A study of galaxy formation across cosmic time from cosmological hydrodynamical simulations

MICHELLE FURLONG

A study of galaxy formation across cosmic time from cosmological hydrodynamical simulations

Michelle Furlong

Abstract

The evolution of galaxies across cosmic time, from the first galaxies to the local Universe, are studied using cosmological hydrodynamical simulations. It is demonstrated that, for the first time, a hydrodynamical simulation can reproduce the observed evolution of galaxy stellar masses and the trends in star formation rates. The success of the simulation in producing galaxies with similar histories to those observed increases the potential for using hydrodynamical simulations to explore galaxy formation physics. With this intention, we consider the effects of the environment and active galactic nuclei (AGN) quenching on the galaxy population and the shape of the galaxy stellar mass function (GSMF). We find environmental processes are effective at quenching galaxies in the simulation and operated on short timescales. AGN feedback, which produces a large passive fraction at high stellar masses, drives the exponential break in the GSMF.

Specific star formation rates (SFRs) in simulations, both hydrodynamical and semi-analytical, have been shown to be discrepant with observations. We investigate proposed solutions to this problem using a suite of cosmological hydrodynamical simulations. The offset in the simulations, at the level of 0.2 to 0.4 dex, can only be resolved by employing an extreme model that does not recover any of the observed trends in stellar mass or the cosmic star formation rate density. This study implies that the observed star formation rates across comic time are inconsistent with the growth of stellar mass.

Two galaxy populations are then explored in more detail. We examine the first galaxies and their potential to reionize the Universe. We find that low-mass ($< 10^8 \mathrm{M}_\odot$) galaxies, undergoing extreme star formation for their stellar mass, pro-

duce the majority of ionizing photons at redshifts 6 and above. The second study considers the most highly star forming galaxies in the simulation, which represent an extreme population. These galaxies have similar stellar and gas masses to the observed sub-mm population, however, their SFRs are lower. On further investigation, the selection of galaxies based on SFRs is not adequate to compare the simulation to the sub-mm population, in particular at high redshift. These results highlight the importance of dust temperature in the selection.

A study of galaxy formation across cosmic time with cosmological hydrodynamical simulations

by Michelle Furlong

A thesis submitted to the University of Durham
in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy.

Department of Physics
University of Durham
September 2014

Contents

List of Figures

List of Tables

List of Acronyms

AGB asymptotic giant branch.

AGN active galactic nuclei.

AMR adaptive mesh refinement.

CDM cold dark matter.

CMB cosmic microwave background.

DMO dark matter only.

FoF friends-of-friends.

GSMF galaxy stellar mass function.

HI neutral hydrogen.

HOD Halo Occupation Distribution.

ICL intra-cluster light.

IMF initial mass function.

ISM interstellar medium.

PM particle-mesh.

SAM Semi-analytical model.

SED spectral energy distribution.

SFR star formation rate.

SPH smooth particle hydrodynamics.

SPS stellar population synthesis.

SSFR specific star formation rate.

Declaration

The work described in this thesis was undertaken between 2010 and 2014 while the author was a research student under the supervision of Prof. Richard G. Bower and Dr. Tom Theuns in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or any other University.

Chapter 3 has been submitted for publication in the form of a paper to Monthly Notices of the Royal Astronomical Society,

- *Evolution of galaxy stellar masses and star formation rates in the* EAGLE SIMULATIONS.

  M. Furlong, R. G. Bower, T. Theuns, J. Schaye, R. A. Crain, M. Schaller, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. Helly, A. Jenkins, and Y. M. Rosas-Guevara, MNRAS submitted, arXiv:1410.3485

Chapters 4, 6 and 7 will be submitted soon in the form of a paper to Monthly Notices of the Royal Astronomical Society.

All figures in this thesis have been produced by the author.

Acknowledgements

I am very grateful to Richard for the endless enthusiasm, support and time he has given me throughout this process. I am also hugely grateful to Tom for sharing his extensive knowledge of computing, Gadget in particular, and interesting ideas. It is also thanks to Tom that my writing skills have drastically improved! Of course I would not have made it here without the whole Eagle team, I am delighted to have worked with Yetli, Joop, Craig, Ian, Claudio, Yen and James during my PhD. Rob and Matthieu deserve a special mention for their added moral support on top of the project support. I am very privileged to have carried out my PhD at Durham, due to the stimulating academic environment it harbours, so a general thanks to all staff and students in astronomy for this. There is one staff member though, without whom I would have been lost (both computationally and generally) on many many occasions, so to Lydia I am eternally grateful!

Is it possible to have too many great people around you? There is a very long list of friends I have accumulated at Durham who have lit up the PhD years. Friends gone but not forgotten are Claudia, Alex I, Nikos II, Mark and Anais, I miss you all. The ones that haven't left me yet, Violeta, Mathilde, Flora, Tamsyn, Alex II, Matthieu, Peter, Will, Charles, James, Yetli, Rob, Alice, Jascha, Lilian, Danny, Jack who make every day so much brighter. Many fantastic people I have met in Durham, from Ustinov and dancing in particular, have made Durham such an enjoyable place to live. Lizzie, Anna, Emma and G deserve a special mention for putting up with me, even on the bad days!

Without the support from across the sea I would not have settled so well in Durham. My friends at home, Laura, Joey, Matt, Bryan, Mairead, etc., thank you all for always being there. That brings me to Paddy. I'm not sure what to say here, thank you doesn't quite seem like enough, no words could express my appreciation for the support you have given me over the last couple of years! Hopefully some day I can pay you back. And finally, my family, my backbone, Mam, Dad, Marita, Jaymie, Nana and D.J., I am who I am and have achieved all I have only because of your belief in me. I love you all.

Chapter 1 | *Introduction*

While we now have a general understanding of the theory of galaxy formation from combining the Λ-cold dark matter (CDM) paradigm and baryonic processes, there remain many unanswered questions and unexplained observations in the field of galaxy formation. In this work we use state-of-the-art cosmological hydrodynamical simulations to test our knowledge and understanding of galaxy formation. The simulations are then used to expand our understanding, by connecting the galaxy population across cosmic time and by considering the behaviour specific galaxy populations. We begin by outlining the processes important to galaxy formation and the role of simulations in this field.

1.1 An Overview of the Theory of Galaxy Formation

To study theoretical galaxy formation we first require a framework to capture the underlying matter density distribution of the Universe and how it evolves. The modern framework is based on the cosmological principle, which assumes that the Universe is homogeneous and isotropic on large scales, and the equations of general relativity of Einstein. In the ΛCDM paradigm, a cosmological constant, Λ, is employed to describe the late-time acceleration of the Universe and a cold dark matter component accounts for the 'missing matter' in the Universe. Observational probes of the Cosmic Microwave Background then provide constraints on the energy and matter density in the Universe. While the nature of dark energy and dark matter remain a mystery, constraining the energy density fractions of dark energy and matter, the index of the power spectrum, the primordial fluctuations of the matter density and the Hubble parameter, the growth of structure can be described. The Wilkinson Microwave Anisotropy Probe (Hinshaw et al., 2013) and

Planck (Planck Collaboration et al., 2013) missions have provided good constraints for these parameters to date, and so far have found no strong evidence against the assumption that the Universe can be described by the ΛCDM paradigm. In a ΛCDM cosmology small density fluctuations in the early Universe increase with time, become self-gravitating and collapse to form halos. The hierarchical growth of this model is now well established as a valid description of the Universe, based on the available constraints from the power spectrum of galaxies (Tegmark et al., 2004; Cole et al., 2005; Reid et al., 2010), galaxy clusters (Vikhlinin et al., 2009) and type Ia supernovae (Reiss et al., 1998; Kowalski et al., 2008; Perlmutter et al., 1998).

The most recent measurements of Planck suggest dark energy accounts for 69.3% of the energy density of the Universe, with dark matter, the second largest component, accounting for 25.9%. Baryons contribute the remaining 4.8%. While baryons make up only 4.8% of the total matter in the Universe, this component is the only one that is visible. As well as being visible, they constitute all the 'ordinary' matter in the Universe, including the stars, the planets and us. Although this component is familiar to us, it is governed by more complex physics than the pure gravitational processes that impact the dark matter, and so the evolution of gas is not fully described by the equations of general relativity. As baryons collapse into halos they are shock heated. The ability of baryons to radiate heat enables further collapse within the halo. The denser environment in turn increases the radiation rate. Gas at high temperatures, T $> 10^7$ K, is fully collisionally ionised and predominantly cools through Bremsstrahlung. For intermediate temperature gas, $10^4 <$ T $< 10^7$ K, collisionally ionised atoms can decay to their ground state, emitting photons. Lower temperature gas, T $< 10^4$ K, cools through metal-line and molecular cooling. As the gas collapses, molecular clouds are formed, becoming the birth places of stars.

The process continues after star formation, when stellar mass loss from asymptotic giant branch (AGB) stars, type Ia supernovae and type II supernovae enrich the interstellar medium (ISM) with metals formed during nucleosynthesis. The rising metal abundance increases the cooling rate of the gas in the ISM. However, this

process of cooling and collapsing to form star can be disrupted. Stars formed return energy to the ISM through radiation pressure, stellar winds and supernovae. This energy has the potential to disrupt the gas in the ISM and drive galactic scale winds, as seen observationally (Heckman et al., 2000; Schwartz et al., 2006; Kornei et al., 2012). These winds create a feedback cycle, slowing the rate at which baryons are converted to stars.

Observationally, massive galaxies are seen to contain super massive black holes (see Kormendy and Ho, 2013, for a review), although the formation processes of the black hole seeds is poorly understood (see Volonteri and Bellovary, 2012, for a review). The black holes in galaxies are seen to eject significant amounts of energy into their environment, through energetic jets and winds (King and Pounds, 2003; Murray et al., 2005; Di Matteo et al., 2005) and possibly, radiation (Fabian et al., 2006; Thompson et al., 2014). This energy resource could also have significant impact on the formation of galaxies.

Clearly the physics associated with baryons is complex; understanding galaxy formation is not a trivial problem. The field of observational galaxy formation has seen significant progress in the last decade. Large area spectroscopic surveys, such as GAMA, SDSS and 2dF (Driver et al., 2009; York et al., 2000; Colless et al., 2003, respectively), have enhanced our view of the local Universe. Increased wave band coverage has improved the accuracy of photometric redshift estimates and wide field surveys covering up to redshift 4 are underway (e.g. UltraVISTA Ilbert et al., 2013; Muzzin et al., 2013). Ultra deep surveys have extended the redshift range observed up to $z \sim 7$, while some higher redshift candidates have been probed using broadband dropouts (Bouwens et al., 2010; Oesch et al., 2014). Not only is the redshift range and fields of view improving, new techniques are providing a more detailed view of galaxies. For example, integral field units open up the possibility to study the internal structure of galaxies (e.g. Stott et al., 2014). This wealth of data produces new challenges to theoretical galaxy formation to explain the interaction of the large scale structure with the detailed baryonic physics on small scales.

Some examples of the observations that require an explanation from theoretical

galaxy formation are

- Are scaling relations between stellar mass and star formation rates, black hole masses, atomic gas masses consequences of the models?
- Can the observed passive fractions of galaxies, at high masses and in dense environments be reproduced through known physical mechanisms?
- Can the evolution of the galaxy stellar mass function be explained by our understanding of galaxy formation, in particular the normalisation of the low mass slope?
- Is the phenomenon of cosmic downsizing consistent with a ΛCDM description of the Universe?
- Can the increase in galaxy sizes of an order of magnitude for passive galaxies between redshift two and zero be explained?

The understanding of baryonic processes and how they interact is crucial for explaining current galaxy observations. The process of galaxy formation is also becoming increasingly important for understanding observations probing the nature of dark matter. The study of van Daalen et al. (2014) have shown the significant impact, up to 20%, that baryonic physics can have on the matter power spectrum, while Schaller and et al. (in prep.) show the effect of baryons on halo profiles. Some of the recent challenges to the cold dark matter paradigm have, indeed, been resolved by the inclusion of baryons, for example the missing satellite problem (Cooper et al., 2010; Sawala et al., 2013).

With a framework for the evolution of the matter density field in the Universe, and an overview of the baryonic processes considered important for the formation of galaxies, we can consider ways of testing our understanding of galaxy formation.

1.2 The Role of Simulations in Understanding Galaxy Formation

There are many simple models to describe galaxy formation; a recent example is the bath tub model (Lilly et al., 2013; Peng et al., 2010). Such models summarise the processes considered to be most crucial in galaxy formation in simple terms. They provide a means of gaining a conceptual understanding of galaxy formation theory. However to fully test and validate our understanding of galaxy formation, more sophisticated models, built on the evolution of the matter density field, are required.

Very useful tools for exploring the unanswered questions of galaxy formation, and, indeed, the validating theoretical work on galaxy formation, are simulations. With the availability of computational resources increasing over the last two decades and the machinery to exploit these resources following a similar trend, simulations are playing an increasingly important role in the understanding of galaxy formation.

Simulations in the area of cosmology probe the large scale structure of Universe in the non-linear regime of halo formation, testing the ΛCDM cosmology. Initially simulations considered only the effect of gravity, simulating only the gravitational interactions, which appears to be the only force acting on most of the matter in the Universe. These simulations are called dark matter only (DMO) simulations. With more sophisticated computers and codes, DMO simulations are now a reliable source for addressing many questions in the field of cosmology, such as what is the large scale structure of the Universe and what is the mass distribution of halos in the Universe. DMO simulations are run at many scales and enable an understanding of dark matter structures (modulo baryonic effects) across Gpcs (Angulo et al., 2012) and within individual halos (Springel et al., 2008).

The primary uncertainty in DMO simulations is the underlying cosmology, which determines the initial conditions. Historically poor constraints on cosmological parameters, assuming a ΛCDM Universe, were a limiting factor in the pre-

cision of galaxy formation models. With the advent of WMAP and Planck, the initial conditions and background cosmology are well known, shifting the dominant uncertainty in galaxy formation simulations towards the modelling of baryons.

Dark matter only simulations trace the filaments, knots and voids in the matter density field, including the accretion of mass onto halos. However the processes that affect baryons and the formation of galaxies are not considered. While the accretion of gas onto halos is the first step in galaxy formation, there are many more processes at work, as discussed in Section 1.1. Once the gas is accreted, how much of the gas in the galaxy is available for star formation? If the star formation is assumed to be proportional to the accreted gas, the distribution of galaxies with stellar mass would be proportional to the predicted halo mass function. Observationally, however, the galaxy stellar mass function (GSMF) is found to have a very different shape to the halo mass function (Benson et al., 2003; Baugh, 2006, for a review). The cooling of the gas, the impact of winds produced following star formation and the disruption of gas inflow and cooling by active galactic nuclei (AGN) heating all play a role in shaping the GSMF.

There are two primary methods to model baryonic processes in simulations, semi-analytically and hydrodynamically. Before discussing these methods, it is worth noting that there are other methods applied to DMO simulations that do not attempt to model the baryons, but instead apply statistical methods to relate observed galaxies to simulated dark matter halos. Halo Occupation Distribution (HOD) models assume a functional form to relate observed galaxies to the simulations. A common variation on HOD models is abundance matching, where observed galaxies are assigned to halos based on their masses. These methods reveal trends of galaxy properties with halo mass and redshift (e.g. Behroozi et al., 2013; Moster et al., 2013) enabling an understanding of the importance of the environment on galaxy formation. However, as HOD models do not invoke any physical processes to produce the models, an understanding of why trends arise can not be provided.

Semi-analytical models (SAMs) take the approach of approximating physical

processes affecting the baryons in the Universe using a set of equations (Cole et al., 2000; Baugh, 2006; Benson and Bower, 2010). The equations are applied in post-processing to DMO simulations to describe the flow of baryons through the dark matter halos. The free parameters in the model are typically adjusted to reproduce observations. As SAMs do not require significant computational expense, vast parameter spaces can be explored. In particular, the importance, or otherwise, of various physical processes can be determined (Bower et al., 2010).

However, while these models are a fantastic tool for gaining insights into the physics of galaxy formation, caution must be applied due to the approximate nature of the method. Processes such as the recycling of gas or the effect of dense environments on galaxies require parameterisation. Furthermore, the dependence of processes, such as feedback, can only vary with global galaxy or halo properties (Bower et al., 2012). As a result of these limitations, to the modelling it is possible to over or under estimate the impact of some processes, while others may be neglected.

In any method depending on dark matter only simulations, there are limitations to the scope of galaxy properties and the underlying baryonic structure that can be studied. For example, diffuse gas in and around galaxies, in the CGM and IGM, can not be traced. Equally, gas outflows are typically approximated, and the interaction of the feedback with the interstellar and circumgalactic media can not be studied. A further limitation is that the reaction of the dark matter to baryons is neglected when using these methods (see Pontzen and Governato, 2014, for a recent review), which could prove important in this era of precision cosmology.

A different approach to simulating baryons is to trace them and the dark matter simultaneously in hydrodynamic simulations. Such simulations provide a more self-consistent method for following the growth of galaxies and their environments than dark matter only simulations. The inclusion of baryons in the simulation allows for the study of galaxy structure, as well as the diffuse gas around galaxies, while accounting for any impact of the gas on the collisionless dark matter. While hydro-simulations also require approximations, there are significantly fewer. For

example the effects of ram pressure stripping and strangulation[1] on satellite galaxies follows from the gas dynamics, as does the recycling of gas ejected from galaxies by feedback processes.

The ultimate goal of theoretical galaxy formation is to apply simulations to produce ab initio predictions of the galaxy population. However, this remains unattainable because of the extreme dynamic range and complex physics involved in galaxy formation. For example, cosmological simulations on scales of 50 to 100 comoving Mpc are required to provide a representative galaxy sample, and to represent large scale modes of the underlying matter power spectrum. In contrast, star formation occurs within molecular clouds of the inter-stellar medium, requiring resolution well below 1 pc, with similar scales (at least) required for the modelling of black hole accretion disks. In terms of mass scales, Milky Way mass galaxies of around $10^{10.5}\mathrm{M}_{\odot}$ dominate the stellar mass content of the Universe, while stars with masses of 0.1 to a few hundred solar masses form in molecular clouds of around $10^{5}\mathrm{M}_{\odot}$. In order to simulate the evolution of the Universe multi-scale methods (i.e. subgrid models) are necessary.

The subgrid models present some freedom in the hydro-simulations, in the choice of models to include and the implementation of the models. There is further freedom in these simulations beyond the subgrid schemes, for example the method applied to solve the hydrodynamic equations, e.g. smooth particle hydrodynamics (SPH), adaptive mesh refinement or moving mesh. However, Scannapieco et al. (2012) show, through a comparison of a number of codes implementing different hydro-solvers and subgrid implementations, that the subgrid modelling is the dominant uncertainty in the resulting galaxy properties.

Some of the subgrid physics prescriptions can be constrained using observations, for example the empirical Kennicutt-Schmidt law (Kennicutt, 1998) can set the parameters for the star formation law in the subgrid model (Schaye and Dalla Vecchia, 2008). However, other elements of the subgrid schemes are less well con-

[1]Strangulation is the depletion of cold gas through star formation after the fresh gas supply is removed (e.g. McCarthy et al., 2008).

strained. Of the subgrid physics implementations, the stellar and AGN feedback schemes contain the most uncertainty. In particular in their efficiency to couple the injected energy to the surrounding gas and the numerical losses associated with the feedback. The uncertainty in these models translates into freedom in the subgrid implementation and parameters. Within the choices for feedback implementation there are two aspects of the scheme to consider; how is energy injected and if/how the amount of energy injected varies in different environments. Because of spurious numerical cooling in early thermal feedback models (Springel and Hernquist, 2003), energy from stellar feedback is typically injected in kinetic form, e.g. Springel and Hernquist (2003), OWLS (Schaye et al., 2010), Oppenheimer et al. (2010), ILLUSTRIS (Vogelsberger et al., 2014), MASSIVEBLACK-II (Khandai et al., 2014), specifying a wind velocity and mass loading. (Spurious numerical cooling may still affect many thermal AGN implementations.) Dalla Vecchia and Schaye (2012) however developed a thermal heating method that distributes energy over a smaller mass than in other thermal feedback schemes, increasing the temperature jump resulting from the energy injection. When the heating temperature is high, the cooling time is longer, allowing the dense gas to expand before the energy is radiated away. This process reduces the spurious numerical radiative losses. There is a further element to distributing stellar feedback that is included in some simulations in an attempt to improve convergence when going to higher resolution (see Schaye et al., 2015, for a discussion). Resolution elements experiencing feedback can be decoupled from the hydrodynamics to allow them to escape the star-forming region (e.g. ILLUSTRIS, Oppenheimer et al. (2010), MASSIVEBLACK-II) or cooling can be switched off in an attempt to avoid numerical radiative losses (e.g. Stinson et al., 2013). However, Dalla Vecchia and Schaye (2008) show that the decoupling of the hydrodynamics affects the star formation rate efficiency, the height of the galactic disc and the distribution of winds. Some implementations (e.g. Dalla Vecchia and Schaye, 2008, 2012) have been developed so as not to require such artificial treatment of feedback events.

Due to the typical mass resolution of $\mathcal{O}(10^{6-8}\ \mathrm{M}_{\odot})$ per resolution element, any

potential environmental dependence of the energy injection can not be resolved. As a result, different scalings of the feedback energy input can be applied in the subgrid model, e.g. dark matter velocity dispersion (Oppenheimer et al. (2010), ILLUSTRIS). Other choices of scaling have also been considered, as we will discuss below.

To study galaxy formation in a statistical way, large galaxy samples are required from cosmological simulations, although the details of the physical processes clearly require approximations in such large simulations. To understand the physical processes further, hydrodynamical simulations can also be used, although at different scales. For example, extremely high resolution isolated galaxies have been used to explore the effects of molecular cooling, going beyond the atomic cooling considered in cosmological simulations (Richings et al., 2014a,b), or the impact of stellar feedback has been explored in detail in isolated galaxies (Hopkins et al., 2012) or cosmological zoom simulations (Governato et al., 2012). These studies are crucial to our understanding of important physical processes, however the computational expense of such simulations limit their use in producing large statistical samples of galaxies and explaining the global galaxy population.

Indeed it could be argued that the 100pc resolution, typical of isolated galaxy simulations, is not sufficient and still requires significant approximations. Simulations on smaller scales are needed to fully resolve the ISM physics, including star formation and feedback processes, on molecular cloud scales, e.g. (Walch et al., 2012; Creasey et al., 2013). Such small scale studies require even better resolution and more computational time. Ideally small scale simulations will predict the behaviour of physical processes on larger scales, thereby bridging the gap between simulation scales and providing more motivation for subgrid models. Such models would truly be multi-scale, and would better inform and motivate subgrid models used in cosmological simulations.

1.3 Why the EAGLE simulations?

Recently a number of groups have produced hydrodynamical simulations showing reasonable agreement with observational data at redshift zero, both for individual galaxy simulations (Aumer et al., 2013) and cosmological simulations (EAGLE, ILLUSTRIS). As many different implementations for hydrodynamics and subgrid physics exist, it is reassuring that different methods applying the same physical processes are converging towards similar virtual universes. The EAGLE simulation suite used in this work supersedes the list of cosmological simulations producing virtual universes, showing unprecedented agreement with local Universe observations greatly surpassing other hydro-simulations (Schaye et al., 2015).

There are many elements of the EAGLE simulation implementation that differ from others in the literature, that make it unique. We apply advanced SPH techniques and state-of-the-art subgrid models to capture the unresolved physics. Cooling, metal enrichment, energy input from stellar feedback, black hole modelling and feedback from AGN are included. The key differences in our simulation relate to the feedback implementations. Strong stellar and AGN feedback are ubiquitous in simulations that can successfully reproduce observed galaxy stellar masses, in EAGLE the implementation of this feedback differs from others in the literature. The thermal feedback recipe of Dalla Vecchia and Schaye (2012) is used as opposed to kinetic feedback, without any decoupling of the hydrodynamics or switching off of cooling. Furthermore, in EAGLE we choose to restrict all feedback dependencies to the local gas properties, thereby providing a more physical motivation for the feedback scaling (many variations of this dependency are considered in Crain et al., 2015). The fraction of energy from stellar feedback available to heat particles depends on the local gas density and metallicity. The motivation for the metallicity dependence comes from unresolved radiative losses that would be expected in high metallicity regions compared to lower metallicity regions. These losses can not be captured by the simulation due to the unresolved multi-phase ISM. It is also plausible that the ISM density would affect the efficiency of stellar

feedback, for example chimneys could form in high density regions resulting in gas escaping more easily when impacted by stellar feedback. As a result our feedback efficiency depends on the local gas density around a newly formed star, such that feedback is stronger in high density regions. Indeed studies of the impact of supernova on very small scales have shown variations with density and metallicity, such as those of Creasey et al. (2013), although the details of the variations do not apply on the scales we consider here. In summary, the implementation of thermal feedback, without any decoupling, and local gas property dependencies, are the major differences between the EAGLE simulations and other projects.

Another key element to the success and potential of the EAGLE simulation is the decision to calibrate the stellar and AGN feedback parameters. Although there is observational evidence for gas outflows in star forming galaxies (Veilleux et al., 2005, for a review) and galaxies with AGN (Harrison et al., 2014; Cicone et al., 2014), it is difficult to constrain both the velocities and mass loadings of these winds. While a given energy may be available from feedback processes, how that energy couples to the gas on the scales that can be probed in the simulation is unclear. A further concern in simulations is the extent of numerical radiative losses. One solution to these unknowns in the simulations is to calibrate to some galaxy property, as is carried out in SAMs (e.g. Benson et al., 2003) and as is argued in relation to hydrodynamical subgrid models in Schaye et al. (2015). In EAGLE we choose to calibrate to the local Universe GSMF, which is dependent on stellar feedback for low mass galaxies and AGN feedback for high mass galaxies. This choice results in a galaxy population at redshift zero that lives in reasonable mass halos, as predicted by abundance matching methods. As a result, in comparing with observations, stellar mass selected samples will have accretion histories similar to those of the observed galaxies.

Schaye et al. (2015) have presented an overview of the Eagle project, focusing on a comparison with observational data in the local Universe. The GSMF from the simulation was shown to lie within 0.2 dex of the observed GSMF of Li and White (2009) and Baldry et al. (2012) at redshift zero. This measure is not a prediction of

the model, due to the calibration, however other local universe quantities such as the Tully-Fisher relation, the mass metallicity relation, group and cluster properties and the column density distribution function for intergalactic CIV and OVI, which were not used for calibration, were also shown to produce results comparable to the observations. It is the EAGLE simulations that will be used throughout this work to test our understanding of galaxy formation and to expand it.

1.4 Thesis Outline

The outline of the remainder of the thesis is as follows. In *Chapter 2* we describe the detail of the EAGLE code, used throughout this work. We include a discussion of the subgrid physics applied, as these prescriptions are key to the galaxy population produced by the simulation. We present an overview of the EAGLE project results in *Chapter 3*, focusing on the evolution of galaxy properties and a comparison with observations. *Chapter 4* continues the study of the evolution of galaxy properties, here we carry out a comparison between different variations of the EAGLE subgrid physics. Models are selected based on their redshift zero galaxy properties and the impact of the variations on the evolution in the simulations is considered. *Chapter 5* looks to shed some light on the evolutionary trends seen in *Chapter 3*.

Our study of the evolution of galaxy properties begins at redshift 7 in previous chapters, below which many galaxy properties can be determined observationally. In *Chapter 6* we extend our study to the first galaxies reporting on the first study of reionization by galaxies from a simulation with a representative galaxy population evolved to redshift zero. *Chapter 7* presents a work in progress on the most highly star forming galaxies in the simulation, with the intention of understanding these galaxies relative to sub-mm galaxies. Finally in *Chapter 8* we summaries the findings reported in this thesis and end with some concluding remarks and future directions.

Chapter 2

Producing and post-processing the EAGLE *simulation suite*

The exploration of galaxy evolution embarked upon in the following chapters relies on the cosmological hydrodynamical simulations of the EAGLE project. In this chapter we present the details of the EAGLE code, following the motivation presented in the introduction. We begin with an overview of the simulation covering the main elements of the code. We focus specifically on the implementation of subgrid physics in the simulation, which currently encapsulate significant uncertainties in cosmological hydrodynamical simulations, as discussed in *Chapter 1*. The idea is to highlight and motivate some of the choices made for the EAGLE simulations, while providing a thorough understanding of the models implemented. We also include details of the post-processing procedures used to produce galaxy catalogues, which are relevant to the explanation of galaxy properties presented in the following chapters.

As a 'builder' of the EAGLE simulation suite the author has been extensively involved in the development of the EAGLE software, including porting the subgrid physics prescriptions from the OWLS code, on which this code is based, testing the ported modules, debugging the code, optimising, in particular the metal enrichment routines, extending the galaxy properties computed in post-processing and modifying the input and output (I/O). My responsibilities also included testing variations of the EAGLE code, running and post-processing the simulations and analysing the results. The work was carried out as part of the EAGLE team. The simulations were performed on 3 clusters, Cosma4 and Cosma5 at Durham Uni-

versity and the PRACE Tier 0 machine, Curie, in Paris. The analysis was carried out on the shared memory systems in the ICC at Durham University.

2.1 The EAGLE Code

The EAGLE code contains 3 primary components, the gravity solver, the hydrodynamical solver and the subgrid physics modules. The gravity solver and the framework for the hydro-solver are based on GADGET3 (last described in the literature by Springel (2005)). The GADGET code and an overview of the changes to the hydrodynamics, are described below. The subgrid physics prescriptions, evolved from the OWLS project, are described in Section 2.2.

The EAGLE simulations assume a flat ΛCDM universe, with a set of parameters determined the Planck survey data (Planck Collaboration et al., 2013); $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\sigma_8 = 0.8288$, $n_s = 0.9611$ and $H_0 = 67.77$ km s^{-1} Mpc^{-1}.

CAMB (Lewis et al., 2000, version Jan_12) was used to compute the linear transfer function with the above parameters. From the power spectrum, Gaussian initial conditions are produced for a real-space white noise field, as described in Jenkins (2013). Particle displacements and velocities are computed for redshift 127 using second-order Langrangian perturbation theory (Jenkins, 2010).

The dark matter is assumed to be represented by a collisionless fluid, interacting through gravity alone. Its motion is thus governed by the collisionless Boltzmann equation coupled to the Poisson equation in comoving coordinates (Peebles, 1980),

$$\nabla^2\Phi(\mathbf{r}) = 4\pi G\bar{\rho}(x)a^2\delta, \tag{2.1}$$

where $\Phi(\mathbf{r})$ is the gravitational potential, $\bar{\rho}$ is the mean density, a is the expansion factor and $\delta = \rho/\bar{\rho} - 1$ is the overdensity. The parallel N-body solver in GADGET is used to solve this by discretising the dark matter density field into N particles tracking the fluid. For the scales considered in cosmological simulations, the motions can be described by Newtonian mechanics. The long-range gravity force is

given by,

$$\mathbf{F}(\mathbf{r_i}) = \Sigma_{j=1,j\neq i}^{N-1} \frac{Gm_i m_j}{|r_i - r_j|^3}(\mathbf{r_i} - \mathbf{r_j}), \tag{2.2}$$

where m_i is the mass of the particle i.

This problem is computationally expensive for large values of N (scaling as $\mathcal{O}(\mathrm{N}^2)$ when solved by direct summation), so faster methods are applied for large problems. GADGET uses a combination of tree-based and mesh-based methods, called the TreePM (Bode et al., 2000; Bagla, 2002) method. For particle-mesh (PM) methods (Klypin and Shandarin, 1983; White et al., 1983) the density is mapped onto a regular mesh, and the Fourier transform of the mesh is used to solve Poisson's equation. Fast Fourier transform algorithms have optimised this method, however there is a trade-off between the mesh size and the force accuracy. Tree methods, based on multipole moments, recursively subdivide the density field into a hierarchy of cells, or a tree. For long-range calculation, particles are grouped together to compute the forces, at shorter ranges smaller grouping are used. There is no intrinsic resolution limit to this scheme. The criteria for 'walking' the tree in GADGET is given by

$$\frac{GM}{r^2}\left(\frac{l}{r}\right)^2 \geq \alpha|\mathbf{a}|, \tag{2.3}$$

where a cell encloses mass M, with side length l and is at a distance of $\mathbf{r}$. α is a tolerance parameter and $|\mathbf{a}|$ is the magnitude of the acceleration in the previous time step.

As mesh-based methods require large grids for high resolution, but are computationally efficient, these methods are used to compute the long range forces in GADGET. The tree method is applied at short ranges due to the accuracy achievable.

The gas in the simulation is a collisional fluid, described by the Euler equations

$$\frac{d\rho}{dt} = -\nabla.(\rho\mathbf{v}), \tag{2.4}$$

$$\frac{d\mathbf{v}}{dt} = -\frac{\nabla \mathrm{P}}{\rho} - \nabla\phi, \tag{2.5}$$

$$\frac{d\mathrm{u}}{dt} = -\frac{\mathrm{P}}{\rho}.\mathbf{v} \tag{2.6}$$

where ρ is the density, P is the pressure, $\mathbf{v}$ is the velocity and u is the internal energy per unit mass. The pressure, energy and internal energy are connected by an equation of state, $\mathrm{P} = (\gamma - 1)\rho\mathrm{u}$, where γ is the ratio of specific heats.

In the literature, there are 3 methods applied to solve these equations, Lagrangian methods (i.e. using particles), Eulerian methods (i.e. using a grid) and a combination (e.g. AREPO Springel (2010a)). GADGET is a particle-based SPH (Monaghan, 1992; Springel, 2010b) code, and this is the basis of the hydrodynamical technique used in EAGLE.

In SPH methods the fluid is discretised into point mass particles, which represent the properties of the fluid. Some advantages of this method are that it trivially traces high density regions, is Galilean invariant, couples easily to N−body gravity schemes and has excellent conservation properties. Continuous fluid properties, such as the denisty and pressure gradients required for the Euler equations, are described by kernel interpolation techniques (e.g. Monaghan, 1992), where the kernel, $W(|\mathbf{r}_i - \mathbf{r}_j|, h)$, depends on the separation of particles i and j and the smoothing length, h. Adaptive smoothing lengths are used so the mass within the kernel volume remains constant, resulting in a mass resolution that remains constant independently of the density flow. For particles of equal mass (which is approximately true in these simulations [1]) a constant mass per kernel volume is equivalent to a constant number of neighbours within the kernel. The kernel is chosen to be symmetric, with an integral of 1 over the volume and in the limit of $h \to 0$, it acts as a δ-function.

To compute the gas properties in SPH, such as density and pressure, they are smoothed over the kernel surrounding a particle. For example for the density,

$$\rho_i = \Sigma_j m_j W(|\mathbf{r}_i - \mathbf{r}_j|, h), \tag{2.7}$$

where ρ_i and $\mathbf{r}_i$ are the density and position of particle i, m_j, and $\mathbf{r}_j$ are the mass and position of neighbouring particles j. As smoothing lengths are adaptive in GADGET it is not necessary that $h = h_i = h_j$. The choice of h_i or h_j depends on the

[1] Varying particle masses result from stellar evolution.

choice of a *scatter* ($h = h_j$) or *gather* ($h = h_i$) method, as detailed in Hernquist and Katz (1989). In Hernquist and Katz (1989) it is shown that both methods result in similar errors. The *gather* method is applied in GADGET (Springel et al., 2001a). For the computation of dynamical forces the difference in smoothing lengths can result the violation of Netwon's third law, $F_{ij} \neq F_{ji}$. To avoid such issues when forces are computed the smoothing procedure is symmetrised as in Hernquist and Katz (1989), giving

$$\rho_i = \Sigma_j \frac{1}{2} m_j \left[W(|\mathbf{r}_i - \mathbf{r}_j|, h_i) + W(|\mathbf{r}_i - \mathbf{r}_j|, h_j) \right]. \tag{2.8}$$

The SPH formulation implemented in GADGET has been shown to perform poorly in certain canonical hydrodynamics tests. For example, an artificial pressure term arises at contact discontinuities (Agertz et al., 2007) and the entropy profiles in the cores of halos differ between GADGET simulations and adaptive mesh refinement (AMR) codes (Frenk et al., 1999; Power et al., 2014; Mitchell et al., 2009). As a result of these issues, a more recent formulation of SPH is implemented in EAGLE called ANARCHY (Dalla Vecchia, in prep.). This implementation includes

- The pressure-entropy formulation of SPH described in Hopkins (2013), which removes the artificial pressure term that arises in the SPH implemented in GADGET, while conserving energy and entropy.

- The artificial viscosity switch of Cullen and Dehnen (2010) to reduce artificial viscosity away from shocks and an artificial conduction switch described by Price (2008) to mimic numerical thermal diffusion.

- A C2 Wendland (1995) kernel with 58 neighbours, replacing the cubic spline with 48 neighbours, to interpolate SPH properties across neighbouring particles to prevent particle pairing, while the number of neighbours produces a similar resolution as the 48 neighbour cubic spline (Dehnen and Aly, 2012).

- The time step limiter from Durier and Dalla Vecchia (2012) that ensures feedback events are accurately modelled.

While these improvements to SPH are important in hydrodynamical tests, the impact of these changes on galaxy properties in cosmological simulations, as considered here is minimal (Schaller and et al., in prep.).

2.2 Subgrid Physics

The subgrid physics included in the EAGLE simulation suite is broadly based on the OWLS (Schaye et al., 2010) and GIMIC (Crain et al., 2009) projects, although many improvements, in particular to the stellar feedback scheme and black hole growth, have been implemented.

The EAGLE simulation suite includes a number of cosmological simulations with varying subgrid parameters and implementations in a range of box sizes (6-100 cMpc) boxes and 3 levels of resolution (Crain et al., 2015). All simulation variations include subgrid models for cooling, star formation, metal enrichment, stellar feedback and black hole growth and feedback (with the exception of the NOAGN model, that does not include the black hole growth and feedback). In this section we focus on the subgrid implementation and parameters relevant to two of the simulations, Ref-L100N1504 and Recal-L025N0752.

Ref-L100N1504 is a simulation in a $(100$ cMpc$)^3$ periodic box using 2×1504^3 particles. The initial mass for baryonic particles is 1.81×10^6 M$_\odot$ and the mass of dark matter particles is 9.70×10^6 M$_\odot$. Plummer equivalent comoving gravitational softenings are set to 1/25 of the mean inter-particle spacing and are limited to a maximum physical size of 0.70 physical kpc (pkpc). Recal-L025N0752 has 8 times better mass resolution and 2 times better spatial resolution than Ref-L100N1504, but is run in a smaller volume of (25 cMpc)^3. The box sizes, particle numbers and resolutions of these simulations are summarised in Table 2.1. The subgrid physics, and variations in parameters between these simulations are discussed following the description of the subgrid methods.

Table 2.1: Detail of box size, particle number, particle mass and gravitational softening for Ref-L100N1504 and Recal-L025N0752 simulations.

Simulation Name	L	N	$m_{\rm g}$	$m_{\rm dm}$	$\epsilon_{\rm com}$	$\epsilon_{\rm prop}$
	[cMpc]		[$M_\odot$]	[$M_\odot$]	[ckpc]	[pkpc]
Ref-L100N1504	100	$2\times(1504)^3$	1.81×10^6	9.70×10^6	2.66	0.70
Recal-L025N0752	25	$2\times(752)^3$	2.26×10^5	1.21×10^6	1.33	0.35

- Radiative cooling and photo-heating in the simulation are included as in Wiersma et al. (2009a). The element-by-element radiative rates are computed in the presence of the cosmic microwave background (CMB) and the Haardt and Madau (2001) model for UV and X-ray background radiation from quasars and galaxies. Eleven elements are tracked: H, He, C, N, O, Ne, Mg, Si, Fe, Ca and Si. The cooling tables, as a function of density, temperature and redshift are produced using CLOUDY, version 07.02 (Ferland et al., 1998), assuming the gas is optically thin and in photoionization equilibrium.

  Above the redshift of reionization, assumed to be 11.5 based on the Planck survey, the CMB and a Haardt and Madau (2001) background up to 1Ryd are imposed to account for photo-dissociation of H_2. Hydrogen reionization is simulated by switching on the full Haardt and Madau (2001) background at redshift 11.5. Helium reionization occurs at redshift 3.5.

- Star formation is implemented following Schaye and Dalla Vecchia (2008). Gas particles above a density threshold, $\mathrm{n_H^*(Z)}$, have a probability of forming stars, determined by their pressure. The Kennicutt-Schmidt star formation law (Kennicutt, 1998), under the assumption of disks in vertical hydro-static equilibrium, can be written as

$$\dot{m}_* = m_\mathrm{g} A (1\mathrm{M}_\odot \mathrm{pc}^{-2})^{-\mathrm{n}} \left(\frac{\gamma}{\mathrm{G}} \mathrm{f_g P}\right)^{(\mathrm{n}-1)/2}, \tag{2.9}$$

  where m_g is the gas particle mass, A and n are the normalisation and power law index of the Kennicutt-Schmidt star formation law, γ = 5/3 is the ratio of specific heats, G is the gravitational constant, f_g = 1 is the gas fraction and P is the pressure of the gas (Schaye and Dalla Vecchia, 2008). The imposed star formation law uses the observational values for A and n, as measured by Kennicutt (1998) (converted to a Chabrier initial mass function (IMF)). Using the pressure to determine the star formation rate implies that changes to the imposed pressure-density relation, $P = P_\mathrm{eos}(\rho)$, will not require retuning of the implemented star formation law to reproduce the Kennicutt-Schmidt law (Schaye and Dalla Vecchia, 2008).

Above a given density, gas is expected to become thermally unstable and form stars. A star formation threshold, above which gas is expected to form stars is imposed. This star formation threshold is metallicity dependent and given by

$$n_{\rm H}^{*}(Z) = 0.1{\rm cm}^{-3}\left(\frac{Z}{0.002}\right)^{0.64} \tag{2.10}$$

where Z is the metallicity (as in Schaye, 2004). Low metallicity gives a higher threshold to capture the metallicity dependence of the HI-H_2 phase transition.

A pressure-density relation is imposed, $P \propto \rho^{\gamma_{\rm eff}}$, with $\gamma_{\rm eff} = 4/3$, to model the unresolved multi-phase ISM. This value of $\gamma_{\rm eff}$ ensures that the Jeans mass is independent of density. The imposed P-ρ relation prevents spurious fragmentation (Schaye and Dalla Vecchia, 2008).

The star formation law is implemented stochastically. Gas particles selected for star formation are converted to collisionless star particles, which represent a simple stellar population with a Chabrier (2003) IMF.

- Stellar evolution and enrichment is based on Wiersma et al. (2009b) and detailed in S14. Mass loss from AGB stars, winds from massive stars and core collapse supernovae are the three stellar evolutionary channels for element production. EAGLE tracks 11 elements most important to radiative cooling. The yield tables of Marigo (2001) and Portinari et al. (1998) are used. The mass and metals are distributed within the SPH kernel. Stellar evolution is carried out at every time step for the first 100 Myr of a stars lifetime, when over 80% of the mass and metals are lost. For the remainder of the simulation, stellar evolution for the star is done every 10 time steps, to improve the computational speed.

- Stellar feedback is treated stochastically, using the thermal injection method described in Dalla Vecchia and Schaye (2012). The total available energy from core collapse supernova is computed for a Chabrier IMF assuming all stars in the stellar mass range $6 - 100$ M$_\odot$ release 10^{51} erg of energy into the ISM.

This energy is injected after a delay of 30 Myr from the time the star particle is formed. Rather than heating all gas particle neighbours within the kernel, neighbours are selected stochastically based on the available energy, then heated to a fixed temperature of $\Delta T = 10^{7.5}$K. The stochastic heating distributes the energy over less mass than heating all neighbours. This results in a longer cooling time relative to the sound crossing time allowing the thermal energy to be converted to kinetic energy, thereby mitigating spurious losses (Dalla Vecchia and Schaye, 2012).

In EAGLE, the fraction of the available stellar feedback energy actually injected into the ISM depends on the local gas metallicity and density. This is computed using a sigmoid function,

$$f_{\rm th} = f_{\rm th,min} + \frac{f_{\rm th,max} - f_{\rm th,min}}{1 + \left(\frac{Z}{0.1Z_\odot}\right)^{n_Z} \left(\frac{n_{\rm Hbirth}}{n_{\rm H,0}}\right)^{-n_{\rm n}}}, \tag{2.11}$$

where Z is the metallicity of the star particle, $n_{\rm H,birth}$ is the density of the star particles parent gas particle when the star was formed and $Z_\odot = 0.0127$ is the solar metallicity. The stellar feedback fraction is in units of the available core collapse supernova energy.

The values for $f_{\rm th,max}$ and $f_{\rm th,min}$, parameters for the maximum and minimum fractions, are fixed at 3 and 0.3 for simulations analysed here. At low Z and high $n_{\rm H,birth}$, $f_{\rm th}$ asymptotes towards $f_{\rm th,max}$ and at high Z and low $n_{\rm H,birth}$ towards $f_{\rm th,min}$. Applying up to 3 times the available energy can be justified by appealing to the different forms of stellar feedback, e.g. supernova, radiation pressure, stellar winds, which are not treated separately here as we do not have the resolution to resolve the details of the stellar feedback. This also offsets numerical radiative losses, as shown by Crain et al. (2015).

