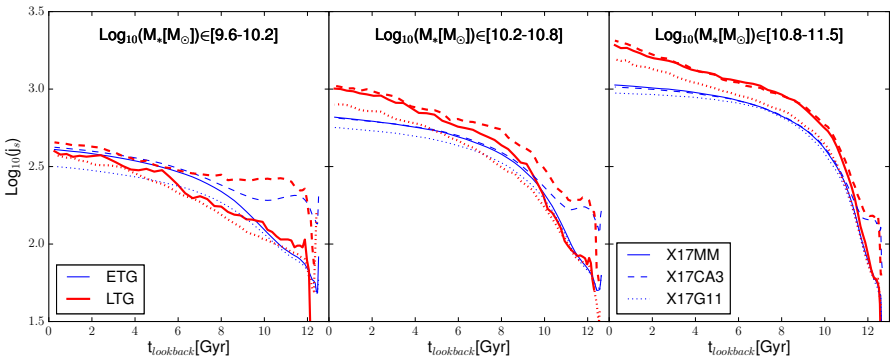

Figure 5.18: Median evolution as a function of look-back time of the stellar disk specific angular momentum, for galaxies selected in stellar mass bins: $M_* \in [10^{9.6-10.2}]$ (left panel), $[10^{10.2}, 10^{10.8}]$ (central panel) and $[10^{10.8}, 10^{11.5}]\mathrm{M}_\odot$ (right panel). Different colors and linestyles are as in Fig. 5.17.

at earlier times compared to X17MM, for all the mass and morphology bins considered. The feedback scheme adopted in the X17G11 run allows ejected gas to reaccrete earlier than in X17MM, resulting in higher quantities of gas cooling onto the model galaxies in the first 1-2 Gyrs. This translates in significant early star formation. After the initial peak, star formation gradually decreases, while in the fiducial model it remains almost constant to the present day. This is because in the X17MM run, ejected gas is reaccreted more gradually to present day. Therefore, most of the stars in the disk of X17G11 galaxies form from gas with rather low specific angular momentum.
In Sec. 5.3.3, we have shown that ET galaxies in the X17G11 run have, on average, a larger size than those of similar mass in the X17MM run. This seems in contrast with what we found for LT galaxies, that have systematically lower sizes, due to the early peak of star formation. In fact, ET galaxies follow a similar evolution, as shown in Fig. 5.17 and 5.18. The reson for this size difference is the different nature of the bulge. In the X17G11 run, disk instabilities are less frequent than in X17MM at late times, because of the early peak of star formation. In contrast, mergers occur at all times. We find that in X17G11 the bulges formed mainly during mergers are more numerous than those in X17MM. The average bulge size in the X17G11 run is, as a consequence, larger than in X17MM run, shifting upwards the size-mass relation. The specific angular momentum of X17G11 ET galaxies is influenced by the larger bulge sizes, because, at fixed mass, the bulge mass profile is shallower. This enhances the importance of rotational velocity at large radii in the integration performed to calculate j_*. The rotational velocity increases with radius, and the integrated final specific angular momentum is thus higher than in the X17MM model.

5.6 Conclusions

In this work, we have analyzed the dynamical properties of galaxies in the framework of a state-of-the-art semi-analytic model. We take advantage of the GAEA semi-analytic model (Hirschmann et al., 2016), as updated in Xie et al.

(2017). This model includes prescriptions to follow the angular momentum exchanges among galactic components, and evaluates disk sizes from specific angular momenta, as in Guo et al. (2011). We extend our analysis to the specific angular momentum of galaxies, a topic poorly studied in the semi-analytic framework, with the exception of dedicated models (Stevens et al., 2016).

Previous studies focused on the size-mass relation at relatively high masses ($M_* > 10^{10}$ M$_\odot$), highlighting the necessity for a specific treatment for cold gas dissipation during mergers to obtain realistic bulge sizes. We have included this additional process in our model, and estimated the half-mass radius of model galaxies by projecting the bulge and the disk mass profiles. The mass profiles adopted are those assumed by the model. To estimate the specific angular momentum, we evaluate the rotational velocity profile of the galaxy disk assuming it is symmetric and fully rotationally supported. Finally, we obtain the specific angular momentum integrating the mass and velocity profiles out to two half-mass radii. These estimates of size and specific angular momentum can be compared to available observational measurements.

For late type galaxies, our predicted size-mass relation is in fairly good agreement with recent observational estimates. On the other hand, dissipation during mergers is necessary to correctly reproduce the size of early type galaxies, especially at low stellar masses ($M_* < 10^{10}$ M$_\odot$). We find that, in our model, dissipation must be limited to major mergers, otherwise the bulge size would be under-predicted. This assumption is reasonable since the adopted prescription was extracted from dedicated binary merger simulations, with a quite large mass ratio ($> 1/6$). Our results are in agreement with those by Shankar et al. (2013), who applied the same treatment for dissipation to the semi-analytic model by Guo et al. (2011). Shankar et al. (2013) also considered dissipation only during major mergers and found that early type model galaxies have sizes in good agreement with observations. Similar results were obtained by Tonini et al. (2016), who accounted for dissipation in their semi-analytic model, using a different prescription, but applied to all mergers.

Interestingly, the selection of late/early type galaxies influences strongly the predicted size-mass relation. A selection based on specific star formation rate ($\log_{10}(sSFR) = -10.6$) leads to too large early type galaxies. This is is due to the large number of quiescent disky satellites selected as ET galaxies: they have large stellar disks, but are not star forming because of the stripping of their hot gas halo, that prevents re-fueling of cold gas and new star formation.
When galaxies are selected on the basis of their morphology, using the bulge-over-total mass ratio ($B/T = 0.5$ or 0.7), the size-mass relations for late and early type galaxies are well separated, and are in good agreement with observations. Different B/T thresholds lead to slightly different relations for early type galaxies. We analyze the different contributions to B/T from bulges formed mainly through disk instabilities and those formed mainly through mergers, finding the former are prevalent for galaxies with $0.5 < B/T < 0.7$, and the latter for $B/T > 0.7$. The sizes of model bulges formed through mergers are in nice agreement with observational estimates, while disk instabilities produce systematically smaller bulges. When we analyze the evolution of galaxies undergoing disk instabilities, we find that these events are typical of halos that suffer, during their life-time, a decrease of specific angular momentum. As the hot gas halo is assumed to have the same specific angular momentum of the

DM halo, cooling transfers this loss of specific angular momentum to the cold gas disk. Cold gas transfers the acquired lower specific angular momentum to the stellar disk through star formation, eventually triggering a disk instability event. In satellite galaxies, disk instabilities tend to result in a net increase of stellar disk sizes. For central galaxies, the size of the stellar disk component is not significantly affected by recursive events.

The unrealistic sizes obtained for bulges formed through disk instability suggest a revision of the adopted prescriptions. The current instability criterion is based on old simulations (Efstathiou et al., 1982), and the treatment of the instability is not consistent with the results of recent N-body simulations, as discussed in several studies (Athanassoula, 2008; Benson and Devereux, 2010; De Lucia et al., 2012a). At the moment, no alternative disk instability prescription that produces more realistic galaxy properties has been proposed.

Accounting explicitly for two distinct bulge components would alleviate, in part, the problem with the size of disk instability bulges. In fact, observations highlight the existence of two categories of bulges. Classical bulges are dispersion dominated spheroids, that are believed to form during mergers, while pseudo-bulges are rotationally supported thick disks, likely formed during secular processes, as disk instabilities. An explicit separation between classical and pseudo bulges in the model (as done by Tonini et al., 2016) would allow the creation of larger bulges during disk instabilities, because disk-like pseudo-bulges would have a shallower profile.

Furthermore, a rotating pseudo-bulge would allow for an explicit treatment of fast and slow rotators. Such a treatment is important to estimate the relation between specific angular momentum and mass for early type galaxies. The relation obtained for our model is slightly below observational measurements, likely because our model does not account for rotating bulges. We find that the cold gas content strongly correlates with the specific angular momentum, with gas-rich galaxies having also a high specific angular momentum. The reason is in the different evolution of gas-poor and gas-rich galaxies. Gas-poor galaxies form in halos that acquire most of their mass earlier than those hosting gas-rich galaxies. Therefore, gas-poor galaxies have a peak of star formation at early times, during which they acquire most of their specific angular momentum. Gas-rich galaxies, instead, continue to form stars gradually to present day, acquiring their specific angular momentum more recently, with respect to gas-poor galaxies.

Our analysis confirms that the dynamical properties of galaxies depend strongly on the galaxy star formation and assembly history. We have also tested the dependence of the sizes and the angular momenta of model galaxies on the adopted stellar feedback scheme. We consider, in particular, an alternative scheme where the ejected gas is re-accreted earlier than in our fiducial model. In this case, the bulk of star formation occurs at early times, and the specific angular momentum is mostly acquired at that moment, when it is still low. We analyze also the influence of a higher specific angular momentum for gas cooled in the cold mode. In this case, the star formation history of model galaxies is not strongly modified. Only small galaxies, those most influenced by cold accretion, are affected: they have slightly larger sizes and angular momenta compared to predictions from our reference model.

In conclusion, our model is able to reproduce both the sizes and the specific angular momenta of galaxies in the local Universe. Considered that the model

does not evolve self-consistently the spatial distribution and the internal dynamics of galaxies, this is an engaging result, that demonstrates the feasibility of this kind of studies in the semi-analytic framework.

A further development of our work will include an explicit treatment for pseudo-bulges, allowing a direct comparison of model results with recent integral field spectroscopic surveys, as ATLAS^{3D} and SAMI.

Chapter 6

Conclusions and future perspectives

In this Thesis work, I have analyzed the role of cold gas in the evolution of galaxies, in the framework of state-of-the-art semi-analytic models of galaxy formation.

Using six independently developed models, based on the same cosmological simulation, I have investigated the prescriptions that mostly influence the galaxy final HI content. With this purpose in mind, I have compared the HI content of model galaxies to observational data, analyzing several statistics: the HI mass function, the HI-stellar mass relation, the 2-point correlation function of HI-selected galaxies, and the dependence of the HI content on halo mass and spin parameter. This analysis has shown that the HI amount at high stellar masses depends strongly on the star formation and stellar feedback prescriptions adopted. These processes also affect the HI content in satellite galaxies, that was for long believed to be mainly regulated by the adopted treatment for gas stripping. In this work, I have demonstrated that there are plausible feedback schemes in which the balance between star formation and stellar feedback is such to retain large quantities of cold gas in galaxies at high redshift, making the hot gas stripping treatment less relevant in the determination of the final amount of HI. This analysis has also highlighted the importance of HI-poor galaxies, as they strongly influence the statistics and the clustering analysis. The simulations adopted in this Thesis have a resolution that does not allow a complete sampling of galaxies at low HI masses. To provide further constraints on the prescriptions that better reproduce observational measurements, it would be useful to apply the analyzed models to higher resolution simulations.

My comparison work on the HI content has indicated two models that are able to produce realistic HI statistics: those described in Henriques et al. (2015) and in Xie et al. (2017). The latter is the only one, among the models considered in this work, that includes an explicit partition of cold gas into its HI and H_2 components, and a star formation law based on the molecular gas surface density, in agreement with recent observational data (Wong and Blitz, 2002; Kennicutt et al., 2007; Leroy et al., 2008; Bolatto et al., 2011; Krumholz, 2013). This model has been developed by our Trieste group, and includes also a complex chemical enrichment scheme, based on a non instantaneous recycling approxi-

mation (De Lucia et al., 2014b), an improved stellar feedback scheme, partially parametrized on the results of hydrodynamical simulations (Hirschmann et al., 2016), and a treatment for specific angular momentum evolution of the stellar and gaseous disks (as that implemented in Guo et al., 2011).

I have exploited this latter implementation in my analysis of the dynamical properties of model galaxies. I have included in the Xie et al. (2017) model a specific treatment for gas dissipation in major mergers, that made the model able to reproduce the size-mass and the specific angular momentum-mass relations predicted by the most recent observations, as well as the scatter in these relations due to morphology.

I have analyzed the contribution of each galactic component to the size-mass relation, evidencing that model bulges formed through disk instabilities are unrealistically small, in contrast with observations. This suggests the necessity of a revision of the disk instability treatment and of the model adopted for bulge sizes formed through this channel.

I have also analyzed the dependence of specific angular momentum on the cold gas fraction, finding a strong correlation, especially in disk dominated galaxies. This dependence originates from the different star formation histories of gas-rich and gas-poor galaxies: the former continue producing stars to present time, while the latter have a peak of star formation at early times. The model assumes that the hot gas halo has the same specific angular momentum of the DM halo, and transfers it to the cold gas disk through cooling. This, in turn, transfers its specific angular momentum to the stellar disk through star formation. In this way, the stellar disk acquires its specific angular momentum when most of its star formation takes place. An early star formation means a small specific angular angular momentum, as that of the DM halo at early times.

I have studied how the dynamical properties of model galaxies are affect by an alternative stellar feedback scheme (that described in Guo et al., 2011). This feedback scheme favors an earlier star formation with respect to the fiducial model, because it lets more ejected gas to be reaccreted at early times. As in the case of gas-poor galaxies, an early peak of star formation results in smaller final disks, with lower specific angular momenta.

I have also studied the influence of a different initial specific angular momentum for cold gas during cold accretion, finding that only low mass galaxies are sensitive to this prescription. Higher resolution simulations would be beneficial to extend the analysis to lower mass galaxies and better quantify the influence of this prescription in this mass range.

In the following sub-sections, I discuss more in detail projects and analyses that originated from the work carried out for this PhD Thesis, and that are currently under development. In particular, I explain how to exploit the abilities of the used model to create mock catalogs suitable for survey planning, a project that is currently ongoing. Furthermore, I focus on the issues raised during this work, suggesting strategies adoptable for their solution.

6.1 Ongoing work: mock catalogs

In the first part of this work, I have developed software that creates mock light-cones from the semi-analytic model outputs (see Sec. 4.2.3). This software positions galaxies in the light-cone avoiding structure replications along the line

of sight, and selecting them at the evolutionary stage corresponding to their distance from the observer.

This software allows the creation of light-cones of desired dimensions, starting from simulations of any volume. The combination of these parameters (mock dimensions and simulation volume) must be chosen according to the scientific goal. If the goal is a statistical study of large volumes of the sky, the simulation volume must be large, otherwise the mock light-cone would be replicating many times the same small volume simulation. On the other hand, a study focused on the properties of small galaxies requires the use of high resolution simulations that will have necessarily also a smaller volume.

Semi-analytic models are the ideal tool to mock large volumes of Universe at a relatively high resolution. Also the HOD approach, discussed in Sec. 2.2.3, allows the construction of large simulated volumes at an even lower computational cost. This method is convenient, in particular, for testing observational strategies or analysis tools on controlled mocks constrained to reproduce a specific set of observational data. Conversely, HOD is not the ideal choice when the mocks are used to provide a physical interpretation of the data. In fact, HOD is based on a purely statistical approach that bypasses an explicit treatment of the physical processes at play. These are, instead, explicitly included in semi-analytic models (as well as in hydrodynamical simulations). Finally, in both cases, an extrapolation is necessary in regimes where observational data are not available yet.

In this context, the Xie et al. (2017) model I analyzed in detail during this Thesis work is the ideal tool to create mock light-cones for various survey. It is able to reproduce the observed evolution of the galaxy stellar mass function up to $z \sim 10$, as well as the measured correlation between stellar mass/luminosity and metal content of galaxies in the local Universe down to the scale of Milky Way satellites (De Lucia et al., 2014b; Hirschmann et al., 2016; Fontanot et al., 2017b), and the evolution of the gas metallicity-mass relation up to redshift $z \sim 2$ (Hirschmann et al., 2016; Xie et al., 2017). For this reason, it can be used for the creation of mock catalogs for high redshift surveys. Moreover, the model ability to reproduce the disks dynamics and their cold gas content suggests it as the ideal instrument to create realistic catalogs of the neutral atomic hydrogen 21 cm line, suitable for survey planning and data interpretation in the framework of the Square Kilometre Array, and its precursors MeerKAT and ASKAP (Schilizzi et al., 2008; Johnston et al., 2008; Booth et al., 2009).

I discuss in detail these two model applications in the following subsections.

6.1.1 Mock catalogs for VANDELS

One activity that I have already completed is the creation of mock light-cones for the ongoing survey VANDELS (McLure et al., 2017). For this purpose, I have applied the software for light-cones creation to the outputs of the model described in Hirschmann et al. (2016).

The VANDELS survey is a deep spectroscopic survey of high redshift galaxies, performed at the VLT. It provides multi-wavelength imaging and near-IR grism spectroscopy for 2560 galaxies in the CANDLES UDS and CDFS fields (Grogin et al., 2011; Koekemoer et al., 2011; Xue et al., 2011; Galametz et al., 2013). These data will be used to study metallicities and other galactic properties, with the goal to better understand the physics of galaxies in the early Uni-

verse. Different targets are distributed in different redshift ranges: $2.5 < z < 5.5$ for star forming galaxies, $1.5 < z < 2.5$ for passive galaxies, and $3 < z < 7$ for faint star forming galaxies.

The mock light-cones I have created for VANDELS cover the redshift range from $z = 0$ to $z = 7.1$, and have an angular size compatible with that of the observed fields (with an aperture of 0.5 deg).

I have provided, among various other physical properties of the galaxies in the mocks, the observed apparent magnitudes in a set of filters selected by the VANDELS working group. The semi-analytic model calculates photometric properties of model galaxies assuming a Chabrier initial mass function and using the stellar population synthesis models by Bruzual A. and Charlot (1993), as described in De Lucia et al. (2004). When star formation occurs, the model attributes to the newly formed stellar population a luminosity contribution, that is extrapolated from look-up tables. These tables, created using the population synthesis model, contain tabulated information about the luminosity of a single star-burst of fixed mass, as a function of the age of the stellar population and of its metallicity. Tables are interpolated linearly in time and logarithmically in metallicities. The photometric properties of model galaxies can be evaluated in the rest frame or in the observer frame, using the same interpolation method. For the light-cones, typically observer-frame magnitudes are required. Let us call $M[z(n)]$ the absolute magnitude corresponding to a model galaxy at a snapshot n and redshift $z(n)$. In the light-cone, a galaxy from the snapshot n can be positioned at a redshift z between $[z(n-1) + z(n)]/2$ and $[z(n+1) + z(n)]/2$, the edges of the simulated box corresponding to the snapshot n. This means that the galaxy is not perfectly placed at $z(n)$, and its magnitude must be corrected for this. I use a correction similar to that used in Blaizot et al. (2005), and assume that the magnitude in a selected filter varies linearly with redshift. The apparent magnitude m, measured in the selected filter from the observer at the origin of the light-cone, can thus be expressed as:

$$m = M[z(n)] + \frac{M[z(n+1)] - M[z(n)]}{z(n+1) - z(n)}[z - z(n)] + 5\log_{10}(D_L(z)) \qquad (6.1)$$

where the derivative of magnitude with redshift is evaluated using the magnitudes in two subsequent snapshots, and $D_L(z)$ is the galaxy luminosity distance from the observer.

As an example of the mocks generated, I show in Fig. 6.1 one of the mock light-cones I created, with galaxies color-coded according to their apparent magnitude in the i band.

Magnitudes are accompanied by galaxy position in the cone, as right ascension, declination, observed redshift, real redshift, and by other galaxy properties like absolute magnitudes, stellar mass, cold gas mass, star formation rate, disk radius, galaxy type (central or satellite), and unique IDs that allow galaxy to be linked to its parent substructures and FoF halo. These latter information allows the VANDELS group to analyze galaxy properties as a function of their environment.

Mock catalogs can be used to define the data analysis strategies, and to carry out a direct comparison between observed and real galactic properties, thereby providing a physical interpretation of observational data. At the moment, the VANDELS group is testing the light-cones against basic statistical properties to

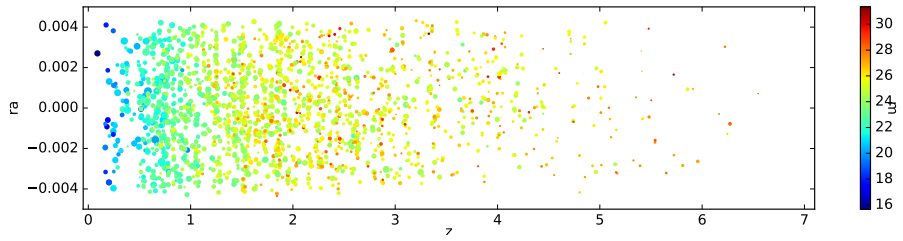

Figure 6.1: One of the mock light-cones created for VANDELS, with galaxies distributed according to observed redshift and right ascension. Galaxies are color-coded according to their apparent magnitude in the i band, and their symbols have a size proportional to their disk radius. For clarity, I show only 2000 randomly selected galaxies, with $M_* > 10^9 \mathrm{M}_\odot \mathrm{h}^{-1}$.

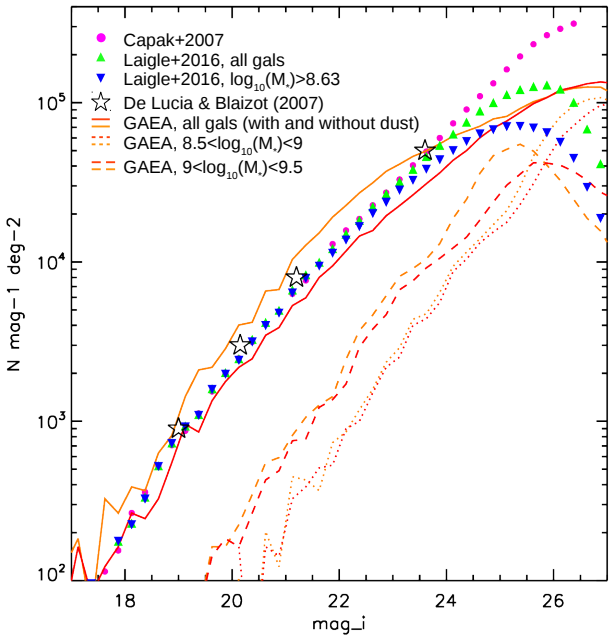

Figure 6.2: Number counts of galaxies in the cone volume, as a function of their i apparent magnitude. Predictions from the GAEA model are shown as red lines (including dust extinction), solid for the total sample, dashed for galaxies with $9 < \log_{10}(M_*[\mathrm{M}_\odot]) < 9.5$, and dotted for $8.5 < \log_{10}(M_*[\mathrm{M}_\odot]) < 9$. Predictions based on the model by De Lucia and Blaizot (2007) are shown, as a comparison, as stars. Model results are compared to the observational data from COSMOS, for the first release (Capak et al., 2007, violet circles) and for the last one (Laigle et al., 2016, triangles). Courtesy of the VANDELS group.

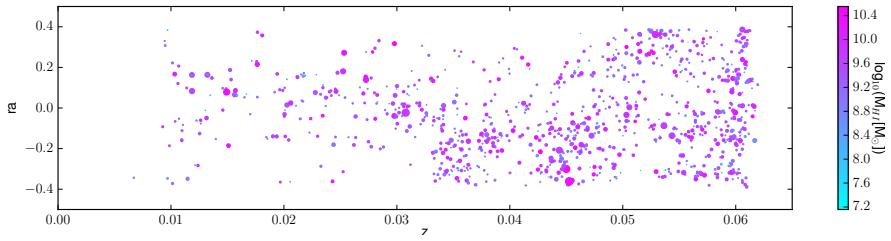

Figure 6.3: One of the mock light-cones used in Sec. 4.4, with galaxies distributed according to observed redshift and right ascension. Galaxies are color-coded according to their HI content, and the symbol size is proportional to the galaxy disk radius. For clarity, I show only 1000 randomly selected galaxies, with $M_* > 0.5 \cdot 10^{10} \mathrm{M}_\odot \mathrm{h}^{-1}$.

understand the model limits, such as mass resolution, and quantify discrepancies between model and real data properties. One of the preliminary tests performed is the comparison of galaxy number counts as a function of magnitude with the most recent observational estimates. The results of this test are shown in Fig. 6.2. The GAEA model (red solid line) slightly under-predicts the observed counts for the survey COSMOS (Capak et al., 2007; Laigle et al., 2016, violet circles and green triangles). This can be due to the incompleteness of model galaxies below the stellar mass resolution $M_* \sim 10^9$ $\mathrm{M}_\odot$, while the observed sample includes galaxies with mass $M_* > 10^{8.6}$ $\mathrm{M}_\odot$.

6.1.2 Mock catalogs for SKA

As described above, the software for mock light-cones creation provides the ideal instrument to build catalogs of galaxies for the SKA (and its precursors). In Fig. 6.3, I show one of the cones I used for the analysis of the 2-point correlation function of HI-selected galaxies, in Sec. 4.4. Galaxies are color coded according to their HI mass, and their symbol size is proportional to their disk radius. The employment of this kind of mock light-cones in the development of survey strategies for SKA requires a further processing of the model outputs, to provide the same quantities measured in observations. Specifically, in this particular case, the missing ingredient is a realistic evaluation of the 21 cm line emitted by the HI in the disk of model galaxies. This can be achieved using the information available in the model outputs. Xie et al. (2017) provide both the HI content and the dynamical properties of the cold gas disk, necessary to calculate the line intensity and the line broadening due to the rotation of the disk. The inclination of the cold gas disk and its scale radius, that further influence the measured 21 cm line, can be consistently evaluated from the vector describing the specific angular momentum of the cold gas disk.

I plan to use model outputs to reconstruct the 21 cm line following Obreschkow et al. (2009a,b). These studies used the semi-analytic model described in De Lucia and Blaizot (2007) to create 21 cm line catalogs. This model, however, does not include a self-consistent partition of cold gas into its HI and H_2 components, and Obreschkow et al. (2009a) evaluated the partition in post-processing, as I have done for other publicly available models. Obreschkow et al. (2009a,b)

assumed specific mass profiles for all components of model galaxies, and used them to evaluate the rotational velocity profiles. For our model, I can use the profiles described in Sec. 5.3 and 5.4.1.

For each galaxy, the 21 cm line profile can be evaluated integrating, over the gas disk radius, the contributions form thin concentric annuli of the HI disk. The larger is the number of these annuli, the higher is the precision of the integration. I will need to evaluate the number of annuli than provides a good compromise between a precise estimate and a fast computation. An annulus with radius r has a rotational velocity $V_{rot}(r)$, and contributes to the total 21 cm line profile in an edge-on galaxy as:

$$\tilde{\psi}(V, V_{rot}(r)) = \begin{cases} \frac{1}{\pi\sqrt{V_{rot}(r)^2 - V^2}} & \text{if } |V| < V_{rot}(r) \\ 0 & |V| \geq V_{rot}(r) \end{cases} \tag{6.2}$$

This equation is normalized to $\int dV \tilde{\psi}(V) = 1$. $\tilde{\psi}(V, V_{rot}(r))$ would diverge for $|V| \to V_{rot}(r)$, but this problem is solved smoothing the rotational velocity profile assuming local turbulent, random motions of the gas. I assume, following Obreschkow et al. (2009a), that the velocity dispersion of the gas is $\sigma_{gas} = 8$ km s^{-1}. The line profile with smoothed velocity is then:

$$\psi(V, V_{rot}(r)) = \frac{1}{\sqrt{2\pi}\sigma_{gas}} \int dV' \exp\left[\frac{(V - V')^2}{-2\sigma_{gas}^2}\right] \tilde{\psi}(V', V_{rot}(r)), \tag{6.3}$$

which preserves the normalization. Integrating this expression over all the annuli, I obtain the edge-on line profile of the 21 cm line $\Psi(V)$:

$$\Psi(V) = \frac{2\pi}{M_{HI}} \int_0^{inf} dr\, r\Sigma_{HI}(r)\psi(V, V_{rot}(r)). \tag{6.4}$$

As explained above, the rotational velocity profile $V_{rot}(r)$ is calculated from the mass distribution, as in Sec. 5.4.1, and the integration is performed numerically, to a limited radius. Observationally, this radius depends on the sensitivity of the instrument adopted, i.e. the luminosity limit achievable by the observation. Therefore, I will also evaluate how the obtained 21 cm line profile depends on the integration radius.

The line profile in Eq. 6.4 must be normalized to the total 21 cm luminosity of the galaxy. Following Obreschkow et al. (2009a), this can be expressed as $L_{HI}[\text{Jy km s}^{-1}] = M_{HI}[\text{M}_\odot]/(1.88 \cdot 10^4)$. This equation results from the integration of the 21 cm line emitted by a mass M_{HI}, and does not account for the HI self-absorption effects, that can affect measurements for large disk galaxies observed edge-on (Rao et al., 1995).

The inclination that can be obtained from the specific angular momentum of the disk affects the measured velocity as $V_{rot}^i = V_{rot} \sin i$. The observed line is, thus, $\Psi^i = \Psi / \sin i$.

We are in contact with groups working on the cosmological applications of the HI large scale observations to be performed with SKA. They are interested in studying optimal survey strategies to adopt in various SKA phases, to create HI intensity maps of large slices of the sky. The mock light-cones I will create will be used to test results expected with different SKA configurations, and their applicability for cosmological purposes.

6.2 Future plans

As explained at the beginning of this section, higher resolution simulations would be beneficial for improving a number of aspects of my work, or extend its validity. For example, low mass galaxies are crucial for the identification of realistic model prescriptions. Mock light-cones that cover large ares of sky require simulations based on large cosmic volumes. At the moment, the Millennium simulation offers a very good compromise between high resolution and large cosmic volumes, but in this section I discuss possible strategies to improve the simulations for applications that require either higher resolution or larger volumes.

Another model aspect that necessitates a strong revision is the prescription adopted for disk instability, that I have shown to create unrealistic bulges. I discuss, in this section, how I plan to upgrade this specific model prescription.

6.2.1 Dedicated cosmological simulations

The Millennium simulation, in a volume of $(500\ \mathrm{Mpc\ h^{-1}})^3$, resolves halos down to a mass of $M_{200} \sim 1.7 \cdot 10^{10}\ \mathrm{M_\odot h^{-1}}$, that correspond to galaxies of stellar mass $M_* \sim 10^9\ \mathrm{M_\odot h^{-1}}$ and HI mass of $M_{HI} \sim 10^{8.8}\ \mathrm{M_\odot h^{-1}}$. The Millennium II improves these mass resolutions by a factor ~ 10, but reducing the simulated volume to $(100\ \mathrm{Mpc\ h^{-1}})^3$.

As I detailed above in this chapter, some studies necessitate of a resolution higher than that of the Millennium Simulations. For example, the analysis of the prescriptions regulating the HI content of satellites and low-mass galaxies, or the implementation of different specific angular momentum prescriptions for the gas cooling in the rapid regime. I plan to address these issues, applying the semi-analytic model to high resolution merger trees, created using dedicated simulations. Cosmological simulations with volumes similar to that of the Millennium II hardly have higher resolution. This must then be achieved improving the resolution of available simulations, and several techniques are adopted in literature for this purpose.

One of these techniques consists in expanding the merger trees of a given simulation, generating merger histories of low mass halos using a Monte Carlo scheme, usually based on the extended Press-Schechter formalism or its modifications (see, for example Parkinson et al., 2008; Zhang et al., 2008; Jiang and van den Bosch, 2014). With this approach, one can generate merger histories of halos well below the resolution limit of the initial cosmological simulation, at a relatively low computational cost. This approach, however, does not provide information on the internal dynamics of the halos, ingredients required by the semi-analytic model to calculate galactic properties, such as size or angular momentum. One possibility is to extract these randomly from distributions built from N-body simulations. For a detailed analysis of low-mass galaxies properties and their dependence on specific model prescriptions, I plan to employ a cosmological simulation that is able to evolve explicitly the internal dynamics of small halos.

One possibility is that of using high-resolution zoom-in simulations. This method is based on improving resolution of specific sub-volumes of a lower-resolution cosmological simulation. The number of particles in these sub-volumes is increased to reach the desired resolution, and outside this region the resolution

is progressively degraded, so as to save computational time, while preserving a correct description of the large-scale tidal field. Zoom-in simulations are widely used to perform, for example, well resolved hydrodynamical re-simulations in a cosmological context (some recent works: Springel et al., 2008; Bonafede et al., 2011; Oser et al., 2012; Hopkins et al., 2014; Sawala et al., 2015). This approach would provide all necessary halo information required as input by our semi-analytic model, including self consistently evolved dynamical properties. The main limitation of this method is the small volume considered at high resolution. A zoom-in simulation is computationally more convenient than a full N-body simulation at the same resolution, because of the typically limited volume of zoomed regions. Due to their nature, zoom-in simulations are not useful to create mock light-cones.

In this case, especially when the required light-cone must cover a large volume of the sky, as in the case of the large surveys planned for SKA and its precursors, a large volume simulation is strongly required.
An approach that allows large cosmic volumes to be simulated at a low computational cost is that of approximated methods. These are based on approximated solutions of the large scale linear structure growth, i.e. the Zeldovich (1972) approximation at the first order, or the Lagrangian Perturbation Theory at higher orders. For example, PINOCCHIO (Monaco et al., 2002) is based on Lagrangian Perturbation Theory, and can be run on very short time-scales, allowing the production of a very large number of halo catalogs at low computational cost. Nevertheless, PINOCCHIO is not able to reproduce well the non-linear regime, and an application to the semi-analytic model would require further post-processing work on the merger tree, for example using Monte Carlo generated subhalo merger trees, based on the extended Press-Schechter formalism or modified versions.
To combine the approximated evolution of large scale structures to a realistic treatment of small scale dynamics, these methods are often coupled with a full N-body solution at non-linear scales. This is the case, for example, of COLA (Tassev et al., 2013). This simulation is slower than PINOCCHIO, but much faster than a standard N-body simulation. The main disadvantage is that much of the computational convenience comes from the low number of time-steps adopted. This can be a problem for our work, because the semi-analytic model requires a certain number of time-steps to evolve the baryons, to achieve convergence of the results (Benson et al., 2012).
There are also large volume non-approximated cosmological simulations, as MICE (Fosalba et al., 2015), that covers $(3\ \mathrm{Gpc\ h^{-1}})^3$ with a minimum particle mass $m_p = 2.93 \cdot 10^{10}\ \mathrm{M_\odot h^{-1}}$. The resolution is slightly lower than that of the Millennium, but for the purposes of large mock light-cones it can be enough.

Whatever the DM simulation, the semi-analytic model requires a suitable halo merger tree as input. Its construction consists of several steps: halos and sub-halos are identified at each snapshot, and the descendant/progenitors are linked together from a snapshot to the other. A merger tree collects all halo properties necessary to feed the semi-analytic model. For the Xie et al. (2017) model, these properties include mass, radius, position, velocity, internal velocity dispersion, maximum rotational velocity, and spin parameter. The latter is particularly important for the analysis of galaxy dynamics, because it determines the disk specific angular momentum and its size. Furthermore, as highlighted in Sec. 4.6.2, it is correlated with the HI content of model galaxies,

because of the dependence of the cold gas disk radius on its value.
Once the halo merger tree is available, the model free parameters must be retuned to reproduce some chosen relations involving galaxy properties, as the stellar/HI mass function.

In conclusion, the halo properties required by our semi-analytic model put strong constraints on the simulations that can be used in our analyses. The study of low-mass galaxies requires resolved subhalo internal dynamics, to have realistic subhalo sizes and spins. Monte Carlo halo merger trees are thus not ideal for this kind of analysis, because they do not include consistently these properties. Zoom-in simulations would offer a fast alternative to full high resolution N-body simulations, provided that the zoomed regions are accurately selected to reproduce to overall galaxy population. The construction of large volume mock light-cones, or the creation of a large set of independent small light-cones, requires large volume simulations. Full N-body high-resolution simulations are computationally costly, and are usually limited in volume. For this reason, I plan to apply our semi-analytic model to cosmological simulations realized with approximated methods. This approach will require post-processing, in order to adapt the approximated merger trees to be used by the semi-analytic model.

6.2.2 A consistent treatment for disk instabilities

In the second part of this Thesis, I have analyzed the sizes and the specific angular momenta of model galaxies. I have demonstrated that bulges formed mainly through disk instabilities have unrealistic small sizes. Furthermore, the assumption of non-rotating bulges leads to an underestimation of the j_*-M$_*$ relation of early type galaxies with respect to observational data.