The power law indexes are $n_{\rm Z} = n_{\rm n} = 2/\ln(10)$ for the Ref model, with $n_{\rm n}$ changed to $1/\ln(10)$ for the Recal model, resulting in a weaker dependence of the high resolution model on the density. The normalisation of the density term, $n_{\rm H,0}$, is set to 0.67 cm^{-3} for the Ref model and 0.25 cm^{-3} for the Recal

model. To match the same observational constraints at higher resolution we require a weaker density and metallicity dependence.

- Black hole seeding and growth is implemented as described in S14. Halos with a mass greater than 10^{10} h^{-1} M$_\odot$ are seeded with a black hole particle of 10^5 h^{-1} M$_\odot$, using the method of Springel (2005). Black holes then grow through mergers and accretion. If two black holes are within three gravitational softening lengths, $\epsilon_{\rm BH}$, of each other, with a relative velocity less than the circular velocity at $3\epsilon_{\rm BH}$, the black holes merge. Accretion of ambient gas onto black holes follows a modified Bondi-Hoyle formula, that accounts for the angular momentum of the accreting gas (Rosas-Guevara et al., 2013). Note that the black hole accretion rate is not increased relative to the standard Bondi accretion rate in high density regions as, in Rosas-Guevara et al. (2013) and Booth and Schaye (2009), nor is it multiplied by an ad-hoc factor as in Springel (2005).

  For the black hole growth there is one free parameter, $C_{\rm visc}$, which is used to determine the accretion rate.

$$\dot{m}_{\rm accr} = \min(\dot{m}_{\rm bondi}\left[C_{\rm visc}^{-1}(\mathrm{c_s}/V_\Phi)^3\right], \dot{m}_{\rm bondi}), \tag{2.12}$$

  where $\mathrm{c_s}$ is the sound speed of the gas, V_Φ is the circular velocity of the gas around the black hole and $\dot{m}_{\rm bondi}$ is the classic Bondi-Hoyle accretion rate,

$$\dot{m}_{\rm bondi} = \frac{4\pi G^2 m_{\rm BH}^2 \rho}{(\mathrm{c_s}^2 + \mathrm{v}^2)^{3/2}}, \tag{2.13}$$

  where v is the relative velocity of the black hole and the gas. The accretion rate is limited to the Eddington rate, given by

$$\dot{m}_{\rm Edd} = \frac{4\pi G m_{\rm BH} m_{\rm p}}{\epsilon_{\rm r}\sigma_{\rm T} c}, \tag{2.14}$$

  where $m_{\rm p}$ is the proton mass, $\sigma_{\rm T}$ is the Thompson scattering cross section and $\epsilon_{\rm r}$ is the radiative efficiency of the accretion disc. The free parameter $C_{\rm visc}$ relates to the viscosity of the (subgrid) accretion disc and $\left(\frac{c_{\rm s}}{V_\Phi}\right)^3/C_{\rm visc}$ relates the Bondi and viscous time scales.

Table 2.2: Parameters variations for Ref-L100N1504, Ref-L025N0376 and Recal-L025N0752.

Simulation Prefix	$n_{\rm H,0}$	$n_{\rm n}$	$C_{\rm visc}$	$\Delta {\rm T}_{\rm AGN}$
	[cm^{-3}]			[K]
Ref	0.67	$2/\ln(10)$	2π	$10^{8.5}$
Recal	0.25	$1/\ln(10)$	$2\pi \times 10^3$	10^9

- AGN feedback follows the accretion rate onto the black hole. A fraction of the accreted gas rest mass energy is released as thermal energy into the surrounding gas. Stochastic heating, similar to the supernova feedback scheme, is implemented with a fixed heating temperature $\Delta {\rm T}_{\rm AGN}$, where $\Delta {\rm T}_{\rm AGN}$ is a free parameter. The method used is based on that of Booth and Schaye (2009).

2.3 Resolution tests

To test the numerical convergence of the simulations we draw a distinction between strong and weak convergence, as defined and motivated in S14. Strong convergence implies similar results are produced in simulations of different resolutions without any change to the subgrid parameters. S14 argues that strong convergence is not expected from current simulations, as higher resolution often implies changes in the effective physics of the model, for example energy from feedback events often scales with the mass of the star particle formed. S14 also argues that with higher resolution the physical conditions of the ISM and the radiative losses resolved will change. Without turning off radiative cooling or the hydrodynamics (which could be sensitive to the point at which they are turned back on) the changes to the ISM and radiative losses are expected to impact the strong convergence of the simulation.

The EAGLE project instead focuses on demonstrating that the simulation shows good weak convergence (although S14 shows that the strong convergence of the simulation is on par with other hydrodynamical simulations). As current simula-

tions can not make ab initio predictions for galaxy properties, and calibration is required, the high resolution EAGLE simulation subgrid parameters are recalibrated to the same observables as the standard resolution simulations. This recalibrated high resolution model, Recal-L025N0752, enables us to test the weak convergence behaviour of the simulation and to push our results for galaxy properties to 8 times lower stellar mass. The parameters chosen for the Recal model are based on the recalibration of the simulation to the redshift zero GSMF and galaxy sizes, as for the standard resolution simulations. In Table 2.2 we highlight the parameters that are varied between the Ref and Recal models.

As a simulation with a factor of 8 better mass resolution requires a minimum of 8 times the CPU time (in practice the increase in time is longer due to the higher density regions resulting in shorter time steps and difficulties in producing perfectly scalable algorithms), we compare the (100 cMpc)^3 intermediate resolution simulation to a (25 cMpc)^3 high resolution simulation. Note that for volume averaged properties the (25 cMpc)^3 box differs from the (100 cMpc)^3 box not only due to the resolution but also due to the absence of larger objects and denser environments in the smaller volume. The (25 cMpc)^3 simulation has ~ 64 times fewer objects, with lower masses which limits the comparison particularly at high redshifts.

2.4 Simulation data and analysis

On-the-fly and post-processing analysis plays a crucial role in simulations as this enables us to analyse the evolution of the particles. Many changes have been implemented to the I/O in EAGLE relative to OWLS, on which the original code is based. This includes new types of output files, new output variables and galaxy properties. There are five primary output styles and two sets of output times, with a small amount of information output at every time step. We record complete particle properties in snapshots from the simulation at 29 points in time between redshift 20 and 0. At each snapshot time we produce a friends-of-friends halo catalogue and a

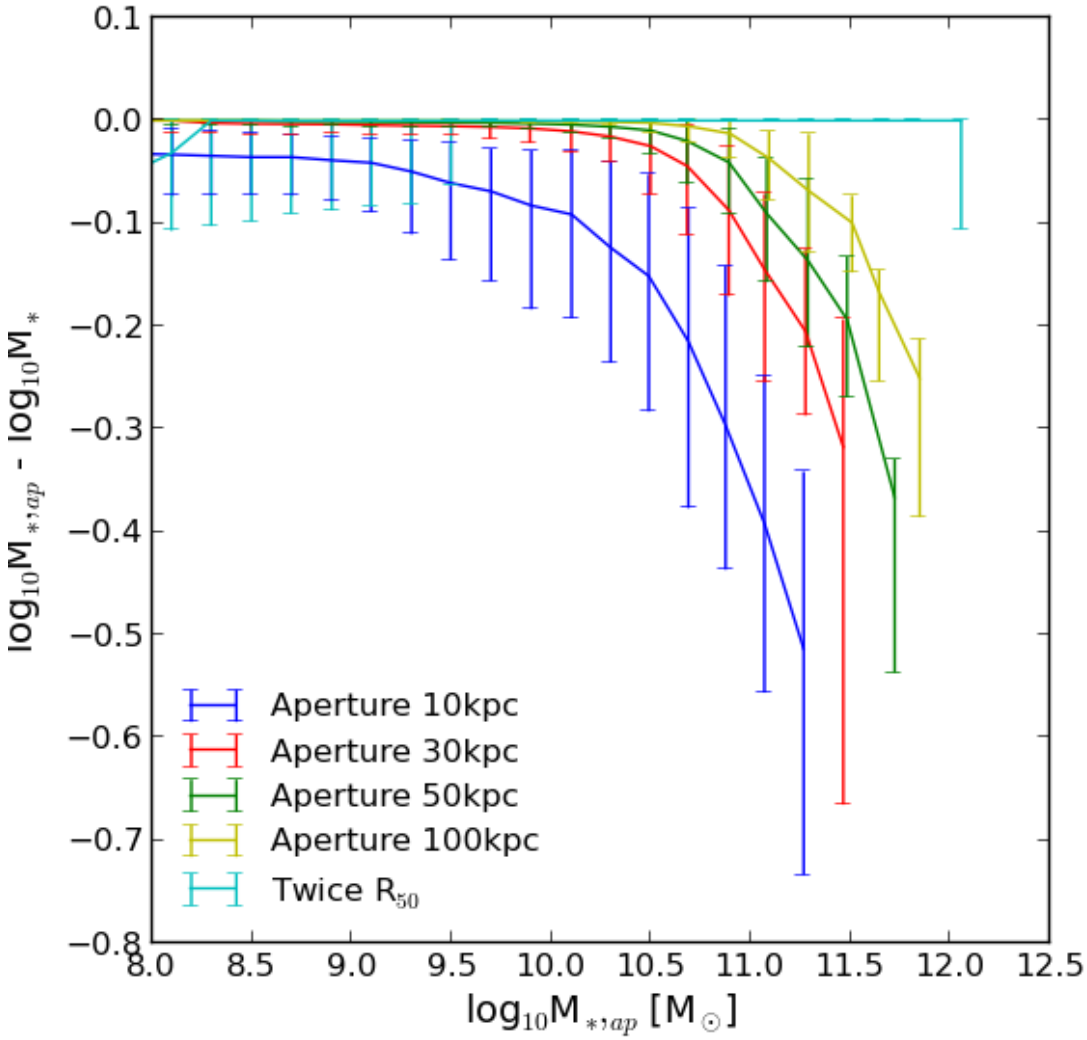

Figure 2.1: The fraction of stellar mass inside an aperture relative to the total stellar mass in a subhalo as a function of stellar mass for 3D apertures of radii listed in the legend, at redshift 0.1. The error bars enclose the 10^{th} to 90^{th} percentiles. The 10 pkpc aperture does not recover the total stellar mass at any mass, implying that most galaxies have radii greater than 10 pkpc. The apertures from 30 to 100 pkpc all recover the total stellar mass for galaxies of stellar mass smaller than 10^{10} $M_{\odot}$. For higher masses the apertures do not recover all the stellar mass. The mass at which the total stellar mass is no longer fully recovered is proportional to the size of the aperture. The stellar mass outside the aperture is due to star particles in the halo of the galaxy.

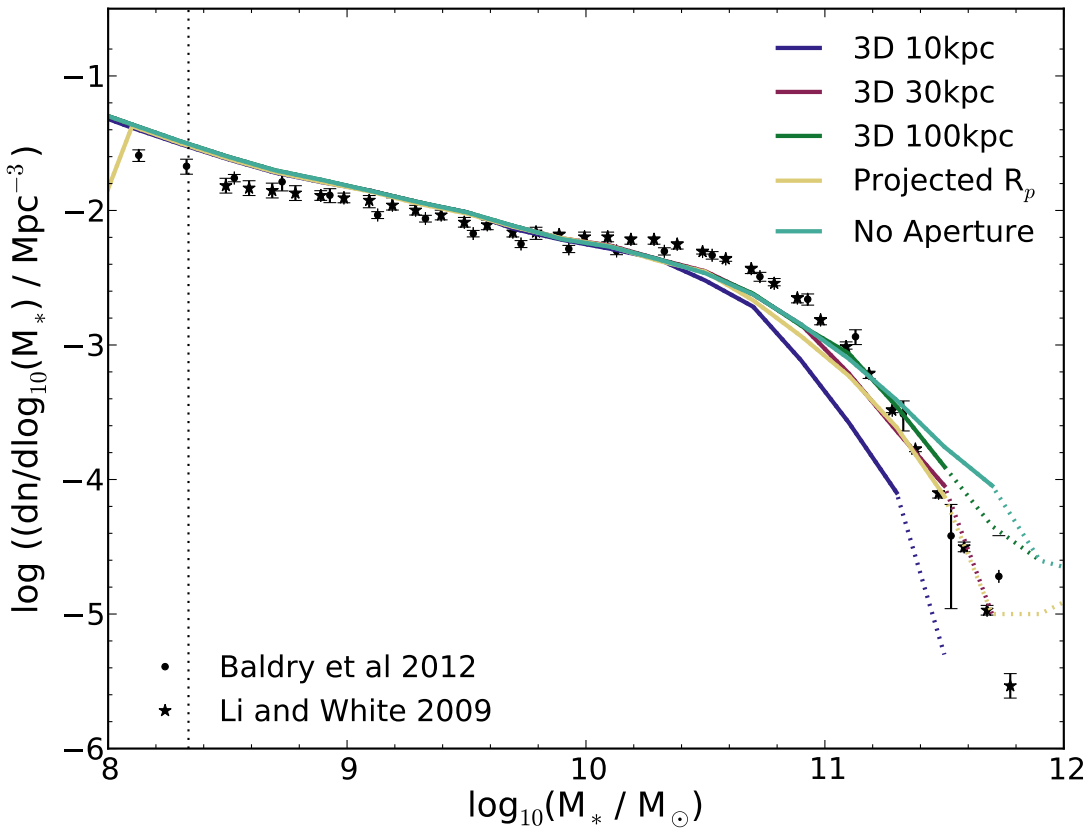

Figure 2.2: The galaxy stellar mass function at redshift 0.1 using different definitions of galaxy stellar mass. When there are fewer than 10 galaxies per mass bin, curves are dotted. The vertical dotted line show the mass of a galaxy with 100 baryonic particles, as a resolution guide. Observational data from Baldry et al. (2012) and Li and White (2009) are shown for comparison. The galaxy stellar mass function for a projected Petrosian aperture, R_p, as defined in the text, is shown in yellow. The 30 pkpc aperture most closely resembles the Petrosian aperture, which is applied to SDSS observations, as such we will use this aperture to define galaxy stellar mass and stellar mass outside this aperture is considered to contribute to the intra-cluster light.

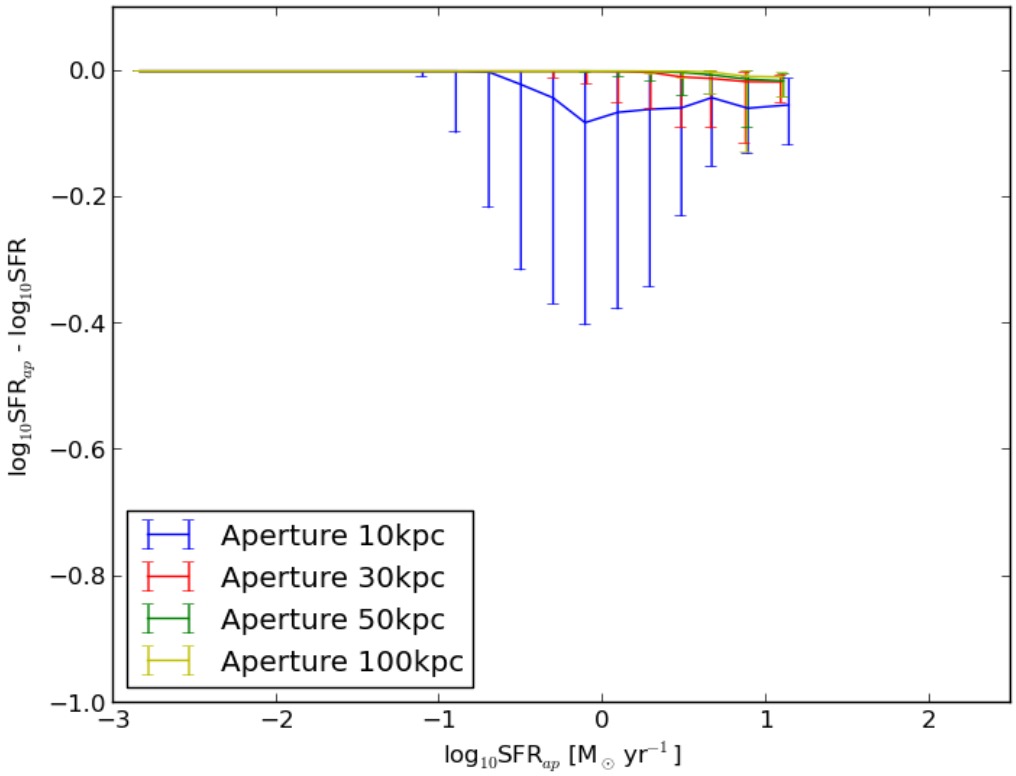

Figure 2.3: The fraction of star formation inside an aperture relative to the total star formation rate as a function of star formation rate for 3D apertures of radii listed in the legend, at redshift 0.1. Unlike the stellar mass, using apertures of between 30 and 100 pkpc the median total star formation rate recovered is approximately equal to the total star formation rate. The star formation in galaxies is much more centrally concentrated than the stellar mass. We apply 30 pkpc apertures to determine the star formation rates of galaxies, for consistency with the stellar mass, however in practice the difference is negligible.

subhalo (galaxy) catalogue, as described below. After finding all halos (and spherical over densities) snapshot files containing only the particles within halos and their spherical over densities are produced, resulting in files significantly reduced in size, and hence accessible on machines with more limited memory restrictions. These files are referred to as particle data files. We complement the snapshots with outputs of a reduced set of particle properties at 400 redshifts between 20 and 0, referred to as snipshots, providing improved time resolution.

Halo finding is carried out using the friends-of-friends (FoF) method (Davis et al., 1985), with a linking length of 0.2 times the mean inter-particle separation. Dark matter particles within a linking length of each other are assigned to the same structure, any structure with more than 20 particles is considered a halo. Baryonic particles are assigned to the group of their nearest dark matter particle. Self-bound overdensities within the group are identified using SUBFIND (Springel et al., 2001b; Dolag et al., 2009); these substructures are the galaxies in our simulation. A 'central' galaxy is the substructure with the largest mass within a given halo. All other galaxies within a halo are 'satellites'. Note that any FoF particles not associated with satellites are assigned to the central object, thus the (total) mass profile of a central galaxy may extend throughout its halo.

A galaxy's stellar mass is defined as the stellar mass associated with the subhalo within a 3D 30 pkpc radius, centred on the subhalo centre of potential. In Fig. 2.1 we show the fraction of the total stellar mass recovered for apertures between 10 pkpc and 100 pkpc at redshift 0.1 in the simulation. Using the galaxy stellar mass within 30 pkpc is equivalent to the total subhalo mass for galaxies with masses less than 10^{10} $M_{\odot}$, but excludes diffuse mass around larger subhalos, which would contribute to the intra-cluster light (ICL). At higher redshifts the mass outside 30 pkpc decreases, as galaxy stellar masses decrease and galaxies are more compact. The mass recovered in an aperture of radius twice the half mass radius, R_{50}, is also shown in Fig. 2.1. This is the aperture applied in Vogelsberger et al. (2014). In the EAGLE simulation, this aperture does not recover the total stellar mass for low mass galaxies but does for high mass galaxies, so does not provide a good

approximation to the observed ICL.

In Fig. 2.2, the galaxy stellar mass function is shown for the different apertures. The exponential break is sensitive to the choice of aperture, where smaller apertures result in a steeper break at lower stellar masses than larger apertures. We also compute the stellar masses based on a 2D Petrosian aperture, as is often used in observations, e.g. Li and White (2009). The Petrosian radius is the radius at which the mean surface brightness (or stellar mass density, in this case) within an annulus around radius r is 0.2 times the mean surface brightness within r. The 3D 30 pkpc aperture produces stellar masses equivalent to this more observationally motivated aperture.

The same 3D 30 pkpc aperture is applied when computing the star formation rates in galaxies. The fraction of the total subhalo star formation rate recovered within apertures between 10 and 100 pkpc is shown in Fig. 2.3. Apertures greater than 10 pkpc have only a minimal effect on the star formation rates because the vast majority of star formation occurs in the central few pkpc, even for massive galaxies.

Merger trees have been produced from the 29 EAGLE simulation snapshots for Ref-L100N1504. The merger trees are based on the subhalo catalogues and are produced by tracing the most bound particles in a subhalo. Using this method the descendants of subhalos at later times can be identified as well as the main progenitors of a subhalo at each snapshot. The search for subhalo descendants is carried out across multiple snapshots, to avoid loosing subhalos as they cross the dense cores of other structures. A full description of the merger tree code is available in Qu and et al. (in prep.).

All analysis presented in this work was produced using a python code written by the author to read, manipulate and plot data from the EAGLE simulations, unless otherwise stated.

Chapter 3

Eagle: A Virtual Universe

3.1 Introduction

One of the most important aims of modern cosmology is to understand the growth and evolution of the Universe's galaxy population. What physical processes determine galaxy masses and sizes? How do these properties evolve throughout cosmic history? How do stars and AGN regulate the evolution of galaxy properties? Theoretical models are a valuable tool for exploring these questions and expanding our understanding of the Universe around us. The focus of this chapter is to compare the evolution of galaxy properties in EAGLE, a new suite of hydrodynamical simulations, to the observed evolution. To determine if the simulations can be used for the purpose of understanding the evolution of the galaxy population, we test if the simulations reproduce observations across cosmic time and if the galaxies in the simulation are representative of the evolution in the observed Universe.

Hydrodynamical simulations track baryonic and dark matter simultaneously, in contrast to other methods that evolve dark matter only simulations and apply baryonic physics as a post-processing step. Due to the tracing of baryonic matter, hydrodynamical simulations require fewer assumptions to model galaxy formation. Assumptions are restricted to the modelling of subgrid processes that can not be resolved directly. The inclusion of baryons in the simulation enables the study of structure within galaxies, as well as the diffuse gas around galaxies, in the circumgalactic and intergalactic media. Any gravitational impact of the gas on the collisionless dark matter is also accounted for. This results in a more detailed and self-consistent way of following the growth of galaxies and their environments.

As discussed in *Chapter 1*, the extreme range of physical scales required in cosmological simulations of galaxy formation imply subgrid schemes are necessary. The EAGLE simulations include subgrid models for star formation, stellar feedback, metal enrichment, radiative cooling, black hole growth and AGN feedback. The subgrid models present some freedom in the hydrodynamical simulations, in the choice of models to include and in the implementation of the models, and therefore need to be calibrated.

S14 and *Chapter 2* present an overview of the EAGLE subgrid physics. Of the subgrid physics implementations, the stellar and AGN feedback schemes contain the most uncertainty, in their efficiency to couple the injected energy to the surrounding gas and the numerical losses associated with the feedback. The stellar feedback is implemented as stochastic thermal heating, without any decoupling of particles from hydrodynamic forces or temporarily disabling cooling. The available feedback energy depends on the local gas properties. The AGN feedback scheme also uses stochastic thermal heating, while the accretion onto black holes (on which the feedback depends) accounts for the angular momentum of the local gas. The subgrid parameters for the feedback are calibrated to the redshift zero observed GSMF, with consideration for galaxy sizes. For motivation of these choices, see S14.

In S14, which focuses on a comparison with observational data in the local Universe, the GSMF from the simulation was shown to lie within 0.2 dex of the observed GSMF of Li and White (2009) and Baldry et al. (2012) at redshift ~ 0. Other local Universe quantities such as the Tully-Fisher relation, the mass metallicity relation, group and cluster properties and the column density distribution function for intergalactic CIV and OVI were also shown to produce results comparable to the observations. As these properties were not considered in the calibration of the simulations, they are predictions by the model. In this chapter we extend the comparison from the local Universe to high redshift, exploring the evolution of galaxy properties. The evolution was not considered when calibrating the model and is thus also a prediction of the simulation.

It is only with this current generation of simulations, due to improved computational resources and more complete subgrid physics (e.g. including metal-line cooling and implementation of stellar and AGN feedback that do not suffer from excessive radiative losses), that calibration of models to the observed redshift ~ 0 GSMF is feasible. As a result, the agreement between EAGLE and the local Universe GSMF is unprecedented. Having confirmed that the simulation is representative of the observed local Universe, it is useful to determine if the redshift zero galaxies have histories similar to those observed. In the case where simulated galaxies are representative of the observed Universe across cosmic time, the simulation can be used as a powerful tool to interpret observational data and the connection between galaxy populations at different redshifts. Given the detailed level of information about galaxy structure, the circumgalactic medium and the intergalactic medium available from hydrodynamical simulations, the availability of a simulation that can produce both a representative sample of redshift zero galaxies and an accurate growth of galaxies opens up new paths to expand our knowledge of galaxy formation.

In this chapter we focus on the evolution of galaxy stellar masses and star formation rates. In recent years there have been a number of surveys extending observations of galaxy stellar masses to higher redshifts and lower stellar masses, e.g PRIMUS (Moustakas et al., 2013), UltraVISTA (Ilbert et al., 2013; Muzzin et al., 2013), ZFORGE (Tomczak et al., 2014). These surveys enable a comparison of the simulated galaxies to observations up to redshift 4. UV observations extend the comparison to even higher redshift, with inferred GSMFs available up to redshift 7 (González et al., 2011; Duncan et al., 2014). Observations of star formation rates also span the redshift range 0 to 7, with many tracers of star formation (e.g. IR, radio, UV) used. With so many observational data sets available there is a wealth of data to compare to, providing a stringent test for the evolution in simulations.

This chapter is organised as follows: In Section 3.2 we present the evolution of the stellar mass growth in the simulation, focusing on the stellar mass density and the GSMF from redshift 0 to 7. We follow this with an analysis of the star formation

rate density and specific star formation rates in Section 3.3. In Section 3.4 we discuss the results and we summarise in Section 3.5. We find that the galaxy properties presented are in agreement with the observations to the level of the observational systematic uncertainties across all redshifts.

The EAGLE simulation suite adopts a flat ΛCDM cosmogony with parameters from Planck (Planck Collaboration et al., 2013); $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\sigma_8 = 0.8288$, $n_s = 0.9611$ and $H_0 = 67.77$km s^{-1} Mpc^{-1}. A Chabrier (2003) IMF is used in the simulations. Where necessary observational stellar masses and volumes are rescaled to the Planck cosmology and renormalised to a Chabrier IMF. Stellar masses from the simulation are quoted in solar masses and computed within spherical apertures with 30 proper kpc radii, except where stated otherwise. Star formation rates are computed within the same aperture. Distances and volumes are quoted in comoving units (e.g. cMpc), unless stated otherwise. Proper distance units will be quoted, for example, as pMpc. Note that values are not given in h^{-1} units.

3.2 Evolution of galaxy stellar masses

We will begin this section by comparing the growth in stellar mass density across cosmic time in the largest EAGLE simulation, Ref-L100N1504, to a number of observational data sets. This is followed with a comparison of the GSMF from redshift 0 to 7 and a discussion on the impact of stellar mass errors. We also consider the convergence of the GSMF in the simulation at different redshifts. Finally, we will provide the Schechter function parameters for the GSMF, to enable easy reproduction of the GSMFs from the simulation.

3.2.1 The stellar mass density

We begin the evaluation of the evolution in the primary EAGLE simulation, Ref-L100N1504, by considering the build up of stellar mass. In Fig. 3.1, we present the stellar mass density (ρ_*) as a function of lookback time, with redshift on the

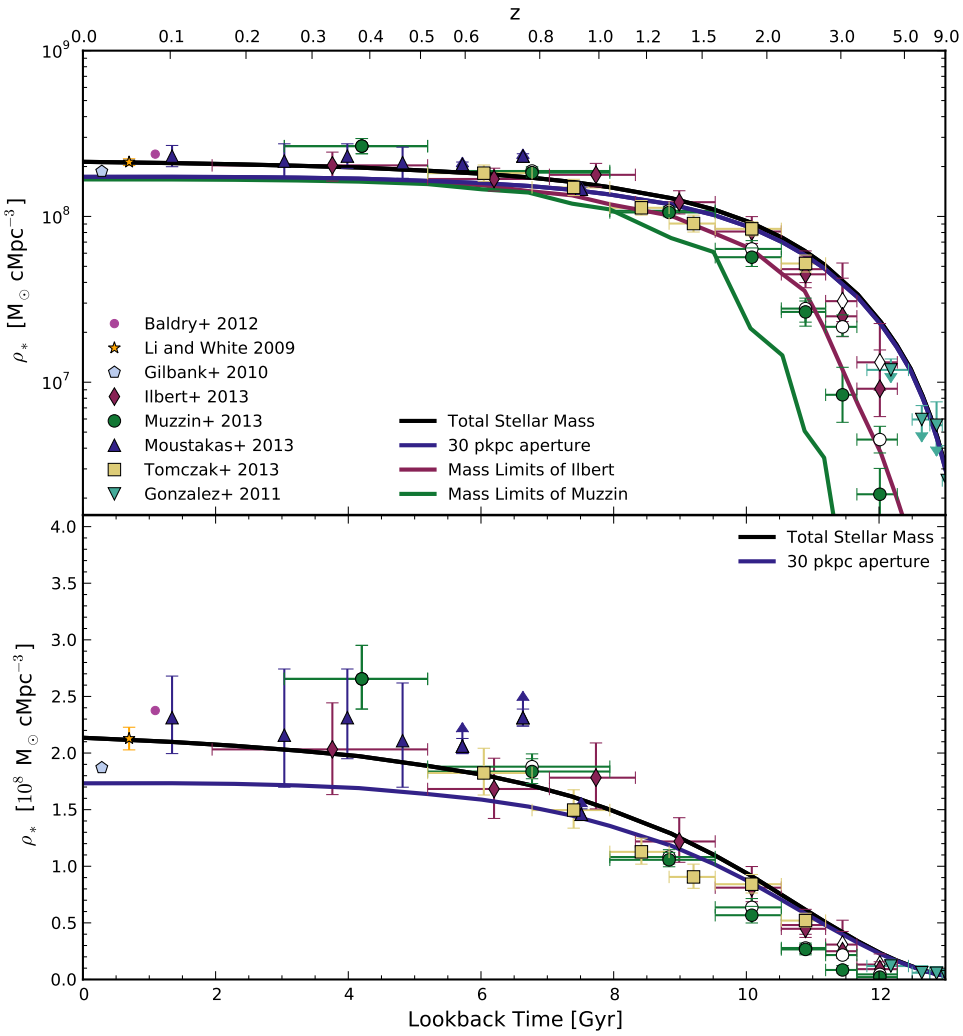

Figure 3.1: The stellar mass density in simulation Ref-L100N1504 as a function of time on a log and linear scale (top and bottom panels, respectively). The black solid curve is the total stellar mass density from the simulation and the blue curve is the stellar mass density in galaxies (i.e. excluding the ICL). Observational data is shown as data points and indicated in the legend. Open symbols show estimates from observations that include extrapolations of the GSMF below the mass completeness of the survey. The **top panel** shows ρ_* for all galaxies in the simulation in *blue* and ρ_* for galaxies above the completeness limit of observations by Ilbert et al. (2013) and Muzzin et al. (2013) in red and green respectively. The data sets or Ilbert et al. (2013) and Muzzin et al. (2013) are coloured accordingly, and lines should be compared to filled red and green symbols (no extrapolation). The **bottom panel** shows ρ_* on a linear scale.

upper x-axis. The stellar mass density is plotted as a function of time, as opposed to redshift or $\log_{10}$(1+z), to show the growth rate of the stellar mass. When ρ_* is plotted as a function of redshift, it is less clear how rapid the growth is at high redshift, or how flat the growth is below redshift one, which accounts for over half the age of the Universe.

We plot recent observational estimates of ρ_* from a number of galaxy surveys. Around redshift 0.1 we show data from Baldry et al. (2012) (GAMA survey), Li and White (2009) (SDSS), Gilbank et al. (2010a) (Stripe82 - SDSS) and Moustakas et al. (2013) (PRIMUS). The values agree to within 0.55×10^8 $M_\odot$cMpc^{-3}, which is better than 0.1 dex. The Moustakas et al. (2013) data set extends to redshift one, providing an estimate for ρ_* for galaxies with masses greater than $10^{9.5}$ $M_\odot$ in 7 redshift bins between redshift zero and one. Note, however, that above redshift 0.725 the Moustakas et al. (2013) measurements of ρ_* are a lower limit as they only include galaxies with stellar masses of 10^{10} $M_\odot$ or above. Ilbert et al. (2013) and Muzzin et al. (2013) estimate ρ_* from redshifts 0.2 to 4 in 8 and 7 redshift bins, respectively, from the UltraVISTA survey. These two data sets use the same observations but apply different signal-to-noise limits and analysis to infer stellar masses. We include both studies to assess the intrinsic systematics. Both data sets extrapolate the observations to 10^8 $M_\odot$ to estimate the total stellar mass density. The data sets are consistent within the estimated error bars up to redshift 3. Above redshift 3 they differ, primarily because of the strong dependence of ρ_* on extrapolations below the mass completeness limit of the survey. The estimated ρ_* from observed galaxies can be compared to the extrapolated ρ_* for each data set by comparing the filled and open symbols in Fig. 3.1. Tomczak et al. (2014) estimate stellar mass densities for 6 redshift bins between redshift 0.5 and 2.5 from the ZFOURGE survey. The mass completeness limits for this survey are below $10^{9.5}$ $M_\odot$ at all redshifts, probing lower masses than other data sets at the same redshifts. For this data set no extrapolation is carried out in estimating ρ_*. At redshift 2 galaxies with masses below 10^9 $M_\odot$ in the simulations contribute 12% to the stellar mass density and their contribution decreases with decreasing redshift due to the flattening of the GSMF

Table 3.1: Mass completeness limit at redshifts 0.2 to 4 for GSMF observations of Ilbert et al. (2013) and Muzzin et al. (2013).

Redshift	**Ilbert et al. (2013)**	**Muzzin et al. (2013)**
	$\log_{10}(M_*)$ $[M_\odot]$	$\log_{10}(M_*)$ $[M_\odot]$
0.2 - 0.5	7.93	8.37
0.5 - 0.8[a]	8.70	8.92
0.8 - 1.1	9.13	-
1.1 - 1.5	9.42	9.48
1.5 - 2.0	9.67	10.03
2.0 - 2.5	10.04	10.54
2.5 - 3.0	10.24	10.76
3.0 - 4.0	10.27	10.94

[a]Muzzin et al. (2013) use redshift ranges 0.5 to 1.0 and 1.0 to 1.5.

(see Section 3.2.2).

At redshifts below two the various observational measurements show agreement on the total stellar mass density to better than 0.1 dex. From redshift 2 to 4 the disagreement increases to up to 0.4 dex primarily as a result of applying different extrapolations to correct for the mass completeness. At redshifts above four only the UV observations of González et al. (2011) are shown. Note that these do not include corrections for nebular emission lines and may overestimate the masses (e.g. Smit et al., 2014). We therefore show these estimated ρ_* as upper limits.

The solid black line in each panel of Fig. 3.1 shows the build up of ρ_* in the simulation. The log scale used in the upper panel emphasises the rapid fractional increase at high redshift, while the linear scale in the lower panel emphasises the slow growth at late times. There is a rapid growth in ρ_* from the early universe until 8 Gyr ago, around redshift 1, by which point 70% of the present day stellar mass has formed. The remaining 30% forms in the 8 Gyr from redshift 1 to 0. We find that 50% of the present day stellar mass was in place 9.75 Gyr ago, by redshift

1.6.

The simulation is in good agreement with the observed growth of stellar mass across the whole of cosmic time, falling within the error bars of the observational data sets. We find that 3.5% of the baryons are in stars at redshift zero, which is consistent with the 3.5% and 4% reported by Li and White (2009) and Baldry et al. (2012), respectively.

However, it should be noted that observed stellar mass densities are determined by integrating the GSMF, thereby excluding stellar mass associated with ICL. To carry out a fairer comparison, we apply a 3D 30 pkpc aperture to the simulated galaxies to mimic a Petrosian aperture, as applied to many observations (see *Chapter 2* and S14). The aperture masses more accurately represent the stellar light that can be detected in observations. The result of the aperture correction is shown as a solid blue line in both panels [1].

In this more realistic comparison of the model to observations, which excludes the ICL, we find that from high redshift to redshift 2 there is little difference between the total ρ_* and the aperture stellar mass density associated with galaxies. At these high redshifts the simulation curve lies within the scatter of the total stellar mass density estimates from the observations of González et al. (2011) (inverted triangles) and Ilbert et al. (2013) (open diamonds), although the simulation data is above the estimates of Muzzin et al. (2013) (open circles) above redshift 2. Between redshifts 2 and 0.1 the simulation data lies within the error bars from different observational estimates, although it is on the lower side of all observed values below redshift 0.9. At redshift 0.1, where ρ_* can be determined most accurately from observations, the simulation falls below the observations by less than 0.1 dex, or 20%, Although recall that the GSMF at redshift 0.1 was used to calibrate the feedback parameter, so this agreement at redshift 0.1 is by construction. We will return to the source of this deficit in stellar mass at low redshift when studying the shape of the GSMF.

[1] Note the mass in the simulation associated with the ICL resides in the largest halos, as will be shown in *Chapter 5*.

Returning to the agreement between redshifts 2 and 4, above redshift 2 the stellar mass density estimated from observations requires extrapolation below the mass completeness limit of the survey. To compare the simulation with the stellar mass density that is observed, without extrapolation, the red and green lines in the top panel show ρ_* from the simulation after applying the mass completeness limits of Ilbert et al. (2013) and Muzzin et al. (2013), respectively. The mass completeness limits applied are listed in Table 3.1. The red and green lines should be compared to the filled red diamonds and filled green circles, respectively, showing ρ_* from the observed galaxies without extrapolating below the mass completeness limit. Note that 30 pkpc apertures are still applied to the simulated galaxies for this comparison. When comparing with Ilbert et al. (2013), we find agreement within the observational error bars from redshifts 0.2 to 4. However, Muzzin et al. (2013) find more stellar mass than the simulation between redshifts 1.5 and 4. This can be understood by noting that the estimated mass completeness limit of Muzzin et al. (2013) is higher than that of Ilbert et al. (2013) (although both groups use the same survey data), resulting in only the most massive objects at a given redshift being detected. These objects are not sufficiently massive in the simulation when compared with the inferred GSMF from observations (without accounting for random or systematic mass errors), as will be shown in Section 3.2.2.

3.2.2 The evolution of the galaxy stellar mass function

The evolution of the stellar mass density of the Universe provides a good overview of the growth of stellar mass in the simulation. However, it does not allow us to compare the breakdown of the galaxy population by mass with observational data. This is important, as it is possible to have a realistic ρ_* as a function of time, but for stars to be distributed in a fashion that is inconsistent with the observed GSMF. We now carry out a full comparison of the GSMFs in the simulation with those inferred from observations at different epochs.

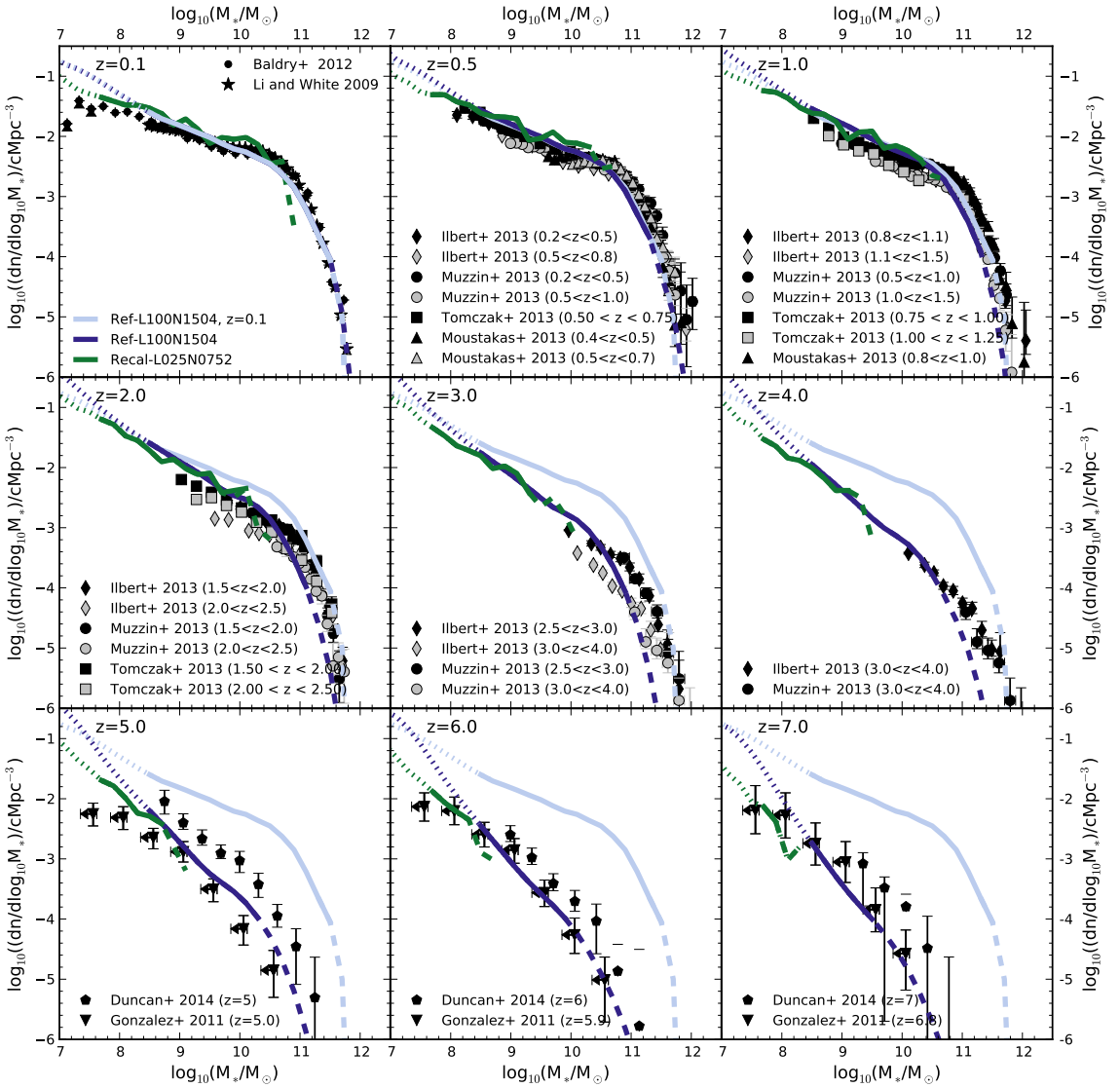

Figure 3.2: The galaxy stellar mass function at the redshifts shown in the upper left of each panel for simulation Ref-L100N1504 and Recal-L025N0752, in blue and green respectively. When the stellar mass falls below the mass of 100 baryonic particles curves are dotted, when there are fewer than 10 galaxies in a stellar mass bin curves are dashed. The redshift 0.1 GSMF is reproduced in each panel as a light blue curve, to highlight the evolution. The data points show observations as indicated in the legends. The black points represent the observational redshift bin below the simulation redshift, while the grey curves are from the redshift bin above the simulation snapshot.

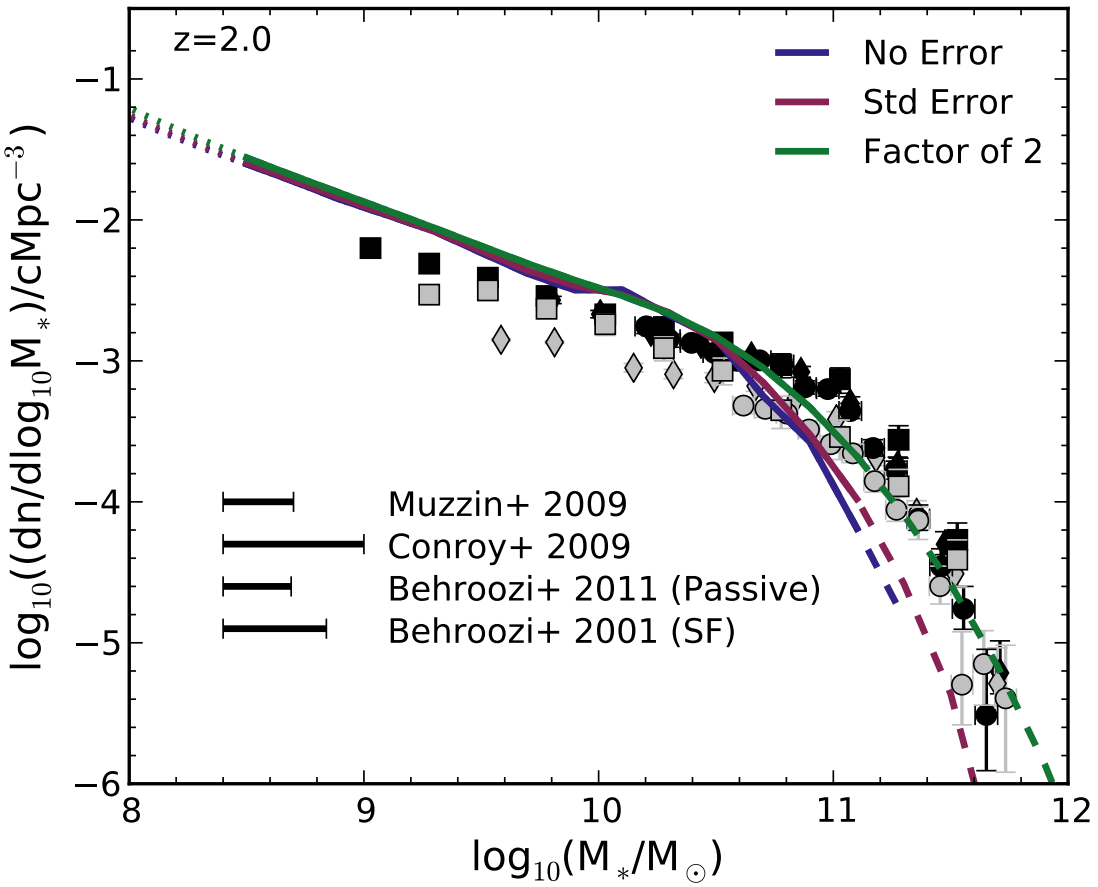

Figure 3.3: The simulated GSMF at redshift two from EAGLE without random mass errors (red), convolved with the stellar mass error of Behroozi et al. (2013), used in Fig. 3.2, (blue) and with random errors of a factor two (green). The random errors have a significant effect on the shape of the massive end of the GSMF, transforming the simulation from mildly discrepant with the observational data to being in excellent agreement with data. The Gaussian convolution with a stellar mass error is motivated by the random errors associated with the Malmquist bias. The horizontal black lines in the lower left of the figure indicate the estimated magnitudes of systematic errors in stellar masses according to Muzzin et al. (2009), Conroy et al. (2009) and Behroozi et al. (2013) at redshift two. Systematic errors are expected to maintain the shape of the GSMF but would shift it horizontally. Within the estimated level of uncertainty in observations, the simulation shows agreement with observations of the GSMF, including the location of the break, although the low-mass slope may be slightly too steep.

The shape of the GSMF is often described by a Schechter (1976) function,

$$\Phi(M)dM = \Phi^* \left(\frac{M}{M_{\rm C}}\right)^{\alpha} e^{-\frac{M}{M_{\rm C}}} dM, \tag{3.1}$$

where M_C is the characteristic mass or "knee", Φ^* is the normalisation and α is the power law slope for $M \ll M_C$. We will refer to α and M_C throughout this comparison. In *Chapter 3.2.3* we fit the simulation GSMFs with Schechter functions to provide a simple way of reproducing the data.

In Fig. 3.2 we compare the GSMF to the observational data sets that were presented in Fig. 3.1 in terms of the total stellar mass density. The GSMFs from these different observations are consistent with each other within their estimated error bars up to redshift two. Between redshifts 0 and 1 there is little evolution seen in the observational data, all show a reasonably flat low-mass slope and a normalisation that varies by less than 0.2 dex at 10^{10} M$_\odot$. From redshift 1 to 2 there is a steepening of the slope at galaxy masses below 10^{10} M$_\odot$ and a drop in normalisation of ~ 0.4 dex. The drop in normalisation appears to continue above redshift two, although the observations do not probe below 10^{10} M$_\odot$ at redshifts two to four.

Observational data at redshifts 5, 6 and 7 from González et al. (2011) and Duncan et al. (2014), based on rest-frame UV observations, are shown in the bottom three panels of Fig. 3.2. There is no clear break in the GSMF at these high redshifts, so it is not clear that the distribution is described by a Schechter function in either data set. Both data sets show similar slopes above 10^8 M$_\odot$. At low masses, below 10^8 M$_\odot$, the data set of González et al. (2011) shows a flattening in the slope at all redshifts shown. These low masses are not probed by Duncan et al. (2014). At redshift 5 the data sets are offset in mass by up to 0.8 dex. The offset reduces to ~ 0.2 dex by redshift 7. A comparison of these data sets provides an impression of the systematic errors in determining the GSMF from observations.

We compare the observations to the GSMFs predicted by Ref-L100N1504 between redshift 0.1 and 7, spanning 13 Gyr. The GSMF for Ref-L100N1504 is shown as a blue curve in Fig. 3.2. Stellar mass bins of 0.2 dex are used, when there are fewer than 10 galaxies per bin the curves are dashed. As an indication of the resolution limit, curves are dotted when the stellar mass corresponds to a mass of less

than 100 baryonic particles. For comparison, the GSMF at redshift 0.1 is repeated in all panels in light blue. To facilitate a comparison with observational data, the GSMF from Ref-L100N1504 is convolved with an estimated observational stellar mass error to account for random errors, where more low-mass galaxies are scattered to higher masses than vice versa. We use the error from Behroozi et al. (2013), $\sigma(z) = \sigma_0 + \sigma_z z$ dex, where $\sigma_0 = 0.07$ and $\sigma_z = 0.04$ for the random error. This gives a fractional error in the galaxy stellar mass of 18% at redshift 0.1 and 40% at redshift 2. Note this error does not account for systematic uncertainties that arise when inferring the stellar mass from observations, which could range from 0.1 to 0.6 dex depending on redshift (see Section 3.2.2).

Recall that the observed GSMF at redshift 0.1 was used to calibrate the free parameters of the simulation. At this redshift, the simulation reproduces the reasonably flat slope of the observed GSMF below $10^{10.5}$ $M_\odot$, with an exponential turnover at higher masses, between $10^{10.5}$ $M_\odot$ and 10^{11} $M_\odot$. Overall, we find agreement within 0.2 dex over the mass range from 2×10^8 $M_\odot$ to over 10^{11} $M_\odot$ and a very similar shape for the simulated and observed GSMF. In our implementation, the interplay between the subgrid stellar and AGN feedback models at the knee of the GSMF, at galaxy masses of around $10^{10.5}$ $M_\odot$, results in a slight underabundance of galaxies relative to observations. As the stellar mass contained in this mass range dominates the stellar mass density of the Universe, this small offset accounts for the shortfall of stellar mass at the 20% level seen at redshift zero in ρ_* in Fig. 3.1 (blue curve).

In the simulation, there is almost no evolution in the GSMF from redshift zero to one, apart from a small decrease of 0.2 dex in galaxy masses at the very high-mass end. This can be seen by comparing the blue and light blue lines in the top panels, where the light blue line repeats the redshift 0.1 GSMF. A similar minimal evolution was reported based on the observational data of Moustakas et al. (2013)(triangles) from redshift 0 to 1, and is also seen in the other data sets shown.

From redshift one to two the simulation predicts strong evolution in the GSMF, in terms of its normalisation, low-mass slope and the location of the break. Between

these redshifts, spanning just 2.6 Gyr in time, the stellar mass density almost doubles, from 0.75 to 1.4 $\times 10^8$ $M_\odot$cMpc^{-3}, and the GSMF evolves significantly. In the simulation, the normalisation of the GSMF for galaxies of mass of $10^{9.5}$ $M_\odot$ drops by 0.3 dex, while the most massive galaxies have decreased in stellar mass by ~ 0.5 dex relative to redshift 0.1. From redshift two to four the normalisation continues to drop and the mass corresponding to the break in the GSMF continues to decrease.

Although the trend of a decrease in normalisation of the GSMF between redshift one and two is qualitatively consistent with what is seen in the observations, the normalisation at redshift two at $10^{9.5}$ $M_\odot$ is too high in the simulation by around 0.2 dex. There is also a suggestion that the normalisation of the GSMF in the simulation is too high at redshift three, although observations do not probe below 10^{10} $M_\odot$ at this redshift. It is therefore difficult to draw a strong conclusion from this comparison without extrapolating the observational data. At redshift two there is also an offset at the massive end of the GSMF. The exponential break occurs at a mass that is around 0.2 dex lower than observed. However, the number of objects per bin in the simulation at redshift two above 10^{11} $M_\odot$ falls below 10 providing a poor statistical sample of the massive galaxy population. Increasing the box size may systematically boost the abundance of rare objects, such as that of galaxies above 10^{11} $M_\odot$ at redshift two and above. The break is also particularly sensitive to any errors in the stellar mass estimates, a point we will return to below.

Comparing the simulated GSMF to observations at redshifts 5, 6 and 7, we find a similar shape to the observational data. The simulation agrees within the observation error bars with González et al. (2011), however it is offset in stellar mass from Duncan et al. (2014). No break in the GSMF is visible, neither in the simulation nor in the observations, at these high redshifts over the mass ranges considered here. Hence, for redshifts above 5 a Schechter fit may not be an appropriate description of the data.

Galaxy stellar mass errors

When comparing the simulation to observations, it is important to consider the role of stellar mass errors, both random and systematic. We begin by considering the random errors. In Fig. 3.3 the GSMF from Ref-L100N1504 is plotted at redshift two assuming no stellar mass error (red), a random mass error of $0.07 + 0.04$z (Behroozi et al., 2013) as in Fig. 3.2 (blue), resulting in an error of 40% in galaxy stellar mass at redshift two, and a mass error of a factor of two (green), i.e. 100%. Where the GSMF is reasonably flat, i.e. at masses below $10^{10.5}$ $M_{\odot}$, the impact of random uncertainty is minimal. However, above this mass the shape of the GSMF depends strongly on the random stellar mass errors in the observations, because more low-mass galaxies are scattered to high masses than vice versa. If we increase the random errors, the exponential break becomes less sharp and the simulation agrees better with the observations.

There are also systematic errors to consider in the determination of stellar masses from observed flux or spectra. Fitting the spectral energy distribution (SED) of a galaxy is sensitive to the choice of stellar population synthesis (SPS) model, e.g. due to the uncertainty in how to treat TP-AGB stars, the choice of dust model and the modelling of the star formation histories (e.g. Mitchell et al., 2013). Systematic variations in the stellar IMF would result in additional uncertainties, which are not considered here. The systematic uncertainties from SED modelling increase with redshift. At redshift zero Taylor et al. (2011) quote ~ 0.1 dex (1σ) errors for GAMA data. At redshift two the estimated systematic error on stellar masses ranges from 0.3 dex (Muzzin et al., 2009) to 0.6 dex (Conroy et al., 2009), based on uncertainties in SPS models, dust and metallicities. Fig. 3.3 gives an impression of the size of these systematic errors by plotting values from Muzzin et al. (2009), Conroy et al. (2009) and Behroozi et al. (2013) in the bottom left corner. The Behroozi et al. (2013) estimate is divided into star forming and passive galaxies due to the reduced sensitivity of passive galaxies to the assumed form of the star formation history. The systematic stellar mass errors are expected to shift the GSMF along the stellar mass axis. Considering the extent of the systematic uncertainties, we find the GSMF

from EAGLE to be consistent with the observational data, although the low-mass slope may be slightly too steep. The observed evolutionary trends in the normalisation and break are reproduced by the simulation, suggesting that the simulation is reasonably representative of the observed Universe.

Numerical convergence

Having found reasonable agreement between the evolution in the Ref-L100N1504 simulation and the observations, it is important to ask if the results are sensitive to numerical resolution. We consider only weak convergence tests, examining the ability of the simulation to reproduce the observed evolution after recalibrating the high-resolution simulation to the same conditions (namely the redshift 0.1 GSMF) as used for the standard resolution simulation. In Fig. 3.2 the high-resolution model, Recal-L025N0752, is shown in green. As for Ref-L100N1504 stellar mass bins of 0.2 dex are used, when there are fewer than 10 galaxies per bin the curves are dashed, and the curves are dotted when the stellar mass corresponds to less than 100 baryonic particles, which occurs at 8 times lower mass than for Ref-L100N1504.

The 25 cMpc box is too small to represent the break in the GSMF accurately. To avoid box size issues, we do not consider the GSMF when there are fewer than 10 galaxies per bin, i.e. where the green curve is dashed. The 25 cMpc box also shows more fluctuations, due to poorer sampling from the reduced number of objects compared to Ref-L100N1504. At masses below 10^8 $M_\odot$, when there are fewer than 100 star particles per galaxies in the Ref-L100N1504 simulation (blue dotted curve), the slope of the high-resolution simulation is flatter than that of Ref-L100N1504. Where the solid part of the blue and green curves overlap, there is excellent agreement, better than 0.1 dex, between the intermediate- and high-resolution simulations across all redshifts. Overall, this amounts to good (weak) numerical convergence in the simulation across all redshifts that can be probed, based on box size limitations.

In summary, we have found the stellar mass density in the simulation to be close

to the estimated values from observations, the maximum offset being $\sim 20\%$ level due to undershooting the GSMF around M_C. The observed evolutionary trends, in the shape and normalisation of the GSMF, from redshift 0.1 to 7 are reproduced, although the evolution in the normalisation is not sufficiently strong from redshift 1 to 2, with an offset in normalisation at redshift 2 of ~ 0.2 dex. The break in the GSMF occurs at too low a mass in the simulation relative to the observations at redshifts 2 to 4. However, the box size limits the number of objects produced in the simulation and we have shown that stellar mass errors play a significant role in defining the observed break of the GSMF. As a result of these uncertainties, the simulation may not be inconsistent with observations.

3.2.3 Schechter function fits

To provide a simple way of reproducing the EAGLE GSMFs and to quantify the trends seen in the evolution of the normalisation and the exponential break, we have fit the EAGLE GSMFs with Schechter functions. We fit the GSMFs of Ref-L100N1504 from redshift 0.1 to 4 that were shown in Fig. 3.2 (blue curves) with single Schechter functions (eq. 3.1) and double Schechter functions,

$$\Phi(M)dM = \left[\Phi_1^* \left(\frac{M}{M_C}\right)^{\alpha_1} + \Phi_2^* \left(\frac{M}{M_C}\right)^{\alpha_2}\right] e^{-M/M_C} dM, \tag{3.2}$$

which is the sum of two Schechter functions with the same characteristic mass, M_C, but different normalisations, Φ_1^* and Φ_2^* and different low mass slopes, α_1 and α_2. Double Schechter fits are increasingly used in observational studies fitting the GSMF. We use least squares fitting with bins of width 0.2 dex in stellar mass. Bins are weighted by their Poisson error, thereby down-weighting the poorly sampled galaxies in the most massive stellar mass bins. The fits over the mass range 10^8 M$_\odot$ to 10^{12} M$_\odot$ are presented in Table 3.2.

Table 3.2: Single (eq. 3.1) and double (eq. 3.2) Schechter function parameters for the EAGLE Ref-L100N1504 GSMFs presented in Fig. 3.2, fitting over the mass range 10^8 to 10^{12} $M_\odot$. The Schechter function parameters provide a simple way of reproducing the GSMFs from the EAGLE simulation over the range where the fitting is carried out.

Redshift	$\mathbf{log_{10}(M_C)}$	Φ^*	α		
	$[M_\odot]$	$[\,10^{-3}$ cMpc$^{-3}]$			
0.1	11.14 ± 0.09	0.84 ± 0.13	-1.43 ± 0.01	-	-
0.5	11.11 ± 0.08	0.84 ± 0.12	-1.45 ± 0.01	-	-
1.0	11.06 ± 0.08	0.74 ± 0.10	-1.48 ± 0.01	-	-
2.0	10.91 ± 0.08	0.45 ± 0.07	-1.57 ± 0.01	-	-
3.0	10.78 ± 0.11	0.22 ± 0.05	-1.66 ± 0.01	-	-
4.0	10.60 ± 0.15	0.12 ± 0.04	-1.74 ± 0.02	-	-
Redshift	$\mathbf{log_{10}(M_C)}$	Φ_1^*	α_1	Φ_2^*	α_2
	$[M_\odot]$	$[\,10^{-3}$ cMpc$^{-3}]$		$[\,10^{-3}$ cMpc$^{-3}]$	
0.1	10.95 ± 0.03	1.45 ± 0.12	-1.31 ± 0.03	0.00 ± 0.00	-2.22 ± 0.22
0.5	10.88 ± 0.04	1.61 ± 0.13	-1.24 ± 0.08	0.08 ± 0.10	-1.79 ± 0.15
1.0	11.13 ± 0.21	0.00 ± 0.36	$1.47 \pm 2.8 \times 10^5$	0.68 ± 0.21	-1.48 ± 0.01
2.0	10.44 ± 0.08	1.06 ± 0.24	-0.25 ± 0.29	0.80 ± 0.14	-1.58 ± 0.02
3.0	10.19 ± 0.09	0.63 ± 0.13	0.23 ± 0.37	0.61 ± 0.13	-1.64 ± 0.02
4.0	10.00 ± 0.11	0.24 ± 0.08	0.89 ± 0.58	0.43 ± 0.12	-1.69 ± 0.03

To understand the dependence of the Schechter function parameters on the fitted mass range, we applied our fitting routine over three mass ranges, from 10^8, 10^9 and 10^{10} to 10^{12} $M_\odot$. Fig. 3.4 shows the evolution of the Schechter function parameters M_C, Φ^* and α for the single Schechter function fits. For the single Schechter fit M_C drops over the redshift range zero to four for all mass ranges. However, the extent of the decrease depends on the fitting range. For example, there is a decrease of 0.5 dex when fitting above 10^8 $M_\odot$ compared to a 0.3 dex decrease for fits above 10^9 and 10^{10} $M_\odot$. Φ^* also decreases with redshift for all fits to the GSMFs, with the values of Φ^* at a given redshift being lower for the fit above 10^8 $M_\odot$ than 10^9 and 10^{10} $M_\odot$. The opposite changes in M_C and Φ^* for the different mass ranges highlight the degeneracy between these two parameters.

The α parameter becomes more negative with increasing redshift for fits above 10^8 and 10^9 $M_\odot$, showing that the low mass slope steepens with redshift. However, different behaviour is seen for fits above 10^{10} $M_\odot$ where α increases to redshift 1, then decreases. This is not unexpected given that fitting for stellar masses above 10^{10} $M_\odot$ does not provide enough information to constrain the slope for masses $\ll M_C$.

We find larger differences between different mass ranges, and in particular larger error bars, when fitting double Schechter functions than what is presented for single Schechter functions in Fig. 3.4. Due to the sensitivity of the Schechter fitting to the mass range over which it is done, it is very difficult to compare the fitting parameters directly to observations. This is especially true when we consider the evolving mass completeness limit for observations. Any trends with redshift could easily be a result of the changing mass range. The degeneracy between M_C and Φ^* also makes a comparison of Schechter parameters difficult to interpret. The final issue with directly comparing Schechter parameters from observations and/or simulations is the sensitivity of the break in the Schechter function to stellar mass errors, as shown in Section 3.2.2. As a result of these issues, we choose not to compare the Schechter function parameters to those determined from observations and consider the comparison of the data presented in Fig. 3.2, from which Schechter pa-

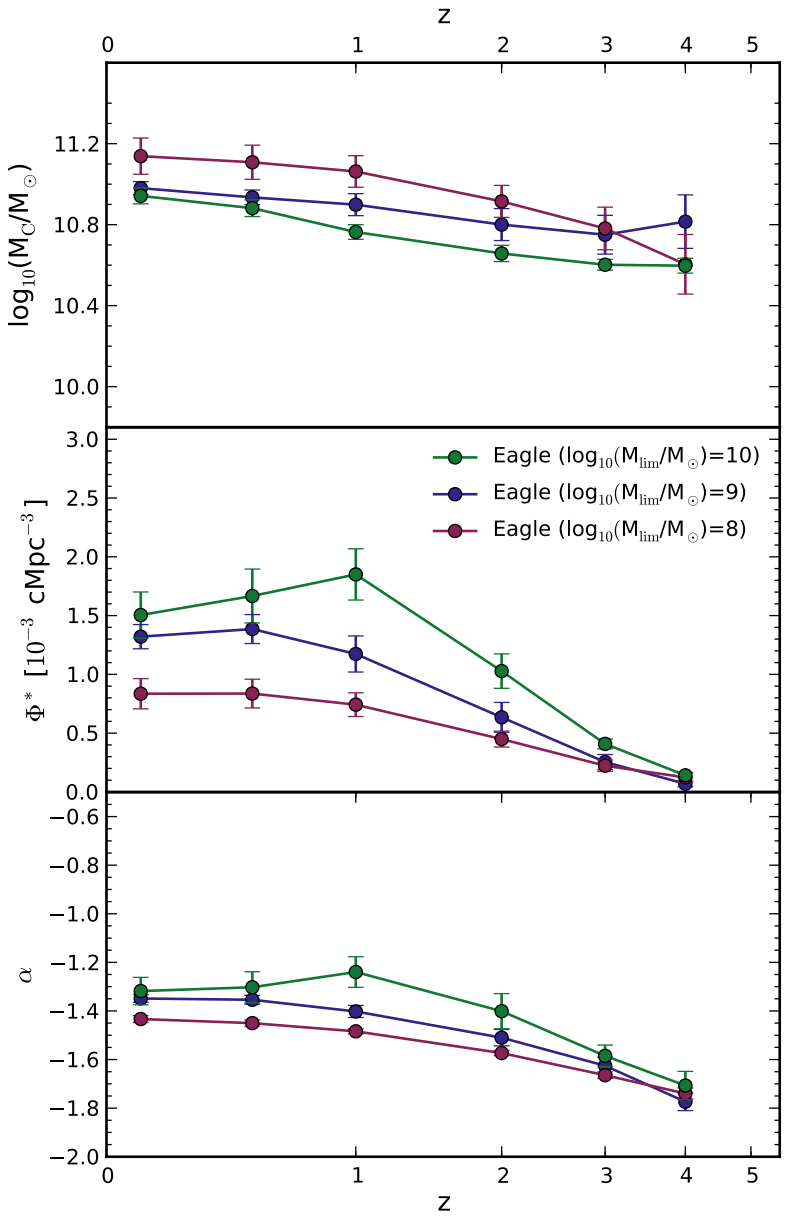

Figure 3.4: The Schechter function parameters, M_C, Φ^* and α for the EAGLE GSMFs (as shown in Fig. 3.2) as a function of redshift. These panels show single Schechter function parameters fit from 10^8, 10^9 and 10^{10} $M_\odot$ to 10^{12} $M_\odot$ in red, blue and green respectively, with 1-σ error bars from the fitting. The Schechter function fitting is sensitive to the mass range over which the fitting is done and the values for both M_C and Φ^* are degenerate. For double Schechter function parameters the agreement between different stellar mass ranges is worse due to the increased freedom (not shown).

rameters are derived, to be sufficient to determine the agreement between observations and simulations. However, the Schechter function parameters do provide a simple way of representing the GSMFs from the Eagle simulation over the range where the fitting is carried out.

3.3 Evolution of star formation rates

3.3.1 The cosmic star formation rate density

The star formation rate density ($\rho_{\rm SFR}$) as a function of redshift is plotted in Fig. 3.5 on a log scale. For comparison, observations from Gilbank et al. (2010a) [Hα], Rodighiero et al. (2010) [24μm], Karim et al. (2011) [Radio], Cucciati et al. (2012) [FUV], Bouwens et al. (2012) [UV] , Robertson et al. (2013) [UV] and Burgarella et al. (2013) [FUV + FIR] are shown across a range of redshifts. This compilation of data covers a number of SFR tracers, providing an overview of $\rho_{\rm SFR}$ estimates from the literature, as well as an indication of the range of scatter and uncertainty arising from different methods. There are differences in the measured $\rho_{\rm SFR}$ of around 0.2 dex at redshifts less than two, while the estimated $\rho_{\rm SFR}$ include error bars of about ± 0.15 dex, with larger error bars above redshift two. Although there are uncertainties in estimating $\rho_{\rm SFR}$, this quantity provides a simple way to view the global evolution of SFRs and the trends within the observations and simulation.

At high redshift the simulation data (solid black line) shows an increase in $\rho_{\rm SFR}$ with the age of the Universe. This increase peaks around redshift two, and is followed by a decline of almost an order of magnitude to redshift zero. The simulation reproduces the shape of the observed $\rho_{\rm SFR}$ very well but is low by a small offset of 0.2 dex.

The integral of $\rho_{\rm SFR}$ gives the stellar mass density of stars formed. However, note that, $\int \rho_{\rm SFR}(t)dt \neq \rho_*$, where ρ_* is shown by the black curve in Fig. 3.1, because the stellar mass density, ρ_*, does not include stellar mass lost due to stellar evolution. We find that $\int \rho_{\rm SFR}(t)dt$ is around 40% higher than ρ_* at redshift zero.

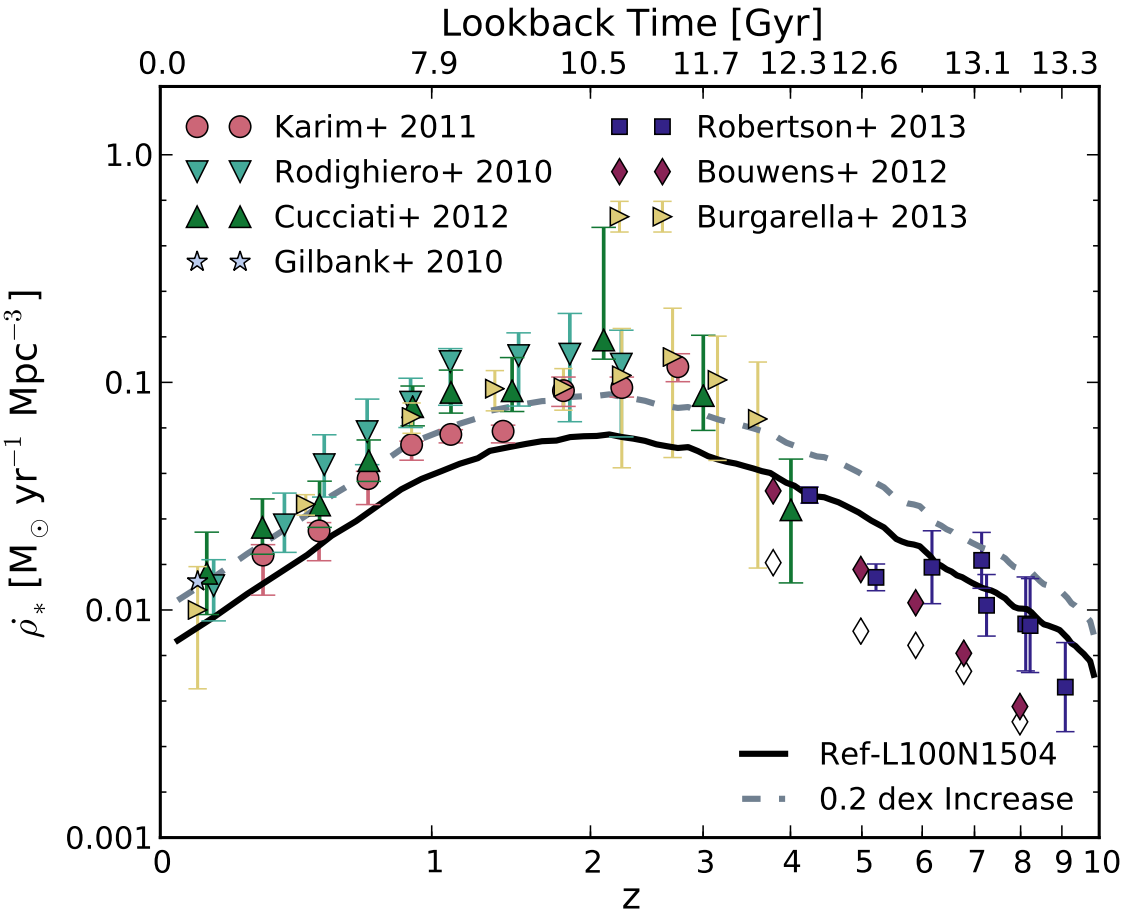

Figure 3.5: Evolution of the cosmic star formation rate density in the EAGLE simulation Ref-L100N1504. The simulation data is plotted as a solid black curve. The grey dashed curve increases the simulation data by 0.2 dex, which corresponds to the offset in ρ_* relative to the observations at redshift zero after accounting for stellar mass loss. Data points correspond to observation measurements of the cosmic star formation rate density from a variety of star formation tracers from redshift zero to four as indicated in the legend. Open symbols from Bouwens et al. (2012) exclude a dust correction to the SFRs, which indicates the uncertainty in the measurement. The predicted shape of $\rho_{\rm SFR}$(z) is consistent with the observations, but there is an offset in normalisation below redshift three of $\sim$0.2 dex.

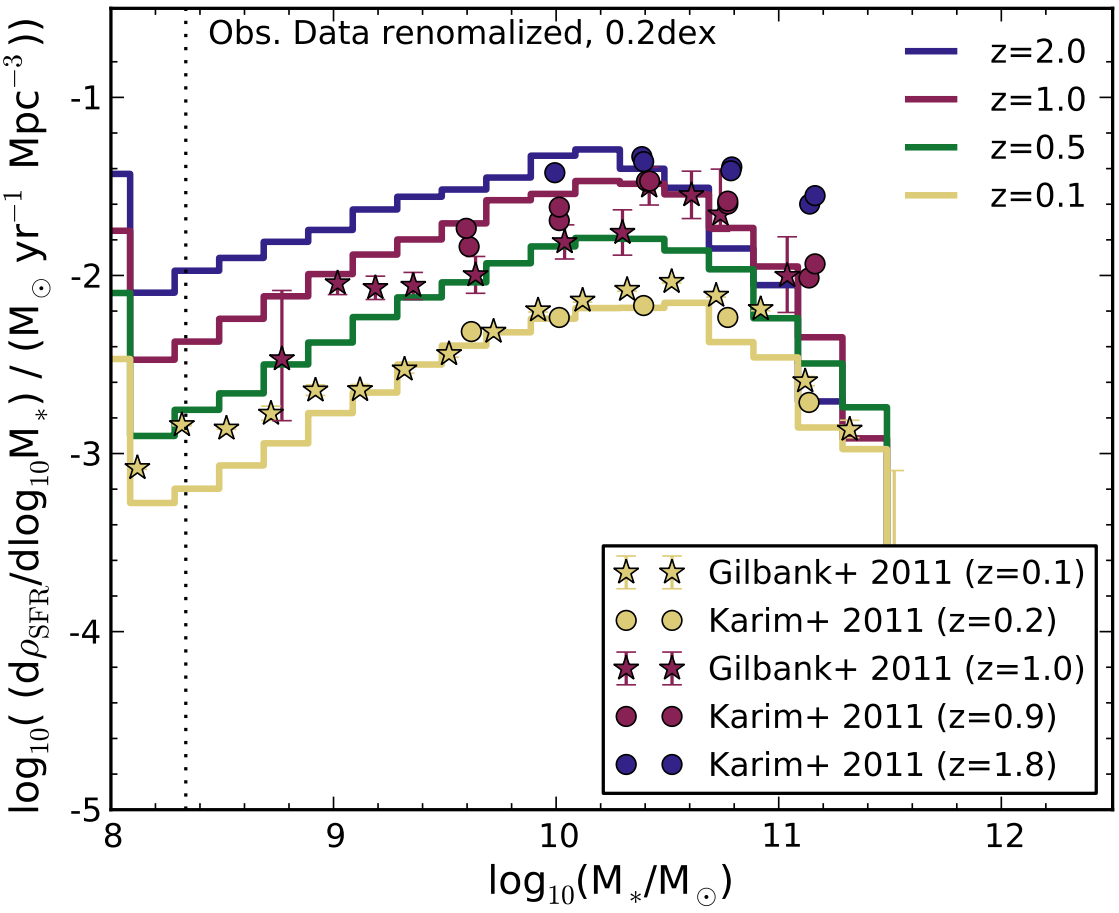

Figure 3.6: The star formation rate density at different redshifts as a function of stellar mass from EAGLE simulation Ref-L100N1504. Simulation data at redshifts 0.1, 0.5, 1 and 2 are shown as green, yellow, red and blue lines respectively. The vertical dotted line denotes the mass of a galaxy with 100 baryonic particles, as a resolution guide. Observational data from Gilbank et al. (2010a), Gilbank et al. (2010b) and Karim et al. (2011) are shown, coloured by redshift. The observational data has been scaled down by 0.2 dex, to make an easier comparison with the trends seen in the simulation. We find that $\rho_{\rm SFR}$ is dominated by galaxies with mass $M_* \sim 10^{10.5}$ $M_\odot$ at all epochs, with an increase in the normalisation of $\rho_{\rm SFR}$ across all masses at higher redshifts. The shape of $\rho_{\rm SFR}$ with stellar mass and the trend with redshift are similar in the simulation and observations, although the simulation may underestimate the relative contribution of very massive galaxies at redshift two.

When we compare ρ_* to observations in Fig. 3.1, we found a shortfall of 0.1 dex at redshift zero. Accounting for stellar mass loss of 40%, the star formation required to account for the shortfall in ρ_* amounts to a ~ 0.2 dex boost in $\rho_{\rm SFR}$. In other words, a boost by 0.2 dex in $\rho_{\rm SFR}$ would increase ρ_* from the simulation by around 20%, improving the agreement seen in Fig. 3.1 between the simulation and observations. (A boost at all redshifts assumes the missing stellar mass follows the average star formation history.) The dashed grey line shows the effect of a constant boost by 0.2 dex, which leads to general agreement with the observations. Reconciling the offset from observational values of $\rho_{\rm SFR}$ with the offset from ρ_* shows the observational data for these two quantities to be consistent. The integral of the observed $\rho_{\rm SFR}$ produces the observed ρ_* after accounting for stellar mass loss.

In Fig. 3.6 $\rho_{\rm SFR}$ is shown as a function of galaxy stellar mass from redshift zero to two and compared to observational estimates. We show data from Gilbank et al. (2011) (stars) and Karim et al. (2011) (circles) coloured by redshift as for the simulation data. The observed $\rho_{\rm SFR}$ are decreased by a constant factor of 0.2 dex to account for the offset of this amount found in the total $\rho_{\rm SFR}$ in Fig. 3.5. This renormalisation facilitates a comparison of the trends between observations and simulation.

The simulation reproduces the shape of the observed $\rho_{\rm SFR}$ values as a function of stellar mass. The distribution peaks in all data sets up to redshift 2 at stellar masses $\rm M_* \sim 10^{10.5}\ M_\odot$. For all masses the normalisation of $\rho_{\rm SFR}$ increases from redshift zero to two by almost a decade. This trend is seen in both the observations and the simulation.

There is a potential discrepancy with observational data at redshift two, where the peak in $\rho_{\rm SFR}$ moves to slightly lower mass objects in the simulation, of $10^{10.2}\ \rm M_\odot$, but remains at $10^{10.5}\ \rm M_\odot$ in the observations. This could be a result of the overly efficient cut off in star formation in massive galaxies in the simulation between redshifts two and four, as discussed in Section 3.2.2. Although it could equally result from observational data at high redshift pre-selecting highly star forming objects. Overall, however, the trends in the simulation are very similar to those

observed.

3.3.2 Specific star formation rates

Observationally, a well defined star forming sequence as a function of stellar mass has been found in the local Universe and up to a redshift of 3 (e.g Noeske et al., 2007; Karim et al., 2011). The star forming sequence is described by a relation of the form

$$\frac{\dot{M}_*}{M_*} = \beta \left(\frac{M_*}{10^{10} M_\odot} \right)^\gamma \tag{3.3}$$

where γ is the logarithmic slope, β is the normalisation and $\dot{M}_*/M_*$ is the SSFR. Observations indicate that γ is negative but close to zero, and is often assumed to be constant with stellar mass. Plotting the SSFR-M_* relation rather than SFR-M_* has the advantage of separating the star forming and passive populations, highlighting the bimodality of the population. As the observed SSFR is reasonably flat and does not vary strongly with stellar mass, a comparison of the slope between the simulation and observations is easier than comparing slopes of the SFR-M_* relation.

Fig. 3.7 shows the SSFR for star forming galaxies as a function of galaxy stellar mass at redshifts 0.1, 1 and 2. The observational data sets for the SSFRs we compare to at redshift 0.1 are from Gilbank et al. (2010a) (stars) and Bauer et al. (2013) (squares). These data sets show similar values for the normalisation and slope and a similar scatter above 10^9 $M_\odot$. Below 10^9 $M_\odot$ only Gilbank et al. (2010a) data is available. This data shows an increase in the SSFR with decreasing stellar mass below $10^{8.5}$ $M_\odot$. Rodighiero et al. (2010) (inverted triangles), Karim et al. (2011) (circles) and Gilbank et al. (2010b) (stars) are shown at higher redshifts. Comparing these data sets, Rodighiero et al. (2010) and Karim et al. (2011) have similar slopes and normalisation at redshifts one and two. However, the Gilbank et al. (2010b) data is substantially (0.8 dex) lower in normalisation over the mass ranges where it overlaps with Rodighiero et al. (2010) and Karim et al. (2011). The ROLES data used by Gilbank et al. (2010b) probes faint galaxies down to masses below

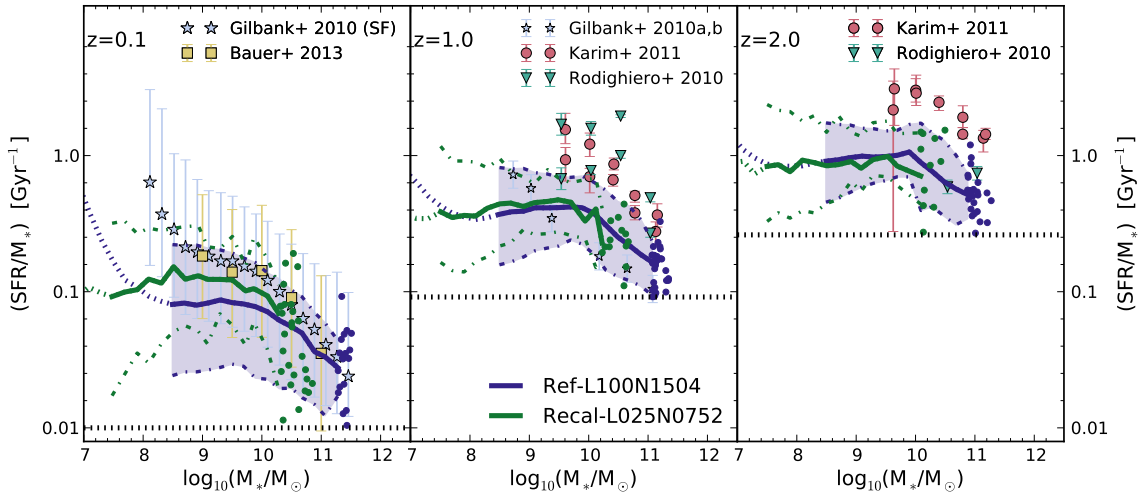

Figure 3.7: The SSFR, $\dot{M}_*/M_*$, as a function of galaxy stellar mass for Ref-L100N1504 and Recal-L025N0752 from left to right at redshifts 0.1, 1 and 2. The solid curves show the median relation for star forming galaxies, defined as those with a SSFR above the limit specified by the horizontal dotted line in each panel. The shaded region (dot dashed curves) encloses the $10^{\rm th}$ to $90^{\rm th}$ percentiles for Ref-L100N1504 (Recal-L025N0752). Where there are fewer than 10 galaxies per bin, individual data points are shown. Lines are dotted when the stellar mass falls below that corresponding to 10 star forming particles for the SSFR limit (dotted line) and the mass of 100 baryonic particles, to indicate that resolution effects may be important. At redshift 0.1 the observational data of Gilbank et al. (2010a) and Bauer et al. (2013) are shown as light blue stars and yellow squares, respectively. Error bars enclose the $10^{\rm th}$ to $90^{\rm th}$ percentiles. At higher redshift, data from Gilbank et al. (2010b), Karim et al. (2011) and Rodighiero et al. (2010) are shown as light blue stars, pink circles and turquoise inverted triangles respectively. The observed flat slope with stellar mass and the increase in normalisation with redshift are reproduced by the simulations, but the simulation is lower in normalisation by 0.2 to 0.4 dex, depending on redshift and the observational data set.

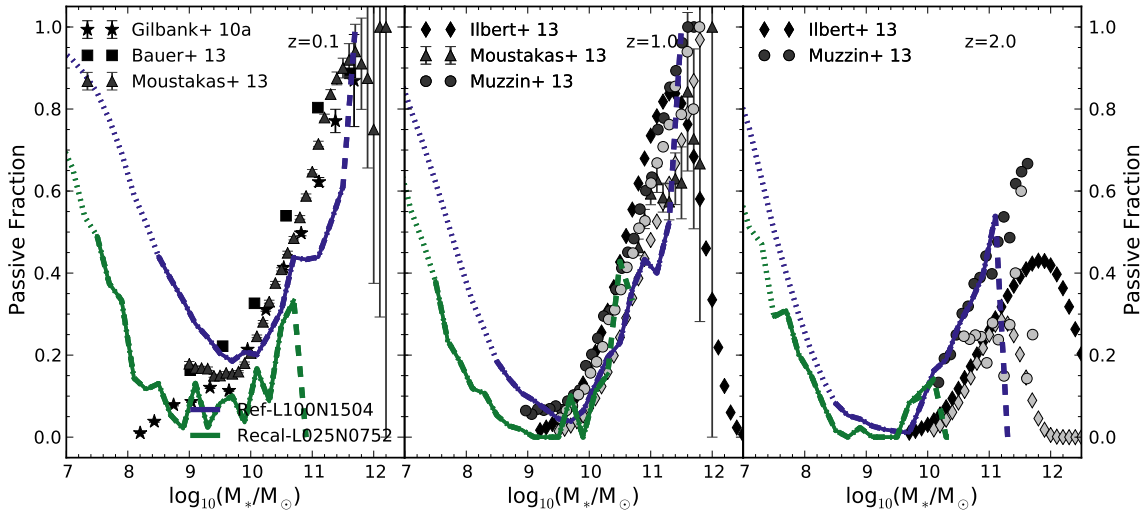

Figure 3.8: The passive fraction as a function of galaxy stellar mass for Ref-L100N1504 and Recal-L025N0752 in blue and green, respectively, where galaxies with a SSFR below the horizontal dotted lines in Fig. 3.7 are defined as passive. Where there are fewer than 10 galaxies per stellar mass bin curves are dashed. Lines are dotted when the stellar mass falls below that corresponding to 10 star forming particles for the SSFR limit (dotted line) and the mass of 100 baryonic particles. Data points show observations as indicated in the legends. Above 10^9 $M_\odot$, the simulated passive fractions show similar normalisation and slope with stellar mass to observations at all redshifts, with a small deficit of passive galaxies of around 15% in the mass range $10^{10.5}$ to $10^{11.5}$ $M_\odot$. The upturn at low masses, below 10^9 $M_\odot$ is a numerical artefact.

10^9 $M_\odot$, but this deep survey covers only a small area of sky. The resulting small number statistics of massive galaxies may be driving this offset in SSFR from the other observational data sets.

The median SSFRs for star forming galaxies from Ref-L100N1504 and Recal-L025N0752 are shown as blue and green curves, respectively. The horizontal dotted lines correspond to the SSFR cut (~ 1 dex below the observational data) used to separate star forming from passive galaxies. Stellar mass bins of 0.2 dex are used to compute the median, when the number of points per bin falls below 10 galaxies individual data points are shown. When the stellar mass falls below the mass that corresponds to a SSFR of 10 star forming particles, curves are dotted to indicate that resolution effects are expected to be severe. The $10^{\rm th}$ to $90^{\rm th}$ percentiles are enclosed by the dot dashed curves (and the shaded region for Ref-L100N1504).

At redshift 0.1 the SSFR in the simulations is reasonably independent of stellar mass (where well resolved) up to masses of 10^{10} $M_\odot$. Above this mass the SSFR decreases slowly with stellar mass. The simulations show a scatter of around 0.6 dex across the stellar mass range resolved by Ref-L100N1504. The normalisation of the Recal-L025N0752 simulation lies 0.2 dex above that of Ref-L100N1504, as was already shown in S14. At low masses, when there are fewer than 10 star forming particles per galaxy, there is an increase in SSFR with stellar mass in Ref-L100N1504. However, by comparing with Recal-L025N0752 we see that this is resolution driven , where a similar upturn is seen at 8 times lower stellar mass.

The trend with stellar mass above 10^9 $M_\odot$ is similar in the simulations and the observations. However, there is an offset in the normalisation from observations, where Recal-L025N0752 and Ref-L100N1504 are low by ~ 0.1 and 0.3 dex respectively. The increase in SSFR at a stellar mass of $10^{8.5}$ $M_\odot$ reported by Gilbank et al. (2010a) is not seen in the Recal-L025N0752 simulation, which has sufficient numerical resolution to compare to observations at these low masses. This could indicate that stellar feedback is too strong in low-mass galaxies, or perhaps that the observational data is not volume-complete due to the difficulty in detecting low-mass galaxies with low star formation rates owing to their low surface brightness (see

S14 for more discussion of the redshift 0.1 properties).

At higher redshifts the simulation SSFRs increase in normalisation, maintaining a flat slope below 10^{10} $M_{\odot}$, with a shallow negative slope above this stellar mass. At redshifts between one and two the Recal-L025N0752 and Ref-L100N1504 SSFRs lie within 0.1 dex of each other across the stellar mass range for which both are resolved. The increase in normalisation seen in the simulations reproduces the observed trend, although the offset in normalisation increases to 0.3 dex when comparing to the data sets of Rodighiero et al. (2010) and Karim et al. (2011). Relative to the Gilbank et al. (2010b) data at redshift one, the median SSFR from the simulation agrees to within around 0.1 dex. Comparing the slope of the SSFR-M_* relation of Gilbank et al. (2010b) to the simulations, the slope of the relation in the simulation is flatter below 10^{10} $M_{\odot}$, in agreement with the slopes of Karim et al. (2011) and Rodighiero et al. (2010).

Observationally the galaxy population exhibits a bimodal colour distribution, which may imply a bimodality in the SSFR. To study this bimodality in the simulation, we show in Fig. 3.8 the passive fraction of galaxies as a function of mass at redshifts 0.1, 1 and 2. In the simulation we define passive galaxies by a cut in SSFR that is an order of magnitude below the median observed SSFR (dotted horizontal line in Fig. 3.7). Varying this limit, while keeping it below the main star forming sequence has negligible impact on the recovered median SSFR, although it can increase or decrease the passive fractions by around 10%.