Both these problems can be partially solved including a specific treatment for the formation of pseudo-bulges. Observed bulges are typically divided in two categories: classical, kinetically hot spheroids believed to form during mergers (Renzini, 1999), and pseudo-bulges, that are dynamically cold, and have Sérsic index, stellar population and velocity dispersion intermediate between classical bulges and disks (de Jong, 1996; Peletier and Balcells, 1996; MacArthur et al., 2003; Kormendy and Kennicutt, 2004; Athanassoula, 2005; Drory and Fisher, 2007). The pseudo-bulge disky characteristics suggest a secular formation and an evolution alternative to that driven by mergers (Kormendy, 1982; Kormendy and Kennicutt, 2004; Fisher et al., 2009).

These different channels for bulge formation are included in semi-analytic models since long time. Our model, for example, forms bulges from both mergers and disk instabilities. Our model, however, does not differentiate the bulge characteristics according to the formation channel, leading to the same dispersion dominated spheroid. Recently, an explicit separation between classical and pseudo bulges has been included in the semi-analytic model described in Tonini et al. (2016). In that model, the classical bulge is formed only through major mergers and through minor mergers on central galaxies dominated by a classical bulge component. The classical bulge size is evaluated from energy conservation, as in our model, but with a specific energy component accounting for dissipation. The pseudo-bulge, instead, is formed through minor mergers and disk instabilities. Its radius is assumed to grow proportionally to the acquired

mass δM and to the current disk radius $R_{*,disk}$, following the formula:

$$R^f_{PB} = \frac{R^0_{PB} M^0_{PB} + \alpha_{DI} \delta M R_{*,disk}}{M^0_{PB} + \delta M}, \tag{6.5}$$

where R^0_{PB} and R^f_{PB} are the radius of the pseudo-bulge before and after the mass growth, and M^0_{PB} is the initial mass of the pseudo-bulge. $\alpha_{DI} = 0.2$ is a dissipation parameter chosen to reproduce the observational measurements by Fisher and Drory (2008, 2010), where pseudo-bulges have a median half-mass radius proportional to the scale radius of the disk: $R_{PB} = 0.2 R_{disk}$.

I am currently implementing a pseudo-bulge component in our model with a similar approach, but with some substantial differences. I assume that the unique channel for pseudo-bulge formation is disk instabilities.
In detail, during a disk instability, the model adds the stars moved to restore disk stability to the pseudo-bulge component. When stars are moved there, they loose angular momentum via dissipation processes, and the pseudo-bulge specific angular momentum changes proportionally to the moved mass:

$$\vec{j}^f_{PB} = \frac{\vec{j}^0_{PB} M^0_{PB} + \alpha_{DI} \vec{j}_{*,disk} \delta M}{M^0_{PB} + \delta M}, \tag{6.6}$$

where $\vec{j}^0_{PB}$ and $\vec{j}^f_{PB}$ are the specific angular momentum of the pseudo-bulge before and after mass accretion. I assume $\alpha_{DI} = 0.2$, as in Tonini et al. (2016), but I plan to test the influence of other possible values. During major mergers the pseudo-bulge is disrupted as the stellar disk, and its content is added to the classical bulge.

I assume that the pseudo-bulge is a rotating disk, with an exponential surface density profile. Its scale radius can be estimated from its specific angular momentum, through a formula that I still need to define. One possibility is to assume the same rotational velocity of the stellar disk, and consequently evaluate the radius using an equivalent formula: $R_{PB} = |\vec{j}_{PB}|/(2V_{max})$, where V_{max} is the maximum rotational velocity of the dark matter halo, assumed to be equal to that of the disk. This assumption is not necessarily valid for the pseudo-bulge, that may be a transient component, whose rotational velocity is not necessarily aligned with that of the DM halo. However, if this assumption is valid, the obtained radius would scale similarly to that in Tonini et al. (2016). More realistically, both the radius and the velocity of the pseudo bulge are smaller than those of the disk, resulting in a lower specific angular momentum. For this reason, I can assume another dissipative factor that affects the rotational velocity of a pseudo-bulge, and calculate its radius using the formula above. The neat effect would be the same of including velocity dissipation directly into α_{DI}. I plan to test alternative treatments for pseudo-bulge radius calculation, based on energy conservation during disk instability, in future work. The obtained formula is similar to that for bulge size during mergers, but includes also energy terms that account for the internal kinetic energy. I will test these radius estimates comparing the obtained size-mass relations to available observational data.

The presence of rotating pseudo-bulges allows a consistent implementation of fast rotators in our model. Taking advantage of the detailed treatment for specific angular momentum performed in this Thesis work, I plan to analyze in

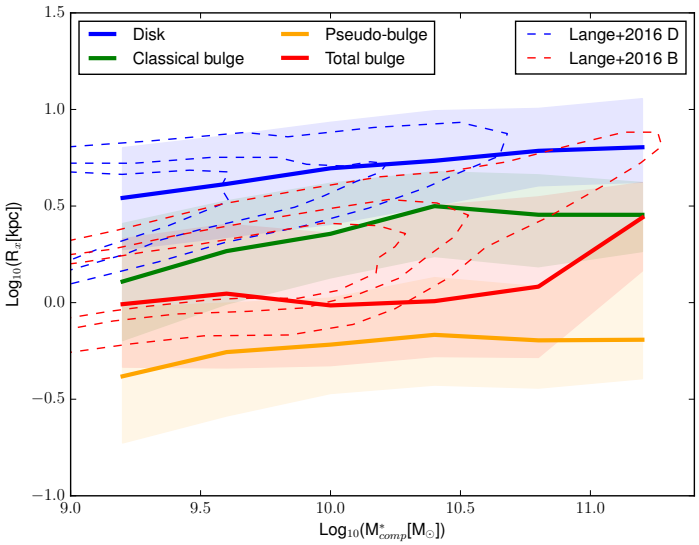

Figure 6.4: The size-mass relation of individual components of galaxies from the model with the implementation of the pseudo-bulge. The analyzed components are the disk (blue), the classical bulge (green), the pseudo-bulge (orange) and the combination of classical and pseudo bulge (red solid line). As a reference I show the observational density distribution of Lange et al. (2016) as dashed blue (disks) and red (bulges) lines.

detail the fast rotators distribution in the j_*-M$_*$ relation, comparing it to the available data from ATLAS^{3D} or SAMI (Cappellari et al., 2011; Bryant et al., 2015).

I have already implemented a pseudo-bulge treatment in the model by Xie et al. (2017), and I am currently analyzing model outputs. I show one preliminary result in Fig. 6.4, that shows the size-mass relation of individual galactic components, as I did for the standard model in Fig. 5.4. In this model I assume that the pseudo-bulge has the same maximum rotational velocity V_{max} of the halo, and $\alpha_{DI} = 0.2$. In this figure, pseudo-bulges and classical bulges correspond to the DI bulges and the merger bulges of the fiducial model, respectively. Pseudo-bulges are considerably larger, though still not enough, especially at high masses.

Of course, the introduction of a pseudo-bulge component does not solve the problem of adopting an unrealistic instability criterion. As explained above, the currently adopted instability prescription is limited to the stellar disk, and does not take into account the cold gas disk and possible star bursts. The current criterion is not able to distinguish among bar stable and bar unstable disks (Athanassoula, 2008; Benson and Devereux, 2010; De Lucia et al., 2012a). At the moment, no alternative treatment for disk instabilities has been proposed. Of course, this is a key ingredient of the model, and necessitates a better description and implementation. I plan to include gas in the stability evaluation, and to move some fraction of it to the center of the galaxy, where it is consumed in a star burst. The details of this project are still under definition, and I am

studying possible implementations on the basis of published dedicated simulations (Noguchi, 1999; Elmegreen et al., 2008; Dekel et al., 2009; Krumholz and Burkert, 2010; Genzel et al., 2011; Bournaud et al., 2011; Forbes et al., 2012; Cacciato et al., 2012; Forbes et al., 2014).

The combination of a realistic disk instability prescription and a pseudo-bulge component will let us study the relations involving Early Type galaxies, with a better control over their origin and their properties.

Chapter 7

Bibliography

Abadi, M. G., Navarro, J. F., Steinmetz, M., and Eke, V. R. (2003). Simulations of Galaxy Formation in a Λ Cold Dark Matter Universe. I. Dynamical and Photometric Properties of a Simulated Disk Galaxy. *ApJ*, 591:499–514.

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., and et al. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. *ApJS*, 182:543–558.

Amorín, R., Muñoz-Tuñón, C., Aguerri, J. A. L., and Planesas, P. (2016). Molecular gas in low-metallicity starburst galaxies: Scaling relations and the CO-to-H_2 conversion factor. *A&A*, 588:A23.

Anderson, M. E. and Bregman, J. N. (2011). Detection of a Hot Gaseous Halo around the Giant Spiral Galaxy NGC 1961. *ApJ*, 737:22.

Angulo, R. E. and Hilbert, S. (2015). Cosmological constraints from the CFHTLenS shear measurements using a new, accurate, and flexible way of predicting non-linear mass clustering. *MNRAS*, 448:364–375.

Angulo, R. E. and White, S. D. M. (2010). One simulation to fit them all - changing the background parameters of a cosmological N-body simulation. *MNRAS*, 405:143–154.

Anninos, P., Norman, M. L., and Clarke, D. A. (1994). Hierarchical numerical cosmology with hydrodynamics: Methods and code tests. *ApJ*, 436:11–22.

Arrigoni, M., Trager, S. C., Somerville, R. S., and Gibson, B. K. (2010). Galactic chemical evolution in hierarchical formation models - I. Early-type galaxies in the local Universe. *MNRAS*, 402:173–190.

Athanassoula, E. (2005). On the nature of bulges in general and of box/peanut bulges in particular: input from N-body simulations. *MNRAS*, 358:1477–1488.

Athanassoula, E. (2008). Disc instabilities and semi-analytic modelling of galaxy formation. *MNRAS*, 390:L69–L72.

Bacon, R., Copin, Y., Monnet, G., Miller, B. W., Allington-Smith, J. R., Bureau, M., Carollo, C. M., Davies, R. L., Emsellem, E., Kuntschner, H., Peletier, R. F., Verolme, E. K., and de Zeeuw, P. T. (2001). The SAURON project - I. The panoramic integral-field spectrograph. *MNRAS*, 326:23–35.

Bahé, Y. M., Crain, R. A., Kauffmann, G., Bower, R. G., Schaye, J., Furlong, M., Lagos, C., Schaller, M., Trayford, J. W., Dalla Vecchia, C., and Theuns, T. (2016). The distribution of atomic hydrogen in EAGLE galaxies: morphologies, profiles, and H I holes. *MNRAS*, 456:1115–1136.

Bardeen, J. M., Bond, J. R., Kaiser, N., and Szalay, A. S. (1986). The statistics of peaks of Gaussian random fields. *ApJ*, 304:15–61.

Barnes, J. and Hut, P. (1986). A hierarchical O(N log N) force-calculation algorithm. *Nature*, 324:446–449.

Barnes, J. E. (2002). Formation of gas discs in merging galaxies. *MNRAS*, 333:481–494.

Basilakos, S., Plionis, M., Kovač, K., and Voglis, N. (2007). Large-scale structure in the HI Parkes All-Sky Survey: filling the voids with HI galaxies? *MNRAS*, 378:301–308.

Baugh, C. M. (2006). A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156.

Baugh, C. M., Lacey, C. G., Frenk, C. S., Benson, A. J., Cole, S., Granato, G. L., Silva, L., and Bressan, A. (2004). Predictions for the SKA from hierarchical galaxy formation models. *New Astron. Rev.*, 48:1239–1246.

Baum, S. A. and O'Dea, C. P. (1991). Multifrequency VLA observations of PKS 0745 - 191 - The archetypal 'cooling flow' radio source? *MNRAS*, 250:737–749.

Behroozi, P. S., Wechsler, R. H., and Wu, H.-Y. (2013). The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. *ApJ*, 762:109.

Benson, A. J. (2010). Galaxy formation theory. *Phys. Rep.*, 495:33–86.

Benson, A. J. (2012). G ALACTICUS: A semi-analytic model of galaxy formation. *New Astron.*, 17:175–197.

Benson, A. J., Borgani, S., De Lucia, G., Boylan-Kolchin, M., and Monaco, P. (2012). Convergence of galaxy properties with merger tree temporal resolution. *MNRAS*, 419:3590–3603.

Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., and Cole, S. (2003). What Shapes the Luminosity Function of Galaxies? *ApJ*, 599:38–49.

Benson, A. J. and Devereux, N. (2010). The origin of the Hubble sequence in ΛCDM cosmology. *MNRAS*, 402:2321–2334.

Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., and Lacey, C. G. (2001a). The clustering evolution of the galaxy distribution. *MNRAS*, 327:1041–1056.

Benson, A. J., Pearce, F. R., Frenk, C. S., Baugh, C. M., and Jenkins, A. (2001b). A comparison of semi-analytic and smoothed particle hydrodynamics galaxy formation. *MNRAS*, 320:261–280.

Berger, M. J. and Colella, P. (1989). Local adaptive mesh refinement for shock hydrodynamics. *Journal of computational Physics*, 82(1):64–84.

Bernardi, M., Meert, A., Sheth, R. K., Vikram, V., Huertas-Company, M., Mei, S., and Shankar, F. (2013). The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile. *MNRAS*, 436:697–704.

Bigiel, F., Leroy, A., Walter, F., Brinks, E., de Blok, W. J. G., Madore, B., and Thornley, M. D. (2008). The Star Formation Law in Nearby Galaxies on Sub-Kpc Scales. *AJ*, 136:2846–2871.

Binney, J. (1977). The physics of dissipational galaxy formation. *ApJ*, 215:483–491.

Binney, J. and Tremaine, S. (2008). *Galactic Dynamics: Second Edition*. Princeton University Press.

Blaizot, J., Wadadekar, Y., Guiderdoni, B., Colombi, S. T., Bertin, E., Bouchet, F. R., Devriendt, J. E. G., and Hatton, S. (2005). MoMaF: the Mock Map Facility. *MNRAS*, 360:159–175.

Blanton, M. R., Hogg, D. W., Bahcall, N. A., Baldry, I. K., Brinkmann, J., Csabai, I., Eisenstein, D., Fukugita, M., Gunn, J. E., Ivezić, Ž., Lamb, D. Q., Lupton, R. H., Loveday, J., Munn, J. A., Nichol, R. C., Okamura, S., Schlegel, D. J., Shimasaku, K., Strauss, M. A., Vogeley, M. S., and Weinberg, D. H. (2003). The Broadband Optical Properties of Galaxies with Redshifts 0.02¡z¡0.22. *ApJ*, 594:186–207.

Blitz, L. and Rosolowsky, E. (2006). The Role of Pressure in GMC Formation II: The H_2-Pressure Relation. *ApJ*, 650:933–944.

Bolatto, A. D., Leroy, A. K., Jameson, K., Ostriker, E., Gordon, K., Lawton, B., Stanimirović, S., Israel, F. P., Madden, S. C., Hony, S., Sandstrom, K. M., Bot, C., Rubio, M., Winkler, P. F., Roman-Duval, J., van Loon, J. T., Oliveira, J. M., and Indebetouw, R. (2011). The State of the Gas and the Relation between Gas and Star Formation at Low Metallicity: The Small Magellanic Cloud. *ApJ*, 741:12.

Bonafede, A., Dolag, K., Stasyszyn, F., Murante, G., and Borgani, S. (2011). A non-ideal magnetohydrodynamic gadget: simulating massive galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 418(4):2234–2250.

Bond, J. R., Cole, S., Efstathiou, G., and Kaiser, N. (1991). Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460.

Booth, R. S., de Blok, W. J. G., Jonas, J. L., and Fanaroff, B. (2009). MeerKAT Key Project Science, Specifications, and Proposals. *ArXiv e-prints*.

Bottrell, C., Torrey, P., Simard, L., and Ellison, S. L. (2017). Galaxies in the Illustris simulation as seen by the Sloan Digital Sky Survey - II. Size-luminosity relations and the deficit of bulge-dominated galaxies in Illustris at low mass. *MNRAS*, 467:2879–2895.

Bournaud, F., Dekel, A., Teyssier, R., Cacciato, M., Daddi, E., Juneau, S., and Shankar, F. (2011). Black Hole Growth and Active Galactic Nuclei Obscuration by Instability-driven Inflows in High-redshift Disk Galaxies Fed by Cold Streams. *ApJ*, 741:L33.

Bower, R. G. (1991). The evolution of groups of galaxies in the Press-Schechter formalism. *MNRAS*, 248:332–352.

Bower, R. G., Benson, A. J., Malbon, R., Helly, J. C., Frenk, C. S., Baugh, C. M., Cole, S., and Lacey, C. G. (2006). Breaking the hierarchy of galaxy formation. *MNRAS*, 370:645–655.

Boylan-Kolchin, M., Ma, C.-P., and Quataert, E. (2008). Dynamical friction and galaxy merging time-scales. *MNRAS*, 383:93–101.

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., and Lemson, G. (2009). Resolving cosmic structure formation with the Millennium-II Simulation. *MNRAS*, 398:1150–1164.

Brown, T., Catinella, B., Cortese, L., Lagos, C. d. P., Davé, R., Kilborn, V., Haynes, M. P., Giovanelli, R., and Rafieferantsoa, M. (2017). Cold gas stripping in satellite galaxies: from pairs to clusters. *MNRAS*, 466:1275–1289.

Brüggen, M., Heinz, S., Roediger, E., Ruszkowski, M., and Simionescu, A. (2007). Shock heating by Fanaroff-Riley type I radio sources in galaxy clusters. *MNRAS*, 380:L67–L70.

Bruzual A., G. and Charlot, S. (1993). Spectral evolution of stellar populations using isochrone synthesis. *ApJ*, 405:538–553.