For comparison, passive fractions from Gilbank et al. (2010a), Bauer et al. (2013) and Moustakas et al. (2013) are shown at redshift 0.1 and from Moustakas et al. (2013), Muzzin et al. (2013) and Ilbert et al. (2013) at higher redshifts. For most observational data sets shown, the passive fraction is determined based on a colour or SSFR cut as applied in the published data sets. Gilbank et al. (2010a) provide tabulated stellar masses and SFRs for each galaxy and we therefore apply the same SSFR cut as we use for the simulation data. At redshift 0.1 the dependence of passive fraction on stellar mass is similar for all observational data sets. At redshift one, each observational data set shows the same trend, but there is a difference of

up to 20% in the passive fraction for $M_* \lesssim 10^{11.5}$ $M_\odot$ for different data sets, and higher above this mass. At redshift two agreement between data sets is poor.

The passive fraction from Ref-L100N1504 and Recal-L025N0752 are shown in blue and green, respectively. Where there are fewer than 10 galaxies per bin, the curves are dashed. Where the stellar mass is less than the maximum of 100 baryonic particles and 10 gas particles for the mass that corresponds to the SSFR cut, lines are dotted. As the SSFR cut evolves with redshift, this resolution guide evolves with redshift. At all redshifts and above $\sim 10^9$ $M_\odot$ both simulations show a similar increase in passive fraction with stellar mass, where there are more than 10 galaxies per bin. Below 10^9 $M_\odot$ for Recal-L025N0752 there is an increase in passive fraction with decreasing stellar mass. However this increase in passive fraction occurs at a stellar mass ~ 8 times lower in Recal-L025N0752, implying the upturn in the passive fraction is an artefact of the finite resolution of the simulation.

Over the resolved mass range, the passive fraction at redshift 0.1 follows a similar trend to the observational data, although there are too few passive galaxies between $10^{10.5}$ and $10^{11.5}$ $M_\odot$ by around 15%. At redshift 1 the passive fraction is lower than at redshift 0.1 in the simulation. This is consistent with what is seen in observational studies, although, there are again fewer passive galaxies in the range of $10^{10.5}$ to $10^{11.5}$ $M_\odot$ than observed. At redshift two there is a further drop in the passive fraction of galaxies, both in the simulation and the observations. Summarising, the passive fractions show the same trend as observations when galaxy masses and SFRs are resolved, although there are too few passive galaxies by $\sim 15\%$ in the stellar mass range $10^{10.5}$ to $10^{11.5}$ $M_\odot$.

3.3.3 Specific star formation rate evolution

To better study the evolution of the SSFR, we show in Fig. 3.9 the SSFR as a function of lookback time in three different stellar mass bins. Similar trends are found when considering other mass bins of 0.5 dex between $10^{8.5}$ and $10^{11.5}$ $M_\odot$. We compare the simulation data with the observations presented in Fig. 3.7, adding González et al. (2012) and Stark et al. (2013) at redshifts 4 and above.

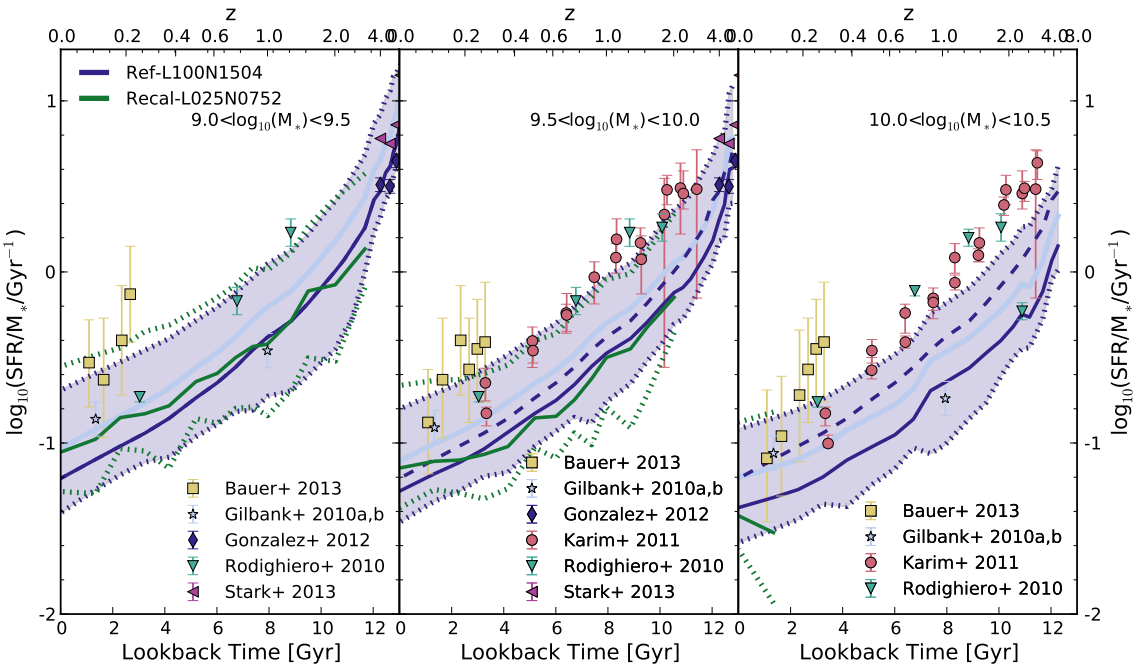

Figure 3.9: Evolution of the SSFR as a function of lookback time for stellar mass bins $10^{9.0}$ <M_* < $10^{9.5}$ $M_\odot$ (left), $10^{9.5}$ <M_* < $10^{10.0}$ $M_\odot$ (middle) and $10^{10.0}$ <M_* < $10^{10.5}$ $M_\odot$ (right) for Ref-L100N1504 and Recal-L025N0752, in blue and green, respectively. Solid curves show the median SSFR from the simulation for star forming galaxies, the shaded region (dotted curves) enclose the 10^{th} and 90^{th} percentile values for Ref-L100N1504 (Recal-L025N0752). Medians are only shown when there are more than 10 galaxies per bin. The dashed curve reproduces the median SSFR for Ref-L100N1504 from the first panel on the remaining panels. The light blue curve shows the median SSFR boosted by 0.2 dex. Observational data from Gilbank et al. (2011), Bauer et al. (2013), Karim et al. (2011), Rodighiero et al. (2010), González et al. (2012) and Stark et al. (2013) are shown. The simulation shows good agreement with the observed shape of the SSFR evolution, but there is an offset in normalisation of 0.2 to 0.3 dex, as seen in Fig. 3.7.

The median SSFR for star forming galaxies from Ref-L100N1504 and Recal-L025N0752 are shown in blue and green, respectively. Values are only shown when there are more than 10 galaxies per bin. The $10^{\rm th}$ to $90^{\rm th}$ percentiles are enclosed by the shaded region (dotted curves) for Ref-L100N1504 (Recal-L025N0752). The median SSFR from Ref-L100N1504 in the first panel is reproduced in subsequent panels as a dashed curve to emphasise trends with stellar mass.

For galaxies in the mass range 10^9 to $10^{9.5}$ $\rm M_\odot$ the intermediate- and high- resolution simulations produce similar median SSFRs above redshift 1, but below redshift 1 Recal-L025N0752 has a higher SSFR by up to 0.2 dex. The scatter is similar in the two simulations. In the mass range $10^{9.5}$ to 10^{10} $\rm M_\odot$ the simulations agree to within 0.1 dex above redshift 0.2, at which point the median Recal-L025N0752 SSFR increases to 0.2 dex above the Ref-L100N1504 value. In the highest mass bin shown, 10^{10} to $10^{10.5}$ $\rm M_\odot$, it is only below redshift 0.2 that there are sufficient galaxies in the Recal-L025N0752 simulation to compare to Ref-L100N1504. At redshift zero the median SSFRs are consistent, although at redshift 0.1 they are offset by 0.1 dex. Above redshift one the SSFRs of the two simulations are converged to within 0.1 dex. At lower redshifts, for stellar masses below $10^{10.5}$ $\rm M_\odot$ Recal-L025N0752 has a slightly higher SSFR, by up to 0.2 dex.

In all mass bins the SSFR increases with lookback time. The dashed blue curves in the second and third panels reproduce the simulation SSFR from the first panel from Ref-L100N1504, for stellar masses between 10^9 and $10^{9.5}$ $\rm M_\odot$. Focusing on the middle panel, the dashed curve is within 0.1 dex of the solid curve, showing how flat the slope of the SSFR with stellar mass is in the simulation from 10^9 to 10^{10} across the entire redshift range, i.e. γ=0 from 10^9 to 10^{10} $\rm M_\odot$ across all epochs. The offset between the dashed and solid blue curves in the right hand panel shows that the normalisation in SSFR has decreased for galaxies in the stellar mass bin of 10^{10} to $10^{10.5}$ $\rm M_\odot$ relative to lower mass bins by ~ 0.2 dex. This corresponds to the shallow negative slope seen in the SSFR-$\rm M_*$ relation in Fig. 3.7. Note however the offset between the $10^{9-9.5}$ $\rm M_\odot$ and $10^{10-10.5}$ $\rm M_\odot$ bins remains constant with time. The slope of the SSFR-$\rm M_*$ relation remains constant over time in the simulation,

although varying weakly across stellar masses for star forming galaxies. Only the normalisation evolves strongly.

Comparing to the observations, we see that the trend with redshift is reproduced. There is, however, an offset in normalisation of $0.2 - 0.4$ dex at all times, across all mass ranges, as seen in Fig. 3.7. We found previously that the global star formation rate density was low by ~ 0.2 dex across all redshifts relative to the values estimated from observations (Section 3.3.1). This offset in $\rho_{\rm SFR}$ thus can not fully account for this offset in SSFR. An offset in $\rho_{\rm SFR}$ does not convert directly into an offset in SSFR, due to the potential increase in stellar mass if SFRs were to increase. None the less, we boost our SSFR by the 0.2 dex offset in $\rho_{\rm SFR}$ that was required in Fig. 3.5, as an upper limit to the resulting increase in SSFR. Even when including this 0.2 dex boost (light blue curve) the model fails to reproduce the normalisation of SSFRs in most observations. If the SFRs were boosted by 0.3 dex across all mass ranges, as required to be consistent with the observational data, the agreement for $\rho_{\rm SFR}$ and ρ_* would be broken. A possible solution to the low SSFRs is that the star formation in the simulated galaxies is not sufficiently bursty. We will return to this possibility in the discussion.

As for the stellar mass, there are also uncertainties in the SFRs inferred from observations. Differences in the measured star formation rate density from different star formation tracers are of order 0.2 dex (as in Fig. 3.5), while Utomo et al. (2014) claim that SFRs inferred from UV and IR observations may be overestimated relative to those obtained by simultaneously modelling of stellar and dust emission simultaneously. Attempting to quantify the level of uncertainty in SFRs is difficult owing to the different sensitivity of each star formation tracer. UV observations require a large correction for the light that is absorbed. IR observations require information about the peak of the SED to constrain the total infrared luminosity and must assume all star formation is shrouded in dust if information from the UV is unavailable. Radio (and IR) observations can suffer from contamination by AGN and rely on an empirical calibration between the flux and SFR. At high redshift, where stacking is often necessary due to decreased ability to detect objects, there is

a risk that the sample is incomplete, biasing results towards higher star formation rates. Pérez-González et al. (2008) quote a factor of two (0.3 dex) in the uncertainty of IR SFRs due to dust, Muzzin et al. (2009) find a scatter of a factor of 2.8 (0.45 dex) depending on the bands available for fitting the SED. Folding uncertainties into the comparison in Fig. 3.9, it is unclear whether the differences in SSFRs between the simulation and observations are significant. This is particularly true in light of the bimodal distribution of star formation rates, reasonable passive fractions and the correct behaviour of the normalisation with redshift reproduced by the simulation.

The systematic offset in SSFRs between models and observations has been noted before. Weinmann et al. (2012) reported this issue for two hydrodynamical simulations, while recent studies such as Mitchell et al. (2014) and White et al. (2014) revisited the issue with semi-analytic models. White et al. (2014) propose two plausible solutions to the issue based on their semi-analytic modelling. In the first solution star formation in low-mass galaxies forming at early times is preferentially suppressed, delaying star formation and providing further fuel for stars to form at later times. In the simulations presented here, Ref-L100N1504 and Recal-L025N0752, the dependence of the feedback on local gas metallicity and density does indeed result in preferential suppression of low mass galaxies at early times and this does improve the behaviour of the SSFRs relative to models with constant feedback or velocity dispersion dependent feedback (*Chapter 4*). However, to fully resolve the issue, much stronger feedback is required in low-mass, high-redshift galaxies than the already strong feedback that is implemented here. (Although the requirement for stronger feedback may in part be a result of numerical radiative losses.) The second solution that White et al. (2014) appeal to, with a similar solution proposed by Mitchell et al. (2014), is limiting the cold gas available for star formation by reducing the accretion of gas from hot and ejected reservoirs onto halos (see also Bower et al., 2012). As our simulation follows the gravity and hydrodynamics of the gas, it is not a reasonable solution to apply ad hoc recipes to the accretion of gas in our simulation.

In summary, the simulation reproduces the shape of the evolution of $\rho_{\rm SFR}$ with redshift seen in observations with an 0.2 dex offset. The bimodality in SSFR, the slope with mass and the shape of the evolution of the SSFRs as a function of time are also reproduced by the simulation. However, the normalisation is 0.2-0.4 dex too low at all redshifts and across all masses. This offset cannot be resolved by a simple systematic shift in SFRs in the simulation due to the implications such a shift would have for $\rho_{\rm SFR}$ and ρ_*.

3.4 Discussion

We have presented the evolution of the stellar masses and star formation rates in two of the EAGLE cosmological hydrodynamical simulations. We have focused on Ref-L100N1504, a $(100$ cMpc$)^3$ box with baryonic mass particles of 1.81×10^6 M$_\odot$, and Recal-L025N0752, a (25 cMpc)^3 box with baryonic mass particles of 2.26×10^5 M$_\odot$. These simulations use advanced SPH techniques and state-of-the-art subgrid models, including cooling, metal enrichment, energy input from stellar feedback, black hole growth and feedback from AGN. The subgrid parameters depend only local gas properties. The free parameters of the model have been calibrated to reproduce the observed local Universe GSMF, with consideration given to galaxy sizes (S14). The resulting model has been shown to reproduce many redshift ~ 0 observations, including the Tully-Fisher relation, specific star formation rates, the mass metallicity relation, black hole masses and the column density distribution functions of intergalactic CIV and OVI (S14).

In this chapter we extend the comparison with observations of galaxy stellar masses and star formation rates from redshift zero to redshift seven. This comparison with observations enables us to carry out a multi-epoch verification of the EAGLE galaxy formation model, where the galaxy properties in this comparison are predictions of the model, i.e. evolution histories were not considered during the calibration of model parameters.

Given the calibration of the simulation at redshift zero, the local Universe GSMF

and stellar mass density is not a prediction of the model. However reproducing many of the local Universe properties implies the resulting galaxies in the simulation are comparable to present day observed galaxies and this enables us to confidently apply the simulation to many questions concerning physical processes of galaxy formation. An important question is whether the simulation provides a plausible description of the observed history of the galaxy population across cosmic time. Without the precondition of reproducing redshift zero properties the interpretation of a comparison to the evolution is limited by the resulting unrealistic galaxy population. While a comparison with observational data is extremely useful to determine the accuracy of the EAGLE model relative to the observed universe, it is important to consider the uncertainties in converting observed fluxes or spectra to intrinsic galaxy properties when comparing the simulation to observations.

We began our comparison by finding a better than 20% agreement with the evolution of the stellar mass density across all epochs (Fig. 3.1). For the GSMF, good agreement was typically found for the evolution of the normalisation and break when comparing the simulation to observationally inferred data (Fig. 3.2). The normalisation remains reasonably constant from redshift 0.1 to 1 and then decreases to redshift 2. The decrease continues at higher redshifts. Although this behaviour is qualitatively consistent with observations, at redshift 2 the normalisation below $10^{10.5}$ $M_\odot$ is too high by ~ 0.2 dex. In the current implementation of stellar feedback, galaxies with low metallicity and high density, typical in the early universe, experience strong feedback. As we incorporate all stellar feedback, e.g. core collapse supernovae, stellar winds, radiation pressure, in one form, the available feedback energy can be up to three times that available from core collapse supernovae. The requirement for excess feedback could also be a result of residual numerical radiative losses. A comparison with the normalisation of the observed GSMF at redshift 2 suggests that even more efficient stellar feedback is required in low mass objects at redshifts above two. More efficient feedback at high redshift could provide surplus gas at later times through recycling, helping to boost the SSFRs ($= \dot{M}_*/M_*$), as is required based on the comparison with observational data in

Fig. 3.9.

The break in the GSMF in the simulation evolves in a similar way to that observed, however, between redshifts 2 and 4 there is too little mass in simulated galaxies above 10^{11} $M_{\odot}$, suggesting that less efficient AGN feedback (or stellar feedback in high mass objects) at high redshift is required to produce the observed evolution of the break in the GSMF. Less efficient AGN feedback at high redshift would also result in more star formation around the epoch of peak star formation, at redshift two, as favoured by current observational data for the star formation rate density. The requirement for weaker AGN feedback however, is very sensitive to the stellar mass errors that arise from inferring the GSMF from observations. While recent observations of the GSMF are typically consistent with each other within their error bars, it is important to consider both random and systematic uncertainties in inferring stellar mass from observed flux, as shown in Fig. 3.3. As a result of the sensitivity of the exponential break in the GSMF to the stellar mass errors it is difficult to determine if the AGN are indeed overly effective in the simulation.

The largest discrepancy we find with observational data is in the SSFRs of star forming galaxies, which are 0.2 to 0.4 dex below values inferred from observations across all of cosmic time (Fig. 3.9). This discrepancy cannot be explained as a simple systematic offset in the simulation, as we have shown the stellar mass density to be consistent with observations to within 0.1 dex. Applying a systematic boost to the star formation rates by 0.3 dex would undo the agreement in the stellar mass density. It is puzzling that the SSFRs are systematically low, yet the stellar mass growth is consistent with the observational data. However, we have also found that the galaxy passive fractions appear too low by up to 15% between $10^{10.5}$ and $10^{11.5}$ $M_{\odot}$ (Fig. 3.7). Assuming that the observed star formation rates are accurate, a potential solution to the low SSFRs is that the star formation is not sufficiently bursty. More bursty episodes of star formation could produce the same stellar mass with higher SFRs over shorter time periods than in the current simulation. This solution has the advantage that it would also increase the passive fractions,

as galaxies would be star forming for a smaller fraction of the time. Within our current exploration of the subgrid parameter space, no model has achieved better agreement with the observed SSFRs, while producing the present day stellar mass density, than those presented here.

Observed stellar masses and star formation rates are uncertain at the 0.1 to 0.3 dex level across all observed redshifts. Until recently hydrodynamical simulations have struggled to reproduce redshift zero galaxy populations within the observational uncertainties, not to mention the evolution of the galaxy population. The simultaneous comparison to stellar masses and star formation rates across cosmic time thus provides a stringent test for the evolution of galaxy properties in our galaxy formation model. The EAGLE Ref-L100N1504 simulation performs relatively well in this test, verifying that the simulation produces galaxies with reasonable formation histories, for a redshift zero galaxy population that is representative of the observed Universe. The agreement with observational data from redshifts 0 to 7 is at the level of the systematic uncertainties and follows the observed evolutionary trends. This gives us confidence that the model can be used as a reliable tool for interpreting observations and to explore the physics of galaxy formation. To give further confidence, our simulation shows numerical convergence of the GSMF to within 0.1 dex for galaxies of stellar masses greater than 100 baryonic particles and SSFRs to within 0.1 dex when star formation rates are resolved by a minimum of 10 star forming particles when going to a factor of 8 higher resolution. This level of convergence enables us to extend the galaxy population to lower stellar masses, by a factor of 8, using the higher-resolution simulation.

While there is scope to improve agreement with observational data, it is not clear that this is necessary for a number of reasons. Given that the level of systematic uncertainty in the observations are similar to the level of agreement with the simulation, better agreement with observations would not automatically translate into more confidence in the model. Secondly, as hydrodynamical simulations are computationally expensive, full parameter space searches are unfeasible using current technology. Finally, it is likely that achieving better agreement with ob-

servations would require more complex parameterisation of the subgrid models, which would be better motivated if they were supported by small scale simulations modelling ISM physics and smoothed to the resolution of current cosmological simulations. While many studies of this kind are underway (e.g. Creasey et al., 2013), they do not yet model all the relevant physics and they currently require too much computational time to be incorporated into full cosmological simulations.

3.5 Summary

Here we have presented the evolution of stellar masses and star formation rates in the EAGLE simulations Ref-L100N1504 and Recal-L025N0752. These simulations use (100 cMpc)^3 and (25 cMpc)^3 boxes respectively, with baryonic particle masses of 1.81×10^6 and 2.26×10^5 M$_\odot$, respectively. The simulations use advance SPH techniques and subgrid models for star formation, cooling, chemical enrichment, stellar feedback, black hole growth and AGN feedback. The feedback parameters were calibrated to reproduce the observed redshift ~ 0 galaxy stellar mass function and galaxy sizes. Many other local Universe properties were shown to be reproduced by the simulation in Schaye et al. (2015). In this paper we have extended the comparison with observations from redshift 0 to 7. Below we summarise our main findings.

- The stellar mass density in the simulation is within 20% of the observed values across cosmic time (Fig. 3.1) and the trends in the evolution of the galaxy stellar mass function are reproduced from redshifts 0.1 to 7 (Fig. 3.2), showing that the growth of stellar mass in the simulation is similar to that observed.

- The shape of the evolution of the star formation rate density (Fig. 3.5) and the trends in specific star formation rate with stellar mass and lookback time (or redshift) (Fig. 3.7, 3.9) are reproduced by the simulation, showing the evolution of star formation rates is similar to that observed.

- Below $10^{10.5}$ M$_\odot$ the normalisation of the galaxy stellar mass function is too

low by 0.2 dex at redshift 2. There is also a similar offset in the normalisation of the SSFRs of 0.2-0.4 dex across all redshifts. Semi-analytic models show a similar discrepancy Mitchell et al. (2014); White et al. (2014). If these issues are not due to systematic uncertainties in the observations, then they imply that even stronger feedback is required at high redshift than what is currently implemented in the simulation, although burstier star formation histories could possibly also resolve the offset in the star formation rates.

- Finally, we have presented convergence tests for the galaxy properties considered here and shown that the stellar masses and star formation rates are reasonably well converged across all redshifts at which the convergence can be tested.

In a companion paper we will show how the evolution of the global statistics presented here depends on galaxy masses and we will address the issue of galaxy downsizing (*Chapter 5*).

Chapter 4

Variations of the EAGLE *universe*

4.1 Introduction

As was discussed in *Chapter 1*, hydrodynamical simulations track baryonic and dark matter particles simultaneously, in contrast to other methods that evolve dark matter only simulations and apply baryonic physics as a post-processing step. Due to the tracing of baryonic particles, hydrodynamical simulations require fewer assumptions to modelling galaxy formation, restricting the assumptions to the modelling of subgrid processes. The inclusion of baryons in the simulation enables the study of structure within galaxies, as well as the diffuse gas around galaxies, in the circumgalactic and intergalactic media. Of particular interest to this study is the inclusion of hydrodynamical forces to trace the accretion, expulsion and reaccretion of gas onto galaxies, without the need to parameterise these processes. These processes occur on a macroscopic scale and so are well captured in the simulation.

The primary uncertainty in cosmological hydrodynamical simulations is the subgrid physics (Scannapieco et al., 2012). The subgrid physics captures the baryonic processes that are necessary for the formation of galaxies, but which are unresolved in the simulations. While the subgrid prescriptions for some of the baryonic processes can be constrained by observational data (e.g. star formation, Schaye and Dalla Vecchia, 2008), other subgrid schemes contain more freedom, e.g. stellar and AGN feedback (Schaye et al., 2015, hereafter S14). As star and gas particles are typically of order 10^6 $M_\odot$ in simulations, individual feedback and accretion events can not be distinguished. It is not known, on the scales resolved in the simulation, how the feedback depends on the local environment or how the energy couples

to its surroundings. In this paper we consider the impact of varying the stellar and AGN feedback implementations and parameters on the evolution of galaxy properties in the simulation.

In order to carry out a comparison of models we constrain them using observations. The selected models have parameters that are calibrated to reproduce the normalisation of the redshift 0.1 GSMF below $10^{10.5}$ $M_{\odot}$. As the galaxies in the various simulations have similar galaxy populations in the local Universe, a comparison of their evolution reveals differences in the histories of the galaxies, not differences that arise from the galaxy populations differing at redshift zero, or galaxies of similar masses across the simulations living in different halos. Having constrained the galaxy populations at redshift zero, we ask if the evolution of the models is degenerate, or if the galaxy histories change with the feedback changes. In cases where the galaxy histories vary, are the variations sufficiently significant to rule out certain models using observations? Can we determine anything about the feedback mechanisms in the real Universe?

In this comparison we consider a number of stellar feedback models. The most simplistic model for stellar feedback is for a constant energy fraction, based on the available energy from Type II supernova, to be injected into the IMF. This method has been employed in many cosmological hydrodynamical simulations (e.g. Springel and Hernquist, 2003; Crain et al., 2009; Schaye et al., 2010). Scaling of the energy injected with halo properties, such as the dark matter velocity dispersion, are also implemented in hydrodynamical simulations (e.g. Oppenheimer et al., 2010; Vogelsberger et al., 2014). Similar scaling were first implemented in semi-analytic models (Cole et al., 1994), where the feedback varies with the circular velocity. This was motivated by the expectation that the gas can escape from a shallow potential well with more ease than from a deeper one, and was necessary to reproduce the flat low mass slope of the galaxy stellar mass function. Recently, Schaye et al. (2015) used a model that scaled the energy injected with the gas density and metallicity of the ISM. We consider all of these options in the stellar feedback variations considered here.

Many studies of cosmological simulations also vary the energy injection method, e.g. as thermal or kinetic energy, and the parameters associated with these methods, e.g. mass loading, velocity, heating temperature. However, these variations have been considered in previous studies (Schaye et al., 2010) and are beyond the scope of this work.

As well as stellar feedback variations, we consider AGN feedback variations. There are a number of subgrid black hole accretion and feedback models available. The sub parsec accretion disks of black holes can not be resolved in simulations, so subgrid recipes are required. For example Springel (2005) and Booth and Schaye (2009) use modified Bondi-Hoyle accretion models. Rosas-Guevara et al. (2013) uses the model of Booth and Schaye (2009), but include a dependence on the angular momentum of the in-falling gas. Here we consider both a Bondi-Hoyle accretion model (without the modifications of Booth and Schaye (2009) or Springel (2005)) and this model with the correction to the accretion timescale from the Rosas-Guevara et al. (2013) model. We also consider the effect of varying the heating temperature of the AGN feedback. Finally, we include NOAGN feedback to determine the importance of the AGN feedback. Note that the simulations analysed in this work are limited to 25 cMpc boxes and so the number of massive halos and galaxies is small. This is a limiting factor in determining differences between the models, although any significant difference will be realised.

A summary of this chapter is as follows. In Section 4.2 the simulations of the Eagle suite considered here are described. First we describe the physics and parameters used in the reference model and then we describe the variations. We justify our choice of variations in Section 4.3. To compare the variations, we begin by looking at global evolution properties, such as the star formation rate density in Section 4.4. Then we look at galaxy properties as a function of stellar mass in Section 4.5. In Section 4.5.3 we aim to understand the differences in each model that produce the resulting differences in evolution of galaxy properties. Finally in Section 4.6 we embark on a discussion of our finding and present our conclusions.

The EAGLE simulation suite adopts a flat ΛCDM universe with Planck cosmol-

ogy (Planck Collaboration et al. (2013)); $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\sigma_8 = 0.8288$, $n_s = 0.9611$ and $H_0 = 67.77$km s^{-1} Mpc^{-1}. Observational stellar masses and volumes are rescaled to the Planck cosmology. A Chabrier (2003) IMF is used, with observational stellar masses renormalised to this IMF where necessary. Stellar masses from the simulation are quoted in solar masses except where otherwise stated and computed within 3D apertures with 30 proper kpc radii. Star formation rates are computed within the same apertures. Distances and volumes are quoted in comoving units (e.g. cMpc), unless otherwise stated. Proper distance units will be quoted, for example, as pMpc. Note that values are not quoted in h^{-1} units.

4.2 Simulations

The EAGLE simulation suite includes a number of cosmological simulations with varying subgrid parameters and implementations in 25-100 cMpc boxes (Crain et al., 2015). Here we focus on a variety of (25 cMpc)^3 boxes with baryonic mass resolution of 1.81×10^6 M$_\odot$ and dark matter mass resolution of 9.70×10^6 M$_\odot$. Plummer equivalent comoving gravitational softenings are set to 1/25 of the initial mean inter-particle spacing and are limited to a maximum physical size of 0.70 pkpc. These simulations use advanced SPH techniques and state-of-the-art subgrid models to capture the unresolved physics. Cooling, metal enrichment, energy input from stellar feedback, black hole growth and feedback from AGN are included. The stellar and AGN feedback are key elements of the subgrid scheme and it is variations of these that are considered here. The models considered are calibrated to reproduce the observed flat slope and normalisation of the redshift zero GSMF. A complete description of the code, subgrid physics and parameters for the Reference model can be found in *Chapter 2*, while the details of all EAGLE variations can be found in Crain et al. (2015). In Section 4.2.1 we describe the subgrid parameter and implementation variations, which relate to the stellar feedback and AGN routines.

4.2.1 Simulation Variations

We restrict the comparison in this paper to models with good agreement to the normalisation and slope of the local Universe GSMF, with two exceptions. For the exceptions, one model over shoots the normalisation of the GSMF and one undershoots it. This selection allows us to analyse the degeneracy in different models with redshift, for the models that can reproduce the present day GSMF below $10^{10.5}$ $M_{\odot}$. Of the models considered, three models vary the stellar feedback subgrid implementation, three vary the AGN accretion model and the final two vary the stellar feedback parameters to over and undershoot the normalisation of the GSMF. Here we describe the variations relative to the Reference model, with a summary of changes provided in Table 4.1.

- **SNConstant**: A constant energy fraction from stellar feedback, with a value of one is applied. Feedback energy that is constant is often applied in cosmological simulations, e.g. Crain et al. (2009), Schaye et al. (2010), Khandai et al. (2014). As we are interested in models that best reproduce the present day GSMF, the black hole accretion effective viscosity parameter, α, is changed from 2π to $2\pi \times 10^2$ to better reproduced the knee of the GSMF. The α-parameter parameterises the unresolved thin accretion disc, although Rosas-Guevara et al. (2013) show in an Appendix that the black hole accretion model depends only weakly on the value for α. The overlap between the AGN feedback and stellar feedback is discussed in Crain et al. (2015).

- **SNZ**: The energy fraction available from stellar feedback depends on the local gas metallicity, whereas in the Reference model it depends on both the metallicity and density. The metallicity dependence of the feedback accounts for excessive radiative losses, which are unresolved in our simulation, in high metallicity regions, which would reduce the effectiveness of the feedback from star formation in these regions. The energy fraction has a maximum value of three times the energy available form Type II supernova in very low metallicity regions, decreasing to 0.3 in high metallicity regions. The same

limits are applied in the Reference model. The resulting median and mean energy fraction from this model for stellar feedback are 0.6 and 0.7 respectively. As reported in Crain et al. (2015), the introduction of a further density dependence of the stellar feedback energy fraction is due to the compact galaxy cores produced by the SNZ model that produce galaxies sizes (and hence surface density star formation rates) that are in tension with observational data. Again we vary the black hole accretion α-parameter relative to the Reference model to produce a model with a good match to the redshift 0.1 GSMF; it is set to $2\pi \times 10^2$.

- **SNsigma**: The fraction of energy available from stellar feedback depends on the local dark matter velocity dispersion. Halo property dependence of the stellar feedback is employed in many subgrid prescriptions in hydrodynamical simulations (Oppenheimer et al., 2010; Vogelsberger et al., 2014). As in the Reference model and metallicity dependent model the energy fraction can vary from 3.0 to 0.3. In this simulation the resulting median and mean values are again 0.6 and 0.7 respectively. The α-parameter is set to $2\pi \times 10^2$.

- **AGNdT9**: The AGN heating temperature is increased from $10^{8.5}$ K, as in the Reference model, to 10^9 K, and the black hole α-parameter is changed to $2\pi \times 10^2$, while the stellar feedback model remains fixed as in the Reference simulation. This variation was used to be in better agreement with group and cluster properties in S14 than the Reference model. As this model shows agreement across the widest range of redshift zero observations, it is useful to include it in the comparison.

- **AGNBH**: To study the impact of the AGN accretion implementation, a Bondi-Hoyle accretion model is shown, replacing the angular momentum accretion model of Rosas-Guevara et al. (2013) used in the Reference simulation. The standard Bondi-Hoyle formula, as in eq. 2.2 determines the accretion onto the black hole, while black holes are seeded at the same halo mass, with the same size black hole seeds as in the Reference model. This is equivalent to

the standard accretion model with a very low α value. Note that, as in the Reference model, the Bondi-Hoyle formula is not multiplied by the density dependent term of Booth and Schaye (2009), as in Rosas-Guevara et al. (2013). This model allows us to investigate the effect of changing the subgrid accretion prescription on global evolution properties.

- **NOAGN**: Black hole seeding and growth are omitted from this model, resulting in no AGN feedback, which allows us to isolates the effect of the AGN on the star formation and stellar mass growth.

- **SNWeak, SNStrong**: The final two models vary the stellar feedback parameters relative to the Reference model, to produce results that over and under shoot the total stellar mass density and the GSMF normalisation at redshift zero. The weak feedback model reduces the allowed energy fraction range by half, giving values of 1.5 to 0.15, resulting in half the energy being available per supernova event as in the Reference models. The strong feedback model doubles the energy fraction.

In Section 4.3 we will present the GSMFs for these variations to justify our choice of these models.

4.2.2 Resolution Tests

For the variations considered here we do not consider resolution tests. Resolution tests for the Reference model were considered in S14 for redshift zero galaxy properties and in *Chapter 3* for the evolution of galaxy properties. The numerical convergence of the EAGLE Reference simulation was found to be very good. Galaxy stellar masses are typically converged where the galaxy mass was greater than that of 100 baryonic particles. Star formation rates are converged within 0.1 dex when resolved by more than 10 star forming particles for galaxies at redshifts above 1 or for galaxies with masses above 10^{10} $M_{\odot}$.

Model	**SN**	$\mathbf{f}_{\mathrm{E,min}}$	$\mathbf{f}_{\mathrm{E,max}}$	$\mathbf{dT}_{\mathrm{SN}}$	**Accretion**	$\mathbf{dT}_{\mathrm{AGN}}$	α
Ref	Z and ρ	3	0.3	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	2π
SNConstant	Constant	1	1	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	$2\pi \times 10^2$
SNZ	Z	3	0.3	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	$2\pi \times 10^2$
SNsigma	σ	3	0.3	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	$2\pi \times 10^2$
AGNdT9	Z and ρ	3	0.3	$10^{7.5}$ K	Angular Momentum	10^{9} K	$2\pi \times 10^2$
AGNBH	Z and ρ	3	0.3	$10^{7.5}$ K	Bondi-Hoyle	$10^{8.5}$ K	N/A
NoAGN	Z and ρ	3	0.3	$10^{7.5}$ K	N/A	N/A	N/A
SNWeak	Z and ρ	1.5	0.15	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	2π
SNStrong	Z and ρ	6	0.6	$10^{7.5}$ K	Angular Momentum	$10^{8.5}$ K	2π

Table 4.1: List of key parameters for model variations considered.

For low-mass, low-redshift galaxies the star formation rates are converged within 0.2 dex, the high resolution simulation increases the star formation rates. We will use the convergence limits that apply for the Reference model as a guide for the simulations considered here.

To highlight the convergence of the simulations with box size, which is particularly relevant for volume weighted properties, we show the 100 cMpc simulation for the Reference model.

4.2.3 Halo and galaxy definition

Halo and subhalo finding is carried out as in *Chapter* 2. We recall that a galaxy's mass is defined by the mass that is gravitationally bound to the subhalo within a 30 pkpc aperture. Star formation rates are defined in the same way. Note that the integrated galaxy properties presented, however, include all mass or star formation; no aperture is applied.

4.3 Model selection and calibration

We present the GSMFs and the stellar mass-halo mass relation at redshift 0.1 for the models described in Section 4.2.1 to justify the choice of variations considered. Fig. 4.1 shows the GSMF at redshift 0.1 for each of the nine models in 25 cMpc boxes in separate panels. Observational data from Li and White (2009) and Baldry et al. (2012) are shown as stars and dots respectively. To get a clearer understanding of the impact of the boxsize, the top left panel shows the Reference model in 25 and 100 cMpc boxes (Ref-L025N0376 and Ref-L100N1504) as solid and dot-dashed lines respectively. At masses below $10^{10.5}$ $M_\odot$ the GSMFs from the boxes are within 0.2 dex of each other. Poorer sampling in the 25 cMpc box causes the fluctuations. The largest discrepancy between the boxes is between $10^{10.2}$ and $10^{10.8}$ $M_\odot$, where the normalisation of the 25 cMpc box is 0.2 dex above the 100 cMpc box. This could be due to cosmic variance, or the box being so small that the periodicity of the box affects the growth at late times. The large scale modes are absent from

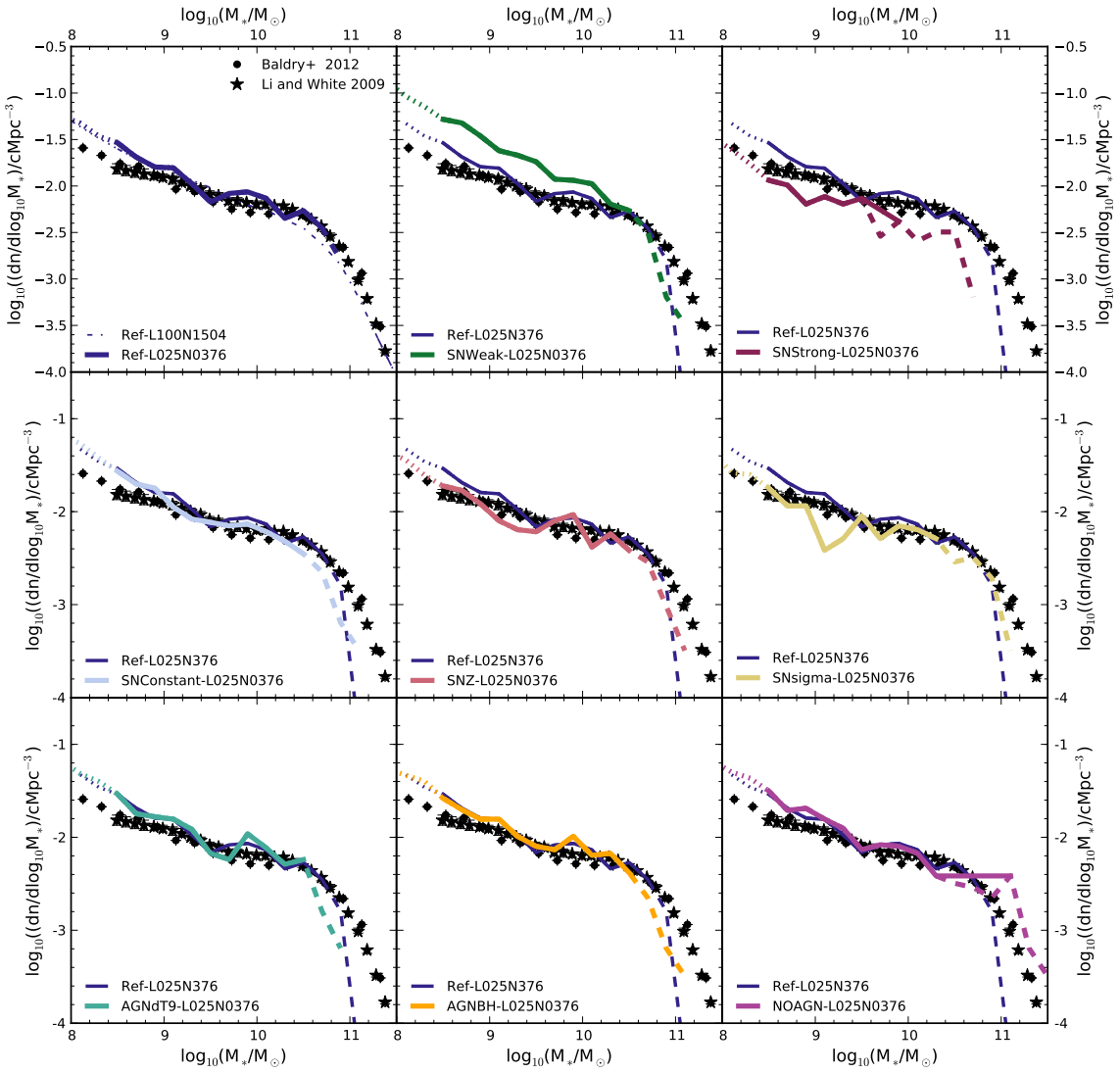

Figure 4.1: The galaxy stellar mass function at redshift 0.1 in 9 of the EAGLE simulations in 25 cMpc boxes. Dashed curves show where stellar mass bins have fewer than 10 objects, dotted curves show where a galaxy stellar mass is resolved by less than 100 baryonic particles. The Ref-L025N0376 GSMF is repeated on each panel for comparison. In the top left panel the Ref-L100N1504 GSMF is shown to see the effect of increasing the boxsize. Observational data from Li and White (2009) and Baldry et al. (2012) from the local Universe are shown as stars and dots respectively. 7 of the 9 simulations reproduce the normalisation and slope of the GSMF at masses between 10^8 $M_\odot$ and $10^{10.5}$ $M_\odot$. In the top middle and right panels, two models that do not reproduce the observations are shown. They over and under shoot the normalisation, bracketing the observed answer.

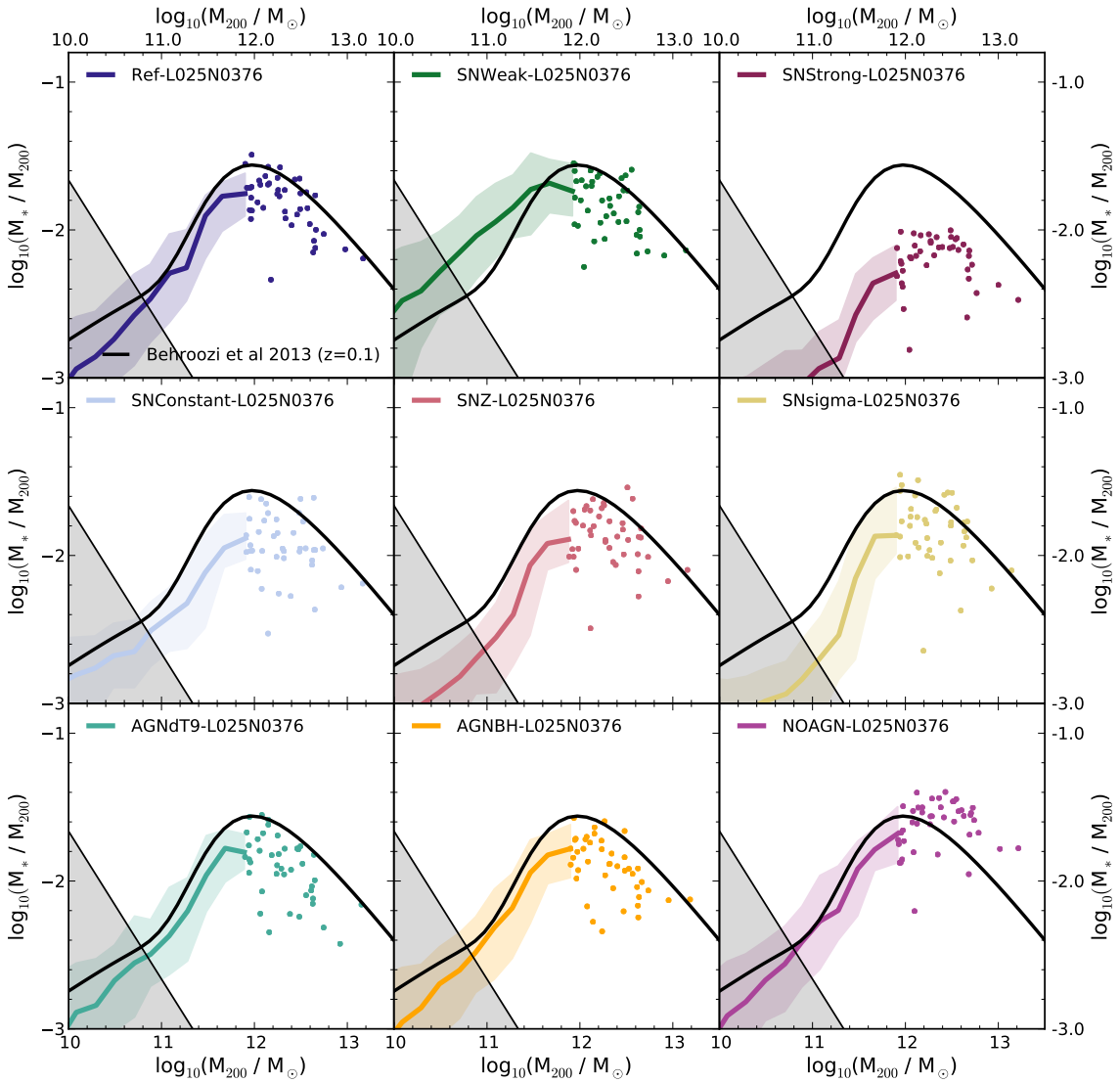

Figure 4.2: The stellar mass over halo mass relation at redshift 0.1 in 9 of the EAGLE simulations in 25 cMpc boxes. The stellar mass is the central galaxy mass, measured within a 30 pkpc aperture. The solid line shows the median relation, the shaded coloured region encloses the $10^{\rm th}$ to $90{\rm th}$ percentiles. When there are fewer than 10 objects per halo mass bin each galaxy is shown. The shaded grey region shows where the galaxy stellar masses are resolved by fewer than 100 baryonic particles. The solid black curve shows the abundance matching results of Behroozi et al. (2013). The relation shown here provides more information on the most massive galaxies in the simulation, which are poorly sampled in the GSMF. It is clear that the NOAGN simulation overestimates the stellar masses relative to the prediction from abundance matching in halos above 10^{12} $\rm M_{\odot}$.

a 25 cMpc box, resulting in no objects with masses greater than $10^{10.8}$ $M_{\odot}$. The Reference model lies within 0.3 dex of the observations in the stellar mass range of 10^{8} to $10^{10.5}$ $M_{\odot}$, showing a good fit to the normalisation and slope of the redshift 0.1 GSMF.

The two variations in the top middle and right panels are offset from the observations by up to 1 and 0.5 dex in normalisation respectively, showing much poorer agreement than the Reference model. These models are SNWeak and SNStrong, which use weak and strong feedback respectively, through varying the Reference model parameters. These models will allow us to explore the evolution in simulated universes that over and under produce stellar mass, and the normalisation of the GSMF.

The stellar feedback variations are shown in the middle panels of Fig. 4.1. These models all have similar normalisation and slope to the observed GSMF over the mass range 10^{8} to $10^{10.5}$ $M_{\odot}$, although the SNZ and SNsigma models fluctuate more than the Reference model. In the bottom panel the AGN variations are shown, which again have similar normalisation and slope to the observed GSMF over the mass range 10^{8} to $10^{10.5}$ $M_{\odot}$. As the stellar feedback and AGN variations reproduce the GSMF, over the range of stellar masses possible within their box size limitations, we consider these models of interest to determine if the history of the redshift zero galaxy populations across the different models is degenerate. Note that the exact level of agreement of the models with observations depends on the time invested in calibration.

In Fig. 4.2 we present the stellar mass over halo mass as a function of halo mass for all models. This figure shows what halos the simulated galaxies live in and can be compared to the abundance matching results of Behroozi et al. (2013). Dividing by the halo mass on the y-axis decreases the dynamic range of the plot, highlighting any differences when comparing to abundance matching results. We can also gain more information about the most massive galaxies in the simulation, that are poorly sampled, than is available in Fig. 4.1. All models show a similar scatter of ~ 0.4 dex up to a halo mass of 10^{12} $M_{\odot}$, above this halo mass there are

too few objects in the simulation to comment on the scatter.

The Reference model reproduces the abundance matching results reasonable well over the halo mass range $10^{10.5}$ to 10^{12} $M_\odot$. Above 10^{12} $M_\odot$ the simulation stellar masses are lower than predicted from abundance matching. Note that we only use the abundance matching results as a guide, and have never found exact agreement, in spite of the agreement of the GSMF with observations. As expected, the SNWeak and SNStrong models over and under shoot the relation predicted by abundance matching, respectively.

The stellar feedback models in the middle panels show similar slopes to the abundance matching predictions over the halo mass range, although the stellar masses are up to 0.3 dex lower at a given halo mass. The AGN models in the bottom panels are similar to the Reference model over the halo mass range $10^{10.5}$ to 10^{12} $M_\odot$. The NOAGN model has more stellar mass in halos greater than 10^{12} $M_\odot$, as expected without AGN feedback to quench star formation in massive objects.

Having chosen a selection of models that can reproduce the present day GSMF and the predicted stellar mass-halo mass relation from abundance matching, we now explore the differences in evolution in these models to reach this present day galaxy stellar mass distribution.

4.4 Global evolution properties

We will compare the global evolution properties of galaxies to see if all are equally plausible descriptions of the observed Universe. If we can distinguish between the models using their evolution, this could be useful in constraining the potential validity of different galaxy formation subgrid implementations.

4.4.1 Star formation rates

We begin by studying the star formation rate density ($\rho_{\rm SFR}$) as a function of redshift for each of the nine models, as shown in the panels of Fig. 4.3. The Reference model, presented in each panel, shows a growth in $\rho_{\rm SFR}$ from redshift 10 to

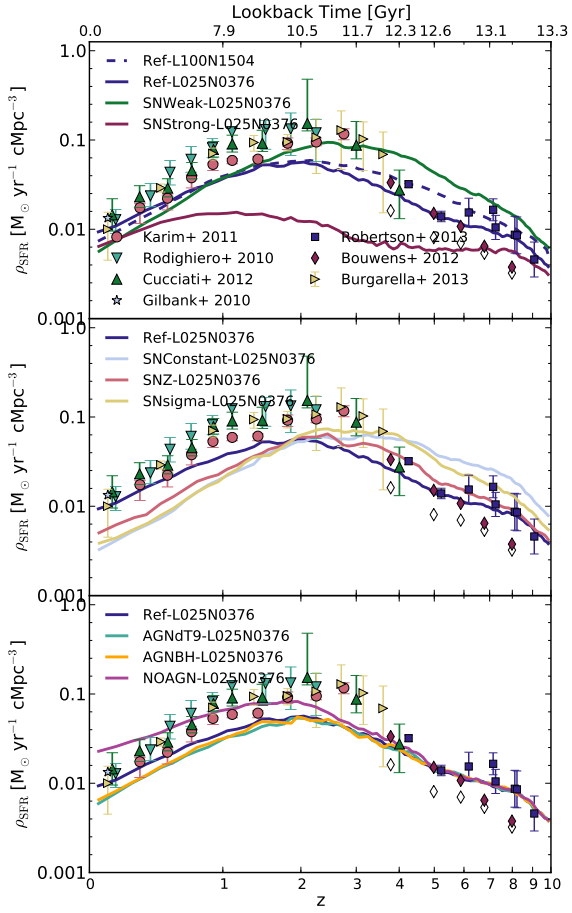

Figure 4.3: The star formation rate density as a function of redshift for 9 simulations from the EAGLE simulation in 25 cMpc boxes. In the top panel Ref-L100N1504 is shown as a dashed line. A comparison of Ref-L100N1504 and Ref-L025N0376 shows the effect of boxsize, increasing the boxsize increases $\rho_{\rm SFR}$ at redshifts higher than 2. For comparison, observational data from Gilbank et al. (2010a), Rodighiero et al. (2010), Cucciati et al. (2012), Karim et al. (2011) , Burgarella et al. (2013), Robertson et al. (2013) and Bouwens et al. (2012) are shown. From the middle panel we see that changes to the stellar feedback can have a significant effect on the shape of $\rho_{\rm SFR}$. In the bottom panel we see that changing the AGN physics has a much smaller effect, unless it is omitted completely. In the NOAGN case $\rho_{\rm SFR}$ is similar to Ref-L025N0376 at redshifts higher than 2, but show significantly more star formation at redshifts below 2.

2. Around redshift 2 $\rho_{\rm SFR}$ peaks, at a value of 0.6 $\rm M_{\odot} yr^{-1} Mpc^{-3}$, before a 0.8 dex decline to redshift zero. At the present day, $\rho_{\rm SFR}$ has a value of 0.01 $\rm M_{\odot} yr^{-1} Mpc^{-3}$.

The top panel shows the Reference model in a 100 cMpc box as a dashed curve. The increased box size results in a boost in star formation between redshift 10 and 2, due to more massive objects forming earlier, from larger scale modes. The offset at these high redshifts is up to 0.2 dex between Ref-L100N1504 and Ref-L025N0376. By redshift two, at the peak in the star formation history, both boxes have converged. Below redshift 2 the star formation histories are similar, with a maximum difference of less then 0.1 dex. Overall, the main difference we anticipate with an increased boxsize is more star formation at high redshifts, above the peak in the star formation history, due to star formation at high redshift occurring in rare massive halos.

The top panel of Fig. 4.3 also shows the two bracketing cases for the GSMF normalisation at redshift zero. The SNWeak model, which overshoots the normalisation, has significantly more star formation at redshifts greater than two. The peak in star formation for this model is at redshift three and the peak value is higher than in the Reference model by 0.2 dex. There is a steeper decline in star formation below the peak than in the Reference model. The weaker feedback results in more stars forming in the simulation, although the excess in star formation occurs above redshift 2, with the peak in star formation 1 Gyr earlier than in the Reference model. This results in more stellar mass being in place at high redshift and an older galaxy population. In spite of the weaker feedback, the low redshift universe in this model has a lower global star formation rate than the Reference model.

The SNStrong model, also in the top panel, overly suppresses star formation, with too few stars formed by redshift zero. In this model $\rho_{\rm SFR}$ never exceeds 0.02 $\rm M_{\odot} yr^{-1} Mpc^{-3}$. The shape of the star formation history is very different to that of the Reference model, it is almost flat with redshift. The $\rho_{\rm SFR}$ begins to rise at redshift two, but is quickly suppressed. The increase in star formation occurs when halos have grown to 10^{12} $\rm M_{\odot}$, where the conversion of gas to stars is most efficient (see Fig. 4.2). As higher halo masses are formed however the star formation is shut

down by the AGN feedback in the simulation, resulting in the suppression of $\rho_{\rm SFR}$.

The middle panel of Fig. 4.3 shows the stellar feedback variations that reproduce the normalisation in the local Universe GSMF. All of these models produce similar present day galaxy populations, based on the population's stellar mass, however the star formation history to produce this stellar mass varies significantly. All 3 variations in this panel have a higher $\rho_{\rm SFR}$ above redshift 2 than the Reference model. In the case of SNConstant, which shows the most extreme increase relative to Reference, $\rho_{\rm SFR}$ is 0.4 dex higher between redshift 3 and 8. The peak in $\rho_{\rm SFR}$ is also moved to higher redshifts for all 3 models relative to the Reference model. As all models produce a similar present day stellar mass density, the star formation below redshift 2 in the stellar feedback variations is below that of the Reference model. At redshift zero there is between 0.2 and 0.4 dex offset in $\rho_{\rm SFR}$ between the Reference model and the stellar feedback variations. Clearly the evolution of the global star formation in the simulation is sensitive to the employed feedback scheme. The differences in $\rho_{\rm SFR}$ for the stellar feedback variations result in these models forming stellar mass earlier and having older galaxy populations than the Reference model, as with the SNWeak model.

In the bottom panel of Fig 4.3 the Reference model is compared to the 3 AGN variations. In the case of AGNdT9 and AGNBH, where the heating temperature of the AGN feedback and the accretion rate onto the black hole are varied respectively, the resulting change in the star formation history is less than 0.1 dex across all redshifts. Note however that the box size of 25 cMpc contains very few large halos, where these AGN variations would have an impact. While changing the AGN feedback parameters or accretion model make little difference, excluding AGN feedback has a significant impact on the star formation history below redshift 2. The peak in $\rho_{\rm SFR}$ for the NOAGN model is 0.2 dex higher than the Reference model and the decline to redshift 0 is less steep. The present day $\rho_{\rm SFR}$ in the NOAGN model is 0.4 dex above the Reference model. While the NOAGN model can reproduce the normalisation of the low mass end of the GSMF it overproduces stellar mass in large halos, greater than 10^{12} $M_{\odot}$(see Fig, 4.2). The hier-

archical growth of structure in the ΛCDM model explains the delay until redshift two of the deviation from the Reference model.

Observational data from Gilbank et al. (2010a) [Hα], Rodighiero et al. (2010) [24μm], Karim et al. (2011) [Radio], Cucciati et al. (2012) [FUV], Bouwens et al. (2012) [UV] , Robertson et al. (2013) [UV] and Burgarella et al. (2013) [FUV + FIR] are shown on each panel of Fig. 4.3 for comparison. In *Chapter 3* it was shown that the Ref-L100N1504 simulation could reproduce the shape of the observations, although the normalisation was low by 0.2 dex. Comparing the models shown here to the observational data, only the NOAGN model can reproduce the normalisation of the observations below redshift 2. However, this model has too much stellar mass in halos greater than 10^{12} M$_\odot$ (Fig. 4.2). The Reference model and other AGN variations are low by ~ 0.2 dex, as for Ref-L100N1504. The stellar feedback variations have a higher normalisation than observations above redshift 3 (this offset will increase in a larger box), with the peak in $\rho_{\rm SFR}$ at higher redshift than observations and a steeper decline to redshift zero. Based on the data, the SNsigma and SNConstant models do not show evolution histories consistent with observed galaxies.

As the growth of stellar mass is clearly affected by these variations in star formation, we now look at the stellar mass density and compare the growth to observations.

4.4.2 Stellar mass density

In Fig. 4.4 we show the stellar mass density (ρ_*) as a function of redshift, with a linear y axis, for each of the 9 models in 25 cMpc boxes. The Reference model has a $\rho_*$of less than 10^7 M$_\odot$cMpc^{-3} above redshift 4. From redshift 4 to 0 the stellar mass increases by an order of magnitude. The present day stellar mass density is 2×10^8 M$_\odot$cMpc^{-3}. The dashed curve in the top panel shows the stellar mass density in Ref-L100N1504, to show the impact of increasing the box size. There is $\sim 10\%$ more stellar mass in Ref-L100N1504 relative to Ref-L025N0376 from redshift 0 to 4. This increase in stellar mass results from the higher $\rho_{\rm SFR}$ in Ref-L100N1504 at redshifts

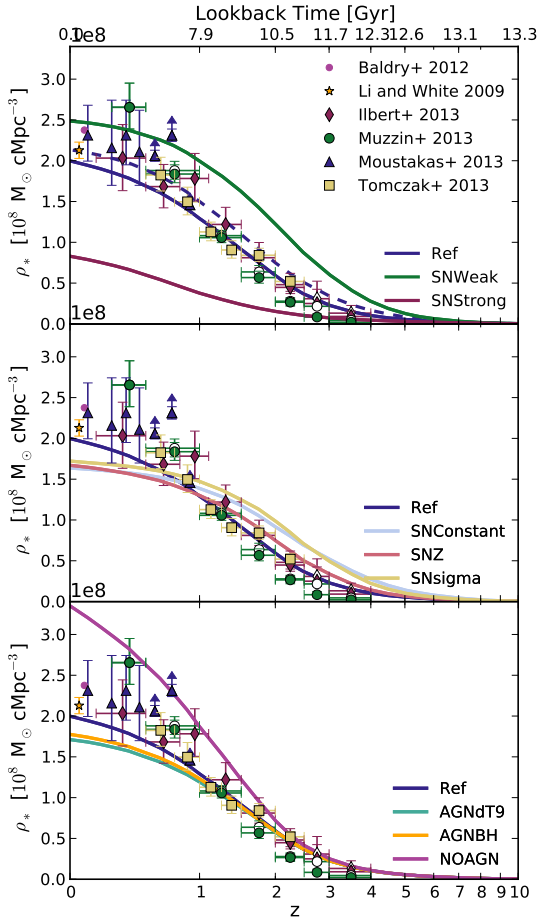

Figure 4.4: The stellar mass density as a function of redshift, with a linear y-axis for 9 EAGLE simulations in 25 cMpc boxes. The dashed line shows Ref-L100N1504, to see the effect of increasing the box size. Observational data from Li and White (2009), Baldry et al. (2012), Moustakas et al. (2013), Tomczak et al. (2014), Ilbert et al. (2013), Muzzin et al. (2013) and González et al. (2011) are shown. 6 of the models have a redshift 0 ρ_* within 0.4 $M_\odot$cMpc^{-3} (20%). Of the 6 models that have reasonable redshift zero values for ρ_*, when considering the boxsize, the primary difference in the models is the difference in growth high redshift, where some models show much steeper growth than the Reference model. These models with steeper growth are less consistent with the observational data above redshift 2.

above 2 seen in in Fig. 4.3.

The top panel also shows ρ_* for SNWeak and SNStrong. Relative to the Reference model, SNWeak and SNStrong over and under produce the redshift 0 ρ_* by 25% and 55%, respectively. The SNWeak model has a similar shape to the Reference model, although with more stellar mass above redshift 3. The SNStrong has a much flatter evolution in ρ_*.

In the middle panel the Reference model is compared to the stellar feedback variations. At redshift zero, the stellar mass densities of these variations are $\sim 20\%$ below the stellar mass density in the Reference model. These variations depend on the extent of the calibration to the local Universe GSMF. What is of more interest than the offset is the variation in shape of the stellar mass density with redshift relative to the Reference model. The stellar mass in all stellar feedback variations builds up quicker than in the Reference model, as expected from the $\rho_{\rm SFR}$. At redshift 4 the SNConstant and SNsigma models have twice as much stellar mass as the Reference model. At redshift 2 there is still almost twice as much stellar mass in SNConstant and SNsigma. The SNZ model however has no excess stellar mass at redshift 4 relative to the Reference model, and there is only $\sim 50\%$ more at redshift 2. Below redshift 2 ρ_* begins to flatten in the stellar feedback variations, so as they have less stellar mass than the Reference model at redshift 0. If these models were recalibrated to produce a stellar mass density in closer agreement with the Reference model, the offset in stellar mass at redshift 2 and above would increase.

In the bottom panel we compare the Reference model to the AGN variations. The stellar mass density in all variations is very similar to the Reference model at redshifts above 1. Below redshift 1 the AGNdT9 and AGNBH models have 10% less stellar mass than the Reference model, as expected from $\rho_{\rm SFR}$. The small difference primarily arises from the break in the GSMF, seen in Fig. 4.1, where the AGN and stellar feedback schemes overlap. The NOAGN model produces almost 75% more stellar mass than the Reference model by redshift 0, this excess mass is in and around the most massive galaxies in the simulation, as is apparent in Fig. 4.2.

On each panel of Fig. 4.4 we show observations from Baldry et al. (2012), Li

and White (2009), Ilbert et al. (2013), Muzzin et al. (2013), Moustakas et al. (2013), Tomczak et al. (2014) and González et al. (2011). In *Chapter 3* it was shown that the stellar mass density from Ref-L100N1504 agrees with observations at the 20% level. While the box size available for the variations is small, a comparison of Ref-L100N1504 and Ref-L025N0376 revealed a $\sim 10\%$ increase in ρ_* below redshift 4, which provides a guide for the impact of increasing the box size. The increased ρ_* between redshifts 2 and 4 in the stellar feedback variations SNConstant and SNsigma, are higher than the observations. Increasing the box size will only increase the tension with observational data, as for $\rho_{\rm SFR}$. While the SNZ model also builds up stellar mass at higher redshift than the Reference model, the increase is not significant and ρ_* for this model remains consistent with observations.

Again in a small box it is difficult to distinguish between variations of the AGN feedback or accretion parameters. However the NOAGN clearly overproduces stellar mass relative to expectations from observations in the local Universe.

If we consider the comparison of the simulation variations to observations, the star formation histories and stellar mass growth of SNConstant, SNsigma and NOAGN are not consistent with expectations from observations. Although, particularly at high redshifts, where SNConstant and SNsigma are too high, the uncertainties in the observational data may be larger than indicated by the error bars.

4.4.3 Gas mass density

In Fig. 4.5 the density of gas in galaxies is shown as a function of redshift for the 9 models considered here. The dashed curves show the total gas density in galaxies, while the solid curves show the density of star forming (i.e. cold) gas. For the Reference model, the total gas density increases with decreasing redshift, the gas density then peaks between redshift 3 and 4. There is a decline from the peak to redshift 0 of 0.8 dex. The increase in the density of star forming gas ($\rho_{\rm sfgas}$) is less steep than the increase in the density of all gas ($\rho_{\rm gas}$), with the peak shifted to redshift 2. This difference results from a feedback scheme that suppresses star formation in high density, low metallicity environments, which are typical in the

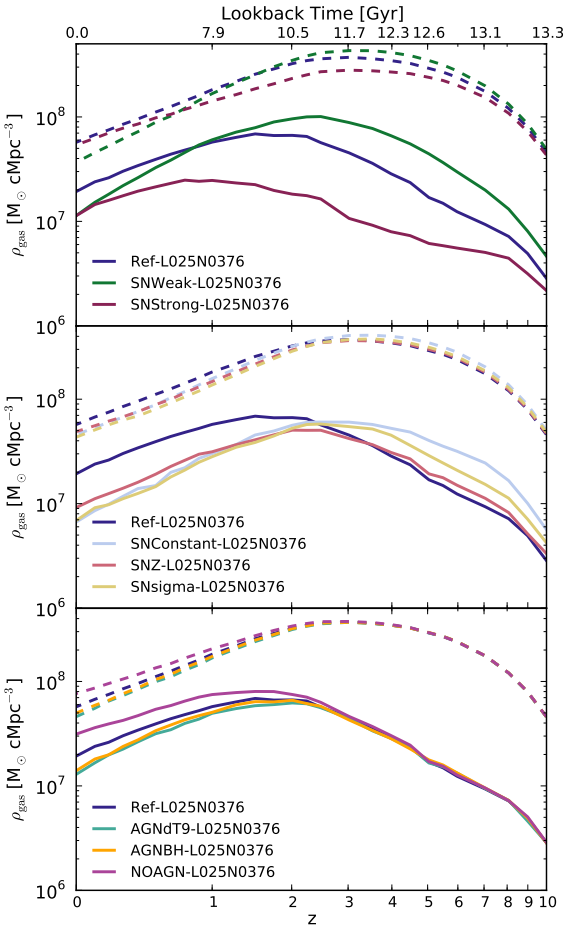

Figure 4.5: The gas mass density in galaxies as a function of redshift for 9 EAGLE simulations in 25 cMpc boxes. Dashed curves show the total gas density in halos, solid curves show the density of star forming gas in galaxies. The $\rho_{\rm gas}$ in galaxies is very similar in shape and normalisation for all models, however the shape and normalisation for star forming gas differs significantly. The $\rho_{\rm sfgas}$ for the Reference models rises slowly and peaks around redshift 1.5 before dropping gradually. In the middle panel, we see the stellar feedback variations all peak at higher redshifts and have a steeper decrease in density from the peak to redshift zero. From high redshift to redshift 2 the AGN variations trace the Reference model. From redshift 2 to 0 they deviate, with more star forming gas in the NOAGN model, and less in the AGNBH and AGNdT9 models.

early Universe. The star forming gas that is heated at early times remains within the galaxies, but with star formation delayed relative to the in-fall of gas.

The largest differences across different models in $\rho_{\rm gas}$ are in the SNWeak and SNStrong models. The SNWeak model has less gas at redshift zero, this is due to the increased stellar mass density at this redshift seen in Fig. 4.4, more of the gas is converted into stars. The SNStrong model has less gas than the Reference model between redshift 7 and 1, this is due to the very strong feedback in SNStrong that can remove more gas from the galaxy. Both SNWeak and SNStrong also differ in $\rho_{\rm sfgas}$, as expected from the different star formation histories and stellar mass densities. SNWeak has more star forming gas than the Reference model above the peak and less below the peak. The shape of $\rho_{\rm sfgas}$ for SNWeak follows that of $\rho_{\rm gas}$, peaking between redshift 3 and 4. The star forming gas in SNStrong is lower than the Reference model by up to 0.4 dex, has a flatter shape and peaks around redshift 1. These changes to the gas densities are driven by the difference in the available feedback energy.

In the middle panel of Fig. 4.5, the total density of gas in galaxies for all models is within 0.1 dex of the Reference model. The density of star forming gas differs more significantly. All stellar feedback variations have a peak in $\rho_{\rm sfgas}$ at higher redshift than the Reference model. These variations also have lower star forming gas density at redshift zero by ~ 0.2 dex, as expected from the difference in $\rho_{\rm SFR}$. The $\rho_{\rm sfgas}$ for the stellar feedback variations are more similar to the $\rho_{\rm gas}$ curves than the Reference model is, showing accretion onto the halo primarily drives the shape of the star formation history. The delay in star formation in the Reference model is key to the improved agreement with observations of ρ_* and $\rho_{\rm SFR}$ relative to the stellar feedback variations.

In the bottom panel, $\rho_{\rm gas}$ and $\rho_{\rm sfgas}$ for AGNdT9 and AGNBH are within 0.1 dex of the Reference model, again in this box size these variation do not affect the results significantly. The NOAGN model has more gas below redshift 1 by up to 0.2 dex, showing the AGN feedback is removing gas from galaxies in the Reference model. There is more star forming gas, by up to 0.4 dex, in the NOAGN model

also, showing that the AGN feedback is suppressing the star formation, although not all the gas has been completely ejected from the halo.

We note that Fig. 4.5 does not include a comparison to observational data of the gas in galaxies. Indeed many observations of the cosmic neutral hydrogen (HI) density, probing cold, star forming gas, at different epochs up to $z \sim 4$ are available (Rao et al., 2006; Prochaska and Wolfe, 2009; Delhaize et al., 2013; Rhee et al., 2013; Zafar et al., 2013), however a comparison to these data should relate the gas in the simulation to the HI density. In Rahmati and et al. (in prep.) the HI density of gas particles in the EAGLERef-L100N1504 simulation is computed using the fitting functions from Appendix A of Rahmati et al. (2013). These fitting functions depend on the gas temperature and total ionization rate (photo-ionization plus collisional ionization) and are based on cosmological simulations coupled with radiative transfer. In Rahmati and et al. (in prep.) the cosmic HI density is compared with observations across redshift 0 to 6. The observations are reasonably flat from redshift 4 to 1, although error bars span up to 0.5 dex, with a similar scatter from different observational data sets. The local Universe observations have a HI density that is ~ 0.4 dex below the redshift 4 to 1 values. The simulation has a similar normalisation to the observations at redshift 4, which is maintained until redshift $\sim$2. Below redshift two the HI density begins to turn over reaching a value ~ 0.6 dex below the redshift 4 to 2 value. The simulation value for the HI density lies within the observation errors above redshift 0.5, although the local Universe values are too low by $\sim$0.2 dex. Overall, the level of agreement recovered for the global HI density is in good agreement with the observations for the reference model. Unfortunately, the error on the observations implies that this measure will not help in determining which models considered here are more or less valid descriptions of the real Universe.

4.5 Galaxy properties

We now look at the evolution of the normalisation of the GSMF and SSFR. These measures of galaxy populations are difficult to reproduce by galaxy formation simulations, both for hydrodynamic simulations (Weinmann et al., 2012) and semi-analytic models (Mitchell et al., 2014; White et al., 2014). Mitchell et al. (2014) and White et al. (2014) both suggest an adjustment to the accretion timescale of the gas as a potential solution to these issues. However, in hydrodynamical simulations the accretion (and reaccretion) of gas onto halos and galaxies is mapped explicitly by the hydrodynamics in the simulation. There is no freedom to adjust accretion rates independently of the stellar and AGN feedback. Furthermore, as these models reproduce the stellar mass - halo mass relation, galaxies at redshift zero live in similar halos to their observed equivalent and should experience similar accretion histories to observed galaxies. Thus, the model variations considered here are a good test of whether a realistic accretion model can resolve the discrepancies with observations.

Another potential solution proposed by White et al. (2014) is for stronger feedback in low mass halos that form early. The stellar feedback schemes implemented here (with the exception of SNConstant) all result in a feedback energy that is stronger at higher redshift, as we will show in Section 4.5.3. We examine what impact the variety of models considered here have on the normalisation of the GSMF and the evolution of SSFRs and if the proposed solutions are plausible.

4.5.1 Normalisation of the galaxy stellar mass function

The normalisation of the GSMF at 10^{10} and 10^{9} $M_{\odot}$ as a function of redshift are shown in Fig. 4.6 and 4.7, respectively, for the 9 EAGLE models. Due to the noise in the GSMFs from fluctuations caused the small number of objects in the box, we fit the GSMF at redshifts 0.1, 0.5, 1, 2, 3 and 4 for all models with a Schechter function, weighting the fit by the Poisson errors per stellar mass bin. Using the Schechter function, we then compute the normalisation at 10^{10} or 10^{9} $M_{\odot}$. At 10^{10} $M_{\odot}$ the

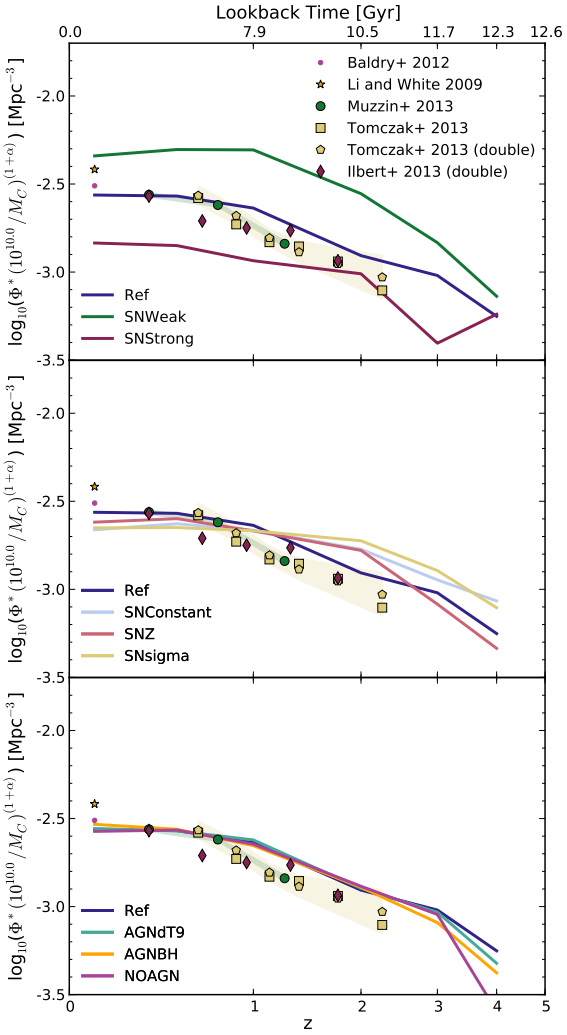

Figure 4.6: The normalisation of the GSMF at a stellar mass of 10^{10} $M_{\odot}$, as determined by Schechter function fitting, as a function of redshift for 9 EAGLE simulation in 25 cMpc boxes. Observational data from Li and White (2009), Baldry et al. (2012), Tomczak et al. (2014), Ilbert et al. (2013) and Muzzin et al. (2013) are shown. The Ref model normalisation remains constant to redshift 1. By redshift 2, the normalisation drops to 0.6 dex below the redshift 0 value. The normalisation continues to drop to redshift 4. The stellar feedback variations in the middle panel have a higher normalisation at redshift 2. The AGN models closely follow the Ref model.

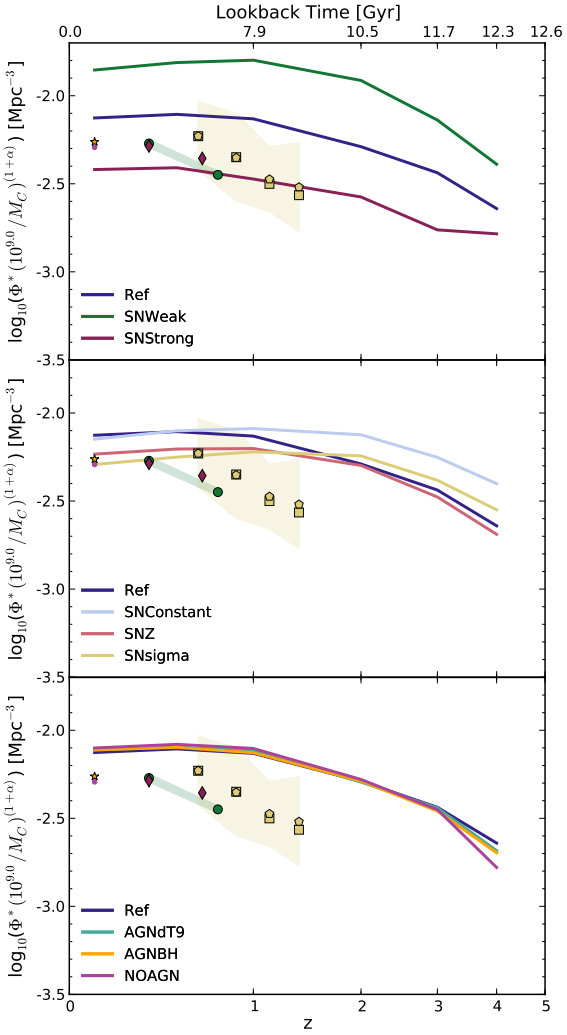

Figure 4.7: The normalisation of the GSMF at a stellar mass of 10^9 $M_\odot$, as determind by Schechter function fitting, as a function of redshift for 9 EAGLE simulation in 25 cMpc boxes. Observational data from Li and White (2009), Baldry et al. (2012), Tomczak et al. (2014), Ilbert et al. (2013) and Muzzin et al. (2013) are shown. The Ref model normalisation remains constant to redshift 1. By redshift 2, the normalisation drops to 0.4 dex below the redshift 0 value. The normalisation continues to drop to redshift 4. The stellar feedback variations in the middle panel show similar slopes to the Ref model, although the initial normalisation varies by 0.2 dex. The AGN models closely follow the Ref model.

Reference model normalisation is constant from redshift 0.1 to 1, it then drops by 0.4 dex between redshift 1 and 2. The normalisation continues to decrease with increasing redshift. At 10^9 $M_\odot$ the normalisation of the Reference model follows a similar trend with redshift, it remains constant from redshift 0.1 to 1 and then decreases. However the decrease at between redshift one and two at 10^9 $M_\odot$ is less pronounced than at 10^{10} $M_\odot$, decreasing by only 0.2 dex.

The SNWeak model, in the top panels, has a similar shape to the Reference model at 10^{10} and 10^9 $M_\odot$, however the normalisation is ~ 0.3 dex higher. The SNStrong model is ~ 0.3 dex lower in normalisation than the Reference model at both stellar masses at redshift zero. At 10^9 $M_\odot$ the normalisation evolves in a similar way to the Reference model, however at 10^{10} $M_\odot$ the normalisation in SNStrong is almost flat from redshift 0.1 to 2. Note however that at 10^{10} $M_\odot$ at redshift 0.1 there are very few objects in this simulation, and the sampling at a given stellar mass decreases with increasing redshift.

In the middle panels of Fig. 4.6 and 4.7 we compare the evolution of the normalisation of the GSMF of the stellar feedback variations. At redshift 0.1 at 10^{10} $M_\odot$ the normalisation of the different models is within 0.1 dex. The evolution of all models is similar from redshift 0.1 to 1. From redshift 1 to 2 the drop in normalisation of the Reference model is not present in any of the feedback variations. The SNZ model does drop in normalisation above redshift 2, while the SNConstant and SNsigma models have a shallow decrease in normalisation above redshift 2. At 10^9 $M_\odot$ the normalisation of SNConstant is within 0.1 dex of the Reference model at redshift 0.1, while the SNZ and SNsigma models are 0.2 dex lower. All 3 models evolve in a similar way at 10^9 $M_\odot$ as at 10^{10} $M_\odot$. As the stellar feedback variations were shown to build up stellar mass earlier than in the Reference model, a higher normalisation at high redshifts is expected.

The normalisation remains constant up to redshift 2 and gradually drops at higher redshift, as for 10^{10} $M_\odot$. The SNZ and SNsigma models are 0.2 dex lower than the Reference model at redshift 0.1 at a mass of 10^9 $M_\odot$. Both models have no evolution in normalisation until redshift 2, at which point the SNZ models de-

creases with redshift with a steeper slope than SNsigma, again similar to the behaviour at 10^{10} $M_{\odot}$.

In the bottom panels of Fig. 4.6 and 4.7 we compare the different AGN variations to the Reference model. As expected, at 10^{10} $M_{\odot}$ and 10^{9} $M_{\odot}$ the AGN variations make no noticeable difference to the normalisation of the GSMF.

In each of the panels of Fig. 4.6 and 4.7 observational data from Li and White (2009), Baldry et al. (2012), Tomczak et al. (2014), Ilbert et al. (2013) and Muzzin et al. (2013) is shown. The normalisation of the observed GSMF is produced using Schechter functions, as for the simulation data. The shaded regions show the estimated errors on the Schechter function parameters. We only include the data from observations at redshifts where the mass completeness limit of the survey fall above 10^{10} or 10^{9} $M_{\odot}$ for Fig. 4.6 and 4.7, respectively. Note that some previous studies, for example Weinmann et al. (2012), not only considered the observational data, but also extrapolated the data below the mass completeness limit of the survey. We refrain from doing this, as even when considering the same survey data, as in Muzzin et al. (2013) and Ilbert et al. (2013), the extrapolations can be significantly different.

Comparing the Reference model to the observations at 10^{10} $M_{\odot}$, between redshift 0.1 and 1 the simulation is within less than 0.1 dex of the observations. Above redshift 1 the simulation follows the same decreasing trend with redshift as the observations, however the normalisation is too high by ~ 0.1 dex. At 10^{9} $M_{\odot}$ the normalisation of the Reference model is too high by ~ 0.2 dex between redshift 0.1 and 1, although the offset from observations with redshift, up to redshift 1, remains almost constant, showing a similar trend. Between redshift 1 and 2 the normalisation in the observations and simulation both drop, although the drop in observations is larger than in the simulation. Above redshift 2 there is no observational data to compare to, making it difficult to determine if the slope of the decrease in normalisation is too shallow or not. As none of the stellar feedback variations have a drop in normalisation for the GSMF between redshifts 1 and 2, they do not reproduce the observed trend.

Previous hydrodynamical simulations have struggled to reproduce the observed evolution of the normalisation of the low mass GSMF, in particular the decrease in normalisation of the GSMF between redshift 1 and 2. In the Reference model presented here however the shape of the normalisation with redshift is similar to that observed, although we recognise that an offset of 0.1-0.2 dex does remain. The agreement of the shape is driven by the strong stellar feedback at high redshifts, as will be discussed in Section 4.5.3. Including somewhat stronger feedback at high redshifts should reconcile the simulation with observations, without jeopardising the agreement with other observations. The Reference model outperforms all the stellar feedback variations considered. The improvement in the Reference model relates to the lower star formation rate density at high redshifts and slower build up of stellar mass presented in Section 4.4.

4.5.2 Specific star formation rates

We now look to see if the discrepancy between the observed and simulated SSFRs can be reconciled by reasonable accretion histories and/or strong feedback at high redshifts. Fig. 4.8 shows the median SSFR for star forming galaxies as a function of redshift for each of the nine models in two stellar mass ranges, $10^{9.0}$ M$_\odot$< M$_*$ < $10^{9.5}$ M$_\odot$ and $10^{9.5}$ M$_\odot$<M$_*$ < $10^{10.0}$ M$_\odot$. The dashed curves enclose the 10 to 90$^{\rm th}$ percentiles for the Reference model, which have a range of about 0.6 dex at all redshifts. For the Reference model the SSFR for both mass bins shown fall from redshift three to zero by ~ 1 decade. Above redshift 3 there are too few galaxies in the mass ranges considered in the box available.

In the top panel we compare the Reference model to SNWeak and SNStrong. While SNWeak was shown to produce more stellar mass by redshift zero, the median SSFRs in the galaxy mass ranges shown falls below the Reference model by 0.2 dex. This is due to the stellar mass forming at early times, in small halos, in this model and depleting the supply of star forming gas at later times. The SNStrong model shows the opposite effect, with higher SSFRs by up to 0.2 dex, particularly in the mass range $10^{9.5}$ to 10^{10} M$_\odot$. The SNStrong model suppresses star formation

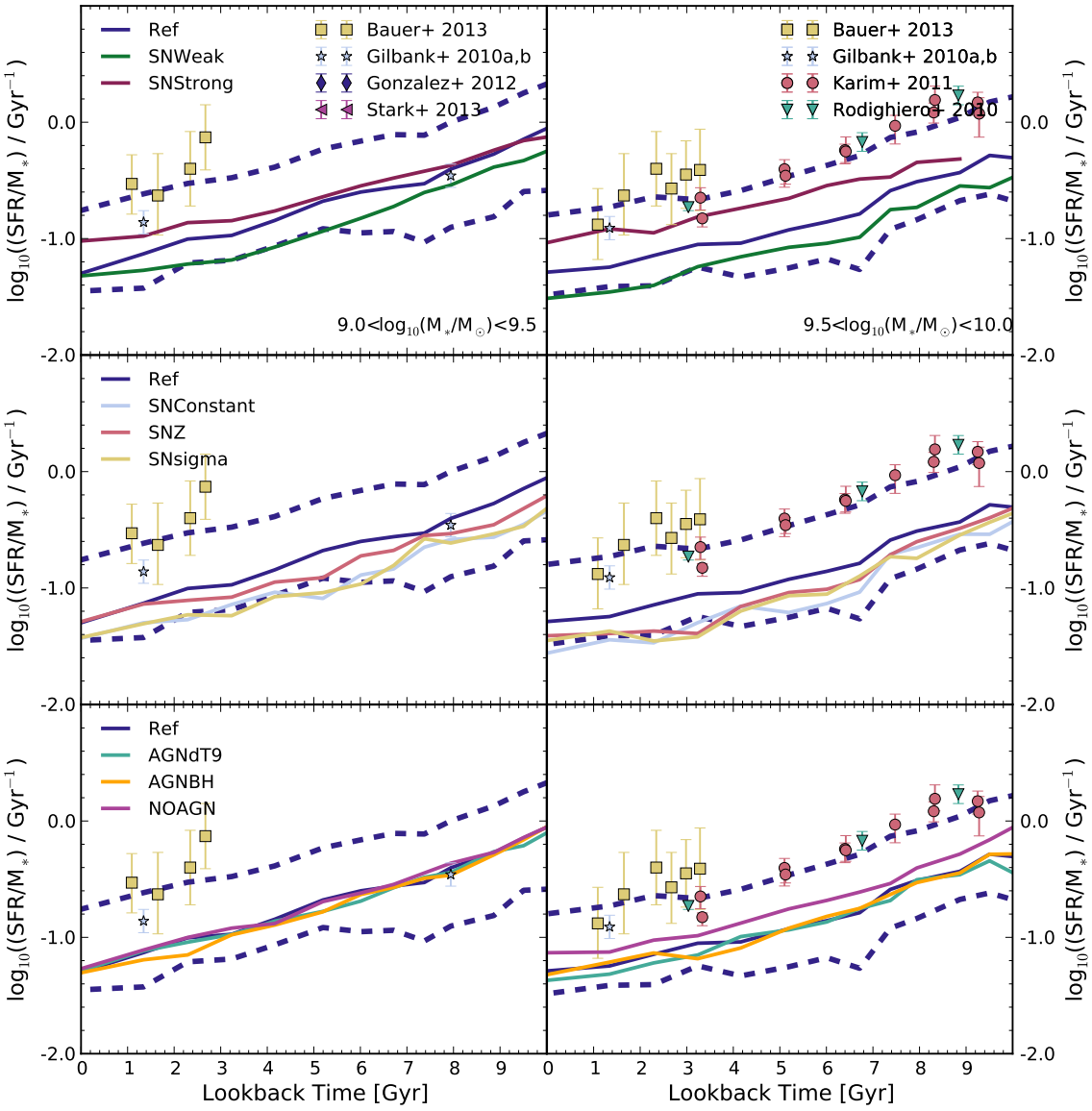

Figure 4.8: The median specific star formation rate for star forming galaxies in 2 mass bins from 9 EAGLE simulations in 25 cMpc boxes. Dashed curves enclose the 10th to 90th percentiles for the Ref-L025N0376 model. Observational data from Gilbank et al. (2011), Bauer et al. (2013), Karim et al. (2011), Rodighiero et al. (2010), González et al. (2012) and Stark et al. (2013) are shown. The median SSFRs for all models lie within 0.4 dex of each other, which is similar to the scatter for the Ref model. None of the models with reasonable redshift 0.1 GSMFs can reproduce the observed SSFR normalisation.

strongly in the early Universe, resulting in lower mass galaxies in a given halo (see Fig. 4.2). Galaxies in the mass range shown are in halos larger than in other models, and hence have larger gas accretion rates. With a larger gas supply, higher SFRs can be produced. Although if the feedback is strong enough to eject the gas from the halo, the SFR will rapidly decrease, which is why in small halos the galaxies can not grow.

In the middle panel the Reference model is compared to the stellar feedback variations. The SSFRs in theses variations are typically lower than the Reference model by up to 0.2 dex in the stellar mass bins shown, as with SNWeak. This is due to the star formation peaking at high redshift.