Bryant, J. J., Owers, M. S., Robotham, A. S. G., Croom, S. M., Driver, S. P., Drinkwater, M. J., Lorente, N. P. F., Cortese, L., Scott, N., Colless, M., Schaefer, A., Taylor, E. N., Konstantopoulos, I. S., Allen, J. T., Baldry, I., Barnes, L., Bauer, A. E., Bland-Hawthorn, J., Bloom, J. V., Brooks, A. M., Brough, S., Cecil, G., Couch, W., Croton, D., Davies, R., Ellis, S., Fogarty, L. M. R., Foster, C., Glazebrook, K., Goodwin, M., Green, A., Gunawardhana, M. L., Hampton, E., Ho, I.-T., Hopkins, A. M., Kewley, L., Lawrence, J. S., Leon-Saval, S. G., Leslie, S., McElroy, R., Lewis, G., Liske, J., López-Sánchez, Á. R., Mahajan, S., Medling, A. M., Metcalfe, N., Meyer, M., Mould, J., Obreschkow, D., O'Toole, S., Pracy, M., Richards, S. N., Shanks, T., Sharp, R., Sweet, S. M., Thomas, A. D., Tonini, C., and Walcher, C. J. (2015). The SAMI Galaxy Survey: instrument specification and target selection. *MNRAS*, 447:2857–2879.

Bundy, K., Bershady, M. A., Law, D. R., Yan, R., Drory, N., MacDonald, N., Wake, D. A., Cherinka, B., Sánchez-Gallego, J. R., Weijmans, A.-M., Thomas, D., Tremonti, C., Masters, K., Coccato, L., Diamond-Stanic, A. M., Aragón-Salamanca, A., Avila-Reese, V., Badenes, C., Falcón-Barroso, J.,

Belfiore, F., Bizyaev, D., Blanc, G. A., Bland-Hawthorn, J., Blanton, M. R., Brownstein, J. R., Byler, N., Cappellari, M., Conroy, C., Dutton, A. A., Emsellem, E., Etherington, J., Frinchaboy, P. M., Fu, H., Gunn, J. E., Harding, P., Johnston, E. J., Kauffmann, G., Kinemuchi, K., Klaene, M. A., Knapen, J. H., Leauthaud, A., Li, C., Lin, L., Maiolino, R., Malanushenko, V., Malanushenko, E., Mao, S., Maraston, C., McDermid, R. M., Merrifield, M. R., Nichol, R. C., Oravetz, D., Pan, K., Parejko, J. K., Sanchez, S. F., Schlegel, D., Simmons, A., Steele, O., Steinmetz, M., Thanjavur, K., Thompson, B. A., Tinker, J. L., van den Bosch, R. C. E., Westfall, K. B., Wilkinson, D., Wright, S., Xiao, T., and Zhang, K. (2015). Overview of the SDSS-IV MaNGA Survey: Mapping nearby Galaxies at Apache Point Observatory. *ApJ*, 798:7.

Cacciato, M., Dekel, A., and Genel, S. (2012). Evolution of violent gravitational disc instability in galaxies: late stabilization by transition from gas to stellar dominance. *MNRAS*, 421:818–831.

Capak, P., Aussel, H., Ajiki, M., McCracken, H. J., Mobasher, B., Scoville, N., Shopbell, P., Taniguchi, Y., Thompson, D., Tribiano, S., Sasaki, S., Blain, A. W., Brusa, M., Carilli, C., Comastri, A., Carollo, C. M., Cassata, P., Colbert, J., Ellis, R. S., Elvis, M., Giavalisco, M., Green, W., Guzzo, L., Hasinger, G., Ilbert, O., Impey, C., Jahnke, K., Kartaltepe, J., Kneib, J.-P., Koda, J., Koekemoer, A., Komiyama, Y., Leauthaud, A., Le Fevre, O., Lilly, S., Liu, C., Massey, R., Miyazaki, S., Murayama, T., Nagao, T., Peacock, J. A., Pickles, A., Porciani, C., Renzini, A., Rhodes, J., Rich, M., Salvato, M., Sanders, D. B., Scarlata, C., Schiminovich, D., Schinnerer, E., Scodeggio, M., Sheth, K., Shioya, Y., Tasca, L. A. M., Taylor, J. E., Yan, L., and Zamorani, G. (2007). The First Release COSMOS Optical and Near-IR Data and Catalog. *ApJS*, 172:99–116.

Cappellari, M., Emsellem, E., Krajnović, D., McDermid, R. M., Scott, N., Verdoes Kleijn, G. A., Young, L. M., Alatalo, K., Bacon, R., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Davies, R. L., Davis, T. A., de Zeeuw, P. T., Duc, P.-A., Khochfar, S., Kuntschner, H., Lablanche, P.-Y., Morganti, R., Naab, T., Oosterloo, T., Sarzi, M., Serra, P., and Weijmans, A.-M. (2011). The ATLAS^{3D} project - I. A volume-limited sample of 260 nearby early-type galaxies: science goals and selection criteria. *MNRAS*, 413:813–836.

Cappellari, M., McDermid, R. M., Alatalo, K., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Crocker, A. F., Davies, R. L., Davis, T. A., de Zeeuw, P. T., Duc, P.-A., Emsellem, E., Khochfar, S., Krajnović, D., Kuntschner, H., Lablanche, P.-Y., Morganti, R., Naab, T., Oosterloo, T., Sarzi, M., Scott, N., Serra, P., Weijmans, A.-M., and Young, L. M. (2012). Systematic variation of the stellar initial mass function in early-type galaxies. *Nature*, 484:485–488.

Carilli, C. L. and Walter, F. (2013). Cool Gas in High-Redshift Galaxies. *ARA&A*, 51:105–161.

Catinella, B. and Cortese, L. (2015). HIGHz: a survey of the most H I-massive galaxies at z $\sim$ 0.2. *MNRAS*, 446:3526–3544.

Catinella, B., Schiminovich, D., Cortese, L., Fabello, S., Hummels, C. B., Moran, S. M., Lemonias, J. J., Cooper, A. P., Wu, R., Heckman, T. M.,

and Wang, J. (2013). The GALEX Arecibo SDSS Survey - VIII. Final data release. The effect of group environment on the gas content of massive galaxies. *MNRAS*, 436:34–70.

Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Hummels, C., Lemonias, J., Moran, S. M., Wu, R., Cooper, A., and Wang, J. (2012). The GALEX Arecibo SDSS Survey. VI. Second data release and updated gas fraction scaling relations. *A&A*, 544:A65.

Catinella, B., Schiminovich, D., Kauffmann, G., Fabello, S., Wang, J., Hummels, C., Lemonias, J., Moran, S. M., Wu, R., Giovanelli, R., Haynes, M. P., Heckman, T. M., Basu-Zych, A. R., Blanton, M. R., Brinchmann, J., Budavári, T., Gonçalves, T., Johnson, B. D., Kennicutt, R. C., Madore, B. F., Martin, C. D., Rich, M. R., Tacconi, L. J., Thilker, D. A., Wild, V., and Wyder, T. K. (2010). The GALEX Arecibo SDSS Survey - I. Gas fraction scaling relations of massive galaxies and first data release. *MNRAS*, 403:683–708.

Cattaneo, A. and Teyssier, R. (2007). AGN self-regulation in cooling flow clusters. *MNRAS*, 376:1547–1556.

Cha, S.-H., Inutsuka, S.-I., and Nayakshin, S. (2010). Kelvin-Helmholtz instabilities with Godunov smoothed particle hydrodynamics. *MNRAS*, 403:1165–1174.

Chabrier, G. (2002). The Galactic Disk Mass Budget. II. Brown Dwarf Mass Function and Density. *ApJ*, 567:304–313.

Chang, T.-C., Pen, U.-L., Bandura, K., and Peterson, J. B. (2010). An intensity map of hydrogen 21-cm emission at redshift z˜0.8. *Nature*, 466:463–465.

Chieffi, A. and Limongi, M. (2004). Explosive Yields of Massive Stars from Z = 0 to Z = Z_{solar}. *ApJ*, 608:405–410.

Christensen, C., Quinn, T., Governato, F., Stilp, A., Shen, S., and Wadsley, J. (2012). Implementing molecular hydrogen in hydrodynamic simulations of galaxy formation. *MNRAS*, 425:3058–3076.

Churazov, E., Forman, W., Jones, C., and Böhringer, H. (2000). Asymmetric, arc minute scale structures around NGC 1275. *A&A*, 356:788–794.

Cole, S. (1991). Modeling galaxy formation in evolving dark matter halos. *ApJ*, 367:45–53.

Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., and Zepf, S. E. (1994). A Recipe for Galaxy Formation. *MNRAS*, 271:781.

Cole, S., Lacey, C. G., Baugh, C. M., and Frenk, C. S. (2000). Hierarchical galaxy formation. *MNRAS*, 319:168–204.

Coles, P. and Lucchin, F. (2002). *Cosmology: The Origin and Evolution of Cosmic Structure, 2nd Edition*. Wiley.

Conroy, C. and van Dokkum, P. G. (2012). The Stellar Initial Mass Function in Early-type Galaxies From Absorption Line Spectroscopy. II. Results. *ApJ*, 760:71.

Conroy, C., Wechsler, R. H., and Kravtsov, A. V. (2006). Modeling Luminosity-dependent Galaxy Clustering through Cosmic Time. *ApJ*, 647:201–214.

Cortese, L., Fogarty, L. M. R., Bekki, K., van de Sande, J., Couch, W., Catinella, B., Colless, M., Obreschkow, D., Taranu, D., Tescari, E., Barat, D., Bland-Hawthorn, J., Bloom, J., Bryant, J. J., Cluver, M., Croom, S. M., Drinkwater, M. J., d'Eugenio, F., Konstantopoulos, I. S., Lopez-Sanchez, A., Mahajan, S., Scott, N., Tonini, C., Wong, O. I., Allen, J. T., Brough, S., Goodwin, M., Green, A. W., Ho, I.-T., Kelvin, L. S., Lawrence, J. S., Lorente, N. P. F., Medling, A. M., Owers, M. S., Richards, S., Sharp, R., and Sweet, S. M. (2016). The SAMI Galaxy Survey: the link between angular momentum and optical morphology. *MNRAS*, 463:170–184.

Covington, M. D., Primack, J. R., Porter, L. A., Croton, D. J., Somerville, R. S., and Dekel, A. (2011). The role of dissipation in the scaling relations of cosmological merger remnants. *MNRAS*, 415:3135–3152.

Cox, T. J., Jonsson, P., Somerville, R. S., Primack, J. R., and Dekel, A. (2008). The effect of galaxy mass ratio on merger-driven starbursts. *MNRAS*, 384:386–409.

Crain, R. A., Bahé, Y. M., Lagos, C. d. P., Rahmati, A., Schaye, J., McCarthy, I. G., Marasco, A., Bower, R. G., Schaller, M., Theuns, T., and van der Hulst, T. (2017). The EAGLE simulations: atomic hydrogen associated with galaxies. *MNRAS*, 464:4204–4226.

Crain, R. A., McCarthy, I. G., Frenk, C. S., Theuns, T., and Schaye, J. (2010). X-ray coronae in simulations of disc galaxy formation. *MNRAS*, 407:1403–1422.

Crenshaw, D. M., Kraemer, S. B., and George, I. M. (2003). Mass Loss from the Nuclei of Active Galaxies. *ARA&A*, 41:117–167.

Croton, D. J., Springel, V., White, S. D. M., De Lucia, G., Frenk, C. S., Gao, L., Jenkins, A., Kauffmann, G., Navarro, J. F., and Yoshida, N. (2006). The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *MNRAS*, 365:11–28.

Dalcanton, J. J., Spergel, D. N., and Summers, F. J. (1997). The Formation of Disk Galaxies. *ApJ*, 482:659–676.

Danovich, M., Dekel, A., Hahn, O., Ceverino, D., and Primack, J. (2015). Four phases of angular-momentum buildup in high-z galaxies: from cosmic-web streams through an extended ring to disc and bulge. *MNRAS*, 449:2087–2111.

Davé, R., Katz, N., Oppenheimer, B. D., Kollmeier, J. A., and Weinberg, D. H. (2013). The neutral hydrogen content of galaxies in cosmological hydrodynamic simulations. *MNRAS*, 434:2645–2663.

Davies, L. J. M., Robotham, A. S. G., Driver, S. P., Alpaslan, M., Baldry, I. K., Bland-Hawthorn, J., Brough, S., Brown, M. J. I., Cluver, M. E., Holwerda, B. W., Hopkins, A. M., Lara-López, M. A., Mahajan, S., Moffett, A. J., Owers, M. S., and Phillipps, S. (2016). Galaxy And Mass Assembly

(GAMA): growing up in a bad neighbourhood - how do low-mass galaxies become passive? *MNRAS*, 455:4013–4029.

Davis, M., Efstathiou, G., Frenk, C. S., and White, S. D. M. (1985). The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394.

Davis, T. A., Alatalo, K., Sarzi, M., Bureau, M., Young, L. M., Blitz, L., Serra, P., Crocker, A. F., Krajnović, D., McDermid, R. M., Bois, M., Bournaud, F., Cappellari, M., Davies, R. L., Duc, P.-A., de Zeeuw, P. T., Emsellem, E., Khochfar, S., Kuntschner, H., Lablanche, P.-Y., Morganti, R., Naab, T., Oosterloo, T., Scott, N., and Weijmans, A.-M. (2011). The ATLAS^{3D} project - X. On the origin of the molecular and ionized gas in early-type galaxies. *MNRAS*, 417:882–899.

de Jong, R. S. (1996). Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. III. The statistics of the disk and bulge parameters. *A&A*, 313:45–64.

De Lucia, G. and Blaizot, J. (2007). The hierarchical formation of the brightest cluster galaxies. *MNRAS*, 375:2–14.

De Lucia, G., Boylan-Kolchin, M., Benson, A. J., Fontanot, F., and Monaco, P. (2010). A semi-analytic model comparison - gas cooling and galaxy mergers. *MNRAS*, 406:1533–1552.

De Lucia, G., Fontanot, F., and Wilman, D. (2012a). What determines the fraction of elliptical galaxies in clusters? *MNRAS*, 419:1324–1330.

De Lucia, G., Fontanot, F., Wilman, D., and Monaco, P. (2011). Times, environments and channels of bulge formation in a Lambda cold dark matter cosmology. *MNRAS*, 414:1439–1454.

De Lucia, G., Kauffmann, G., and White, S. D. M. (2004). Chemical enrichment of the intracluster and intergalactic medium in a hierarchical galaxy formation model. *MNRAS*, 349:1101–1116.

De Lucia, G., Muzzin, A., and Weinmann, S. (2014a). What Regulates Galaxy Evolution? Open questions in our understanding of galaxy formation and evolution. *New Astron. Rev.*, 62:1–14.

De Lucia, G., Tornatore, L., Frenk, C. S., Helmi, A., Navarro, J. F., and White, S. D. M. (2014b). Elemental abundances in Milky Way-like galaxies from a hierarchical galaxy formation model. *MNRAS*, 445:970–987.

De Lucia, G., Weinmann, S., Poggianti, B. M., Aragón-Salamanca, A., and Zaritsky, D. (2012b). The environmental history of group and cluster galaxies in a Λ cold dark matter universe. *MNRAS*, 423:1277–1292.

Decarli, R., Walter, F., Aravena, M., Carilli, C., Bouwens, R., da Cunha, E., Daddi, E., Ivison, R. J., Popping, G., Riechers, D., Smail, I. R., Swinbank, M., Weiss, A., Anguita, T., Assef, R. J., Bauer, F. E., Bell, E. F., Bertoldi, F., Chapman, S., Colina, L., Cortes, P. C., Cox, P., Dickinson, M., Elbaz, D., Gónzalez-López, J., Ibar, E., Infante, L., Hodge, J., Karim, A., Le Fevre, O.,

Magnelli, B., Neri, R., Oesch, P., Ota, K., Rix, H.-W., Sargent, M., Sheth, K., van der Wel, A., van der Werf, P., and Wagg, J. (2016). ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: CO Luminosity Functions and the Evolution of the Cosmic Density of Molecular Gas. *ApJ*, 833:69.

DeFelippis, D., Genel, S., Bryan, G. L., and Fall, S. M. (2017). The Impact of Galactic Winds on the Angular Momentum of Disk Galaxies in the Illustris Simulation. *ApJ*, 841:16.

Dekel, A., Sari, R., and Ceverino, D. (2009). Formation of Massive Galaxies at High Redshift: Cold Streams, Clumpy Disks, and Compact Spheroids. *ApJ*, 703:785–801.

Driver, S. P., Hill, D. T., Kelvin, L. S., Robotham, A. S. G., Liske, J., Norberg, P., Baldry, I. K., Bamford, S. P., Hopkins, A. M., Loveday, J., Peacock, J. A., Andrae, E., Bland-Hawthorn, J., Brough, S., Brown, M. J. I., Cameron, E., Ching, J. H. Y., Colless, M., Conselice, C. J., Croom, S. M., Cross, N. J. G., de Propris, R., Dye, S., Drinkwater, M. J., Ellis, S., Graham, A. W., Grootes, M. W., Gunawardhana, M., Jones, D. H., van Kampen, E., Maraston, C., Nichol, R. C., Parkinson, H. R., Phillipps, S., Pimbblet, K., Popescu, C. C., Prescott, M., Roseboom, I. G., Sadler, E. M., Sansom, A. E., Sharp, R. G., Smith, D. J. B., Taylor, E., Thomas, D., Tuffs, R. J., Wijesinghe, D., Dunne, L., Frenk, C. S., Jarvis, M. J., Madore, B. F., Meyer, M. J., Seibert, M., Staveley-Smith, L., Sutherland, W. J., and Warren, S. J. (2011). Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413:971–995.

Drory, N. and Fisher, D. B. (2007). A Connection between Bulge Properties and the Bimodality of Galaxies. *ApJ*, 664:640–649.

Dubois, Y., Devriendt, J., Slyz, A., and Teyssier, R. (2010). Jet-regulated cooling catastrophe. *MNRAS*, 409:985–1001.

Duffy, A. R., Kay, S. T., Battye, R. A., Booth, C. M., Dalla Vecchia, C., and Schaye, J. (2012). Modelling neutral hydrogen in galaxies using cosmological hydrodynamical simulations. *MNRAS*, 420:2799–2818.

Dunn, R. J. H. and Fabian, A. C. (2006). Investigating AGN heating in a sample of nearby clusters. *MNRAS*, 373:959–971.

Dunn, R. J. H. and Fabian, A. C. (2008). Investigating heating and cooling in the BCS and B55 cluster samples. *MNRAS*, 385:757–768.

Efstathiou, G., Lake, G., and Negroponte, J. (1982). The stability and masses of disc galaxies. *MNRAS*, 199:1069–1088.

Eke, V., Efstathiou, G., and Wright, L. (2000). The cosmological dependence of galactic specific angular momenta. *MNRAS*, 315:L18–L22.