For the AGN models in the bottom panel the variations all have similar SSFRs in the mass bin of 10^{9} to $10^{9.5}$ $M_{\odot}$. In the higher mass bin of $10^{9.5}$ to 10^{10} $M_{\odot}$ the NOAGN model has a higher SSFR than other models by 0.1 to 0.2 dex. This model forms too much stellar mass in massive galaxies, without AGN feedback quenching galaxies.

Observational data from Gilbank et al. (2011), Bauer et al. (2013), Karim et al. (2011), Rodighiero et al. (2010), González et al. (2012) and Stark et al. (2013) is plotted on each panel of Fig. 4.8. In *Chapter 3* it was shown that the SSFRs in Ref-L100N1504 were low compared to observations by 0.2 to 0.4 dex. Looking at all variations that reproduce the observed normalisation of the stellar mass density at redshift zero, none of these models can produce the observed SSFRs. The only model that produces SSFRs significantly higher than the Reference model and closer to the observations is SNStrong. However, this model was shown to under predict the normalisation of the GSMF at redshift 0.1, with galaxies that have too little stellar mass per halo. The star formation rate history is too low in normalisation and much flatter than the observational data and the stellar mass density is very different in shape to the observations. In hydrodynamical simulations we can not appeal to gas accretion/recycling timescales to resolve this issue. As a result, it appears that the observed SSFRs are not compatible with other observations, such as the stellar masses. It is certainly concerning that such a drastic feedback model

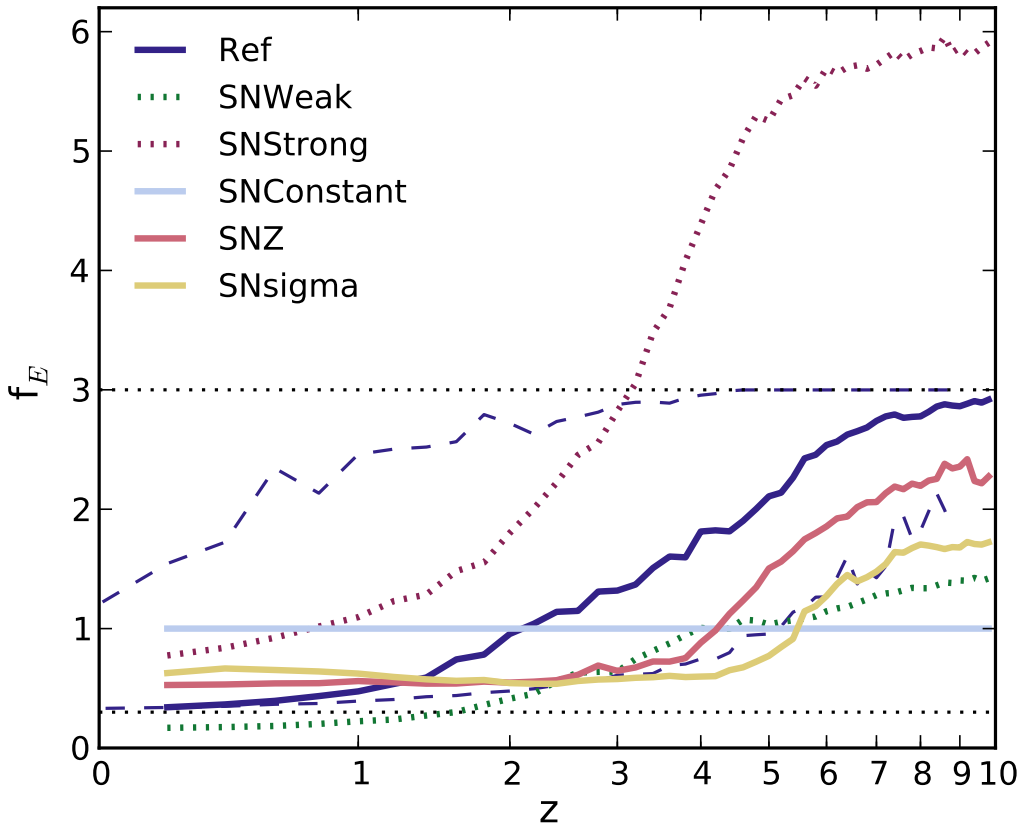

Figure 4.9: The median energy fraction from stellar feedback as a function of redshift for 6 EAGLE simulations with varying stellar feedback in 25 cMpc boxes. The dashed curves enclose the 10^{th} to 90^{th} percentiles for the Ref model. The horizontal dotted lines show the minimum and maximum energy fraction for the Ref model.

is required to reproduce the observed normalisation in SSFRs.

4.5.3 Evolution of the feedback energy

As the primary differences in the outcome of the simulations arise from the different stellar feedback parameters and implementations, we look at how the stellar feedback varies with redshift. In Fig. 4.9 we show the median feedback energy as a function of redshift. The energy fraction is given in units of available energy from Type II supernova and can range between 3 and 0.3 for the Reference model and

stellar feedback variations. The energy fraction ranges between half the Reference model values and double for SNWeak and SNStrong, respectively. The Reference model median energy fraction is close to the maximum of 3 at redshift 10, when star forming gas is metal poor and dense. As the stars enrich the surrounding gas and the mean density decreases, the median energy fraction decreases. The dashed curves enclose the $10^{\rm th}$ to $90^{\rm th}$ percentiles. The scatter of energy fractions at any redshift is large.

Comparing the Reference model to SNZ, above redshift 1.5 the feedback energy fractions are lower, removing less gas from the ISM and allowing extra star formation and an earlier build up of stellar mass. This excess star formation occurs in galaxy cores, as shown in Crain et al. (2015). More star formation and an earlier build up of stellar mass is also expected, to a larger extent, for the SNsigma and SNConstant models. It is these variations in the energy fraction of stellar feedback with redshift that drives the differences in the different models. These variations are the reason for the Reference model performing better than other stellar feedback variations relative to observations. This implies that such an implementation is more realistic on the scales resolvable within the simulation than other choices. However, the normalisation in the GSMF remains ~ 0.1 and ~ 0.2 dex too high at redshift 2 at 10^{10} and 10^{9} $\rm M_{\odot}$, respectively, although the issue of the shape can be resolved. The normalisation in the SSFRs remains a concern and can only be reproduced by an extreme feedback model, that can not reproduce any other observations considered here.

4.6 Discussion and Summary

We have presented the evolution of 6 variations of the EAGLE Reference model, and compared their evolution in 25 cMpc boxes. The 6 variations reproduce the observed redshift 0.1 galaxy stellar mass function over the mass range 10^{8} to $10^{10.5}$ $\rm M_{\odot}$, as does the Reference model. Two further models that over and under shoot the GSMF at redshift 0.1 are also considered, to bracket the observed relation. We

have compared the global galaxy properties of these models, including the star formation rate density, stellar mass density and gas density in galaxies. We also considered the evolution of the normalisation of the GSMF at 10^{10} and 10^{9} $M_{\odot}$ and the SSFRs of galaxies in the mass ranges $10^{9.0}$ $M_{\odot}$< M_{*} < $10^{9.5}$ $M_{\odot}$ and $10^{9.5}$ $M_{\odot}$<M_{*} < $10^{10.0}$ $M_{\odot}$. Both of these measures are typically difficult to reproduce in simulations.

- Stellar feedback variations

  The 3 stellar feedback variations, which change the dependencies of the feedback energy, show differences in the star formation history and stellar mass growth relative to the Reference model. All 3 models have higher star formation rate densities between redshift 2 and 4 and higher stellar mass densities across this redshift range. In the case of two of the models, SNConstant and SNsigma, the increases in these quantities relative to the Reference model result in too high a normalisation relative to observations, ruling out their galaxy properties as having an evolution similar to real galaxies. As these models build up stellar mass at earlier times than the Reference model, the evolutionary trend in the normalisation of the GSMF, in particular the drop in normalisation from redshift 1 to 2, is not reproduced by any of the stellar feedback variations. The differences relative to the Reference model stem from the stronger feedback at high redshift in the Reference model. This results in accreted gas being delayed at forming stars at early times. It is the time delay between the gas in fall and the conversion of gas to stars, stemming from the changes in feedback, that results in the Reference model best reproducing the observational trends with redshift.

- AGN variations

  AGNdT9 and AGNBH vary the AGN feedback parameters and the black hole accretion rate, respectively. These models show no significant difference in their global evolution properties, the normalisation of the GSMF and the SSFRs. However, this finding may be somewhat driven by the small box size,

of 25 cMpc, available for this study. The AGN variations are expected to impact only massive galaxies, in halos greater than 10^{12} $M_{\odot}$, but there are only a small number of halos this massive in 25 cMpc boxes. We aim to repeat this study in 50 cMpc boxes, when the simulation data is available.

The third AGN variation, NOAGN, varies significantly from the Reference model at redshifts below 2, in spite of the small numbers of high mass galaxies, implying that excluding AGN strongly impacts the galaxy population. In this model there is no black hole growth or AGN feedback, so galaxies in halos greater than 10^{12} $M_{\odot}$ continue to grow when they are quenched in the Reference model. This model produces around 50% too much stellar mass relative to observations at redshift zero, and increasing the box size will only increase this discrepancy. Interestingly, the NOAGN model is the only model to reproduce the normalisation of the star formation rate density observations below redshift 2, suggesting less quenching by AGN feedback is required in the simulation at redshift 2. However, without significantly suppressing star formation at higher redshifts (from redshift 2 to 0 covers 10.5 Gyr of the Universe) the stellar mass density at redshift 0 will be too high.

- Evolution of the normalisation of the GSMF

  The normalisation of the GSMF across different redshifts has typically been difficult to reproduce in simulations, both in semi-analytics and hydrodynamical simulations. However, we have shown that the reference model shows a similar trend to the observations, with little evolution from redshift 0 to 1 followed by a drop in normalisation from redshift 1 to 2. It is the strong stellar feedback at high redshift in the Reference model that results in the trend being reproduced.

- Specific star formation rates

  As with the normalisation of the GSMF, the evolution of the normalisation of the SSFRs with redshift is difficult to reproduce in simulations. Two solutions have been suggested based on semi-analytic models (White et al., 2014;

Mitchell et al., 2014). One solution is to adjust the reaccretion timescale of the gas onto halos, however in the hydrodynamical simulations considered here the accretion and reaccretion of gas is determined by the hydrodynamical forces acting on the gas particles. These simulations have no freedom in this regard, but do have reasonable gas accretion histories for the galaxies, assuming the cosmology is accurate. A second proposed solution is to increase stellar feedback in halos that form early. Because of the dependencies of our stellar feedback models on properties that evolve with redshifts (such as metallicity) all the stellar feedback models, except SNConstant, inject more energy per feedback event at high redshift than low redshift. The model that shows the most variation with redshift, while reproducing the observed redshift 0.1 GSMF, is the Reference model. While this model has a similar shape as the observations for the SSFRs as a function of redshift, the normalisation is too low by 0.2 to 0.4 dex. The only model that has a similar normalisation to the observed SSFRs is the SNStrong model. However, this model does not form sufficient stellar mass by redshift zero, the normalisation of the GSMF is too low and neither the shape nor the normalisation of the star formation rate density agree with observations. It is concerning that such an extreme stellar feedback model, that is in tension with all other observations considered, is necessary to reproduce the observed normalisation of the SSFRs. The SSFRs, as observed across cosmic time, can not be explained by current theoretical galaxy formation models and it is not clear how they can be consistent with the observed stellar mass evolution of galaxies.

In summary, the findings of this study show that AGN feedback, so long as it is included, does not significantly change the evolution history of the simulated universe (in volumes of $(25\ \mathrm{cMpc})^3$). Stellar feedback variations, however, impact the time at which stellar mass is formed, and strong suppression of star formation above redshift 2 is necessary to reproduce observed global galaxy evolution properties and the evolution of the normalisation of the GSMF. None of the models that reproduce the redshift 0.1 GSMF can reproduce the observed SSFRs, raising the

question of how the data can be correct.

Chapter 5

Investigating galaxy formation physics in EAGLE

5.1 Introduction

Observational data provides a snapshot in time of galaxy formation. Theoretical models allow us to connect the snapshots across time to interpret the observational results and understand how galaxies form and evolve. Hydrodynamical simulations are a particularly useful way to interpret observational data, because they provide 3D information on galaxies, Milky Way galaxies and above are resolved by $(10^{4.5})$ or more particles . The explicit modelling of environmental processes are also mapped through the hydrodynamics.

As the EAGLE project simulations include subgrid physics for the processes considered to be important for galaxy formation and was shown to reproduce many local Universe observations (Schaye et al., 2015), as well as the evolution of galaxy properties (*Chapter 3*), this simulation suite is a useful tool for understanding the evolution of galaxies. In this analysis we begin to explore the physics of galaxy formation using the Ref-L100N1504 simulation. In particular we look to answer the following questions,

- What galaxies dominate the stellar masses and star formation rates, and how does this evolve with time?
- How do central and satellite galaxies differ? What role does environment play in galaxy star formation rates?

- What shapes the GSMF?
- How are cosmic downsizing and the hierarchical growth of the ΛCDM universe compatible?

The analysis in this chapter is based on the EAGLE Ref-L100N1504 simulations, as has been described in *Chapter 2*. This chapter is laid out as follows: In Section 5.2 we ask what galaxies dominate the stellar mass and star formation in the simulation. The role of different galaxy types in building up the shape of the GSMF is discussed in Section 5.3. In Section 5.4 cosmic downsizing in the simulation is considered. Finally in Section 5.5 we present the conclusions.

5.2 Which galaxies dominate?

We begin by determining the galaxy stellar mass range that dominates the stellar mass and star formation rate densities in the simulation and how this changes with time. The evolution of these fractions provide an insight into the role of different galaxies in determining global properties, such as the stellar mass density. This is followed by a description of the ICL in the simulation, which comprises 20% of the total stellar mass density at redshift 0. We then consider the split by central and satellite galaxies in the simulation and the differences in their properties. As satellite galaxies of a given stellar mass live in halos of higher mass than their central equivalent, this analysis probes the evolving role of galaxy environments.

5.2.1 Stellar mass and star formation rate fractions

In *Chapter 3* it was shown that the stellar mass density in the Ref-L100N1504 simulation reproduced the observed stellar mass density to within 20% as well as the evolution of the GSMF from redshift 0 to 7. Also the observed star formation rate density was shown to be reproduced within 0.2 dex, along with the trends in specific star formation rates. This level of agreement between the simulation and observations implies that the simulated universe is reasonably representative of the

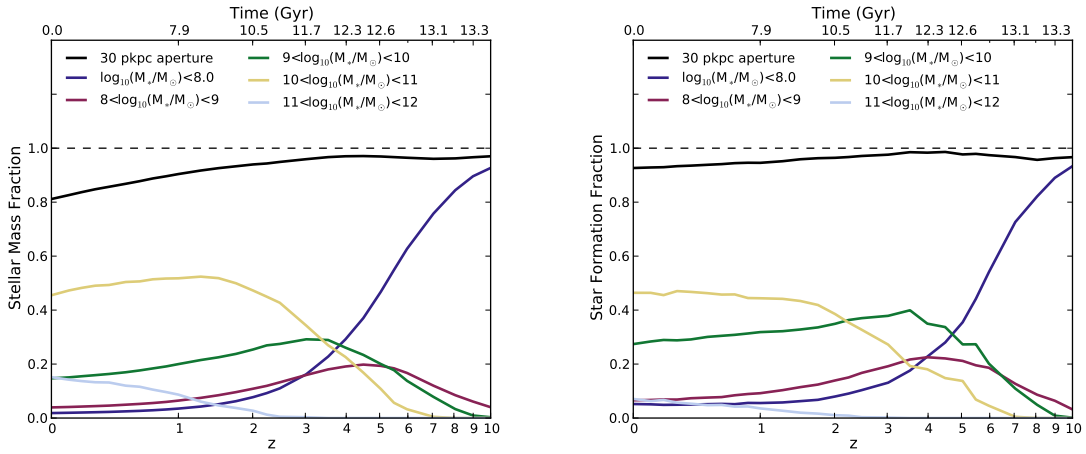

Figure 5.1: The fraction of stellar mass and star formation in galaxies of different stellar masses as a function of redshift in the left and right panels respectively. The total stellar mass and star formation fraction within the inner 30 pkpc of galaxies is shown in black. The fraction within different galaxy mass ranges are also shown, given in the legend. Galaxies of stellar mass between 10^{10} and 10^{11} $M_{\odot}$ dominate the stellar mass fraction and star formation fraction from redshifts 2 and 3 respectively. At higher redshifts lower mass galaxies dominate, as anticipated from a hierarchical growth model. From redshift 2.5 galaxies in the mass range 10^{11} to 10^{12} $M_{\odot}$ contribute to the total stellar mass and star formation, however in both the left and right panels the total contribution is small. The exponential break in the GSMF and strong increase in passive fraction with stellar mass explain this small contribution for the stellar mass and star formation rates respectively, for this population.

observed Universe. We now use the simulation to look at the role of galaxies of different masses in the growth of stellar mass by studying the stellar mass fractions and star formation rate fractions as a function of redshift in Fig. 5.1.

In the left panel the black curve shows the total stellar mass within galaxies (as defined in *Chapter 2*). Above redshift 3, over 95% of the stellar mass is within galaxies. Below redshift 3 the mass outside galaxies, defined as the ICL, grows to 20%. We will discuss the galaxies with which the ICL is associated in Section 5.2.2. Galaxies below 10^8 $M_\odot$ dominate the stellar mass fraction above redshift 4. Although these galaxies are poorly resolved in our simulation, this dominant contribution from small galaxies in the early Universe is expected due to the hierarchical growth in a ΛCDM cosmology, collapsing small objects first, that accrete further mass and merge. Indeed we find these same low mass galaxies are also the major contributors to the stellar mass density at high redshifts in the higher resolution run, Recal-L025N0752, presented in Schaye et al. (2015) and *Chapter 3* (not shown). From redshift 4 galaxies of 10^{10} to 10^{11} $M_\odot$ dominate the stellar mass fraction. Although these are not the most massive galaxies in the simulation, their abundance results in them making the most significant contribution to the stellar mass fraction. Consequently, the evolution of integrated galaxy properties, e.g. stellar mass density and star formation rate density, is quite sensitive to this population. From redshift 2 there is a contribution from the most massive galaxies, 10^{11} to 10^{12} $M_\odot$, growing to 18% by redshift zero. Due to the sharp cut off in the GSMF, this relatively small contribution from high mass galaxies is expected.

In the right panel, the black curve shows the total star formation fraction within galaxies. Unlike the stellar mass fraction, the amount of star formation outside of galaxies is small, reaching a maximum of 7%. This 7% is within the FoF halos, but outside the central 30 pkpc. As with the stellar mass density, the smallest galaxies, less than 10^8 $M_\odot$, dominate the star formation fraction at high redshifts. From redshift 2 galaxies in the mass range 10^{10} to 10^{11} $M_\odot$ dominate. The most massive galaxies, from 10^{11} to 10^{12} $M_\odot$, only contribute below redshift 2 at a maximum of 7%. The low star formation fraction from massive galaxies is again expected,

not only because of the low number density of massive galaxies but also the high passive fraction at high stellar masses, as shown in *Chapter 3*.

5.2.2 Intra-cluster light

The ICL in the simulation, defined as the stellar mass outside a 3D 30 pkpc aperture centred on the galaxies (excluding satellite galaxies), amounts to up to 20% of the stellar mass by redshift zero, as shown in Fig. 5.1. We note that it is difficult to determine the ICL around galaxies observationally, as galaxies do not have sharp edges and the low surface brightness around galaxies is difficult to detect due to uncertainty in the background subtraction. As a result, in this section we only aim to provide a qualitative comparison to the observed ICL. We will show that the majority of stars contributing to the ICL were not formed outside the 30 pkpc aperture that defines galaxies and that the halos and galaxies in the simulation with which the ICL is associated are rare and massive.

We begin by discussing where the mass associated with the ICL is formed. The ICL is negligible above redshift two, but grows from below 5% at redshift two to 20% at the present day. There are three possible sources for the mass in the outskirts of galaxies: in situ star formation, the stripping or destruction of satellite galaxies, or the scattering of stars from inside galaxies. In the simulation only a maximum of 7% of all star formation occurs outside the inner 30 pkpc of galaxies, which can not account for all the mass associated with the ICL. In situ star formation thus is not the dominant source of stellar mass in the ICL. This implies that the stellar mass originates from stars formed within galaxies that are later removed. We leave a more detailed investigation of the origin of the ICL to future work.

In Fig. 5.2 we show the ratio of stellar mass in the ICL to the stellar mass in the halo as a function of halo stellar mass at redshift zero for apertures of 30 and 100 pkpc. For halos with stellar masses below $10^{10.5}$ $\mathrm{M}_{\odot}$ the ICL fraction is typically below 5%. It is only for halo stellar masses above $10^{10.5}$ $\mathrm{M}_{\odot}$ (halos masses of $\sim 10^{12}$ $\mathrm{M}_{\odot}$) that halos have over 5% of their stellar mass associated with the ICL. This reaches almost 40% for the most massive galaxies in the simulation. The ICL

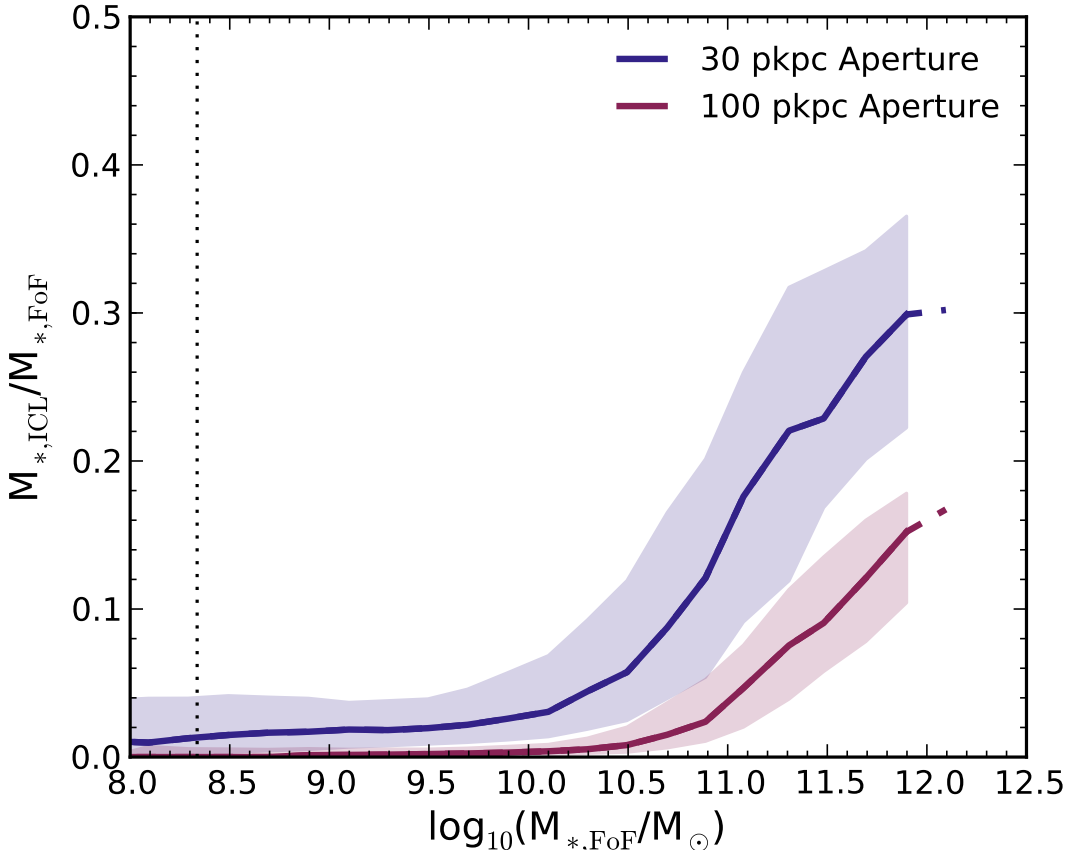

Figure 5.2: The fraction of stellar mass in the ICL relative to stellar mass in the FoF halo as a function of FoF stellar mass at redshift zero. ICL mass is defined as any stellar mass outside a radius of 30 pkpc or 100 pkpc of a galaxy for blue and red curves respectively. The solid curves shows the median relation, while the shaded region shows the 10^{th} to 90^{th} percentiles. A dashed curve is shown for the median values when the number of galaxies per bin falls below 10. The vertical dotted line denotes the mass of a galaxy with 100 baryonic particles, as a resolution guide. Central galaxies below $10^{10.5}$ $M_\odot$ have less than 5% of their halo stellar mass contributing to the ICL, galaxies more massive than this have a maximum 40% of their mass in the ICL. While the ICL fraction decreases for the larger aperture of 100 pkpc, the behaviour with stellar mass is qualitatively similar.

fractions seen here are broadly consistent with values found in observations of galaxy clusters (e.g. Theuns and Warren, 1997; Presotto et al., 2014).

Of course the percentage of mass associated with the ICL is dependent on the definition. For example if we increase the aperture to 100 pkpc the ICL accounts for only 10% of the stellar mass. However, the galaxies with which this mass is associated are still the most massive, as seen in Fig. 5.2. As when using 30 pkpc aperture, most of the stellar mass in the ICL was not formed in situ (not shown).

Given that the ICL is built up from stars born within galaxies, as opposed to in situ star formation, and that the ICL is associated with is the most massive and rare objects, we conclude that the ICL mass in the simulation shows qualitative agreement with observations.

5.2.3 Centrals and satellites

While considering what galaxies dominate, we consider the division of galaxies by centrals and satellites. Central galaxies are those with the largest mass within a friends of friends halo. All other galaxies within the halo are satellites. For a given stellar mass, satellite galaxies probe more dense environments than centrals, as a satellite has at least one more massive neighbouring galaxy.

Fig. 5.3 shows the fraction of satellite galaxies as a function of stellar mass at redshifts from 4 to 0.1. At redshift 2 and above there is a maximum of 35% satellites in the simulation at a given stellar mass, and this decreases with increasing stellar mass. From redshift 1 to 0.1 $\sim 40\%$ of galaxies up to 10^{10} $M_{\odot}$ are satellites, with a reasonably flat relation with stellar mass. At higher stellar masses the satellite fraction decreases sharply. The stellar mass at which this decrease in satellite fraction begins moves to higher stellar masses with decreasing redshift.

The increase in satellite fractions with redshift results from the increased abundance of more massive halos with time and the resulting accretion of small halos (and galaxies). The growth of more massive halos also explains the increase in the galaxy mass at which the satellite fraction decreases; with more massive halos more massive galaxies can grow. The decrease in satellite fraction at high stellar

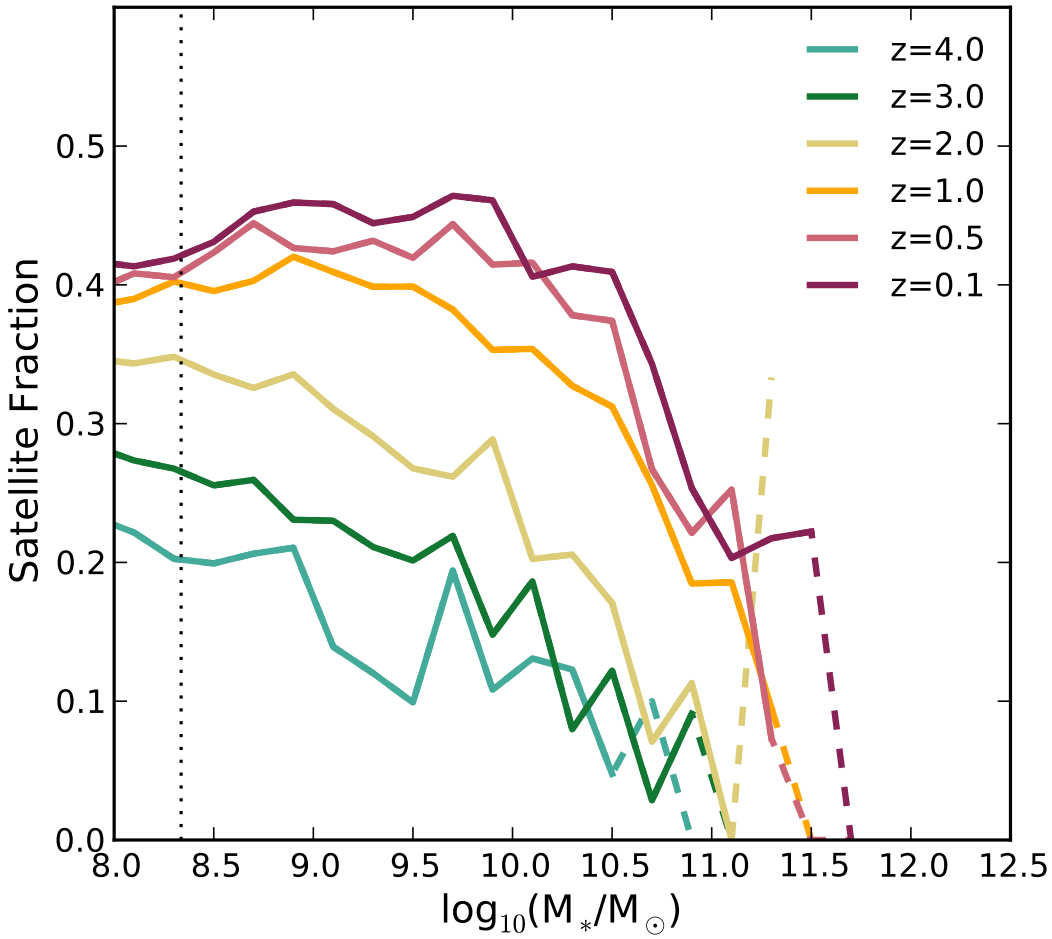

Figure 5.3: The fraction of satellite galaxies as a function of stellar mass for redshifts shown in the legend. Stellar mass bins of 0.2 dex are used, when there are fewer than 10 galaxies per bin curves are dashed. The vertical dotted line shows where a galaxy stellar mass is equal to that of 100 baryonic particles. The satellite fraction increases with decreasing redshift. At redshift 2 and above the satellite fraction falls with stellar mass. Below redshift two the satellite fraction is reasonably flat with stellar mass at low masses and then falls off at higher masses. The mass at which the satellite fraction decreases increases with decreasing redshift.

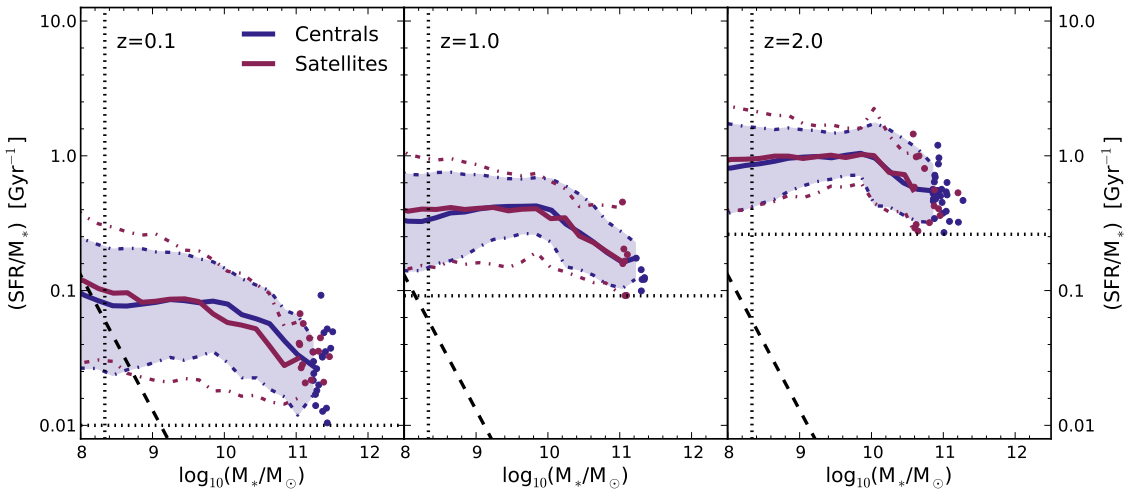

Figure 5.4: The SSFRs for galaxies star forming central and satellite galaxies in blue and red respectively at redshifts 0.1, 1 and 2. Medians are shown by solid curves and the 10^{th} to 90^{th} percentile scatter is enclosed in the dot-dashed curves (shaded for central). When there are fewer than 10 galaxies per bin data points are plotted. The horizontal dotted line shows the SSFR that defines passive galaxies, not galaxies below this line are not included in computing the median and percentiles. The vertical dotted line corresponds to a stellar mass resolved by 100 baryonic particles. The diagonal dashed line represents a SFR of 10 star forming gas particles. We see that at all redshifts the central and satellite galaxies show similar SSFRs as a function of stellar mass, with a similar scatter.

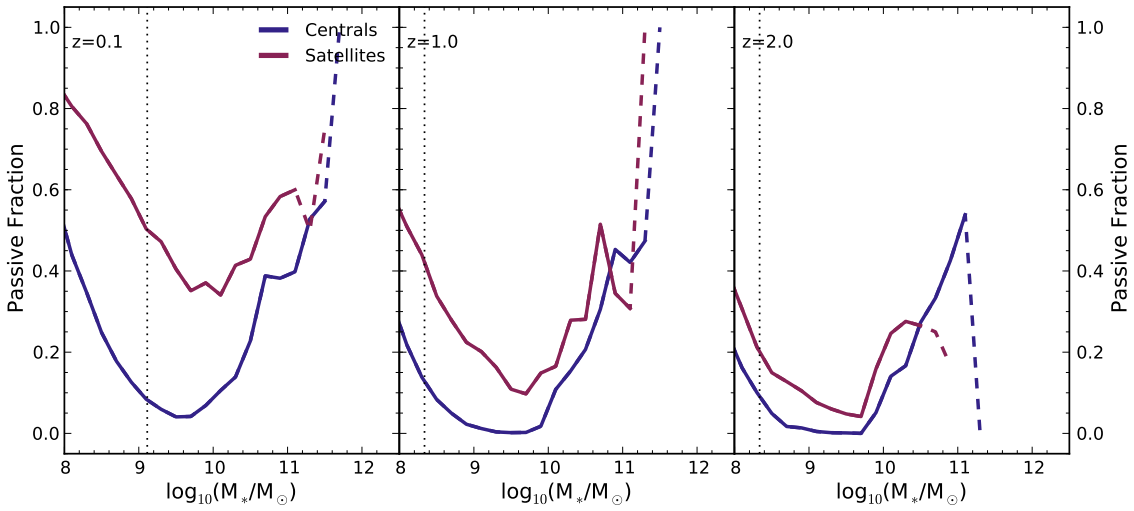

Figure 5.5: The passive fractions of galaxies as a function of stellar mass for central and satellite galaxies in blue and red respectively at redshifts 0.1, 1 and 2. When there are fewer than 10 galaxies per bin curves are dashed. The horizontal dotted line shows the maximum mass corresponding to 100 baryonic particles or 10 star forming gas particles for the SSFR cut (i.e. the maximum mass where the dashed diagonal line or dotted vertical line cross the horizontal dotted line in Fig. 5.4). This acts as a resolution guide. While the SSFRs for satellites and centrals are found to be similar in the simulation, the passive fraction of satellite galaxies is higher than that of centrals.

masses is due to the flattening of the stellar mass to halo mass relation, reducing the growth rate of galaxies as their halos grow. Furthermore, as satellites are defined as not the most massive galaxy in a halo, we expect the satellite fraction to be lower for massive galaxies.

5.2.4 Evidence of environmental quenching

We now briefly look at some of the properties of central and satellite galaxies in the simulation, to determine the role of environment on galaxies. In Fig. 5.4 the median SSFRs for star forming galaxies as a function of stellar mass are shown for centrals and satellites in blue and red respectively at redshifts 0.1, 1 and 2. Star forming galaxies in the simulation are those with a SSFR above a given limit. The limit is chosen to be an order of magnitude below the observational data and shown on each panel as a horizontal dotted line. The SSFRs from centrals and satellites follow very similar trends in the simulation, both in normalisation and scatter. One notable difference is that galaxies within the $10^{\rm th}$ and $90^{\rm th}$ percentiles have lower and higher SSFRs by up to 0.4 and 0.2 dex respectively at redshift 0.1. The lower SSFRs could signal environmental quenching where galaxies are transitioning to the passive population. The higher SSFRs could be induced by interactions, galaxies have been shown to experience enhanced SSFRs when going through a merger (or close encounter) both in observations and simulations (Patton et al., 2013). Aside from these minor differences in the scatter, the two galaxy types are indistinguishable by their SSFRs for star forming galaxies. The SSFRs from the simulation were shown to be low in normalisation by up to 0.3 dex. It is clear that this offset from observations is not specifically the result of either central galaxies or satellites.

Observationally the SSFRs of centrals and satellites are found to be similar (Wetzel et al., 2012). To date these trends, in particular the similarity of the median SSFR for centrals and satellites has proven difficult for many models (e.g. Font et al., 2008; Hirschmann et al., 2014). These previous studies of these trends in centrals and satellites stem from SAMs, where assumptions are required about the stripping of galaxies and the associated timescales. In hydrodynamical simulations, how-

ever, the environmental processes are traced explicitly, and so these simulations are a better probe of quenching due to stripping and strangulation. The simplifications required in the SAMs is likely to be the cause of the satellite SSFRs being too low relative to centrals (see Guo and et al., in prep., for a comparison of EAGLE to SAMs).

In Fig. 5.5 the fraction of central and satellite galaxies that are passive as a function of stellar mass are shown in blue and red respectively at redshifts 0.1, 1 and 2. The dotted horizontal lines act as a resolution guide, so galaxies with stellar masses below this guide are not well resolved (either in stellar mass or their star formation rate). While the SSFRs showed very similar behaviour for satellites and centrals, for the passive fractions we find the behaviour is different. At redshift 2 there are only $\sim 5\%$ more passive satellites than centrals at a given mass. At this redshift, environment appears to only play a minor role in galaxy quenching. By redshift 0.1 there are between 10% and 30% more passive satellites than centrals.

The increased passive fraction for satellite galaxies implies that in the simulated galaxies in more dense environments are more likely to have their star formation quenched. As the star forming satellites and centrals display similar behaviour, the quenching process must be rapid, as it does not result in a population of galaxies with SSFRs below the main sequence that decreases the median satellite SSFRs. A future EAGLE paper will explore the environmental quenching mechanisms in detail.

5.3 The build up of the GSMF

Following the success of the simulation in reproducing the observed GSMF from redshift 0 to 7, as shown in *Chapter 3*, we now look at how the Schechter (1976) shape of the GSMF emerges over time. To do this we split the galaxy population to gain a physical understanding of how the galaxies grow. We show in Fig. 5.6 the GSMF from redshift 4 to 0.1 split by star forming and passive galaxies in the left and right panels. Passive galaxies are determined using a cut in SSFR, discussed

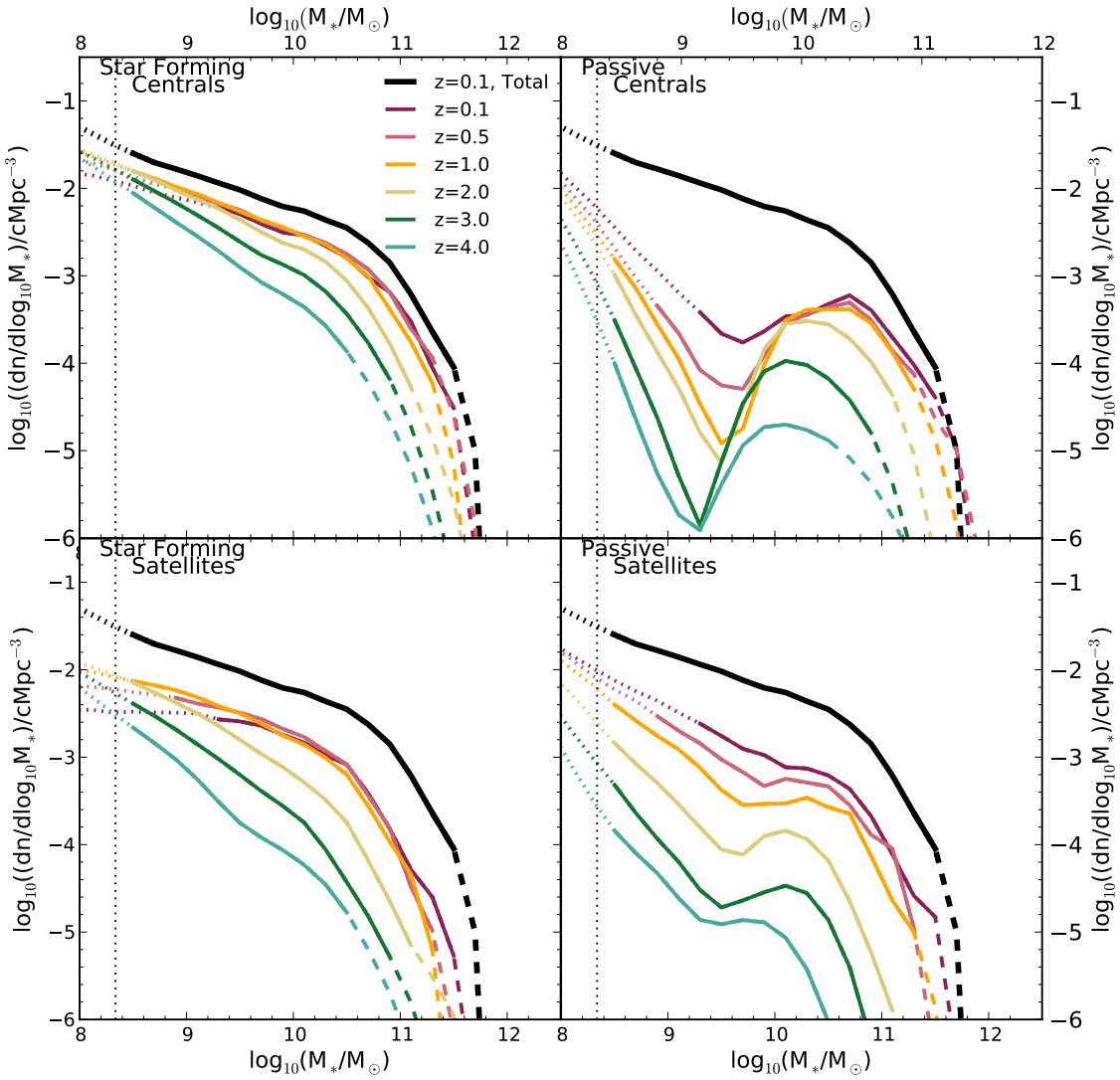

Figure 5.6: The evolution of the GSMF split by star forming (left) and passive (right) galaxies and central (top) and satellite (bottom) galaxies. Curves are coloured by redshift, as seen in the legend. Redshifts from 4 to 0.1 are shown, covering the past 12 Gyr of evolution. The redshift 0.1 GSMF for all galaxies is reproduced on each panel for reference as a solid black line. When there are fewer than 10 galaxies per bin, curves are dashed. The vertical dotted line corresponds to a stellar mass resolved by 100 baryonic particles. The star forming GSMFs are similar in shape, however the passive GSMFs differ, particularly in the mass range $10^{9.5}$ to $10^{10.5}$ $\mathrm{M}_{\odot}$. Note the upturn of the passive GSMF at masses below 10^{9} $\mathrm{M}_{\odot}$is an artefact of the resolution.

in Section 5.2.3. Central galaxies and satellite galaxies are shown on the top and bottom of the figure.

The star forming central and satellite galaxies dominate the total number density of galaxies from 10^9 $M_\odot$ to the break in the GSMF, as can be seen by comparing the normalisation in the left and right panels. Focusing on the star forming GSMF for both central and satellite galaxies (top and bottom left hand panels), there is an increase in the total number density from redshift 4 to 1. Over this redshift range the slope of low mass galaxies flattens. The flat slope of the SSFR with stellar mass for galaxies below $10^{10.5}$ $M_\odot$, as seen in Fig. 5.4, results in all galaxies taking a similar time to increase their stellar mass by a given factor, so galaxies move smoothly to higher masses resulting in a smooth slope for the GSMF, without a pile up at any mass. The flattening of the low mass slope can be explained by faster growth in galaxies relative to their halos; the stellar mass to halo mass relation is very steep below $10^{10.5}$ $M_\odot$, resulting in a wider range of stellar masses for given halo mass range (see Bower et al., 2012). The flat slope ends in an exponential break for the star forming GSMF when galaxy masses reach above $10^{10.5}$ $M_\odot$ from redshift 1 to 0, or at lower stellar masses for higher redshifts.

Comparing satellites and centrals for the star forming GSMFs (bottom and top left hand panels), they have a similar shape and evolution. The primary differences are in the normalisation, which is lower for satellite galaxies. This results from the hierarchical growth in ΛCDM, where small halos, hence galaxies, form first and merge to produce larger halos with satellites. The other difference in the GSMFs of star forming satellites and centrals is the largest stellar masses reached. This difference is a consequence of the definition of centrals, which are the most massive galaxy in a halo, so their maximum mass is larger than the maximum mass for satellites.

The break in the star forming GSMFs creates a bump in the passive GSMFs (right hand panels), where active galaxies become passive. In our simulation this is a result of the onset of black hole feedback heating the gas in galaxies, disrupting gas in flowing from the halo and shutting down star formation. The emergence of

passive central galaxies is not however a sharp function of halo mass, the subgrid accretion model depends on gas properties in the proximity of the black hole, as opposed to global galaxy or halo properties, although these are strongly correlated (Bower and et al., in prep.).

Unlike the star forming GSMF, the passive GSMFs of centrals and satellites are different. For central galaxies (upper right hand panel), there is a bump in the GSMFs at high masses. The normalisation of this bump increases with decreasing redshift, as does the stellar mass at which the bump peaks. As more massive galaxies can grow with time in the simulation, the number density and masses of the massive central galaxies that are quenched increase. Between 10^9 and 10^{10} $M_\odot$ there is a dip in the passive central GSMF, with very few passive central galaxies in this mass range. The dearth of passive central galaxies between masses of 10^9 $M_\odot$ and 10^{10} $M_\odot$ is a result of such low mass galaxies not yet experiencing effective black hole feedback; their black holes are not yet growing efficiently. There is an increase in the normalisation of the passive central GSMF at masses less than 10^9 $M_\odot$. We have already shown this to be a resolution effect in *Chapter 3*.

Looking now at the passive satellite GSMF (bottom right hand panel), above redshift 1 there is an exponential break at high masses, a dip at intermediate masses followed by a rise at low masses. This shape is similar to that for passive centrals, although the dip is less pronounced. Below redshift 1 the dip is filled in, producing a flatter GSMF at intermediate masses than for central passive galaxies. The increased number density of passive satellites from 10^9 to 10^{10} $M_\odot$ at all redshifts shows the important role of environmental suppression, due to processes such as satellite stripping and strangulation, on satellites, as was also seen in Fig. 5.5.

We have described how the shape of the GSMF, a Schechter function with power law slope for masses below the characteristic mass and an exponential break above the characteristic mass, is built up through the growth of galaxies and their quenching. At masses below the exponential break, the shape of the GSMF is primarily determined by star forming central galaxies at masses below the exponential break. Although, as shown in Fig. 5.3, the satellite fraction below 10^{10} $M_\odot$

is around 40% below redshift 1, this implies satellite galaxies also play an important role. However, of the satellite galaxies, the star forming population dominate above the resolution limit, and below 10^{10} $M_{\odot}$ (Fig. 5.5), and these star forming satellites have a similar shape in their GSMF to the centrals.

Quenching of star formation by black hole feedback is key to reproducing the cut off at high masses. Other models have also shown the need for quenching of massive objects to produce the Schechter function form, for example Bower et al. (2006) (semi-analytics) and Peng et al. (2010) (empirical model).

5.4 Cosmic downsizing

One of the important goals of galaxy formation is to understand the role of cosmic downsizing in the Universe, and its compatibility with the hierarchical growth of a ΛCDM cosmology. To determine the existence of cosmic downsizing in the simulation we present the cosmic star formation rate density ($\rho_{\rm SFR}$) in Fig. 5.7. The left and right panels show the contribution of galaxies in different stellar mass ranges based on the galaxy mass at redshift z (i.e observed mass, $M_* = M_*(z)$) and at redshift zero ($M_* = M_*(z=0)$), respectively.

Before we discuss cosmic downsizing, we define what we mean by this term, as there are two subtly different definitions that can be implied. Cosmic downsizing refers to the more significant contribution of present day high mass galaxies to the star formation rate density at high redshift relative to their contribution today. This definition can be extended to imply that low mass galaxies dominate the star formation rate density in the present day however we do not include this in our definition, although we will comment on the role of low mass galaxies. An alternative definition of cosmic down sizing is that massive galaxies make a more significant contribution to the star formation rate density at high redshift than today. The difference in this definition is that galaxies are selected based on their 'observed' mass, not their mass at redshift 0.

The hierarchical growth of structure results in low mass galaxies, less than 10^8

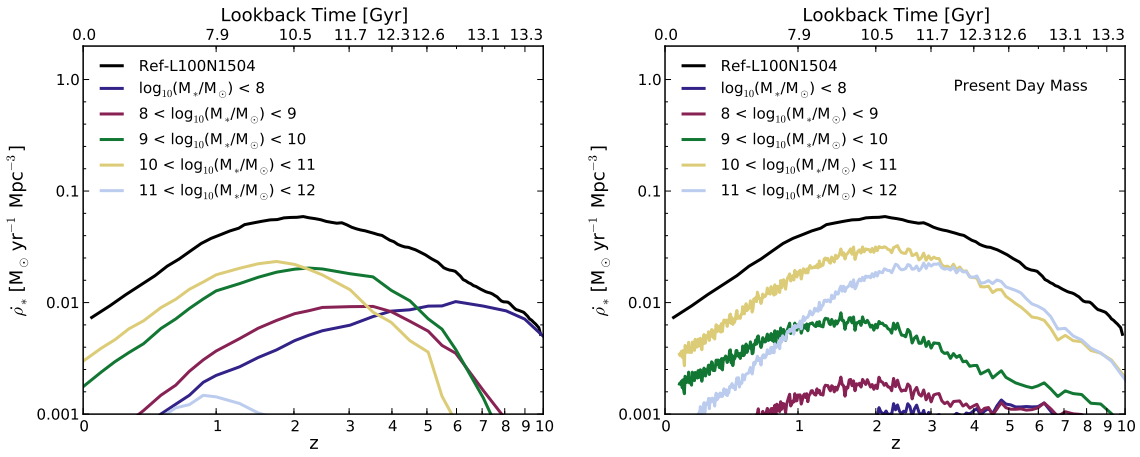

Figure 5.7: The star formation rate density as a function of redshift showing the contribution of galaxies selected by their 'observed' and present-day stellar masses in the left and right panels. The black solid line is the total $\rho_{\rm SFR}$ and the coloured lines show the contribution of different mass galaxies to the total star formation rate density. For the observed masses, the contribution of all galaxies rises at high redshift, but is decreasing since the peak in star formation, with a similar slope to the total. The most massive present-day galaxies in the simulation, 10^{11} M$_\odot$ to 10^{12} M$_\odot$, account for much of the early star formation above redshift 3, dropping steeply relative to other mass bins from redshift three. The lower mass galaxies of 10^{10} M$_\odot$ to 10^{11} M$_\odot$ dominate the star formation rate density over the past 11 Gyr, since redshift 3.

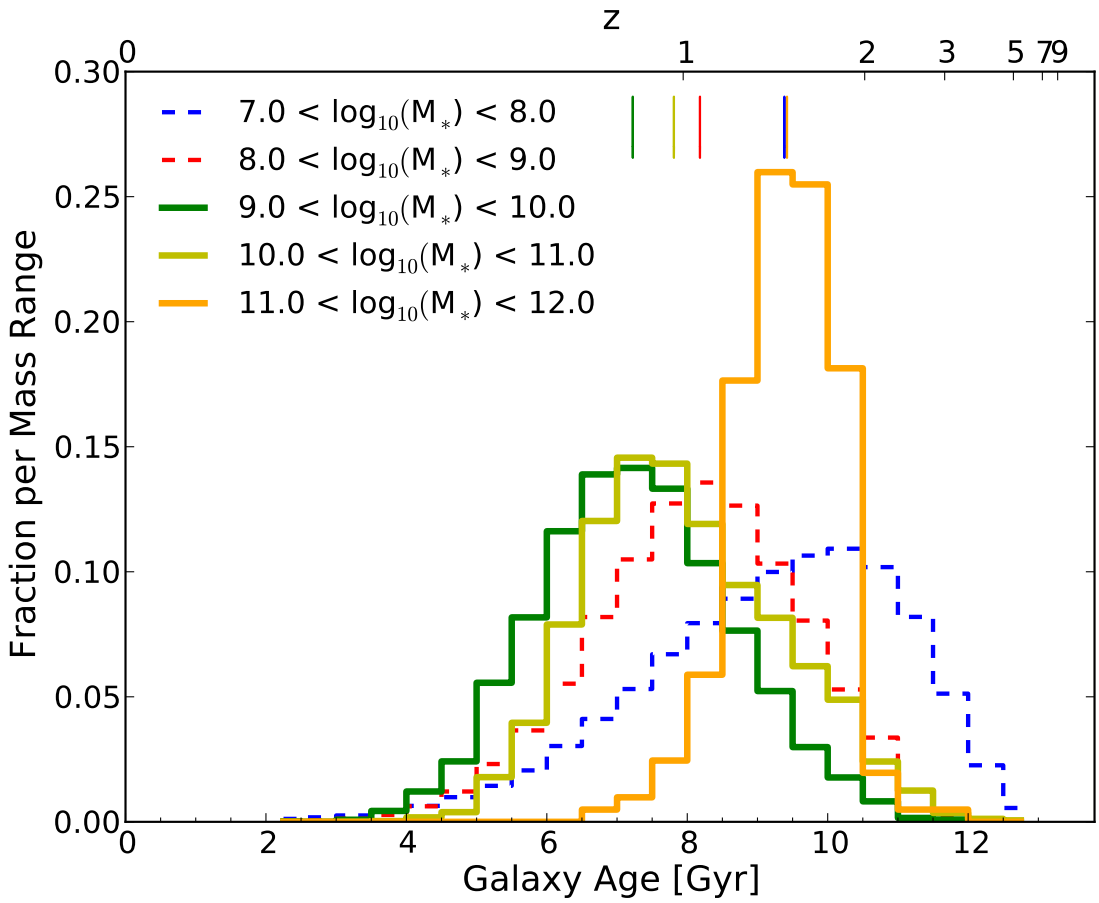

Figure 5.8: The age distribution of galaxies for galaxies in different (redshift zero) stellar mass ranges. Each curve is normalised to one, showing the fraction of galaxies within a given mass range with a given age. Median ages are shown using coloured vertical lines on the upper x-axis. The age of a galaxy is determined by the mass averaged age of the star particles within the galaxy. The most massive galaxies, in the mass bin $10^{11} < M_* < 10^{12}$ $M_\odot$, are typically older than galaxies between 10^9 and 10^{11} $M_\odot$, which is qualitatively in agreement with observations. Sampling issues affect galaxies with masses below 10^9 $M_\odot$, which lead to poorly resolved galaxy histories in low mass objects and hence poorly resolved galaxy ages.

$M_\odot$, dominating the $\rho_{\rm SFR}$ at early epochs above redshift 5, with galaxies of 10^{10} $M_\odot$ to 10^{11} $M_\odot$ dominating from redshift two, as can be seen in the left hand panel of Fig. 5.7. High mass galaxies clearly do not dominate the star formation rate density at high redshift in a hierarchical model, as low mass galaxies must grow to form high mass galaxies. This highlights that the alternative definition of cosmic downsizing stated above is inconsistent with a ΛCDM model. In fact, the highest masses of 10^{11} to 10^{12} $M_\odot$ contribute very little to $\rho_{\rm SFR}$ at any redshift and 10^{10} to 10^{11} $M_\odot$ galaxies are the majority contributor since redshift 3, with a 45% contribution at redshift 0 (see Fig. 5.1, accounts for 50% of star formation in galaxies).

To determine if there is cosmic downsizing in the simulation we look to the right hand panel. We measure the "archaeology" of the present-day galaxy populations, i.e. the age distribution of the stars, to determine the contribution of the progenitors of the most massive galaxies to $\rho_{\rm SFR}$ at high redshift. The past contribution to $\rho_{\rm SFR}$ from present-day galaxies is determined from the simulation using the birth time of star particles in galaxies within a range of redshift zero stellar masses. The birth times are binned in time to produce the $\rho_{\rm SFR}$ at redshifts above zero.

The progenitors of the most massive present-day galaxies, 10^{11} $M_\odot$ to 10^{12} $M_\odot$, peak in their contribution to $\rho_{\rm SFR}$ at a higher redshift than other stellar mass ranges. The peak is followed by a sharp decline, much sharper than that seen in lower galaxy mass ranges. In our model the sharp downturn is due to the quenching experienced by massive galaxies from the onset of AGN feedback, which shuts off the star formation of massive galaxies. Clearly, this is evidence that the most massive present day galaxies were more star forming, and contributed more to $\rho_{\rm SFR}$ in the past. We conclude that cosmic downsizing is seen in the simulation.

Above redshift 3 the progenitors of galaxies with present-day masses of 10^{10} $M_\odot$ to 10^{11} $M_\odot$ have a similar contribution to $\rho_{\rm SFR}$ as the progenitors of the most massive galaxies. Below redshift 3 the progenitors of these galaxies dominate $\rho_{\rm SFR}$, as the progenitors of 10^{11} $M_\odot$ to 10^{12} $M_\odot$ galaxies decrease sharply in their star formation. As the downturn in star formation is seen for galaxies of all masses, the contribution of low mass galaxies has not significantly increased with decreasing

redshift. 10^{10} to 10^{11} $M_\odot$ galaxies remain the majority contributor to the $\rho_{\rm SFR}$ at redshift zero. While these galaxies dominate, they also experience a downturn in the total $\rho_{\rm SFR}$, which is also seen in the lower stellar mass ranges. This downturn is less sharp than for the most massive galaxies.

As mentioned, this analysis is equivalent to studying the age distribution of stars in galaxies of different mass ranges. In Fig. 5.8 the distribution of galaxy ages at redshift zero is shown. Galaxy ages from the simulation are determined by computing the mass averaged age of stars in a galaxy. As we are not applying SPS models to produce luminosity weighted ages, the results are only qualitatively comparable to observations. Here we see that the most massive galaxies of 10^{11} $M_\odot$ to 10^{12} $M_\odot$ typically have stellar population ages of 8 to 10 Gyr, older than galaxies with masses of 10^{10} $M_\odot$ to 10^{11} $M_\odot$, and with a narrower distribution. 10^{9} $M_\odot$ to 10^{10} $M_\odot$ galaxies are younger again. The peak in the stellar age distribution is determined by the peak in the left panel of Fig. 5.7. These results show qualitative agreement with observational data, where more massive galaxies are typically found to be older (e.g. Gallazzi et al., 2005). The trend of decreasing stellar masses having younger stellar populations reverses for galaxies smaller than 10^{9} $M_\odot$. While there is some observational evidence that galaxies of masses around 10^{9} $M_\odot$ are older on average than previously reported (Kauffmann, 2014), in the simulation the older ages of galaxies with masses of 10^{9} $M_\odot$ or less is driven by the artificially bursty star formation histories in these objects due to the poor sampling of star formation in small objects in the simulation (*Chapter 3*).

In summary, the compatibility of cosmic downsizing with hierarchical growth is sensitive to the definition. Based on the definition adopted here, the simulation exhibits cosmic downsizing, in the sense that the most massive galaxies contributed more significantly to $\rho_{\rm SFR}$ in the past. This is not to say, however, that $\rho_{\rm SFR}$ is now dominated by low mass galaxies. Rather, the progenitors of 10^{10} to 10^{11} $M_\odot$ galaxies have made the largest contribution to $\rho_{\rm SFR}$ since redshift 3, 11 Gyr ago.

5.5 Summary and conclusions

The aim of this analysis is to investigate galaxy formation physics using the EAGLE Ref-L100N1504 simulation. As this simulation was shown to reproduce both local and high-redshift observations, it is a useful tool to explore galaxy formation.

- In Section 5.2.1, it was shown that from redshift 2 to the present day, over the past 10 Gyr, galaxies in the stellar mass range 10^{10} to 10^{11} $M_{\odot}$ dominate the stellar mass density and the star formation rate density. This highlights the importance of these galaxies in reproducing global galaxy properties.

- To determine the role of environment on galaxies, satellite and central galaxies were compared in Section 5.2.4. Star forming centrals and satellites are indistinguishable based on their star formation rates, it is only by considering the passive fractions of these galaxy that any differences can be seen in the simulation. The elevated passive fraction for satellites relative to centrals shows the effect of environmental processes on the galaxies, while the similar properties for the star forming populations implies the quenching mechanisms work on short timescales.

- We have considered the role of star formation rates and environment on the shape of the galaxy stellar mass function. Central galaxies are the majority contributor at all redshifts and stellar masses. Star forming galaxies determine the shape for low mass galaxies, although at these masses, below redshift 1 satellites do play some role. The increase in passive galaxies at high masses causes the exponential break.

- We adopt the definition of cosmic downsizing to mean that the present-day most massive galaxies contributed more significantly to the star formation rate density in the past. We have shown cosmic downsizing to exist in the simulation, and its compatibility with hierarchical growth. While the most massive galaxies contributed more significantly to the star formation rate density in the past, the star formation rate density is not dominated by low

mass galaxies at the present day. In fact, 10^{10} to 10^{11} $M_{\odot}$ galaxies are the majority contributors, and have been over the past 11 Gyr.

This investigation into the galaxy population in the EAGLE Ref-L100N1504 simulation is a step towards understanding the physics of galaxy formation. However many more questions remain. For example, while quenching has been shown to be efficient, both through environmental and feedback processes, the mechanisms require further investigation. Related to quenching, is the observed phenomenon of galaxy conformity, where satellite galaxy properties are correlated with the properties of centrals. Semi-analytic models of galaxy formation have, to date, struggled to explain this (Kauffmann et al., 2013). We will readdress this question using a full hydrodynamical simulation.

Chapter 6

Where are the reionizing photons produced?

6.1 Introduction

Different measurements of the epoch of reionization suggest that reionization occurred between redshifts 4 and 12, spanning 1.5 Gyr in time. The Thompson optical depth of electrons, as measured by Planck (Planck Collaboration et al., 2013), suggests the epoch of reionization occurred around redshift 11.5, while measurements of the Ly-α forest constrain the epoch of reionization to above redshift 6 (Mortlock et al., 2011). The nature of the sources that reionized the universe is uncertain. Quasars and very bright galaxies are one potential source of ionizing photons, however, they are too rare to contribute significantly (e.g. Shapiro, 1986; Grissom et al., 2014). Pop III stars produce a significant amount of ionizing photons per star (Schaerer, 2002), however their contribution to to total budget is unknown. The most obvious source of photons to reionize, though, is galaxies, which are the focus of this study.

The galaxies that reionized the Universe are difficult to detect observationally due to their faintness. However, recently observations in the infra-red (e.g. HST and CANDELS) aim to select Lyman break galaxies by detecting dropouts at rest frame wavelength 912Å. This allows the selection of star forming galaxies up to redshift $\sim$7, down to magnitudes of -18, or stellar masses around 10^8 $M_\odot$. There is an indication that the sources detected do not produce sufficient ionizing photons to maintain the ionization of the IGM (Bolton and Haehnelt, 2007), which raises interesting questions about the abundance and nature of these sources. There are many further observational studies underway, or planned, study this epoch, e.g.

the spaced-based mission of the James Webb Space Telescope, or studies to detect the 21cm transition of neutral hydrogen between spin states.

To complement the challenging studies being carried out observationally, simulations anticipate what will be found. In this work we study the galaxies that reionized the Universe from a theoretical point of view, availing of state of the art hydrodynamical simulations, from the EAGLE simulation suite. There are many theoretical approaches to studying the epoch of reionization. One approach is to apply radiative transfer codes to trace the evolution of the photons produced in hydrodynamical simulations (e.g. Wise et al. (2014), So et al. (2014)). Radiative transfer codes trace the paths of photon packets produced by star particles in simulations, and each photon packet can include the energy distribution of the photons. The interaction of the photons with gas particles and the ability of the photons to escape from galaxies can then be directly followed. While such methods are an excellent tool for understanding how the photons progress through the neutral hydrogen and reionize it, the computational expense of radiative transfer methods limits the size of regions in the Universe that can be explored ($\mathcal{O}$(1 cMpc)). Furthermore, as the simulated universe evolves and more stellar mass is produced the number of photons to trace becomes prohibitive to the evolution of the simulation, restricting the simulation to high redshifts, typically above redshift 5. These restrictions imply that for this study, of a relatively large galaxy population evolved to redshift zero, that radiative transfer is not a viable option.

To compliment radiative transfer methods, SAMs are often used to model large scale galaxy populations, e.g. Raičević et al. (2011), Wise et al. (2014). Such models avail of N-body or Monte Carlo simulations to describe the evolution of the underlying dark matter structures. Baryonic processes are modelled through a series of equations approximating the flow of baryons into and out of the dark matter structure. Such methods are very flexible in allowing variations of the physics and parameters to be tested.

A parallel exploration into the reionization of the Universe is underway using dark matter only simulations and radiative transfer, with simple prescriptions for

how stellar mass populates the halos. Such methods allow the exploration of a much larger volume of the Universe, however, they require simple approximations for the evolution of the baryons producing the photons.

To study large galaxy populations in a more self consistent way, we use hydrodynamical simulations. The EAGLE simulations trace particles with baryonic masses between 10^5 and 10^6 $M_\odot$ evolved simultaneously with the dark matter. Subgrid physics recipes are applied locally to the gas particles, on the scale of the SPH kernel. Box sizes range from 25 to 100 cMpc, covering a representative sample of the Universe. The EAGLE simulations were shown to reproduce many redshift zero galaxy properties as well as the evolution of galaxies across all observed redshifts (*Chapter 3 Schaye et al., 2015)*. These results can provide some confidence in the simulation at even higher redshifts, where observational data is limited.

As it is not feasible to run radiative transfer on the EAGLE simulations due to their size, an intricate study of escape fractions and recombinations is not possible. However, we can consider how many photons are produced by the stellar mass formed in the simulation and what galaxy populations are producing them. Under some simplified assumptions relating to the escape fraction and the recombination rate, we can also estimate when sufficient photons are produced by the galaxies to reionize the universe. We can ask what galaxies prove most important to this process and, as the simulations are run to redshift zero, we also study of the descendants of the galaxies that reionized the Universe.

This chapter is organised as follows: In Section 6.2 we discuss the assumptions used to determine the number of ionizing photons produced, and how many are needed to reionize the universe. We begin by justifying the use of the EAGLE simulations in this study in Section 6.3.1 In Section 6.3.2 and 6.3.3 we determine when sufficient ionizing photons are produced by galaxies in the simulation to reionize the Universe and what are the properties of these galaxies. We then look at what becomes of the stellar mass that produced ionizing photons in Section 6.3.4. Finally, we summarise the findings in Section 6.4.

6.2 Method

The EAGLE simulations Ref-L100N1504 and Recal-L025N0752 are used here, both simulations are described in *Chapter 2*.

For this analysis, the stellar mass formed in the simulation is converted to a number of ionizing photons, as detailed in Section 6.2.1. The remaining requirements in determining when galaxies produce sufficient photons to reionize the Universe depends on how many photons can escape from the dense environments of galaxies and the rate at which atoms recombine. Both of these quantities are uncertain, the assumptions applied in this work are discussed in Section 6.2.1.

6.2.1 Counting Ionizing Photons

The first step in considering the reionization of the neutral hydrogen is to compute the number of ionizing photons produced. For this work we have taken a simple approach to estimate the number density of ionizing photons, n_γ and emissivity, ϵ, produced by stars in the simulation. (The emissivity is the number density of ionizing photons produced per unit time.) The values for n_γ and ϵ can, in principle, be computed using a population synthesis code, assuming an IMF and metallicity. The spectral energy distribution of a simple stellar population can be integrated above 1 Ryd, the photon energy required to ionize a hydrogen atom, to determine the number of ionizing photons produced per unit mass. This calculation is carried out in Murray and Rahman (2010) for a Chabrier IMF, giving a value of 5×10^{46} photons $\mathrm{s}^{-1}\mathrm{M}_\odot^{-1}$, which is consistent with the value determined in Schaerer (2002).

The high energy photons required to ionize hydrogen are typically produced by massive stars, which have an average lifetime, $\mathrm{t}_{\mathrm{avg}}$, of 30 Myr. This gives the number of ionizing photons per stellar mass, N_γ. To determine ϵ and n_γ, we use

$$\varepsilon(\mathrm{z}) \quad = \quad \frac{N_\gamma \sum_{N_{\mathrm{gas}}(\mathrm{z})} \dot{m}_*(\mathrm{z})}{V}, \tag{6.1}$$

$$\begin{aligned} n_\gamma(\mathrm{z}) &= \int_\infty^{\mathrm{z}} \varepsilon(\mathrm{z}')\mathrm{d}\mathrm{z}' \\ &= \frac{N_\gamma \sum_{N_{\mathrm{star}}(\mathrm{z})} m_{*,\mathrm{init}}(\mathrm{z})}{V}, \end{aligned} \tag{6.2}$$

where $N_{\mathrm{gas}}(\mathrm{z})$ and $N_{\mathrm{star}}(\mathrm{z})$ are the number of gas and star particles, respectively, in the simulation at redshift z, $\dot{m}_*(\mathrm{z})$ is the star formation rate at redshift z, which gives the stellar mass produced per unit time, and $m_{*,\mathrm{init}}(\mathrm{z})$ is the initial mass of a star particle.

We are interested in comparing the number density of ionizing photons to the number density of neutral hydrogen, $\mathrm{n_H}$. We compute $\mathrm{n_H}$ using

$$n_{\mathrm{H}} = \frac{\bar{\rho}\Omega_{\mathrm{b}} f_{\mathrm{H,init}}}{m_{\mathrm{H}}} \tag{6.3}$$

where $\bar{\rho}$ is the critical density of the Universe, Ω_{b} is the universal baryon fraction, $f_{\mathrm{H,init}} = 0.752$ is the initial hydrogen fraction by mass and m_{H} is the mass of a hydrogen atom. While the number density of hydrogen atoms to be reionized evolves, in particular due to the atoms being locked up in stars and processed to produce heavier elements, in the early Universe the density of stars is negligible. By redshift 4 only 0.006% of the total baryonic matter is contained in stars. As a result, we neglect the evolution of $\mathrm{n_H}$ with redshift.

The Escape Fraction and Recombinations

The reionization of the Universe depends on both the escape fraction of photons, f_{esc}, from the environments of star formation and the recombination rate of electrons with ionized hydrogen. For reionization to occur, the ratio, R of n_γ to $\mathrm{n_H}$ must be

$$R \geq \frac{1 + \mathrm{N_{rec}}}{\mathrm{f_{esc}}}, \tag{6.4}$$

where N_{rec} is the mean number of recombinations per hydrogen atom.

As the ionizing photons are produced by stars, in the interstellar medium, their ability to reionize the Universe depends on what fraction of them can escape to the

neutral hydrogen in the circumgalactic medium and, particularly, the intergalactic medium. This fraction is very uncertain. Radiative transfer simulations have been used to test the escape fraction of the photons. For example, Wise et al. (2014) found that the escape fraction depends on halo mass for low mass galaxies, although the simulation could only test galaxies with masses below 10^9 $M_\odot$, with low number statistics. Because of the uncertainties, and for simplicity, we adopt a constant escape fraction of 10%, which is similar to the values used in the literature (Raičević et al., 2011; Alvarez et al., 2012).

The recombination rate of hydrogen with free electrons also plays a key role in predicting when the Universe was reionized. However, this rate is sensitive to the conditions of the gas, for example the recombination rate is reduced when the IGM is ionized (Pawlik et al., 2009). We adopt a value of 2, from the simulations of Iliev et al. (2006); McQuinn et al. (2007); Trac and Cen (2007).

6.3 Results

6.3.1 Galaxy Properties

We begin by considering the galaxy properties at high redshifts in the EAGLE simulation. In examining the general properties of the galaxies we also compare to observational data, where available, to reiterate the agreement of the simulation with the observed Universe, as shown in Schaye et al. (2015) and *Chapter 3*.

In Fig. 6.1 the stellar mass density (ρ_*) as a function of redshift is shown for Ref-L100N1504 from redshift 4 to 10. The ρ_* increases with decreasing redshift in the simulation. Observational data from González et al. (2011) is shown for comparison. The simulation data is in agreement with the observations. Note however that the stellar masses in González et al. (2011) may be overestimated due to nebular emission lines, that may affect measurements at high redshifts. As a result, the data points are shown as upper limits. (For convergence tests of ρ_* see Appendix A.)

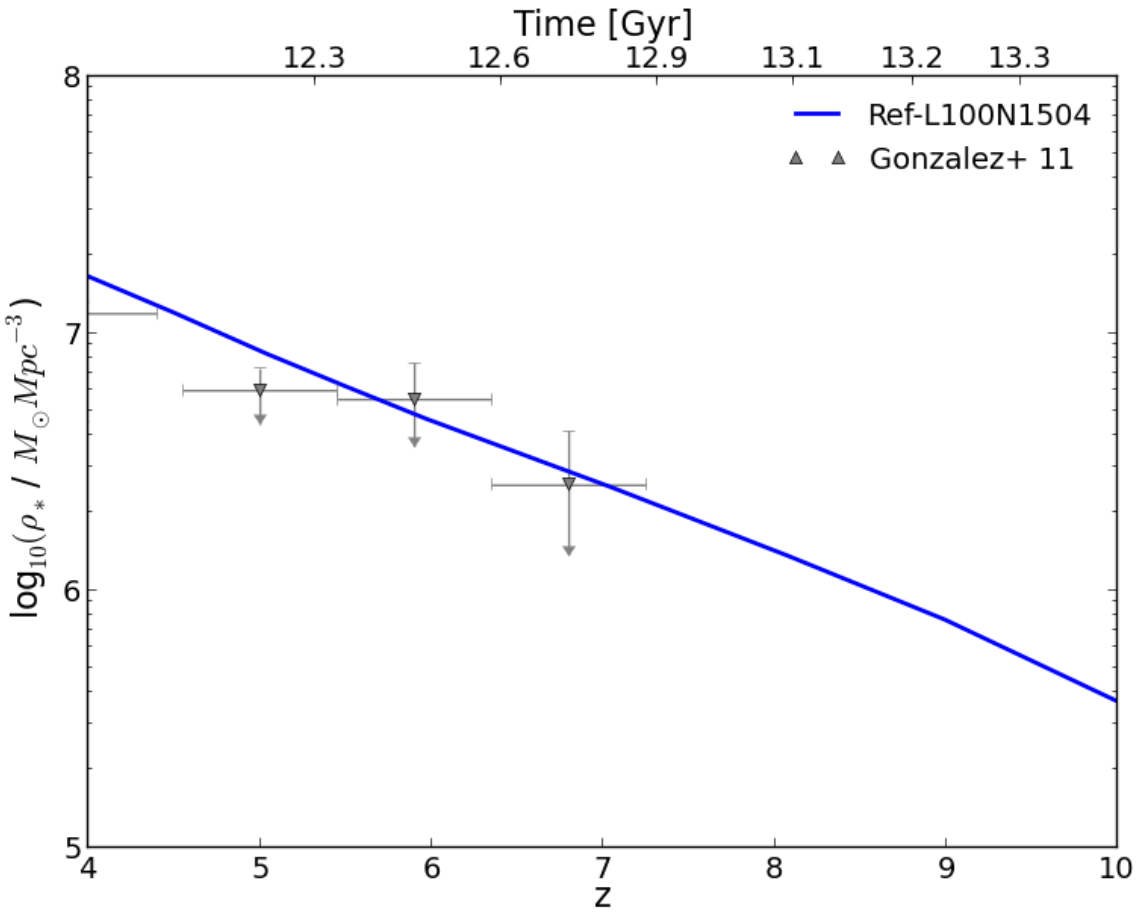

Figure 6.1: The stellar mass density as a function of redshift from Ref-L100N1504. Observational data from González et al. (2011) is shown for comparison. The stellar mass density in the simulation increases with decreasing redshift in the simulation. The simulation is consistent with the observations within the error bars.

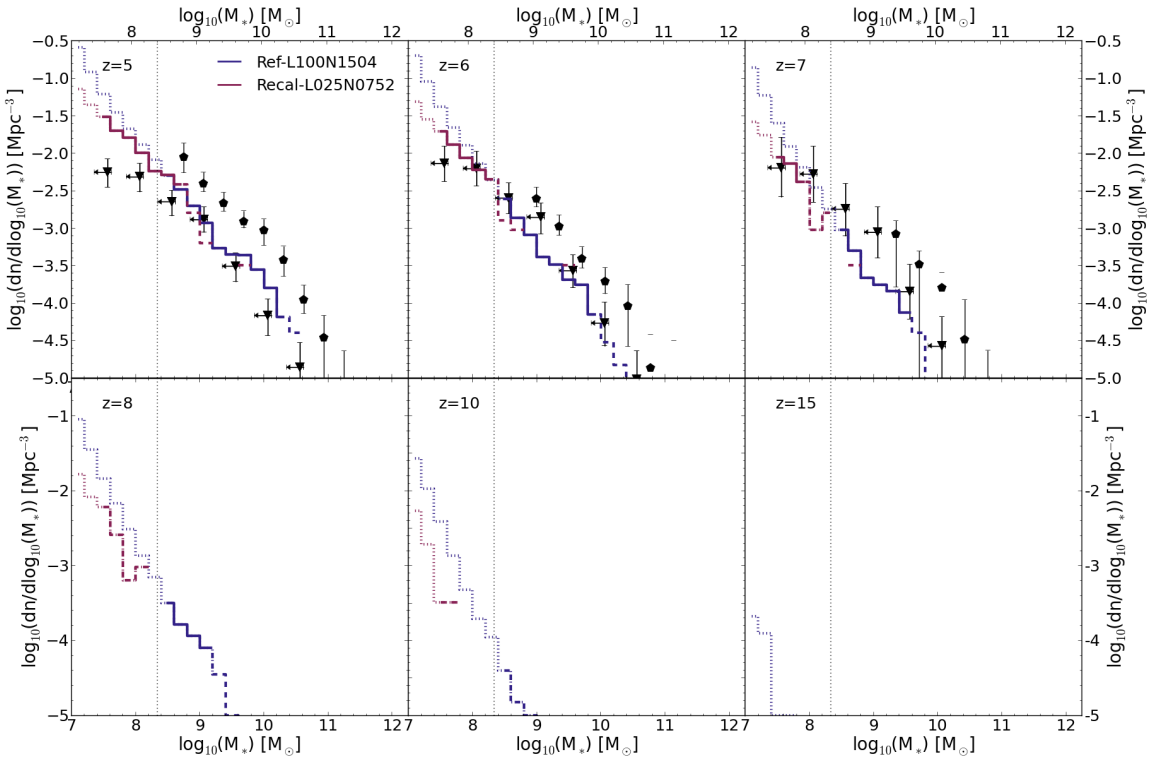

Figure 6.2: The galaxy stellar mass function from the Ref-L100N1504 and Recal-L025N0752 models in blue and red respectively at redshifts given in the upper left of each panel. Where there are fewer than 10 galaxies per stellar mass bin, dashed lines are shown, when the galaxy stellar mass falls below that of 100 baryonic particles the lines are dotted. Over the mass range shown the GSMF has a power law shape up to redshift 10. At redshift 15 there are very limited objects in the simulation to probe the GSMF shape. The normalisation of galaxies is lower in the higher resolution simulation by up to 0.6 dex below 10^8 $M_\odot$. Observational data from González et al. (2011) and Duncan et al. (2014) are shown for comparison, at redshifts 5, 6 and 7. The simulation shows a similar slope to both observational data sets, and lies within the observational error bars of González et al. (2011). However there is an offset in stellar mass of up to 0.4 dex from Duncan et al. (2014).

To provide a more detailed comparison of the stellar mass density, we present the GSMF at redshifts 5, 6, 7, 8, 10 and 15 in Fig. 6.2. The blue and red lines show the GSMF from Ref-L100N1504 and Recal-L025N0752 respectively. Below the limit of 100 baryonic particles, the curves are dotted for both simulations. This was shown to provide a reasonable guide for the mass at which resolution affects the intermediate resolution simulation at lower redshift. When there are fewer than 10 galaxies per mass bin, the curve is dashed. The intermediate resolution simulation (blue) has a steep power law like slope over the mass range probed at all redshifts. The normalisation of the GSMF at a given mass increases with decreasing redshift, while the growth of larger objects results in larger mass galaxies with decreasing redshift.

For the high resolution simulation (red), above redshift 7 there are too few objects in the simulation to draw conclusions about the convergence from a comparison of the simulations. At redshift 7 and below, we use the high resolution simulation to show the convergence of the GSMF. The intermediate and high resolution simulations have similar normalisation to $10^{7.5}$ $M_{\odot}$, at which point they begin to diverge to up to 1 dex, with fewer low-mass objects in the higher resolution simulation.

We compare the Ref-L100N1504 simulation to observational data from González et al. (2011) and Duncan et al. (2014) at redshifts 5, 6 and 7. At redshifts above 7 observational data is not available to compare to. Both data sets use UV data to determine the galaxy stellar masses to produce the GSMF. In the González et al. (2011) data set nebular emission lines are not accounted for, as noted previously. Both observational data sets have similar slopes, although there is an offset in stellar mass of around 0.4 dex between them. This offset is not driven by the nebular emission lines, as these errors are anticipated to decrease stellar masses and increase the offset (Smit et al., 2014). At redshift 5 there is a flattening of the slope at low masses in the observations of González et al. (2011), such low masses are not probed in the observations of Duncan et al. (2014).

The simulation data shows a similar slope to the observational data sets and is

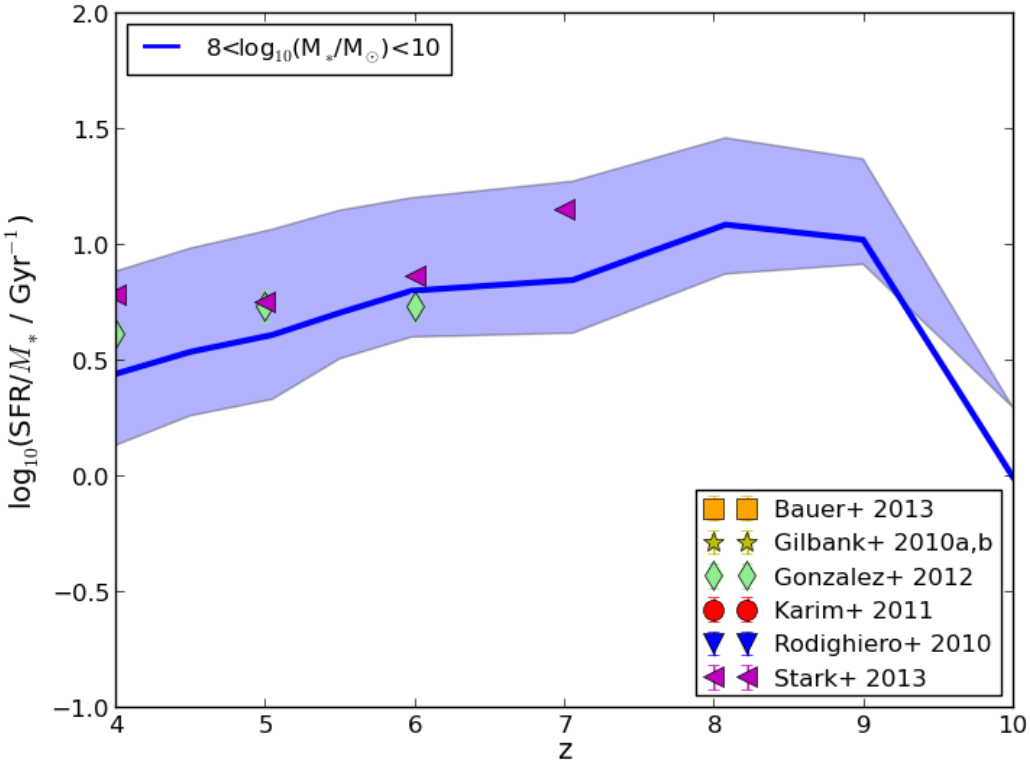

Figure 6.3: The SSFRs as a function of redshift from Ref-L100N1504 for galaxies in the mass range 10^8 to 10^{10} $M_\odot$. The shaded region encloses the 10^{th} to 90^{th} percentiles. Note that above redshift 8 there are few galaxies in this mass range (see Fig. 6.4, left panel). Observational data from González et al. (2012) and Stark et al. (2013) are shown for comparison. The values for SSFR at the redshifts shown are consistent with those observed.

typically consistent within the observational error. The main discrepancy with the data is at the lowest masses, below 10^8 $M_\odot$, when the González et al. (2011) data set shows a flattening, while the simulation data (of both simulations) continues to rise. This could however be an artefact of incompleteness in the observations, as complete low-mass galaxy samples can be difficult to observe.

The SSFR as a function of redshift is shown in Fig. 6.3 from Ref-L100N1504. The solid line shows the median for galaxies in the stellar mass range 10^8 to 10^{10} $M_\odot$, the shaded region encloses the 10^{th} to 90^{th} percentiles. The SSFRs decrease

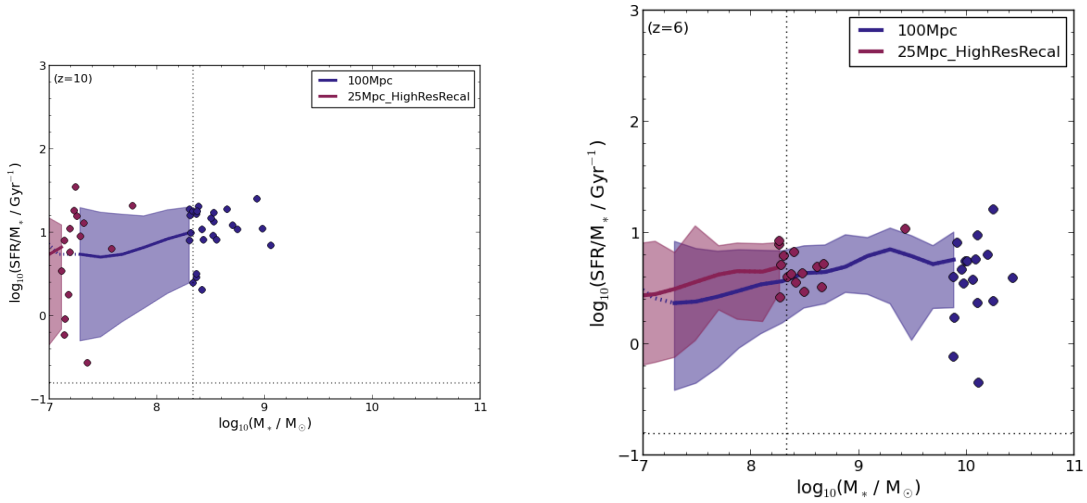

Figure 6.4: The SSFR as a function of stellar mass at redshift 10 and 6 for Ref-L100N1504 and Recal-L025N0752 in blue and red respectively. The median is shown by the solid line, with the 10^{th} to 90^{th} percentiles enclosed within the shaded regions. When the stellar mass falls below the mass of 100 baryonic particles the lines are dotted. When there are fewer than 10 objects per mass bin to determine the median and scatter, all data points are shown. The median SSFRs for both simulation resolutions have similar normalisation and a flat slope with stellar mass. For low mass objects, below 10^8 $M_\odot$, there is a large spread in SSFRs of around an order of magnitude. At high masses the range of SSFRs is smaller, around 0.6 dex.

with redshift, with a scatter of $\sim$0.8 dex at at redshifts below 8. For comparison, we include the observational data of González et al. (2012) and Stark et al. (2013). The observational data are in agreement with each other to better than 0.3 dex, and all fall within the range of SSFRs found in the simulation.

In Fig. 6.4 we show the SSFRs as a function of stellar mass at redshifts 10 and 6. The median values for Ref-L100N1504 and Recal-L025N0752 are shown as solid blue and red lines, with the $10^{\rm th}$ to $90^{\rm th}$ percentile scatter shown as shaded regions. Where the stellar mass falls below the mass of 100 baryonic particles, lines are dotted. When there are fewer than 10 galaxies per stellar mass bin, each data point is shown.

Comparing the different resolutions, at redshift 10 there are few objects in the high resolution simulation, however those found occupy a similar part of the SSFR-M_* plane as the intermediate resolution simulation. At redshift 6 the median values of the two simulations lie within 0.2 dex and have similar scatter, although the intermediate resolution extends to lower SSFRs within the scatter.

Comparing to the local Universe SSFRs, the high redshifts shown are ~ 1.5 dex above the observed redshift zero normalisation. The relation at high redshift has a reasonably flat slope above 10^7 $M_\odot$, similar to the slope found at redshift zero in the EAGLE simulation (Schaye et al. (2015)) and observationally (e.g. Gilbank et al. (2010a)). At stellar masses below 10^8 $M_\odot$ there is a strikingly large scatter of over an order of magnitude in the SSFR-M_* relation. At redshift zero the scatter is around 0.3 dex. At fixed stellar mass, galaxies at redshifts 6 to 10 have higher SFRs and a larger range than in the local Universe.

Overall we find the galaxy population above redshift 5 to have stellar masses and SSFRs are in reasonable agreement with those observed at high redshifts. The galaxies are highly star forming for their stellar mass compared to local Universe counterparts and show a large scatter around the median relation. When comparing two of the simulations with different resolutions we find good agreement in stellar masses above $10^{7.5}$ $M_\odot$ and star formation rates above 10^7 $M_\odot$ at the redshifts we consider here.

6.3.2 The Ionizing Potential of the First Galaxies

Having established that the EAGLE simulations considered here have galaxy properties similar to those observed at redshifts above 4, we now look at the potential of the first galaxies to provide enough photons to reionize the neutral hydrogen in the early universe. In Fig. 6.5 the ration of ionizing photons, n_γ, to neutral hydrogen atoms, n_H, is shown as a function of redshift for Ref-L100N1504.