Elmegreen, B. G. (1989). A pressure and metallicity dependence for molecular cloud correlations and the calibration of mass. *ApJ*, 338:178–196.

Elmegreen, B. G., Bournaud, F., and Elmegreen, D. M. (2008). Bulge Formation by the Coalescence of Giant Clumps in Primordial Disk Galaxies. *ApJ*, 688:67–77.

Elmegreen, B. G., Kimura, T., and Tosa, M. (1995). Sequential Star Formation in OB Associations: The Role of Molecular Cloud Turbulence. *ApJ*, 451:675.

Elmergreen, B. G. (1975). Isolated Stroemgren spheres as a source of galactic H alpha emission. *ApJ*, 198:L31–L35.

Fabian, A. C., Nulsen, P. E. J., and Canizares, C. R. (1991). Cooling flows in clusters of galaxies. *A&ARv*, 2:191–226.

Fall, S. M. (1983). Galaxy formation - Some comparisons between theory and observation. In Athanassoula, E., editor, *Internal Kinematics and Dynamics of Galaxies*, volume 100 of *IAU Symposium*, pages 391–398.

Fall, S. M. and Efstathiou, G. (1980). Formation and rotation of disc galaxies with haloes. *MNRAS*, 193:189–206.

Fall, S. M. and Romanowsky, A. J. (2013). Angular Momentum and Galaxy Formation Revisited: Effects of Variable Mass-to-light Ratios. *ApJ*, 769:L26.

Ferland, G. J., Korista, K. T., Verner, D. A., Ferguson, J. W., Kingdon, J. B., and Verner, E. M. (1998). CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. *PASP*, 110:761–778.

Fernández, X., Gim, H. B., van Gorkom, J. H., Yun, M. S., Momjian, E., Popping, A., Chomiuk, L., Hess, K. M., Hunt, L., Kreckel, K., Lucero, D., Maddox, N., Oosterloo, T., Pisano, D. J., Verheijen, M. A. W., Hales, C. A., Chung, A., Dodson, R., Golap, K., Gross, J., Henning, P., Hibbard, J., Jaffé, Y. L., Donovan Meyer, J., Meyer, M., Sanchez-Barrantes, M., Schiminovich, D., Wicenec, A., Wilcots, E., Bershady, M., Scoville, N., Strader, J., Tremou, E., Salinas, R., and Chávez, R. (2016). Highest Redshift Image of Neutral Hydrogen in Emission: A CHILES Detection of a Starbursting Galaxy at z = 0.376. *ApJ*, 824:L1.

Fisher, D. B. and Drory, N. (2008). The Structure of Classical Bulges and Pseudobulges: the Link Between Pseudobulges and SÉRSIC Index. *AJ*, 136:773–839.

Fisher, D. B. and Drory, N. (2010). Bulges of Nearby Galaxies with Spitzer: Scaling Relations in Pseudobulges and Classical Bulges. *ApJ*, 716:942–969.

Fisher, D. B., Drory, N., and Fabricius, M. H. (2009). Bulges of Nearby Galaxies with Spitzer: The Growth of Pseudobulges in Disk Galaxies and its Connection to Outer Disks. *ApJ*, 697:630–650.

Fontanot, F., De Lucia, G., Hirschmann, M., Bruzual, G., Charlot, S., and Zibetti, S. (2017a). Variations of the stellar initial mass function in semi-analytical models: implications for the mass assembly and the chemical enrichment of galaxies in the GAEA model. *MNRAS*, 464:3812–3824.

Fontanot, F., De Lucia, G., Monaco, P., Somerville, R. S., and Santini, P. (2009). The many manifestations of downsizing: hierarchical galaxy formation models confront observations. *MNRAS*, 397:1776–1790.

Fontanot, F., Hirschmann, M., and De Lucia, G. (2017b). Strong Stellar-driven Outflows Shape the Evolution of Galaxies at Cosmic Dawn. *ApJ*, 842:L14.

Forbes, J., Krumholz, M., and Burkert, A. (2012). Evolving Gravitationally Unstable Disks over Cosmic Time: Implications for Thick Disk Formation. *ApJ*, 754:48.

Forbes, J. C., Krumholz, M. R., Burkert, A., and Dekel, A. (2014). Balance among gravitational instability, star formation and accretion determines the structure and evolution of disc galaxies. *MNRAS*, 438:1552–1576.

Fosalba, P., Crocce, M., Gaztañaga, E., and Castander, F. J. (2015). The MICE grand challenge lightcone simulation - I. Dark matter clustering. *MNRAS*, 448:2987–3000.

Friedmann, A. (1922). Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386.

Fu, J., Guo, Q., Kauffmann, G., and Krumholz, M. R. (2010). The atomic-to-molecular transition and its relation to the scaling properties of galaxy discs in the local Universe. *MNRAS*, 409:515–530.

Furlong, M., Bower, R. G., Theuns, T., Schaye, J., Crain, R. A., Schaller, M., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J., Jenkins, A., and Rosas-Guevara, Y. M. (2015). Evolution of galaxy stellar masses and star formation rates in the EAGLE simulations. *MNRAS*, 450:4486–4504.

Gadotti, D. A. (2009). Structural properties of pseudo-bulges, classical bulges and elliptical galaxies: a Sloan Digital Sky Survey perspective. *MNRAS*, 393:1531–1552.

Galametz, A., Grazian, A., Fontana, A., Ferguson, H. C., Ashby, M. L. N., Barro, G., Castellano, M., Dahlen, T., Donley, J. L., Faber, S. M., Grogin, N., Guo, Y., Huang, K.-H., Kocevski, D. D., Koekemoer, A. M., Lee, K.-S., McGrath, E. J., Peth, M., Willner, S. P., Almaini, O., Cooper, M., Cooray, A., Conselice, C. J., Dickinson, M., Dunlop, J. S., Fazio, G. G., Foucaud, S., Gardner, J. P., Giavalisco, M., Hathi, N. P., Hartley, W. G., Koo, D. C., Lai, K., de Mello, D. F., McLure, R. J., Lucas, R. A., Paris, D., Pentericci, L., Santini, P., Simpson, C., Sommariva, V., Targett, T., Weiner, B. J., Wuyts, S., and the CANDELS Team (2013). CANDELS Multiwavelength Catalogs: Source Identification and Photometry in the CANDELS UKIDSS Ultra-deep Survey Field. *ApJS*, 206:10.

Gao, L. and White, S. D. M. (2007). Assembly bias in the clustering of dark matter haloes. *MNRAS*, 377:L5–L9.

Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., and Springel, V. (2004). The subhalo populations of ΛCDM dark haloes. *MNRAS*, 355:819–834.

Gargiulo, I. D., Cora, S. A., Padilla, N. D., Muñoz Arancibia, A. M., Ruiz, A. N., Orsi, A. A., Tecce, T. E., Weidner, C., and Bruzual, G. (2015). Chemoarchaeological downsizing in a hierarchical universe: impact of a top-heavy IGIMF. *MNRAS*, 446:3820–3841.

Gaspari, M., Brighenti, F., D'Ercole, A., and Melioli, C. (2011). AGN feedback in galaxy groups: the delicate touch of self-regulated outflows. *MNRAS*, 415:1549–1568.

Gaspari, M., Brighenti, F., and Temi, P. (2012). Mechanical AGN feedback: controlling the thermodynamical evolution of elliptical galaxies. *MNRAS*, 424:190–209.

Genel, S., Fall, S. M., Hernquist, L., Vogelsberger, M., Snyder, G. F., Rodriguez-Gomez, V., Sijacki, D., and Springel, V. (2015). Galactic Angular Momentum in the Illustris Simulation: Feedback and the Hubble Sequence. *ApJ*, 804:L40.

Genel, S., Vogelsberger, M., Springel, V., Sijacki, D., Nelson, D., Snyder, G., Rodriguez-Gomez, V., Torrey, P., and Hernquist, L. (2014). Introducing the Illustris project: the evolution of galaxy populations across cosmic time. *MNRAS*, 445:175–200.

Genzel, R., Newman, S., Jones, T., Förster Schreiber, N. M., Shapiro, K., Genel, S., Lilly, S. J., Renzini, A., Tacconi, L. J., Bouché, N., Burkert, A., Cresci, G., Buschkamp, P., Carollo, C. M., Ceverino, D., Davies, R., Dekel, A., Eisenhauer, F., Hicks, E., Kurk, J., Lutz, D., Mancini, C., Naab, T., Peng, Y., Sternberg, A., Vergani, D., and Zamorani, G. (2011). The Sins Survey of z ˜ 2 Galaxy Kinematics: Properties of the Giant Star-forming Clumps. *ApJ*, 733:101.

Geréb, K., Morganti, R., Oosterloo, T. A., Guglielmino, G., and Prandoni, I. (2013). The Lockman Hole project: gas and galaxy properties from a stacking experiment. *A&A*, 558:A54.

Giavalisco, M., Vanzella, E., Salimbeni, S., Tripp, T. M., Dickinson, M., Cassata, P., Renzini, A., Guo, Y., Ferguson, H. C., Nonino, M., Cimatti, A., Kurk, J., Mignoli, M., and Tang, Y. (2011). Discovery of Cold, Pristine Gas Possibly Accreting onto an Overdensity of Star-forming Galaxies at Redshift z ˜ 1.6. *ApJ*, 743:95.

Gill, S. P. D., Knebe, A., and Gibson, B. K. (2004). The evolution of substructure - I. A new identification method. *MNRAS*, 351:399–409.

Gingold, R. A. and Monaghan, J. J. (1977). Smoothed particle hydrodynamics - Theory and application to non-spherical stars. *MNRAS*, 181:375–389.

Giovanelli, R., Haynes, M. P., Kent, B. R., Perillat, P., Saintonge, A., Brosch, N., Catinella, B., Hoffman, G. L., Stierwalt, S., Spekkens, K., Lerner, M. S., Masters, K. L., Momjian, E., Rosenberg, J. L., Springob, C. M., Boselli, A., Charmandaris, V., Darling, J. K., Davies, J., Garcia Lambas, D., Gavazzi, G., Giovanardi, C., Hardy, E., Hunt, L. K., Iovino, A., Karachentsev, I. D., Karachentseva, V. E., Koopmann, R. A., Marinoni, C., Minchin, R., Muller, E., Putman, M., Pantoja, C., Salzer, J. J., Scodeggio, M., Skillman, E., Solanes, J. M., Valotto, C., van Driel, W., and van Zee, L. (2005). The Arecibo Legacy Fast ALFA Survey. I. Science Goals, Survey Design, and Strategy. *AJ*, 130:2598–2612.

Glover, S. C. O. and Clark, P. C. (2012). Approximations for modelling CO chemistry in giant molecular clouds: a comparison of approaches. *MNRAS*, 421:116–131.

Glover, S. C. O. and Mac Low, M.-M. (2011). On the relationship between molecular hydrogen and carbon monoxide abundances in molecular clouds. *MNRAS*, 412:337–350.

Gnedin, N. Y. (1995). Softened Lagrangian hydrodynamics for cosmology. *ApJS*, 97:231–257.

Gnedin, N. Y. (2000). Effect of Reionization on Structure Formation in the Universe. *ApJ*, 542:535–541.

Gnedin, N. Y. and Kravtsov, A. V. (2011). Environmental Dependence of the Kennicutt-Schmidt Relation in Galaxies. *ApJ*, 728:88.

Gnedin, N. Y., Tassis, K., and Kravtsov, A. V. (2009). Modeling Molecular Hydrogen and Star Formation in Cosmological Simulations. *ApJ*, 697:55–67.

Governato, F., Brook, C., Mayer, L., Brooks, A., Rhee, G., Wadsley, J., Jonsson, P., Willman, B., Stinson, G., Quinn, T., and Madau, P. (2010). Bulgeless dwarf galaxies and dark matter cores from supernova-driven outflows. *Nature*, 463:203–206.

Governato, F., Mayer, L., Wadsley, J., Gardner, J. P., Willman, B., Hayashi, E., Quinn, T., Stadel, J., and Lake, G. (2004). The Formation of a Realistic Disk Galaxy in Λ-dominated Cosmologies. *ApJ*, 607:688–696.

Greggio, L. (2005). The rates of type Ia supernovae. I. Analytical formulations. *A&A*, 441:1055–1078.

Grogin, N. A., Kocevski, D. D., Faber, S. M., Ferguson, H. C., Koekemoer, A. M., Riess, A. G., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L. N., Barden, M., Bell, E. F., Bournaud, F., Brown, T. M., Caputi, K. I., Casertano, S., Cassata, P., Castellano, M., Challis, P., Chary, R.-R., Cheung, E., Cirasuolo, M., Conselice, C. J., Roshan Cooray, A., Croton, D. J., Daddi, E., Dahlen, T., Davé, R., de Mello, D. F., Dekel, A., Dickinson, M., Dolch, T., Donley, J. L., Dunlop, J. S., Dutton, A. A., Elbaz, D., Fazio, G. G., Filippenko, A. V., Finkelstein, S. L., Fontana, A., Gardner, J. P., Garnavich, P. M., Gawiser, E., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N. P., Häussler, B., Hopkins, P. F., Huang, J.-S., Huang, K.-H., Jha, S. W., Kartaltepe, J. S., Kirshner, R. P., Koo, D. C., Lai, K., Lee, K.-S., Li, W., Lotz, J. M., Lucas, R. A., Madau, P., McCarthy, P. J., McGrath, E. J., McIntosh, D. H., McLure, R. J., Mobasher, B., Moustakas, L. A., Mozena, M., Nandra, K., Newman, J. A., Niemi, S.-M., Noeske, K. G., Papovich, C. J., Pentericci, L., Pope, A., Primack, J. R., Rajan, A., Ravindranath, S., Reddy, N. A., Renzini, A., Rix, H.-W., Robaina, A. R., Rodney, S. A., Rosario, D. J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Somerville, R. S., Spinrad, H., Straughn, A. N., Strolger, L.-G., Telford, O., Teplitz, H. I., Trump, J. R., van der Wel, A., Villforth, C., Wechsler, R. H., Weiner, B. J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., Yan, H.-J., and Yun, M. S. (2011). CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. *ApJS*, 197:35.

Guedes, J., Callegari, S., Madau, P., and Mayer, L. (2011). Forming Realistic Late-type Spirals in a ΛCDM Universe: The Eris Simulation. *ApJ*, 742:76.

Gunn, J. E. and Gott, III, J. R. (1972). On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, 176:1.

Guo, Q., White, S., Angulo, R. E., Henriques, B., Lemson, G., Boylan-Kolchin, M., Thomas, P., and Short, C. (2013). Galaxy formation in WMAP1 and WMAP7 cosmologies. *MNRAS*, 428:1351–1365.

Guo, Q., White, S., Boylan-Kolchin, M., De Lucia, G., Kauffmann, G., Lemson, G., Li, C., Springel, V., and Weinmann, S. (2011). From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. *MNRAS*, 413:101–131.

Hao, C. N., Mao, S., Deng, Z. G., Xia, X. Y., and Wu, H. (2006). Isophotal shapes of elliptical/lenticular galaxies from the Sloan Digital Sky Survey. *MNRAS*, 370:1339–1350.

Harrison, E. R. (1970). Fluctuations at the Threshold of Classical Cosmology. *Phys. Rev. D*, 1:2726–2730.

Haynes, M. P. and Giovanelli, R. (1984). Neutral hydrogen in isolated galaxies. IV - Results for the Arecibo sample. *AJ*, 89:758–800.

Heckman, T. M., Lehnert, M. D., Strickland, D. K., and Armus, L. (2000). Absorption-Line Probes of Gas and Dust in Galactic Superwinds. *ApJS*, 129:493–516.

Helly, J. C., Cole, S., Frenk, C. S., Baugh, C. M., Benson, A., Lacey, C., and Pearce, F. R. (2003). A comparison of gas dynamics in smooth particle hydrodynamics and semi-analytic models of galaxy formation. *MNRAS*, 338:913–925.

Henriques, B. M. B., White, S. D. M., Thomas, P. A., Angulo, R., Guo, Q., Lemson, G., Springel, V., and Overzier, R. (2015). Galaxy formation in the Planck cosmology - I. Matching the observed evolution of star formation rates, colours and stellar masses. *MNRAS*, 451:2663–2680.

Henriques, B. M. B., White, S. D. M., Thomas, P. A., Angulo, R. E., Guo, Q., Lemson, G., and Springel, V. (2013). Simulations of the galaxy population constrained by observations from z = 3 to the present day: implications for galactic winds and the fate of their ejecta. *MNRAS*, 431:3373–3395.

Hernandez, X., Park, C., Cervantes-Sodi, B., and Choi, Y.-Y. (2007). Empirical distributions of galactic λ spin parameters from the Sloan Digital Sky Survey. *MNRAS*, 375:163–170.

Hernquist, L. (1992). Structure of merger remnants. I - Bulgeless progenitors. *ApJ*, 400:460–475.

Hernquist, L. (1993). Structure of merger remnants. II - Progenitors with rotating bulges. *ApJ*, 409:548–562.

Heyl, J. S., Hernquist, L., and Spergel, D. N. (1996). Structure of Merger Remnants. V. Kinematics. *ApJ*, 463:69.

Hirschmann, M., De Lucia, G., and Fontanot, F. (2016). Galaxy assembly, stellar feedback and metal enrichment: the view from the gaea model. *MNRAS*, 461:1760–1785.

Holwerda, B. W., Blyth, S.-L., and Baker, A. J. (2012). Looking at the distant universe with the MeerKAT Array (LADUMA). In Tuffs, R. J. and Popescu, C. C., editors, *The Spectral Energy Distribution of Galaxies - SED 2011*, volume 284 of *IAU Symposium*, pages 496–499.

Hopkins, P. F., Cox, T. J., Younger, J. D., and Hernquist, L. (2009a). How do Disks Survive Mergers? *ApJ*, 691:1168–1201.

Hopkins, P. F., Hernquist, L., Cox, T. J., Keres, D., and Wuyts, S. (2009b). Dissipation and Extra Light in Galactic Nuclei. IV. Evolution in the Scaling Relations of Spheroids. *ApJ*, 691:1424–1458.

Hopkins, P. F., Kereš, D., Oñorbe, J., Faucher-Giguère, C.-A., Quataert, E., Murray, N., and Bullock, J. S. (2014). Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. *MNRAS*, 445:581–603.

Huang, S., Haynes, M. P., Giovanelli, R., and Brinchmann, J. (2012). The Arecibo Legacy Fast ALFA Survey: The Galaxy Population Detected by ALFALFA. *ApJ*, 756:113.

Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173.

Huertas-Company, M., Shankar, F., Mei, S., Bernardi, M., Aguerri, J. A. L., Meert, A., and Vikram, V. (2013). No Evidence for a Dependence of the Mass-Size Relation of Early-type Galaxies on Environment in the Local Universe. *ApJ*, 779:29.

Hunt, L. K., García-Burillo, S., Casasola, V., Caselli, P., Combes, F., Henkel, C., Lundgren, A., Maiolino, R., Menten, K. M., Testi, L., and Weiss, A. (2015). Molecular depletion times and the CO-to-H_2 conversion factor in metal-poor galaxies. *A&A*, 583:A114.

Ichikawa, T., Kajisawa, M., and Akhlaghi, M. (2012). A universal stellar mass-size relation of galaxies in the GOODS-North region. *MNRAS*, 422:1014–1027.

Jaffe, W. (1983). A simple model for the distribution of light in spherical galaxies. *MNRAS*, 202:995–999.

Jaffé, Y. L., Poggianti, B. M., Verheijen, M. A. W., Deshev, B. Z., and van Gorkom, J. H. (2012). Gas Reservoirs and Star Formation in a Forming Galaxy Cluster at zbsime0.2. *ApJ*, 756:L28.

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., and Yoshida, N. (2001). The mass function of dark matter haloes. *MNRAS*, 321:372–384.

Jiang, C. Y., Jing, Y. P., Faltenbacher, A., Lin, W. P., and Li, C. (2008). A Fitting Formula for the Merger Timescale of Galaxies in Hierarchical Clustering. *ApJ*, 675:1095–1105.

Jiang, F. and van den Bosch, F. C. (2014). Generating merger trees for dark matter haloes: a comparison of methods. *MNRAS*, 440:193–207.

Johnston, S., Taylor, R., Bailes, M., Bartel, N., Baugh, C., Bietenholz, M., Blake, C., Braun, R., Brown, J., Chatterjee, S., Darling, J., Deller, A., Dodson, R., Edwards, P., Ekers, R., Ellingsen, S., Feain, I., Gaensler, B., Haverkorn, M., Hobbs, G., Hopkins, A., Jackson, C., James, C., Joncas, G., Kaspi, V., Kilborn, V., Koribalski, B., Kothes, R., Landecker, T., Lenc, E., Lovell, J., Macquart, J.-P., Manchester, R., Matthews, D., McClure-Griffiths, N., Norris, R., Pen, U.-L., Phillips, C., Power, C., Protheroe, R., Sadler, E., Schmidt, B., Stairs, I., Staveley-Smith, L., Stil, J., Tingay, S., Tzioumis, A., Walker, M., Wall, J., and Wolleben, M. (2008). Science with ASKAP. The Australian square-kilometre-array pathfinder. *Experimental Astronomy*, 22:151–273.

Kaiser, N. (1984). On the spatial correlations of Abell clusters. *ApJ*, 284:L9–L12.

Karakas, A. I. (2010). Updated stellar yields from asymptotic giant branch models. *MNRAS*, 403:1413–1425.

Kauffmann, G. (1996). The age of elliptical galaxies and bulges in a merger model. *MNRAS*, 281:487–492.

Kauffmann, G., Colberg, J. M., Diaferio, A., and White, S. D. M. (1999). Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0. *MNRAS*, 303:188–206.

Kauffmann, G. and Haehnelt, M. (2000). A unified model for the evolution of galaxies and quasars. *MNRAS*, 311:576–588.

Kauffmann, G., White, S. D. M., and Guiderdoni, B. (1993). The Formation and Evolution of Galaxies Within Merging Dark Matter Haloes. *MNRAS*, 264:201.

Kennicutt, Jr., R. C. (1989). The star formation law in galactic disks. *ApJ*, 344:685–703.

Kennicutt, Jr., R. C. (1998). The Global Schmidt Law in Star-forming Galaxies. *ApJ*, 498:541–552.

Kennicutt, Jr., R. C., Calzetti, D., Walter, F., Helou, G., Hollenbach, D. J., Armus, L., Bendo, G., Dale, D. A., Draine, B. T., Engelbracht, C. W., Gordon, K. D., Prescott, M. K. M., Regan, M. W., Thornley, M. D., Bot, C., Brinks, E., de Blok, E., de Mello, D., Meyer, M., Moustakas, J., Murphy, E. J., Sheth, K., and Smith, J. D. T. (2007). Star Formation in NGC 5194 (M51a). II. The Spatially Resolved Star Formation Law. *ApJ*, 671:333–348.

Keres, D., Yun, M. S., and Young, J. S. (2003). CO Luminosity Functions for Far-Infrared- and B-Band-selected Galaxies and the First Estimate for Ω_{HI}+H_2. *ApJ*, 582:659–667.

Kerr, F. J. and Hindman, J. V. (1953). Preliminary report on a survey of 21 CM radiation from the Magellanic Clouds. *AJ*, 58:218.

Kim, H.-S., Baugh, C. M., Benson, A. J., Cole, S., Frenk, C. S., Lacey, C. G., Power, C., and Schneider, M. (2011). The spatial distribution of cold gas in hierarchical galaxy formation models. *MNRAS*, 414:2367–2385.

Kim, H.-S., Wyithe, J. S. B., Baugh, C. M., Lagos, C. d. P., Power, C., and Park, J. (2017). The spatial distribution of neutral hydrogen as traced by low H I mass galaxies. *MNRAS*, 465:111–122.

Kim, H.-S., Wyithe, J. S. B., Power, C., Park, J., Lagos, C. d. P., and Baugh, C. M. (2015). The H I mass function as a probe of photoionization feedback on low-mass galaxy formation. *MNRAS*, 453:2315–2325.

Kim, J.-h. and Lee, J. (2013). How does the surface density and size of disc galaxies measured in hydrodynamic simulations correlate with the halo spin parameter? *MNRAS*, 432:1701–1708.

Klypin, A. A., Trujillo-Gomez, S., and Primack, J. (2011). Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. *ApJ*, 740:102.

Knapp, G. R., Turner, E. L., and Cunniffe, P. E. (1985). The statistical distribution of the neutral-hydrogen content of elliptical galaxies. *AJ*, 90:454–468.

Knollmann, S. R. and Knebe, A. (2009). AHF: Amiga's Halo Finder. *ApJS*, 182:608–624.

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., Grogin, N. A., Kocevski, D. D., Koo, D. C., Lai, K., Lotz, J. M., Lucas, R. A., McGrath, E. J., Ogaz, S., Rajan, A., Riess, A. G., Rodney, S. A., Strolger, L., Casertano, S., Castellano, M., Dahlen, T., Dickinson, M., Dolch, T., Fontana, A., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N. P., Huang, K.-H., van der Wel, A., Yan, H.-J., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L. N., Barden, M., Bell, E. F., Bournaud, F., Brown, T. M., Caputi, K. I., Cassata, P., Challis, P. J., Chary, R.-R., Cheung, E., Cirasuolo, M., Conselice, C. J., Roshan Cooray, A., Croton, D. J., Daddi, E., Davé, R., de Mello, D. F., de Ravel, L., Dekel, A., Donley, J. L., Dunlop, J. S., Dutton, A. A., Elbaz, D., Fazio, G. G., Filippenko, A. V., Finkelstein, S. L., Frazer, C., Gardner, J. P., Garnavich, P. M., Gawiser, E., Gruetzbauch, R., Hartley, W. G., Häussler, B., Herrington, J., Hopkins, P. F., Huang, J.-S., Jha, S. W., Johnson, A., Kartaltepe, J. S., Khostovan, A. A., Kirshner, R. P., Lani, C., Lee, K.-S., Li, W., Madau, P., McCarthy, P. J., McIntosh, D. H., McLure, R. J., McPartland, C., Mobasher, B., Moreira, H., Mortlock, A., Moustakas, L. A., Mozena, M., Nandra, K., Newman, J. A., Nielsen, J. L., Niemi, S., Noeske, K. G., Papovich, C. J., Pentericci, L., Pope, A., Primack, J. R., Ravindranath, S., Reddy, N. A., Renzini, A., Rix, H.-W., Robaina, A. R., Rosario, D. J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Snyder, D., Somerville, R. S., Spinrad, H., Straughn, A. N., Telford, O., Teplitz, H. I., Trump, J. R., Vargas, C., Villforth, C., Wagner, C. R., Wandro, P., Wechsler, R. H., Weiner, B. J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., and Yun,

M. S. (2011). CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey: The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics. *ApJS*, 197:36.

Kormendy, J. (1982). Rotation of the bulge components of barred galaxies. *ApJ*, 257:75–88.

Kormendy, J. and Illingworth, G. (1982). Rotation of the bulge components of disk galaxies. *ApJ*, 256:460–480.

Kormendy, J. and Kennicutt, Jr., R. C. (2004). Secular Evolution and the Formation of Pseudobulges in Disk Galaxies. *ARA&A*, 42:603–683.

Kregel, M., van der Kruit, P. C., and de Grijs, R. (2002). Flattening and truncation of stellar discs in edge-on spiral galaxies. *MNRAS*, 334:646–668.

Krumholz, M. and Burkert, A. (2010). On the Dynamics and Evolution of Gravitational Instability-dominated Disks. *ApJ*, 724:895–907.

Krumholz, M. R. (2013). The star formation law in molecule-poor galaxies. *MNRAS*, 436:2747–2762.

Krumholz, M. R. (2014). The big problems in star formation: The star formation rate, stellar clustering, and the initial mass function. *Phys. Rep.*, 539:49–134.

Krumholz, M. R., McKee, C. F., and Tumlinson, J. (2008). The Atomic-to-Molecular Transition in Galaxies. I. An Analytic Approximation for Photodissociation Fronts in Finite Clouds. *ApJ*, 689:865–882.

Krumholz, M. R., McKee, C. F., and Tumlinson, J. (2009). The Atomic-to-Molecular Transition in Galaxies. II: H I and H_2 Column Densities. *ApJ*, 693:216–235.

Kuhlen, M., Madau, P., and Krumholz, M. R. (2013). Dwarf Galaxy Formation with H_2-regulated Star Formation. II. Gas-rich Dark Galaxies at Redshift 2.5. *ApJ*, 776:34.

Lacey, C. and Cole, S. (1993). Merger rates in hierarchical models of galaxy formation. *MNRAS*, 262:627–649.

Lacey, C. and Silk, J. (1991). Tidally triggered galaxy formation. I - Evolution of the galaxy luminosity function. *ApJ*, 381:14–32.

Lagos, C. D. P., Baugh, C. M., Lacey, C. G., Benson, A. J., Kim, H.-S., and Power, C. (2011a). Cosmic evolution of the atomic and molecular gas contents of galaxies. *MNRAS*, 418:1649–1667.

Lagos, C. d. P., Crain, R. A., Schaye, J., Furlong, M., Frenk, C. S., Bower, R. G., Schaller, M., Theuns, T., Trayford, J. W., Bahé, Y. M., and Dalla Vecchia, C. (2015a). Molecular hydrogen abundances of galaxies in the EAGLE simulations. *MNRAS*, 452:3815–3837.

Lagos, C. d. P., Lacey, C. G., and Baugh, C. M. (2013). A dynamical model of supernova feedback: gas outflows from the interstellar medium. *MNRAS*, 436:1787–1817.

Lagos, C. D. P., Lacey, C. G., Baugh, C. M., Bower, R. G., and Benson, A. J. (2011b). On the impact of empirical and theoretical star formation laws on galaxy formation. *MNRAS*, 416:1566–1584.

Lagos, C. D. P., Padilla, N. D., and Cora, S. A. (2009). Black hole spin and radio loudness in a Λ cold dark matter universe. *MNRAS*, 395:625–636.

Lagos, C. d. P., Padilla, N. D., Davis, T. A., Lacey, C. G., Baugh, C. M., Gonzalez-Perez, V., Zwaan, M. A., and Contreras, S. (2015b). The origin of the atomic and molecular gas contents of early-type galaxies - II. Misaligned gas accretion. *MNRAS*, 448:1271–1287.

Lagos, C. d. P., Theuns, T., Stevens, A. R. H., Cortese, L., Padilla, N. D., Davis, T. A., Contreras, S., and Croton, D. (2017). Angular momentum evolution of galaxies in EAGLE. *MNRAS*, 464:3850–3870.

Lah, P., Chengalur, J. N., Briggs, F. H., Colless, M., de Propris, R., Pracy, M. B., de Blok, W. J. G., Fujita, S. S., Ajiki, M., Shioya, Y., Nagao, T., Murayama, T., Taniguchi, Y., Yagi, M., and Okamura, S. (2007). The HI content of star-forming galaxies at z = 0.24. *MNRAS*, 376:1357–1366.

Laigle, C., McCracken, H. J., Ilbert, O., Hsieh, B. C., Davidzon, I., Capak, P., Hasinger, G., Silverman, J. D., Pichon, C., Coupon, J., Aussel, H., Le Borgne, D., Caputi, K., Cassata, P., Chang, Y.-Y., Civano, F., Dunlop, J., Fynbo, J., Kartaltepe, J. S., Koekemoer, A., Le Fèvre, O., Le Floc'h, E., Leauthaud, A., Lilly, S., Lin, L., Marchesi, S., Milvang-Jensen, B., Salvato, M., Sanders, D. B., Scoville, N., Smolcic, V., Stockmann, M., Taniguchi, Y., Tasca, L., Toft, S., Vaccari, M., and Zabl, J. (2016). The COSMOS2015 Catalog: Exploring the 1¡z¡6 Universe with Half a Million Galaxies. *ApJS*, 224:24.

Landy, S. D. and Szalay, A. S. (1993). Bias and variance of angular correlation functions. *ApJ*, 412:64–71.

Lange, R., Driver, S. P., Robotham, A. S. G., Kelvin, L. S., Graham, A. W., Alpaslan, M., Andrews, S. K., Baldry, I. K., Bamford, S., Bland-Hawthorn, J., Brough, S., Cluver, M. E., Conselice, C. J., Davies, L. J. M., Haeussler, B., Konstantopoulos, I. S., Loveday, J., Moffett, A. J., Norberg, P., Phillipps, S., Taylor, E. N., López-Sánchez, Á. R., and Wilkins, S. M. (2015). Galaxy And Mass Assembly (GAMA): mass-size relations of z ¡ 0.1 galaxies subdivided by Sérsic index, colour and morphology. *MNRAS*, 447:2603–2630.

Lange, R., Moffett, A. J., Driver, S. P., Robotham, A. S. G., Lagos, C. d. P., Kelvin, L. S., Conselice, C., Margalef-Bentabol, B., Alpaslan, M., Baldry, I., Bland-Hawthorn, J., Bremer, M., Brough, S., Cluver, M., Colless, M., Davies, L. J. M., Häußler, B., Holwerda, B. W., Hopkins, A. M., Kafle, P. R., Kennedy, R., Liske, J., Phillipps, S., Popescu, C. C., Taylor, E. N., Tuffs, R., van Kampen, E., and Wright, A. H. (2016). Galaxy And Mass Assembly (GAMA): M_star - R_e relations of z = 0 bulges, discs and spheroids. *MNRAS*, 462:1470–1500.

Larson, R. B., Tinsley, B. M., and Caldwell, C. N. (1980). The evolution of disk galaxies and the origin of S0 galaxies. *ApJ*, 237:692–707.

Lemson, G. and Virgo Consortium, t. (2006). Halo and Galaxy Formation Histories from the Millennium Simulation: Public release of a VO-oriented and SQL-queryable database for studying the evolution of galaxies in the LambdaCDM cosmogony. *ArXiv Astrophysics e-prints*.

Leroy, A. K., Bolatto, A., Gordon, K., Sandstrom, K., Gratier, P., Rosolowsky, E., Engelbracht, C. W., Mizuno, N., Corbelli, E., Fukui, Y., and Kawamura, A. (2011). The CO-to-H_2 Conversion Factor from Infrared Dust Emission across the Local Group. *ApJ*, 737:12.

Leroy, A. K., Walter, F., Brinks, E., Bigiel, F., de Blok, W. J. G., Madore, B., and Thornley, M. D. (2008). The Star Formation Efficiency in Nearby Galaxies: Measuring Where Gas Forms Stars Effectively. *AJ*, 136:2782–2845.

Li, Y., Mo, H. J., and Gao, L. (2008). On halo formation times and assembly bias. *MNRAS*, 389:1419–1426.

Lu, Y., Kereš, D., Katz, N., Mo, H. J., Fardal, M., and Weinberg, M. D. (2011). On the algorithms of radiative cooling in semi-analytic models. *MNRAS*, 416:660–679.

Lu, Y., Mo, H. J., and Wechsler, R. H. (2015). Formation of disc galaxies in preheated media: a preventative feedback model. *MNRAS*, 446:1907–1923.

Lucy, L. B. (1977). A numerical approach to the testing of the fission hypothesis. *AJ*, 82:1013–1024.

MacArthur, L. A., Courteau, S., and Holtzman, J. A. (2003). Structure of Disk-dominated Galaxies. I. Bulge/Disk Parameters, Simulations, and Secular Evolution. *ApJ*, 582:689–722.

Magdis, G. E., Daddi, E., Elbaz, D., Sargent, M., Dickinson, M., Dannerbauer, H., Aussel, H., Walter, F., Hwang, H. S., Charmandaris, V., Hodge, J., Riechers, D., Rigopoulou, D., Carilli, C., Pannella, M., Mullaney, J., Leiton, R., and Scott, D. (2011). GOODS-Herschel: Gas-to-dust Mass Ratios and CO-to-H_2 Conversion Factors in Normal and Starbursting Galaxies at High-z. *ApJ*, 740:L15.

Mann, R. G., Peacock, J. A., and Heavens, A. F. (1998). Eulerian bias and the galaxy density field. *MNRAS*, 293:209.

Marasco, A., Crain, R. A., Schaye, J., Bahé, Y. M., van der Hulst, T., Theuns, T., and Bower, R. G. (2016). The environmental dependence of H I in galaxies in the EAGLE simulations. *MNRAS*, 461:2630–2649.

Martin, A. M., Giovanelli, R., Haynes, M. P., and Guzzo, L. (2012). The Clustering Characteristics of H I-selected Galaxies from the 40% ALFALFA Survey. *ApJ*, 750:38.

Martin, A. M., Papastergis, E., Giovanelli, R., Haynes, M. P., Springob, C. M., and Stierwalt, S. (2010). The Arecibo Legacy Fast ALFA Survey. X. The H I Mass Function and Ω_H I from the 40% ALFALFA Survey. *ApJ*, 723:1359–1374.

McGaugh, S. S. (2005). The Baryonic Tully-Fisher Relation of Galaxies with Extended Rotation Curves and the Stellar Mass of Rotating Galaxies. *ApJ*, 632:859–871.

McKee, C. F. and Ostriker, E. C. (2007). Theory of Star Formation. *ARA&A*, 45:565–687.

McLure, R., Pentericci, L., and VANDELS Team (2017). VANDELS: Exploring the Physics of High-redshift Galaxy Evolution. *The Messenger*, 167:31–35.

McNamara, B. R., Wise, M., Nulsen, P. E. J., David, L. P., Sarazin, C. L., Bautz, M., Markevitch, M., Vikhlinin, A., Forman, W. R., Jones, C., and Harris, D. E. (2000). Chandra X-Ray Observations of the Hydra A Cluster: An Interaction between the Radio Source and the X-Ray-emitting Gas. *ApJ*, 534:L135–L138.

Meyer, M. J., Zwaan, M. A., Webster, R. L., Brown, M. J. I., and Staveley-Smith, L. (2007). The Weak Clustering of Gas-rich Galaxies. *ApJ*, 654:702–713.

Meyer, M. J., Zwaan, M. A., Webster, R. L., Staveley-Smith, L., Ryan-Weber, E., Drinkwater, M. J., Barnes, D. G., Howlett, M., Kilborn, V. A., Stevens, J., Waugh, M., Pierce, M. J., Bhathal, R., de Blok, W. J. G., Disney, M. J., Ekers, R. D., Freeman, K. C., Garcia, D. A., Gibson, B. K., Harnett, J., Henning, P. A., Jerjen, H., Kesteven, M. J., Knezek, P. M., Koribalski, B. S., Mader, S., Marquarding, M., Minchin, R. F., O'Brien, J., Oosterloo, T., Price, R. M., Putman, M. E., Ryder, S. D., Sadler, E. M., Stewart, I. M., Stootman, F., and Wright, A. E. (2004). The HIPASS catalogue - I. Data presentation. *MNRAS*, 350:1195–1209.

Mo, H. J., Mao, S., and White, S. D. M. (1998). The formation of galactic discs. *MNRAS*, 295:319–336.

Mo, H. J. and White, S. D. M. (1996). An analytic model for the spatial clustering of dark matter haloes. *MNRAS*, 282:347–361.

Monaco, P., Benson, A. J., De Lucia, G., Fontanot, F., Borgani, S., and Boylan-Kolchin, M. (2014). A semi-analytic model comparison: testing cooling models against hydrodynamical simulations. *MNRAS*, 441:2058–2077.

Monaco, P., Theuns, T., and Taffoni, G. (2002). The pinocchio algorithm: pinpointing orbit-crossing collapsed hierarchical objects in a linear density field. *MNRAS*, 331:587–608.

Moore, B., Katz, N., Lake, G., Dressler, A., and Oemler, A. (1996). Galaxy harassment and the evolution of clusters of galaxies. *Nature*, 379:613–616.

Moster, B. P., Somerville, R. S., Maulbetsch, C., van den Bosch, F. C., Macciò, A. V., Naab, T., and Oser, L. (2010). Constraints on the Relationship between Stellar Mass and Halo Mass at Low and High Redshift. *ApJ*, 710:903–923.

Moustakas, J., Coil, A. L., Aird, J., Blanton, M. R., Cool, R. J., Eisenstein, D. J., Mendez, A. J., Wong, K. C., Zhu, G., and Arnouts, S. (2013). PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from z = 0-1. *ApJ*, 767:50.

Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., Hopkins, P. F., Quataert, E., and Murray, N. (2015). Gusty, gaseous flows of FIRE: galactic winds in cosmological simulations with explicit stellar feedback. *MNRAS*, 454:2691–2713.

Nagashima, M., Lacey, C. G., Baugh, C. M., Frenk, C. S., and Cole, S. (2005). The metal enrichment of the intracluster medium in hierarchical galaxy formation models. *MNRAS*, 358:1247–1266.

Nakamura, O., Fukugita, M., Yasuda, N., Loveday, J., Brinkmann, J., Schneider, D. P., Shimasaku, K., and SubbaRao, M. (2003). The Luminosity Function of Morphologically Classified Galaxies in the Sloan Digital Sky Survey. *AJ*, 125:1682–1688.

Narayan, R. and Fabian, A. C. (2011). Bondi flow from a slowly rotating hot atmosphere. *MNRAS*, 415:3721–3730.

Narayanan, D., Krumholz, M. R., Ostriker, E. C., and Hernquist, L. (2012). A general model for the CO-H_2 conversion factor in galaxies with applications to the star formation law. *MNRAS*, 421:3127–3146.

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. *ApJ*, 462:563.

Navarro, J. F. and Steinmetz, M. (2000). Dark Halo and Disk Galaxy Scaling Laws in Hierarchical Universes. *ApJ*, 538:477–488.

Neistein, E. and Dekel, A. (2008). Constructing merger trees that mimic N-body simulations. *MNRAS*, 383:615–626.

Neto, A. F., Gao, L., Bett, P., Cole, S., Navarro, J. F., Frenk, C. S., White, S. D. M., Springel, V., and Jenkins, A. (2007). The statistics of Λ CDM halo concentrations. *MNRAS*, 381:1450–1462.

Noguchi, M. (1999). Early Evolution of Disk Galaxies: Formation of Bulges in Clumpy Young Galactic Disks. *ApJ*, 514:77–95.

Obreschkow, D., Croton, D., De Lucia, G., Khochfar, S., and Rawlings, S. (2009a). Simulation of the Cosmic Evolution of Atomic and Molecular Hydrogen in Galaxies. *ApJ*, 698:1467–1484.

Obreschkow, D. and Glazebrook, K. (2014). Fundamental Mass-Spin-Morphology Relation Of Spiral Galaxies. *ApJ*, 784:26.

Obreschkow, D., Klöckner, H.-R., Heywood, I., Levrier, F., and Rawlings, S. (2009b). A Virtual Sky with Extragalactic H I and CO Lines for the Square Kilometre Array and the Atacama Large Millimeter/Submillimeter Array. *ApJ*, 703:1890–1903.

Obreschkow, D. and Rawlings, S. (2009). The Cosmic Decline in the H_2/H I Ratio in Galaxies. *ApJ*, 696:L129–L132.

Okamoto, T., Gao, L., and Theuns, T. (2008). Mass loss of galaxies due to an ultraviolet background. *MNRAS*, 390:920–928.

Omont, A. (2007). Molecules in galaxies. *Reports on Progress in Physics*, 70:1099–1176.

Oser, L., Naab, T., Ostriker, J. P., and Johansson, P. H. (2012). The Cosmological Size and Velocity Dispersion Evolution of Massive Early-type Galaxies. *ApJ*, 744:63.

Padilla, N. D., Salazar-Albornoz, S., Contreras, S., Cora, S. A., and Ruiz, A. N. (2014). Stochastic angular momentum slews and flips and their effect on discs in galaxy formation models. *MNRAS*, 443:2801–2814.

Padovani, P. and Matteucci, F. (1993). Stellar Mass Loss in Elliptical Galaxies and the Fueling of Active Galactic Nuclei. *ApJ*, 416:26.

Papastergis, E., Giovanelli, R., Haynes, M. P., Rodríguez-Puebla, A., and Jones, M. G. (2013). The Clustering of ALFALFA Galaxies: Dependence on H I Mass, Relationship with Optical Samples, and Clues of Host Halo Properties. *ApJ*, 776:43.

Parkinson, H., Cole, S., and Helly, J. (2008). Generating dark matter halo merger trees. *MNRAS*, 383:557–564.

Pedlar, A., Ghataure, H. S., Davies, R. D., Harrison, B. A., Perley, R., Crane, P. C., and Unger, S. W. (1990). The Radio Structure of NGC1275. *MNRAS*, 246:477.

Pedrosa, S. E. and Tissera, P. B. (2015). Angular momentum evolution for galaxies in a Λ-CDM scenario. *A&A*, 584:A43.

Peeples, M. S. and Shankar, F. (2011). Constraints on star formation driven galaxy winds from the mass-metallicity relation at z= 0. *MNRAS*, 417:2962–2981.

Peletier, R. F. and Balcells, M. (1996). Ages of Galaxies Bulges and Disks From Optical and Near-Infrared Colors. *AJ*, 111:2238.

Pelupessy, F. I., Papadopoulos, P. P., and van der Werf, P. (2006). Incorporating the Molecular Gas Phase in Galaxy-sized Numerical Simulations: First Applications in Dwarf Galaxies. *ApJ*, 645:1024–1042.

Peng, Y., Maiolino, R., and Cochrane, R. (2015). Strangulation as the primary mechanism for shutting down star formation in galaxies. *Nature*, 521:192–195.

Penzias, A. A. and Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. *ApJ*, 142:419–421.

Pichon, C., Pogosyan, D., Kimm, T., Slyz, A., Devriendt, J., and Dubois, Y. (2011). Rigging dark haloes: why is hierarchical galaxy formation consistent with the inside-out build-up of thin discs? *MNRAS*, 418:2493–2507.

Pizzolato, F. and Soker, N. (2010). Solving the angular momentum problem in the cold feedback mechanism of cooling flows. *MNRAS*, 408:961–974.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., and et al. (2014). Planck 2013 results. XVI. Cosmological parameters. *A&A*, 571:A16.

Popping, A., Davé, R., Braun, R., and Oppenheimer, B. D. (2009). The simulated H I sky at low redshift. *A&A*, 504:15–32.

Popping, G., Somerville, R. S., and Trager, S. C. (2014). Evolution of the atomic and molecular gas content of galaxies. *MNRAS*, 442:2398–2418.

Porter, L. A., Somerville, R. S., Primack, J. R., and Johansson, P. H. (2014). Understanding the structural scaling relations of early-type galaxies. *MNRAS*, 444:942–960.

Pounds, K. A., Reeves, J. N., King, A. R., Page, K. L., O'Brien, P. T., and Turner, M. J. L. (2003). A high-velocity ionized outflow and XUV photosphere in the narrow emission line quasar PG1211+143. *MNRAS*, 345:705–713.

Power, C., Baugh, C. M., and Lacey, C. G. (2010). The redshift evolution of the mass function of cold gas in hierarchical galaxy formation models. *MNRAS*, 406:43–59.

Press, W. H. and Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438.

Price, D. J. (2008). Modelling discontinuities and Kelvin Helmholtz instabilities in SPH. *Journal of Computational Physics*, 227:10040–10057.

Rafieferantsoa, M., Davé, R., Anglés-Alcázar, D., Katz, N., Kollmeier, J. A., and Oppenheimer, B. D. (2015). The impact of environment and mergers on the H I content of galaxies in hydrodynamic simulations. *MNRAS*, 453:3980–3998.

Rahmati, A., Schaye, J., Bower, R. G., Crain, R. A., Furlong, M., Schaller, M., and Theuns, T. (2015). The distribution of neutral hydrogen around high-redshift galaxies and quasars in the EAGLE simulation. *MNRAS*, 452:2034–2056.

Rao, S. M., Turnshek, D. A., and Briggs, F. H. (1995). The incidence of damped ly alpha systems in the redshift interval 0¡ z¡ 4. *The Astrophysical Journal*, 449:488.

Read, J. I., Hayfield, T., and Agertz, O. (2010). Resolving mixing in smoothed particle hydrodynamics. *MNRAS*, 405:1513–1530.

Rees, M. J. and Ostriker, J. P. (1977). Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters. *MNRAS*, 179:541–559.

Reeves, J. N., O'Brien, P. T., Braito, V., Behar, E., Miller, L., Turner, T. J., Fabian, A. C., Kaspi, S., Mushotzky, R., and Ward, M. (2009). A Compton-thick Wind in the High-luminosity Quasar, PDS 456. *ApJ*, 701:493–507.

Renzini, A. (1999). Origin of Bulges. In Carollo, C. M., Ferguson, H. C., and Wyse, R. F. G., editors, *The Formation of Galactic Bulges*, page 9.

Reynolds, C. S., Ward, M. J., Fabian, A. C., and Celotti, A. (1997). A multi-wavelength study of the Seyfert 1 galaxy MCG-6-30-15. *MNRAS*, 291:403.

Roberts, M. S. (1975). *Radio Observations of Neutral Hydrogen in Galaxies*, page 309. the University of Chicago Press.

Robertson, B. E. and Kravtsov, A. V. (2008). Molecular Hydrogen and Global Star Formation Relations in Galaxies. *ApJ*, 680:1083–1111.

Romanowsky, A. J. and Fall, S. M. (2012). Angular Momentum and Galaxy Formation Revisited. *ApJS*, 203:17.

Roukema, B. F., Quinn, P. J., Peterson, B. A., and Rocca-Volmerange, B. (1997). Merging history trees of dark matter haloes - A tool for exploring galaxy formation models. *MNRAS*, 292:835.

Saintonge, A., Kauffmann, G., Kramer, C., Tacconi, L. J., Buchbender, C., Catinella, B., Fabello, S., Graciá-Carpio, J., Wang, J., Cortese, L., Fu, J., Genzel, R., Giovanelli, R., Guo, Q., Haynes, M. P., Heckman, T. M., Krumholz, M. R., Lemonias, J., Li, C., Moran, S., Rodriguez-Fernandez, N., Schiminovich, D., Schuster, K., and Sievers, A. (2011). COLD GASS, an IRAM legacy survey of molecular gas in massive galaxies - I. Relations between H_2, H I, stellar content and structural properties. *MNRAS*, 415:32–60.

Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., van de Ven, G., Vílchez, J. M., Wisotzki, L., Walcher, C. J., Mast, D., Aguerri, J. A. L., Albiol-Pérez, S., Alonso-Herrero, A., Alves, J., Bakos, J., Bartáková, T., Bland-Hawthorn, J., Boselli, A., Bomans, D. J., Castillo-Morales, A., Cortijo-Ferrero, C., de Lorenzo-Cáceres, A., Del Olmo, A., Dettmar, R.-J., Díaz, A., Ellis, S., Falcón-Barroso, J., Flores, H., Gallazzi, A., García-Lorenzo, B., González Delgado, R., Gruel, N., Haines, T., Hao, C., Husemann, B., Iglésias-Páramo, J., Jahnke, K., Johnson, B., Jungwiert, B., Kalinova, V., Kehrig, C., Kupko, D., López-Sánchez, Á. R., Lyubenova, M., Marino, R. A., Mármol-Queraltó, E., Márquez, I., Masegosa, J., Meidt, S., Mendez-Abreu, J., Monreal-Ibero, A., Montijo, C., Mourão, A. M., Palacios-Navarro, G., Papaderos, P., Pasquali, A., Peletier, R., Pérez, E., Pérez, I., Quirrenbach, A., Relaño, M., Rosales-Ortega, F. F., Roth, M. M., Ruiz-Lara, T., Sánchez-Blázquez, P., Sengupta, C., Singh, R., Stanishev, V., Trager, S. C., Vazdekis, A., Viironen, K., Wild, V., Zibetti, S., and Ziegler, B. (2012). CALIFA, the Calar Alto Legacy Integral Field Area survey. I. Survey presentation. *A&A*, 538:A8.

Sancisi, R., Fraternali, F., Oosterloo, T., and van der Hulst, T. (2008). Cold gas accretion in galaxies. *A&ARv*, 15:189–223.

Sanders, R. H. (1980). Neutral hydrogen in elliptical galaxies - A bimodal distribution. *ApJ*, 242:931–937.

Saro, A., De Lucia, G., Borgani, S., and Dolag, K. (2010). Gas cooling in semi-analytic models and smoothed particle hydrodynamics simulations: are results consistent? *MNRAS*, 406:729–743.

Sawala, T., Frenk, C. S., Fattahi, A., Navarro, J. F., Bower, R. G., Crain, R. A., Dalla Vecchia, C., Furlong, M., Jenkins, A., McCarthy, I. G., Qu, Y., Schaller, M., Schaye, J., and Theuns, T. (2015). Bent by baryons: the low-mass galaxy-halo relation. *MNRAS*, 448:2941–2947.

Scannapieco, C., Tissera, P. B., White, S. D. M., and Springel, V. (2005). Feedback and metal enrichment in cosmological smoothed particle hydrodynamics simulations - I. A model for chemical enrichment. *MNRAS*, 364:552–564.

Scannapieco, C., Tissera, P. B., White, S. D. M., and Springel, V. (2006). Feedback and metal enrichment in cosmological SPH simulations - II. A multiphase model with supernova energy feedback. *MNRAS*, 371:1125–1139.

Scannapieco, C., Tissera, P. B., White, S. D. M., and Springel, V. (2008). Effects of supernova feedback on the formation of galaxy discs. *MNRAS*, 389:1137–1149.

Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D. M., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., and Trayford, J. (2015). The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *MNRAS*, 446:521–554.

Schaye, J., Dalla Vecchia, C., Booth, C. M., Wiersma, R. P. C., Theuns, T., Haas, M. R., Bertone, S., Duffy, A. R., McCarthy, I. G., and van de Voort, F. (2010). The physics driving the cosmic star formation history. *MNRAS*, 402:1536–1560.

Schilizzi, R. T., Dewdney, P. E. F., and Lazio, T. J. W. (2008). The Square Kilometre Array. In *Ground-based and Airborne Telescopes II*, volume 7012 of *Proc. SPIE*, page 70121I.

Schmidt, M. (1959). The Rate of Star Formation. *ApJ*, 129:243.

Serra, P., Oosterloo, T., Morganti, R., Alatalo, K., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Cappellari, M., Crocker, A. F., Davies, R. L., Davis, T. A., de Zeeuw, P. T., Duc, P.-A., Emsellem, E., Khochfar, S., Krajnović, D., Kuntschner, H., Lablanche, P.-Y., McDermid, R. M., Naab, T., Sarzi, M., Scott, N., Trager, S. C., Weijmans, A.-M., and Young, L. M. (2012). The ATLAS^{3D} project - XIII. Mass and morphology of H I in early-type galaxies as a function of environment. *MNRAS*, 422:1835–1862.

Serra, P., Oser, L., Krajnović, D., Naab, T., Oosterloo, T., Morganti, R., Cappellari, M., Emsellem, E., Young, L. M., Blitz, L., Davis, T. A., Duc, P.-A., Hirschmann, M., Weijmans, A.-M., Alatalo, K., Bayet, E., Bois, M., Bournaud, F., Bureau, M., Crocker, A. F., Davies, R. L., de Zeeuw, P. T., Khochfar, S., Kuntschner, H., Lablanche, P.-Y., McDermid, R. M., Sarzi, M., and Scott, N. (2014). The ATLAS^{3D} project - XXVI. H I discs in real and simulated fast and slow rotators. *MNRAS*, 444:3388–3407.

Shankar, F., Lapi, A., Salucci, P., De Zotti, G., and Danese, L. (2006). New Relationships between Galaxy Properties and Host Halo Mass, and the Role of Feedbacks in Galaxy Formation. *ApJ*, 643:14–25.

Shankar, F., Marulli, F., Bernardi, M., Mei, S., Meert, A., and Vikram, V. (2013). Size evolution of spheroids in a hierarchical Universe. *MNRAS*, 428:109–128.

Sharma, S. and Steinmetz, M. (2005). The Angular Momentum Distribution of Gas and Dark Matter in Galactic Halos. *ApJ*, 628:21–44.

Shen, S., Mo, H. J., White, S. D. M., Blanton, M. R., Kauffmann, G., Voges, W., Brinkmann, J., and Csabai, I. (2003). The size distribution of galaxies in the Sloan Digital Sky Survey. *MNRAS*, 343:978–994.

Sheth, R. K., Mo, H. J., and Tormen, G. (2001). Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes. *MNRAS*, 323:1–12.

Sheth, R. K. and Tormen, G. (1999). Large-scale bias and the peak background split. *MNRAS*, 308:119–126.

Silk, J. (1977). On the fragmentation of cosmic gas clouds. I - The formation of galaxies and the first generation of stars. *ApJ*, 211:638–648.

Solomon, P. M. and Vanden Bout, P. A. (2005). Molecular Gas at High Redshift. *ARA&A*, 43:677–725.

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., and Hernquist, L. (2008). A semi-analytic model for the co-evolution of galaxies, black holes and active galactic nuclei. *MNRAS*, 391:481–506.

Somerville, R. S., Popping, G., and Trager, S. C. (2015). Star formation in semi-analytic galaxy formation models with multiphase gas. *MNRAS*, 453:4337–4367.

Somerville, R. S., Primack, J. R., and Faber, S. M. (2001). The nature of high-redshift galaxies. *MNRAS*, 320:504–528.

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2003). First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *ApJS*, 148:175–194.

Springel, V. (2010). E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *MNRAS*, 401:791–851.

Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., and White, S. D. M. (2008). The Aquarius Project: the subhaloes of galactic haloes. *MNRAS*, 391:1685–1711.

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., and Pearce, F. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636.

Springel, V., White, S. D. M., Tormen, G., and Kauffmann, G. (2001). Populating a cluster of galaxies - I. Results at [formmu2]z=0. *MNRAS*, 328:726–750.

Steinmetz, M. and Navarro, J. F. (1999). The Cosmological Origin of the Tully-Fisher Relation. *ApJ*, 513:555–560.

Stevens, A. R. H., Croton, D. J., and Mutch, S. J. (2016). Building disc structure and galaxy properties through angular momentum: the DARK SAGE semi-analytic model. *MNRAS*, 461:859–876.

Stewart, K. R., Kaufmann, T., Bullock, J. S., Barton, E. J., Maller, A. H., Diemand, J., and Wadsley, J. (2011). Observing the End of Cold Flow Accretion Using Halo Absorption Systems. *ApJ*, 735:L1.

Strickland, D. K. and Heckman, T. M. (2009). Supernova Feedback Efficiency and Mass Loading in the Starburst and Galactic Superwind Exemplar M82. *ApJ*, 697:2030–2056.

Sutherland, R. S. and Dopita, M. A. (1993). Cooling functions for low-density astrophysical plasmas. *ApJS*, 88:253–327.

Tabor, G. and Binney, J. (1993). Elliptical Galaxy Cooling Flows Without Mass Drop-Out. *MNRAS*, 263:323.

Tassev, S., Zaldarriaga, M., and Eisenstein, D. J. (2013). Solving large scale structure in ten easy steps with COLA. *J. Cosmology Astropart. Phys.*, 6:036.

Teklu, A. F., Remus, R.-S., Dolag, K., Beck, A. M., Burkert, A., Schmidt, A. S., Schulze, F., and Steinborn, L. K. (2015). Connecting Angular Momentum and Galactic Dynamics: The Complex Interplay between Spin, Mass, and Morphology. *ApJ*, 812:29.

Thielemann, F.-K., Argast, D., Brachwitz, F., Hix, W., Höflich, P., Liebendörfer, M., Martinez-Pinedo, G., Mezzacappa, A., Panov, I., and Rauscher, T. (2003). Nuclear cross sections, nuclear structure and stellar nucleosynthesis. *Nuclear Physics A*, 718:139 – 146.

Thronson, Jr., H. A., Tacconi, L., Kenney, J., Greenhouse, M. A., Margulis, M., Tacconi-Garman, L., and Young, J. S. (1989). Molecular gas, the interstellar medium, and star formation in S0 and SA galaxies. *ApJ*, 344:747–762.

Tombesi, F., Cappi, M., Reeves, J. N., Palumbo, G. G. C., Yaqoob, T., Braito, V., and Dadina, M. (2010). Evidence for ultra-fast outflows in radio-quiet AGNs. I. Detection and statistical incidence of Fe K-shell absorption lines. *A&A*, 521:A57.

Tonini, C., Mutch, S. J., Croton, D. J., and Wyithe, J. S. B. (2016). The growth of discs and bulges during hierarchical galaxy formation - I. Fast evolution versus secular processes. *MNRAS*, 459:4109–4129.

Tonnesen, S. and Bryan, G. L. (2009). Gas Stripping in Simulated Galaxies with a Multiphase Interstellar Medium. *ApJ*, 694:789–804.

Tucker, W. and David, L. P. (1997). A Feedback Model for Radio Sources Fueled by Cooling Flows. *ApJ*, 484:602–607.

Übler, H., Naab, T., Oser, L., Aumer, M., Sales, L. V., and White, S. D. M. (2014). Why stellar feedback promotes disc formation in simulated galaxies. *MNRAS*, 443:2092–2111.

Vale, A. and Ostriker, J. P. (2004). Linking halo mass to galaxy luminosity. *MNRAS*, 353:189–200.

van den Bosch, F. C., Abel, T., Croft, R. A. C., Hernquist, L., and White, S. D. M. (2002). The Angular Momentum of Gas in Protogalaxies. I. Implications for the Formation of Disk Galaxies. *ApJ*, 576:21–35.

van den Bosch, F. C., Abel, T., and Hernquist, L. (2003). The angular momentum of gas in protogalaxies - II. The impact of pre-heating. *MNRAS*, 346:177–185.

van der Wel, A., Franx, M., van Dokkum, P. G., Skelton, R. E., Momcheva, I. G., Whitaker, K. E., Brammer, G. B., Bell, E. F., Rix, H.-W., Wuyts, S., Ferguson, H. C., Holden, B. P., Barro, G., Koekemoer, A. M., Chang, Y.-Y., McGrath, E. J., Häussler, B., Dekel, A., Behroozi, P., Fumagalli, M., Leja, J., Lundgren, B. F., Maseda, M. V., Nelson, E. J., Wake, D. A., Patel, S. G., Labbé, I., Faber, S. M., Grogin, N. A., and Kocevski, D. D. (2014). 3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since z = 3. *ApJ*, 788:28.

Veilleux, S., Cecil, G., and Bland-Hawthorn, J. (2005). Galactic Winds. *ARA&A*, 43:769–826.

Verheijen, M., van Gorkom, J. H., Szomoru, A., Dwarakanath, K. S., Poggianti, B. M., and Schiminovich, D. (2007). WSRT Ultradeep Neutral Hydrogen Imaging of Galaxy Clusters at z ˜ 0.2: A Pilot Survey of Abell 963 and Abell 2192. *ApJ*, 668:L9–L13.

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Bird, S., Nelson, D., and Hernquist, L. (2014a). Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, 509:177–182.

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., and Hernquist, L. (2014b). Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *MNRAS*, 444:1518–1547.

Vollmer, B., Cayatte, V., Balkowski, C., and Duschl, W. J. (2001). Ram Pressure Stripping and Galaxy Orbits: The Case of the Virgo Cluster. *ApJ*, 561:708–726.

Walch, S., Girichidis, P., Naab, T., Gatto, A., Glover, S. C. O., Wünsch, R., Klessen, R. S., Clark, P. C., Peters, T., Derigs, D., and Baczynski, C. (2015). The SILCC (SImulating the LifeCycle of molecular Clouds) project - I. Chemical evolution of the supernova-driven ISM. *MNRAS*, 454:238–268.

Walter, F., Brinks, E., de Blok, W. J. G., Bigiel, F., Kennicutt, Jr., R. C., Thornley, M. D., and Leroy, A. (2008). THINGS: The H I Nearby Galaxy Survey. *AJ*, 136:2563–2647.

Walter, F., Decarli, R., Aravena, M., Carilli, C., Bouwens, R., da Cunha, E., Daddi, E., Ivison, R. J., Riechers, D., Smail, I., Swinbank, M., Weiss, A., Anguita, T., Assef, R., Bacon, R., Bauer, F., Bell, E. F., Bertoldi, F., Chapman, S., Colina, L., Cortes, P. C., Cox, P., Dickinson, M., Elbaz, D., Gónzalez-López, J., Ibar, E., Inami, H., Infante, L., Hodge, J., Karim, A., Le Fevre, O., Magnelli, B., Neri, R., Oesch, P., Ota, K., Popping, G., Rix, H.-W., Sargent, M., Sheth, K., van der Wel, A., van der Werf, P., and Wagg, J. (2016). ALMA Spectroscopic Survey in the Hubble Ultra Deep Field: Survey Description. *ApJ*, 833:67.

Wang, J., De Lucia, G., Kitzbichler, M. G., and White, S. D. M. (2008). The dependence of galaxy formation on cosmological parameters: can we distinguish between the WMAP1 and WMAP3 parameter sets? *MNRAS*, 384:1301–1315.

Wang, L., Li, C., Kauffmann, G., and De Lucia, G. (2007). Modelling and interpreting the dependence of clustering on the spectral energy distributions of galaxies. *MNRAS*, 377:1419–1430.

Warren, M. S., Abazajian, K., Holz, D. E., and Teodoro, L. (2006). Precision Determination of the Mass Function of Dark Matter Halos. *ApJ*, 646:881–885.

Weidner, C. and Kroupa, P. (2005). The Variation of Integrated Star Initial Mass Functions among Galaxies. *ApJ*, 625:754–762.

Weil, M. L., Eke, V. R., and Efstathiou, G. (1998). The formation of disc galaxies. *MNRAS*, 300:773–789.

Weil, M. L. and Hernquist, L. (1994). Kinematic misalignments in remnants of multiple mergers. *ApJ*, 431:L79–L82.

Weiner, B. J., Coil, A. L., Prochaska, J. X., Newman, J. A., Cooper, M. C., Bundy, K., Conselice, C. J., Dutton, A. A., Faber, S. M., Koo, D. C., Lotz, J. M., Rieke, G. H., and Rubin, K. H. R. (2009). Ubiquitous Outflows in DEEP2 Spectra of Star-Forming Galaxies at z = 1.4. *ApJ*, 692:187–211.

Weinmann, S. M., Kauffmann, G., van den Bosch, F. C., Pasquali, A., McIntosh, D. H., Mo, H., Yang, X., and Guo, Y. (2009). Environmental effects on satellite galaxies: the link between concentration, size and colour profile. *MNRAS*, 394:1213–1228.

Weinmann, S. M., van den Bosch, F. C., Yang, X., Mo, H. J., Croton, D. J., and Moore, B. (2006). Properties of galaxy groups in the Sloan Digital Sky Survey - II. Active galactic nucleus feedback and star formation truncation. *MNRAS*, 372:1161–1174.

West, A. A., Garcia-Appadoo, D. A., Dalcanton, J. J., Disney, M. J., Rockosi, C. M., and Ivezić, Ž. (2009). H I-Selected Galaxies in the Sloan Digital Sky Survey. II. The Colors of Gas-Rich Galaxies. *AJ*, 138:796–807.

West, A. A., Garcia-Appadoo, D. A., Dalcanton, J. J., Disney, M. J., Rockosi, C. M., Ivezić, Ž., Bentz, M. C., and Brinkmann, J. (2010). H I-Selected Galaxies in the Sloan Digital Sky Survey. I. Optical Data. *AJ*, 139:315–328.

Westmeier, T., Staveley-Smith, L., Calabretta, M., Jurek, R., Koribalski, B. S., Meyer, M., Popping, A., and Wong, O. I. (2015). On the neutral gas content of nine new Milky Way satellite galaxy candidates. *MNRAS*, 453:338–344.

Wetzel, A. R., Cohn, J. D., and White, M. (2009). The clustering and host haloes of galaxy mergers at high redshift. *MNRAS*, 394:2182–2190.

Wetzel, A. R., Tinker, J. L., Conroy, C., and van den Bosch, F. C. (2013). Galaxy evolution in groups and clusters: satellite star formation histories and quenching time-scales in a hierarchical Universe. *MNRAS*, 432:336–358.

White, C. E., Somerville, R. S., and Ferguson, H. C. (2015). A Parametric Study of Possible Solutions to the High-redshift Overproduction of Stars in Modeled Dwarf Galaxies. *ApJ*, 799:201.

White, S. D. M. and Frenk, C. S. (1991a). Galaxy formation through hierarchical clustering. *ApJ*, 379:52–79.

White, S. D. M. and Frenk, C. S. (1991b). Galaxy formation through hierarchical clustering. *ApJ*, 379:52–79.

White, S. D. M. and Rees, M. J. (1978). Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. *MNRAS*, 183:341–358.

Wiklind, T. and Henkel, C. (1989). The molecular cloud content of early type galaxies. I - Detections and global properties. *A&A*, 225:1–11.

Wilman, D. J., Fontanot, F., De Lucia, G., Erwin, P., and Monaco, P. (2013). The hierarchical origins of observed galaxy morphology. *MNRAS*, 433:2986–3004.

Wong, T. and Blitz, L. (2002). The Relationship between Gas Content and Star Formation in Molecule-rich Spiral Galaxies. *ApJ*, 569:157–183.

Xie, L., De Lucia, G., Hirschmann, M., Fontanot, F., and Zoldan, A. (2017). H2-based star formation laws in hierarchical models of galaxy formation. *MNRAS*, 469:968–993.

Xue, Y. Q., Luo, B., Brandt, W. N., Bauer, F. E., Lehmer, B. D., Broos, P. S., Schneider, D. P., Alexander, D. M., Brusa, M., Comastri, A., Fabian, A. C., Gilli, R., Hasinger, G., Hornschemeier, A. E., Koekemoer, A., Liu, T., Mainieri, V., Paolillo, M., Rafferty, D. A., Rosati, P., Shemmer, O., Silverman, J. D., Smail, I., Tozzi, P., and Vignali, C. (2011). The Chandra Deep Field-South Survey: 4 Ms Source Catalogs. *ApJS*, 195:10.

Yates, R. M., Henriques, B., Thomas, P. A., Kauffmann, G., Johansson, J., and White, S. D. M. (2013). Modelling element abundances in semi-analytic models of galaxy formation. *MNRAS*, 435:3500–3520.

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi,

M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C.-h., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Neilsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sandford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S.-i., Watanabe, M., Weinberg, D. H., Yanny, B., Yasuda, N., and SDSS Collaboration (2000). The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120:1579–1587.

Yoshida, N., Stoehr, F., Springel, V., and White, S. D. M. (2002). Gas cooling in simulations of the formation of the galaxy population. *MNRAS*, 335:762–772.

Young, J. S., Xie, S., Tacconi, L., Knezek, P., Viscuso, P., Tacconi-Garman, L., Scoville, N., Schneider, S., Schloerb, F. P., Lord, S., Lesser, A., Kenney, J., Huang, Y.-L., Devereux, N., Claussen, M., Case, J., Carpenter, J., Berry, M., and Allen, L. (1995). The FCRAO Extragalactic CO Survey. I. The Data. *ApJS*, 98:219.

Zavala, J., Frenk, C. S., Bower, R., Schaye, J., Theuns, T., Crain, R. A., Trayford, J. W., Schaller, M., and Furlong, M. (2016). The link between the assembly of the inner dark matter halo and the angular momentum evolution of galaxies in the EAGLE simulation. *MNRAS*, 460:4466–4482.

Zavala, J., Okamoto, T., and Frenk, C. S. (2008). Bulges versus discs: the evolution of angular momentum in cosmological simulations of galaxy formation. *MNRAS*, 387:364–370.

Zeldovich, Y. B. (1972). A hypothesis, unifying the structure and the entropy of the Universe. *MNRAS*, 160:1P.

Zentner, A. R., Hearin, A. P., and van den Bosch, F. C. (2014). Galaxy assembly bias: a significant source of systematic error in the galaxy-halo relationship. *MNRAS*, 443:3044–3067.

Zhang, J., Fakhouri, O., and Ma, C.-P. (2008). How to grow a healthy merger tree. *MNRAS*, 389:1521–1538.

Zheng, Z., Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., Cole, S., Davé, R., Frenk, C. S., Katz, N., and Lacey, C. G. (2005). Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies. *ApJ*, 633:791–809.

Zoldan, A., De Lucia, G., Xie, L., Fontanot, F., and Hirschmann, M. (2017). H I-selected galaxies in hierarchical models of galaxy formation and evolution. *MNRAS*, 465:2236–2253.

Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., and Webster, R. L. (2005). The HIPASS catalogue: Ω_{HI} and environmental effects on the HI mass function of galaxies. *MNRAS*, 359:L30–L34.