In Fig. 6.5 the dashed horizontal line shows when the ratio of n_γ to n_H is equal, the solid horizontal line is where n_γ is 20 times higher than n_H. When n_γ / n_H crosses the solid horizontal line, there are sufficient photons produced by galaxies in the simulation to reionize the Universe, assuming an escape fraction and recombination rate as discussed in Section 6.2.1. Applying these approximations for the number of photons required to reionize the Universe, reionization from stellar emission by galaxies occurs by redshift 8 in the simulation. The shaded region highlights the region where n_γ / n_H varies between 10 and 30 times, this results in the reionization by stellar emission ranging from redshift 7.5 to 9.5.

Note that n_γ / n_H is shown for both the galaxies in the simulation and the halos. Both lines are within 10% of each other, however the fraction from halos is systematically lower than that from galaxies. This is as a result of the available simulation information for halos and galaxies. For galaxies we store the initial stellar mass of the galaxy as well as the present day stellar mass (after stellar mass loss from stellar evolution), however for the halos we only have the mass after mass loss, which produces the systematic offset. In grey the emissivity, ϵ, as a function of redshift is shown relative to the right hand axis, as determined from the star formation rates. This value increases by almost two orders of magnitude across the redshift range 15 to 4.

In Fig. 6.6 we show n_γ / n_H as a function of redshift again, this time comparing Ref-L025N0376, Ref-L025N0752 and Recal-L025N0752, to show the strong and weak convergence of the simulation for n_γ / n_H. First we note that the high resolution simulations with different physics are identical, showing that the impact of the changes to the subgrid physics do not affect galaxy properties at the high

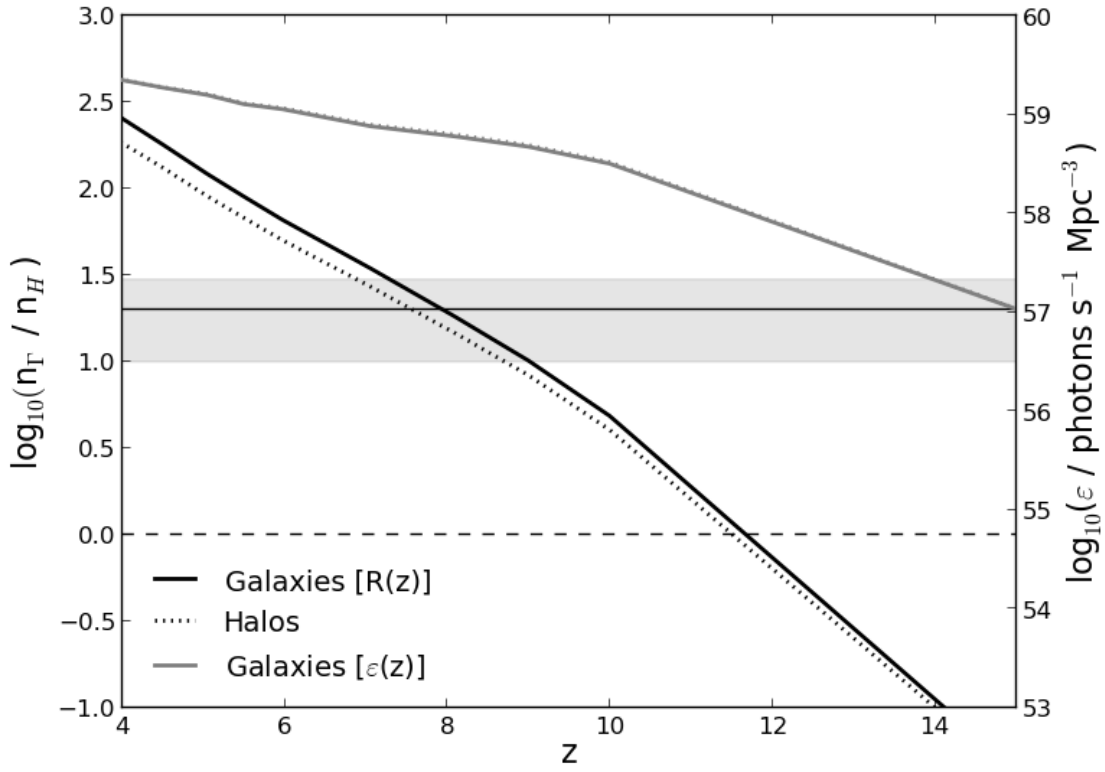

Figure 6.5: The ratio of ionizing photons produced in the simulation to hydrogen atoms as a function of redshift for Ref-L100N1504 determined from subhalo outputs (solid curve) and friends-of-friends outputs (dotted curve). The dashed horizontal line highlights where there are equal ionizing photons to hydrogen atoms. The solid horizontal line shows where there are 20 times more ionizing photons than hydrogen atoms. The shaded region spans 10 to 30. Assuming 20 times more ionizing photons than neutral hydrogen atoms are necessary to reionize the Universe, galaxies produce sufficient photons to reionize the Universe by redshift 8 in the simulation. The grey lines is the emissivity of the simulation at different redshifts, relative to the right hand axis.

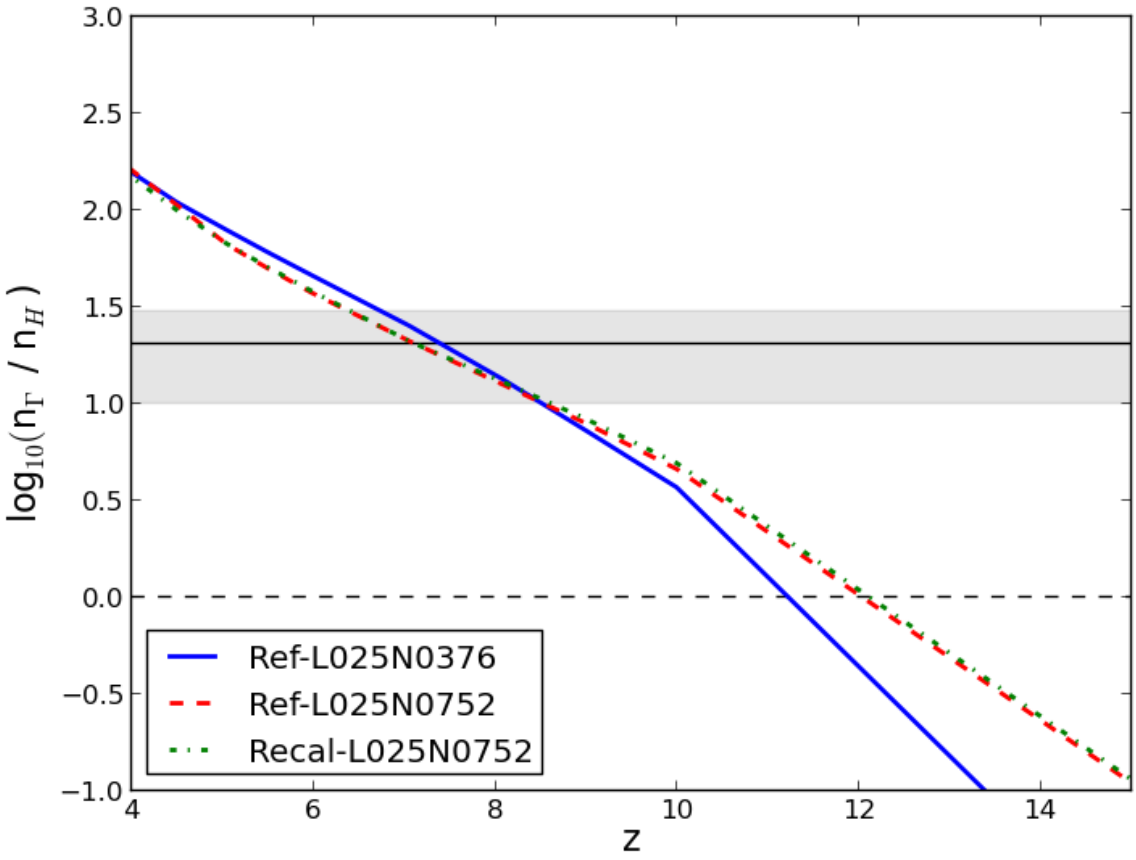

Figure 6.6: As in Fig. 6.5 showing Ref-L025N0376, Ref-L025N0752 and Recal-L025N0752, to show the strong and weak convergence in the simulation in blue, green and red respectively. Below redshift 10, the number of ionizing photons is converged to better than 0.1 dex for all three simulations.

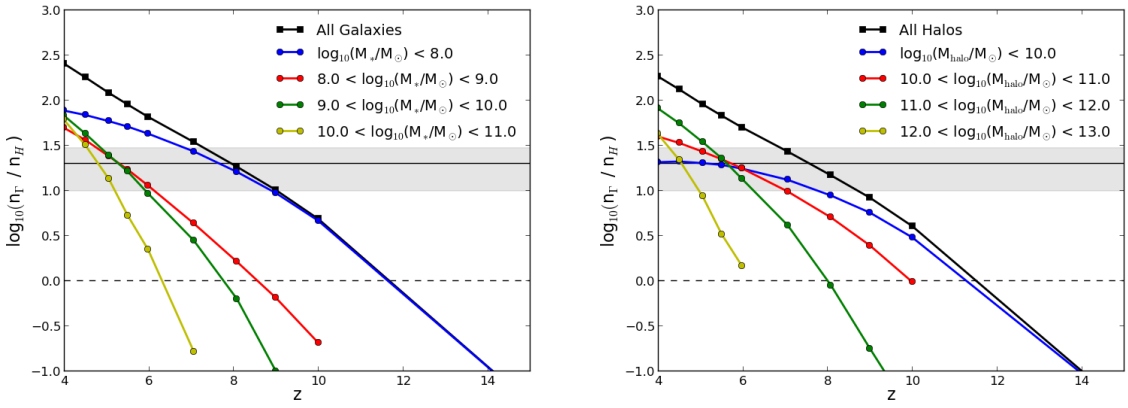

Figure 6.7: The differential contribution of galaxies and halos to the number of ionizing photons in the left and right panels respectively from Ref-L100N1504. The horizontal lines are as in Fig. 6.5. Galaxies of stellar mass less than 10^8 $M_\odot$ and halos of mass less than 10^{10} $M_\odot$ dominate the production of ionizing photons in the early Universe.

redshifts. Indeed, in Appendix A it is shown that the different subgrid parameters only show significant differences below redshift 2. Above redshift 10 we find the high and intermediate resolutions differ by up to 0.5 dex. However, the time resolution of snapshots results in no data points between redshift 15 and 10. It is expected that stars can form earlier in the high resolution simulation than the intermediate resolution simulation due to smaller mass particles resolving lower star formation rates ($\dot{m}_{*,\mathrm{min}} \propto m_{\mathrm{gas}}$, Schaye and Dalla Vecchia (2008)), as a result more star formation at high redshift is unsurprising. From redshift 10 to 4 the simulations are converged to better than 0.1 dex in n_γ/n_H. Increasing the resolution in these simulations does not significantly change the redshift by which reionization by galaxies can occur.

In summary, the galaxies in the simulation produce sufficient ionizing photons by redshift 8 to reionize the neutral hydrogen in the Universe, based on the assumptions in Section 6.2. This result is not significantly affected by resolution.

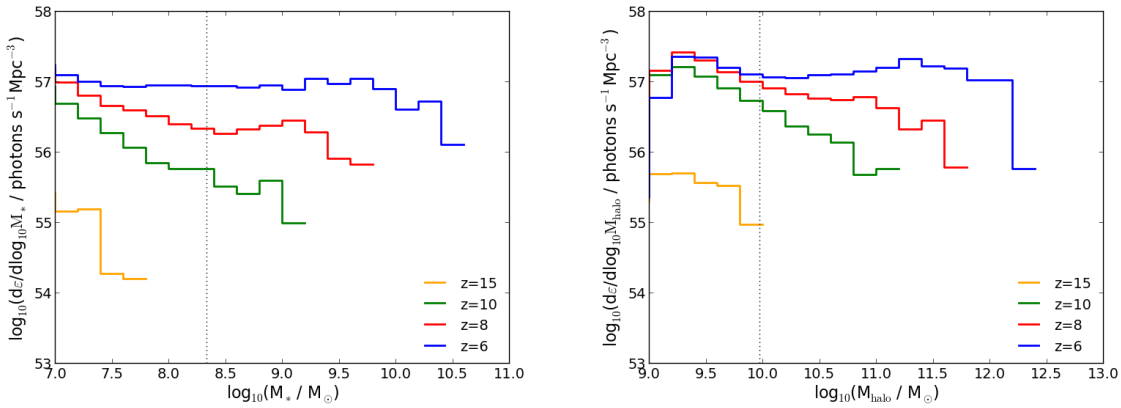

Figure 6.8: The contribution to the emissivity from galaxies of different stellar masses and halo masses in the left and right panels from Ref-L100N1504 at redshifts shown in the legend. The dotted vertical lines are a resolution guides showing 100 star particles and 1000 dark matter particles in the left and right panels. Above redshift 8 there is a decreasing contribution to the emissivity with increasing stellar mass. At redshift 8 the slope with stellar mass is flatter and a peak appears around 10^9 $M_\odot$, although lower mass galaxies continue to dominate. By redshift 6 the contribution of different mass galaxies to the emissivity is approximately equal, falling only for the highest mass galaxies in the simulation, above 10^{10} $M_\odot$. A similar trend with redshift and mass is seen in the case of the halo masses.

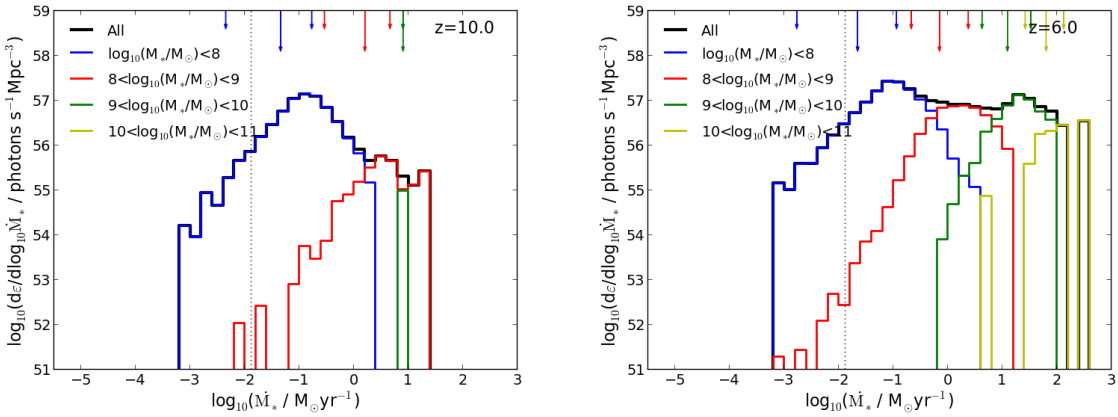

Figure 6.9: The contribution of galaxies of different SFRs to the emissivity, at redshift 10 and 6, showing the contribution from different stellar masses. The dotted vertical line shows the SFR resolved by 10 star forming gas particles in the simulation. Arrows on the upper x-axis show the median SFR (large) and 10th to 90th percentile values (small) for each mass range. The lowest mass galaxies clearly dominate the total emissivity, as seen in Fig. 6.8. Here we see that it is the highly star forming galaxies within a given mass range that contribute the most to the emissivity, with the peak falling to the right of the median SFR in all stellar mass bins (at redshift 10 there is only one object in the 10^9 to 10^{10} $M_\odot$ bin).

6.3.3 The Properties of the Ionizing Galaxies

We now begin exploring the nature of the galaxies that produced the ionizing photons. In Fig. 6.7 again n_γ / n_H is shown as a function of redshift, this time highlighting the contribution of galaxies in different stellar and halo mass ranges in the left and right panels respectively. At all redshifts above 4, galaxies of masses less than 10^8 $M_\odot$ dominate the production of ionizing photons. However, below redshift 8 the contribution of galaxies with masses above 10^8 $M_\odot$ increases. As expected, the low mass galaxies dominating the production of ionizing photons reside in low mass halos. Halos with masses less the 10^{10} $M_\odot$ are the main contributors of ionizing photons, as seen in the bottom panel. Around redshift 7 large mass halos become important, with 10^{11} to 10^{12} $M_\odot$ halos dominating from redshift 7 to redshift 4.

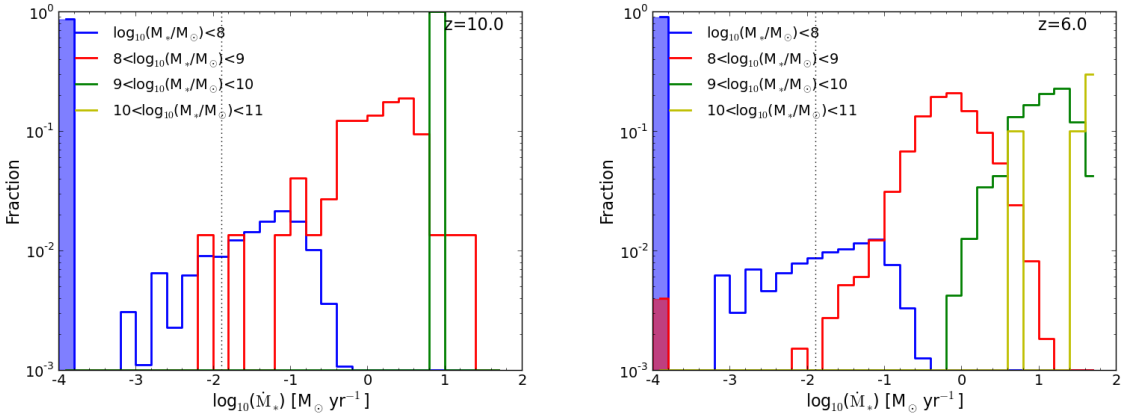

Figure 6.10: The fraction of galaxies in different stellar mass bins as a function of SFR at redshifts 10 and 6 from Ref-L100N1504. The fraction of galaxies with an emissivity of zero are shown at $\log_{10}$(SFR/ $M_\odot$ yr^{-1}) = -4, shaded according to stellar mass range. This is to give an impression of the passive galaxies that are not contributing to the production of ionizing photons at any given time. For each stellar mass bin, the distribution is peaked, with tails to lower and higher SFRs. Note that for the low mass galaxies that dominate the production of reionizing photons, the majority of them are passive at any given time.

Having looked at the number of ionizing photons produced by stars, we now look at the emissivity of galaxies and halos in Fig. 6.8. The contribution to ϵ per dex stellar mass and halo mass are shown in the left and right panels respectively at redshifts 15, 10, 8 and 6. On each panel the dotted line is a resolution guide, showing a galaxy with 100 star particles and 1000 dark matter particles for the top and bottom panels respectively. From the top panel we see the emissivity per dex stellar mass is a decreasing function of stellar mass, with low mass galaxies producing the most emissivity at redshifts 15 and 10. At redshift 8 the lowest mass galaxies continue to dominate but with increasing production from higher mass galaxies. By redshift 6 the emissivity produced across all galaxy stellar masses is approximately equal. In the bottom panel, similar trends with redshift are seen for different halo masses. The emissivity is a decreasing function of halo mass up to redshift 8, however by redshift 6 the relation has flattened across halo masses with a bump around $10^{11.5}$ $M_\odot$.

We have firmly established that galaxies of stellar masses less than 10^8 $M_\odot$ dominate the production of ionizing photons in the early Universe and live in low mass halos. Now we consider the star formation rates of galaxies that contribute to reionization. In Fig. 6.9 the contribution to the emissivity as a function of SFR is shown at redshifts 10 and 6. The dotted line shows a SFR resolved by 10 SF gas particles. At both redshifts shown, and indeed redshifts in between those shown, the emissivity is dominated by galaxies with SFRs of 0.1 $M_\odot$ yr^{-1}, although at redshift 6, higher SFRs become increasingly important.

Fig 6.9 also shows the contribution of galaxies of different stellar masses to the emissivity. As in Fig. 6.8 the lowest mass galaxies clearly dominate. The arrows on the upper x-axis show the median (large) and 10^{th} and 90^{th} percentile values (small), coloured by stellar mass. It is clear that in each stellar mass range, the galaxies contributing most to the emissivity are not those with the median SFR, but the highly star forming galaxies.

In Fig. 6.10 the fraction of galaxies in a given stellar mass range as a function of SFR at redshift 10 and 6 is shown. Galaxies with a SFR of zero are shown in a shaded bin at 10^{-4} $M_\odot$ yr^{-1}. At both redshifts in the lowest mass bin, less than 10^8 $M_\odot$, the galaxy population is dominated by passive galaxies, with only around 10% of galaxies contributing to the emissivity. The average emissivity produced by galaxies of stellar masses greater than 10^8 $M_\odot$ is one to two orders of magnitude higher than the average for galaxies of stellar masses less than 10^8 $M_\odot$, however, there are not sufficient galaxies with larger masses for them to be the dominant source of ionizing photons.

In summary we have found that the lowest mass galaxies in the simulation, less than 10^8 $M_\odot$, dominate the production of ionizing photons around the epoch of reionization. Within this population, at any time, many of the galaxies are not forming stars, implying these galaxies have bursty star formation histories. Of those that are star forming, it is the highly star forming galaxies, as opposed to those with average SFRs, that contribute most to the emissivity.

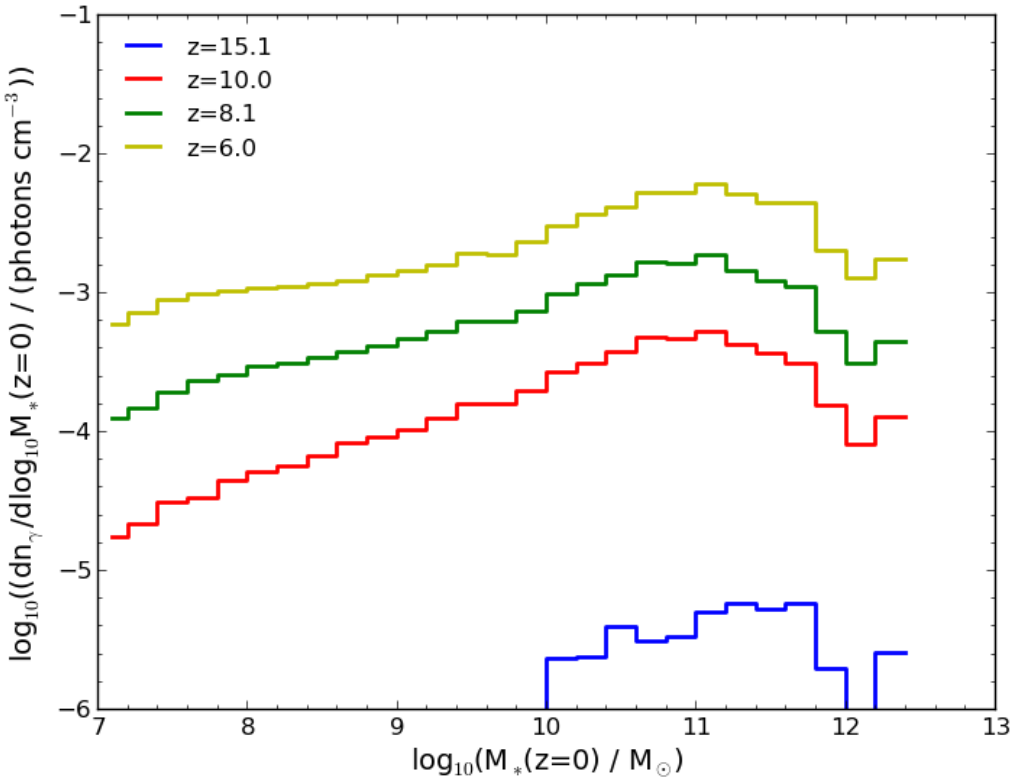

Figure 6.11: The distribution of the number density of ionizing photons as a function of present day stellar mass from progenitor galaxies at redshifts shown in the legend. The contribution from present day galaxies to the number of ionizing photons increases with stellar mass up to $\sim 10^{11}$ $M_{\odot}$, and decreases at higher stellar masses at all redshifts. The stars producing photons to reionize the neutral hydrogen content of the Universe are mostly in 10^{11} $M_{\odot}$ galaxies today.

6.3.4 The Descendants of the First Galaxies

One of the advantages of using the EAGLE simulations for exploring the reionizing potential of galaxies is the ability to study the descendants of galaxies from the epoch of reionization at the present day. We use merger trees to determine all progenitors of redshift zero galaxies at a given redshift. From the progenitors, we compute the number density of ionizing photons We then find the total n_{γ} produced in bins of redshift zero stellar mass. In Fig. 6.11 we show the distribution of

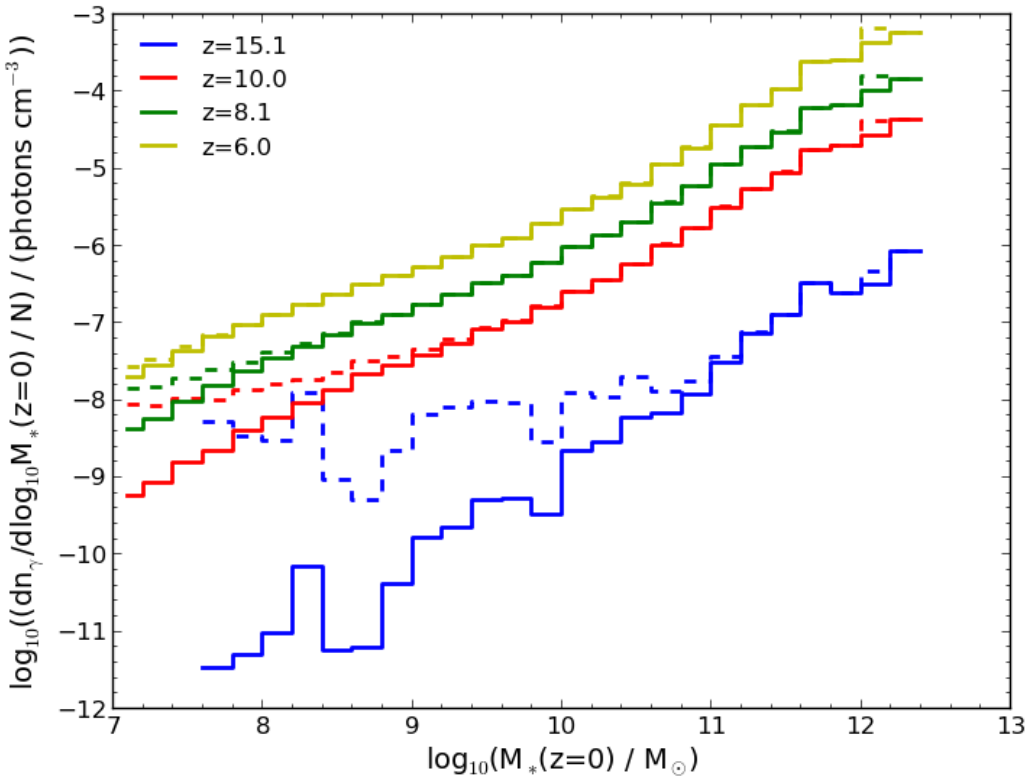

Figure 6.12: As in Fig. 6.11 normalised by the number of objects per bin at redshifts shown in the legend. This shows the efficiency per galaxy at different masses at producing ionizing photons, and is a reflection of the galaxy ages. Solid lines are normalised by the number of present day galaxies per stellar mass bin, dashed are normalised by the number of present day galaxies with progenitors at the redshifts shown. The choice of normalisation primarily affects redshift 15 results, where there are very few present day low mass galaxies that have progenitors at this redshift, hence when including only those with progenitors at redshift 15 in the normalisation, the production of ionizing photons per galaxy increases. What we see at all redshifts shown is that the present day highest mass galaxies are the most efficient at producing ionizing photons per galaxy, with more massive galaxies producing more of the ionizing photons.

the number density of ionizing photons as a function of redshift zero stellar mass. At all redshifts the peak number density of ionizing photons is at a present day stellar mass of $\sim 10^{11}$ $M_{\odot}$. Most of the stars producing photons that reionized the Universe are in galaxies of this mass today.

In Fig 6.12 we show the number density of ionizing photons per redshift zero stellar mass normalised by the number of galaxies per bin. This is equivalent to the efficiency of ionizing photon production per galaxy stellar mass as a function of present day galaxy stellar mass. The trend seen at all redshifts is that the number of ionizing photons increases with present day stellar mass, more massive present day galaxies produced more ionizing photons per galaxy at all redshifts shown.

We normalise by the total number of galaxies in a given mass bin at redshift 0 (solid) and by the number of redshift 0 galaxies with progenitors at a given redshift (dashed). The most significant difference the normalisation makes is at low masses and high redshifts, where some present day low mass galaxies do not have progenitors. At redshift 15, the relation is less steep at low stellar masses, below 10^{11} $M_{\odot}$.

In summary we see that the most massive present day galaxies were the most efficient at producing ionizing photons per galaxy.

6.4 Summary and Conclusions

We have presented results from the EAGLE simulation suite, focusing on Ref-L100N1504 and Recal-L025N0752. These simulations have box sizes of 100 cMpc and 25 cMpc, respectively, with baryonic particle masses of 1.81×10^6 $M_{\odot}$ and 2.26×10^5 $M_{\odot}$. We have considered the potential of high redshift simulated galaxies in the EAGLE Ref-L100N1504 simulation to produce ionizing photons to reionize the Universe. We began by highlighting the agreement of the EAGLE simulations with the observed high redshift Universe. In particular, the stellar mass density agrees well with observations, while the stellar mass function has a similar slope to observations and agrees within the error bars of the González et al. (2011) data. The spe-

cific star formation rates agree with the observations of González et al. (2012) and Stark et al. (2013). Agreement with observational data within the systematic uncertainties of observations to lower redshifts has previously been shown in *Chapter 3*, while agreement at redshift zero was shown by Schaye et al. (2015).

Reionization of the neutral hydrogen in the universe by the galaxy population would occur by redshift 8 in the simulation, assuming an escape fraction of 10% and two recombinations per hydrogen atom. While changes to these assumptions will change the redshift by which galaxies provide enough photons to reionize the universe, they do not affect the properties of the galaxies producing the ionizing photons. The remainder of this study focuses on the galaxy properties.

The dominant source of ionizing photons in the early Universe is from galaxies of masses less than 10^8 $M_\odot$. These galaxies have bursty histories, with many of them not forming stars at any given time. Of those that are star forming, it is the galaxies with SFRs above the median value that contribute most strongly to the production of ionizing photons. This is true for all galaxy mass ranges, it is the most highly star forming objects of a given population that contribute most to the production of ionizing photons.

The majority of the stellar mass producing ionizing photons in the early Universe can be found today in $\sim 10^{11}$ $M_\odot$ galaxies. Although, the number of ionizing photons produced per galaxy is proportional to present day stellar mass.

Having completed an initial analysis of the galaxies that reionized the universe, further resolution tests are required to ensure the conclusions are insensitive to numerical effects. Particularly, we are interested in exploring the bursty nature of low mass galaxies at high redshifts, as they play a crucial role in producing ionizing photons.

Another interest is to compare to other simulations looking at reionization. Many simulations use simple methods to populate dark matter simulations with galaxies, and then run radiative transfer to model the photon interactions. As the EAGLE simulation provides a reasonable description of the observed Universe, a comparison to the galaxy population used in these models would be informative.

Chapter 7

Extreme star formation rates: a comparison to the sub-mm population

7.1 Introduction

In this work we carry out a preliminary comparison between the most highly star forming galaxies in the EAGLE Ref-L100N1504 simulation and the observed sub-mm galaxy population. The sub-mm population consist of rare objects, with a number density of 10^{-5} $h^{-1}Mpc^{-3}$ (Chapman et al., 2005). These galaxies are found to be at redshifts of ~ 2 or above, with high SFRs and a significant dust mass. They are referred as sub-mm as they are detected in sub-mm wave bands due to the reprocessing of the UV light from young stars by the dust (see Casey et al., 2014, for a review). This population is particularly interesting due to the extreme nature of their star formation and the many open questions that remain to be answered about their formation and evolution. Some examples are

- For how long can these galaxies form stars at such extreme rates?
- What triggers such extreme star formation, mergers or disc instabilities?
- How do they evolve following the star burst episode? (Simpson et al., 2014)

These questions, while difficult to determine observationally, can be explored with reasonable ease in simulations.

Previous comparisons of models with observations for this galaxy population have highlighted an apparent problem with the current theory of galaxy formation

(Baugh et al., 2005; Davé et al., 2010). The models fall short of producing the observed SFRs and sub-mm number counts while maintaining agreement with other observations. One proposed solution to this issues is the inclusion of a top heavy IMF, as proposed in Baugh et al. (2005). However the extent of the variation to the IMF has reduced as more complete observational data has become available (Cowley et al., 2015).

In this analysis, we are first interested in testing a new hydrodynamical simulation to determine if the discrepancies with observations are again seen. The EAGLE project simulation Ref-L100N1504 is used to compare to the sub-mm galaxy population. A comparison of this simulation to observations, at redshift zero and across cosmic time, reveals similar trends for the simulated galaxies as those observed. The level of agreement is within the systematic uncertainties from the observations (Schaye et al., 2015, *Chapter 3)*. Because of the reasonable growth histories reported for the simulated galaxies, it is of interest to use this study to look at the sub-mm galaxy population. Furthermore, as the simulation reproduces the observed redshift zero galaxy population, galaxies live in similar mass halos to real galaxies and hence have similar merger and accretion histories (assuming the simulation cosmology is accurate). Note, however, that this simulation only has a box size of (100 cMpc)^3, which is a lower limit on the volume required to study sub-mm galaxies.

This chapter is laid out as follows; In Section 7.2 we describe the selection of galaxies from the simulation to compare to the observed sub-mm population. The SFRs and masses of the selected galaxies are compared to observations in Section 7.3.1. We consider the redshift distribution of the selected galaxies in Section 7.3.2 and consider in some more detail their detectability. In Section 7.3.3 we present a 3D impression of the selected galaxy population in the simulation. Finally, in Section 7.4 we summarise and conclude.

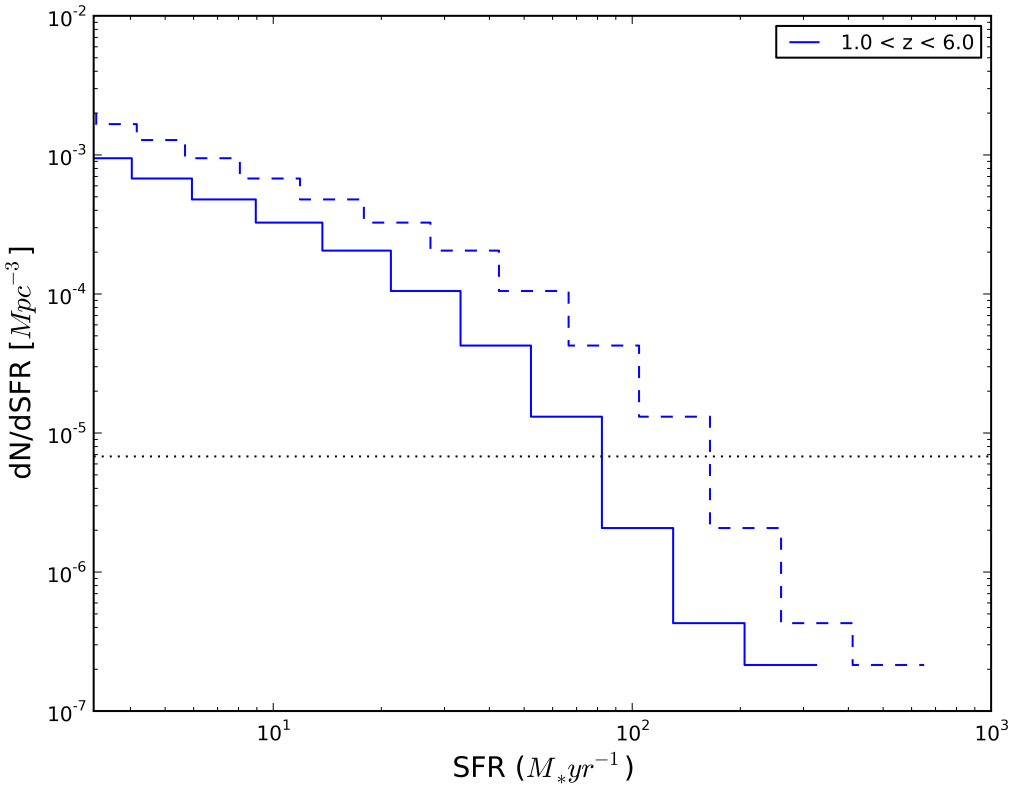

Figure 7.1: The number density of galaxies in the simulation as a function of SFR for a stack of 14 snapshots between redshift one and six. The expected number density of sub-mm galaxies from Chapman et al. (2005), of 10^{-5} h^{-1}cMpc^{-3}, is shown a horizontal black dotted line. This number density implies a SFR of ~ 80 M$_\odot$ yr^{-1} or above to select an equivalent number density of galaxies from the simulation. The dashed curve shows the number density of galaxies when multiplying the SFRs from the simulation by a factor of two.

7.2 Selection

We are interested in comparing simulated galaxies to sub-mm galaxies, which are observed to be highly star forming and dusty. However, to produce sub-mm emission for simulated galaxies requires computationally expensive radiative transfer methods and dust modelling, which are sensitive to the assumptions about the distribution and quantity of the dust. As a result, we look to more basic methods of selecting simulated galaxies for comparison with sub-mm observations. Instead, for this initial investigation, we use the observed number density of sub-mm galaxies to determine a cut in SFR to define the popuation.

The number density of observed sub-mm galaxies is estimated in Chapman et al. (2005) to be 10^{-5} $h^{-1}cMpc^{-3}$. In Fig. 7.1 the number density of galaxies as a function of SFR is shown for a stack of 14 snapshots between redshift 1 and 6. From this figure, the SFR that corresponds to the number density of Chapman et al. (2005) is ~ 80 $M_\odot$ yr^{-1}. In the following analysis we consider galaxies with SFRs greater than 80 and 100 $M_\odot$ yr^{-1}. Although this number density of galaxies results in less than 10 galaxies per snapshot, providing poor statistics, this selection enables a study the physics of the most extreme star forming objects in the simulation.

Having selected the galaxy sample in the simulation to compare with observational data for sub-mm galaxies, we begin by considering the average properties of these galaxies in the simulation. The redshift distribution of the highly star forming galaxies in the simulation is then compared to the shape determined from observations.

7.3 Results

7.3.1 The properties of highly star forming galaxies

In this section we look at the properties of the highly star forming galaxy population in the simulation, and how the properties of these selected galaxies evolve with redshift. The simulation data is divided into galaxies with SFRs greater than

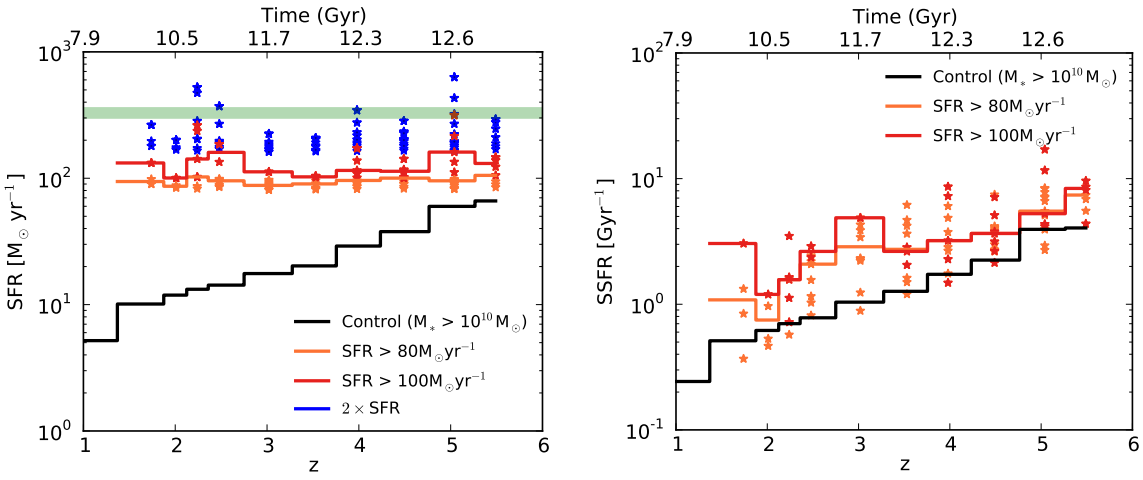

Figure 7.2: The star formation rates and specific star formation rates of the highly star forming galaxies from the simulation as a function of redshift. SFRs greater than 80 $M_\odot$ yr^{-1} are shown in orange, and greater than 100 $M_\odot$ yr^{-1} are in red. Medians are shown by the solid lines and the scatter is shown by the points. In the left panel, the blue points are the SFRs boosted by a factor of two, due to the offset in the median star formation relation of all star forming galaxies relative to observations. The median SFRs and SSFRs of a control sample, selecting galaxies with stellar masses greater then 10^{10} $M_\odot$ are shown in black. The green band shows the median of observed SFRs for sub-mm galaxies from Swinbank et al. (2014), the range is from 20 to 1030 $M_\odot$ yr^{-1}.

80 and 100 $M_\odot$ yr^{-1}, to determine if any strong trend exists with SFR. To understand these galaxies in the context of the galaxy population that are not highly star forming, we compare to a mass selected control sample. The control sample includes galaxies with stellar masses greater than 10^{10} $M_\odot$ at each redshift. Median properties for the control sample galaxies are shown in each panel. Again, we reiterate that this is prelimiary work, however, as the results have proven interesting we include it here.

Star formation rates

The left hand panel of Fig. 7.2 shows the evolution of the SFRs for highly star forming galaxies with redshift. Unsurprisingly, given the selection of galaxies is based on the SFRs, the relation is reasonably constant with redshift. What is interesting however is that the SFRs of the control sample drop by an order of magnitude over this redshift range, implying the highly star forming objects are more extreme at low redshifts than high redshifts. The SSFRs, in the right hand panel, for highly star forming galaxies decreases with redshift. As the SFR cut to select galaxies remains constant, the stellar mass of the highest star forming galaxies increases with redshift, as will be shown in Fig. 7.3. The decrease in SSFRs for the selected galaxies is less steep than the decrease for the control sample, this reflects the relative difference between the SFRs of the highest star forming galaxies and the control sample at high and low redshifts.

The median SFR from the observational study of Swinbank et al. (2014) is shown as a green band in the left hand panel of Fig. 7.2, although the range of SFRs found spans 20 to 1030 $M_{\odot}$ yr^{-1}. Compared to the observations, the simulations' highest star forming galaxies are systematically low. It was reported in *Chapter 3* that SFRs, for the full galaxy population, are low relative to observations by 0.2-0.3 dex (around a factor of 2). The source of this offset with observations could be due to simulated galaxies not exhibiting sufficient burstiness, although this offset is within the typical systematic uncertainty quoted for observations (see *Chapters 3* and 4 for a discussion). To adjust for the offset of the total population SFRs, the effect of boosting all SFRs by a factor of two is shown in Fig. 7.2. Even when boosting the SFRs by a factor of two, a difference of 100 $M_{\odot}$ yr^{-1} remains.

A further discrepancy in the simulation with observations is the most extreme SFRs recovered. In the simulation, the highest SFR is 260 $M_{\odot}$ yr^{-1}, while in the study of Swinbank et al. (2014) it is 1030 $M_{\odot}$ yr^{-1}. From an observational perspective, the SFRs may be over-estimated, for example, if AGN contamination is not accounted for. Another potential issue is the dust modelling used to determine the observed SFRs, where a uniform screen of dust is assumed when converting ob-

served flux to SFR. In reality the dust can be a clumpy medium spread throughout the galaxy. A final issue is the waveband coverage of the observations. As the SFRs are produced using SED fitting, if the peak in the SED is not well constrained the SFR could be over or under estimated.

In the simulation, a simple explanation for the lack of extreme SFRs is the box size. The simulation consists of only a (100 cMpc)^3 box, implying there are less than 10 sub-mm galaxies per snapshot, given the observed number density. As the very highly star forming objects are typically rarer, they may simply not be probed in this box. Another issue could be the timescales of such enhanced SFRs. At a rate of 1000 $M_\odot$ yr^{-1}, a galaxy requires a constant supply of gas to maintain its star formation. Without a sufficient gas supply the high SFRs will be short-lived, and therefore snapshots in the simulation would be less likely to capture such events. Although there are potential explanations for this discrepancy, that require further investigation, the simulation fails to reproduce the observations in this case.

Galaxy masses

In Fig. 7.3 the evolution of mass properties of highly star forming galaxies are presented. In the top left panel the evolution of the stellar mass is shown. Over the redshift range 5.5 to 1.5 the stellar mass of the highly star forming population increases by an order of magnitude from 10^{10} to 10^{11} $M_\odot$. Comparing the galaxies with SFRs greater than 80 $M_\odot$ yr^{-1} to those greater than 100 $M_\odot$ yr^{-1}, the galaxy stellar masses are similar. At redshift 4 and above, higher star forming objects typically have higher stellar mass, although this trend does not hold below redshift 4. It appears that the difference in SFR is not related to the simple SFR - M_* relation. The selected galaxies can be compared to the control sample. As the control sample is mass selected this comparison is not very informative. Indeed the highly star forming galaxies are more massive than the control sample at all redshifts, although the difference is larger at redshifts below 3. The green band shows an estimate of the stellar mass of sub-mm galaxies from the observations of Swinbank et al. (2014). Below redshift 3 the masses of the selected simulated galaxies are consistent with

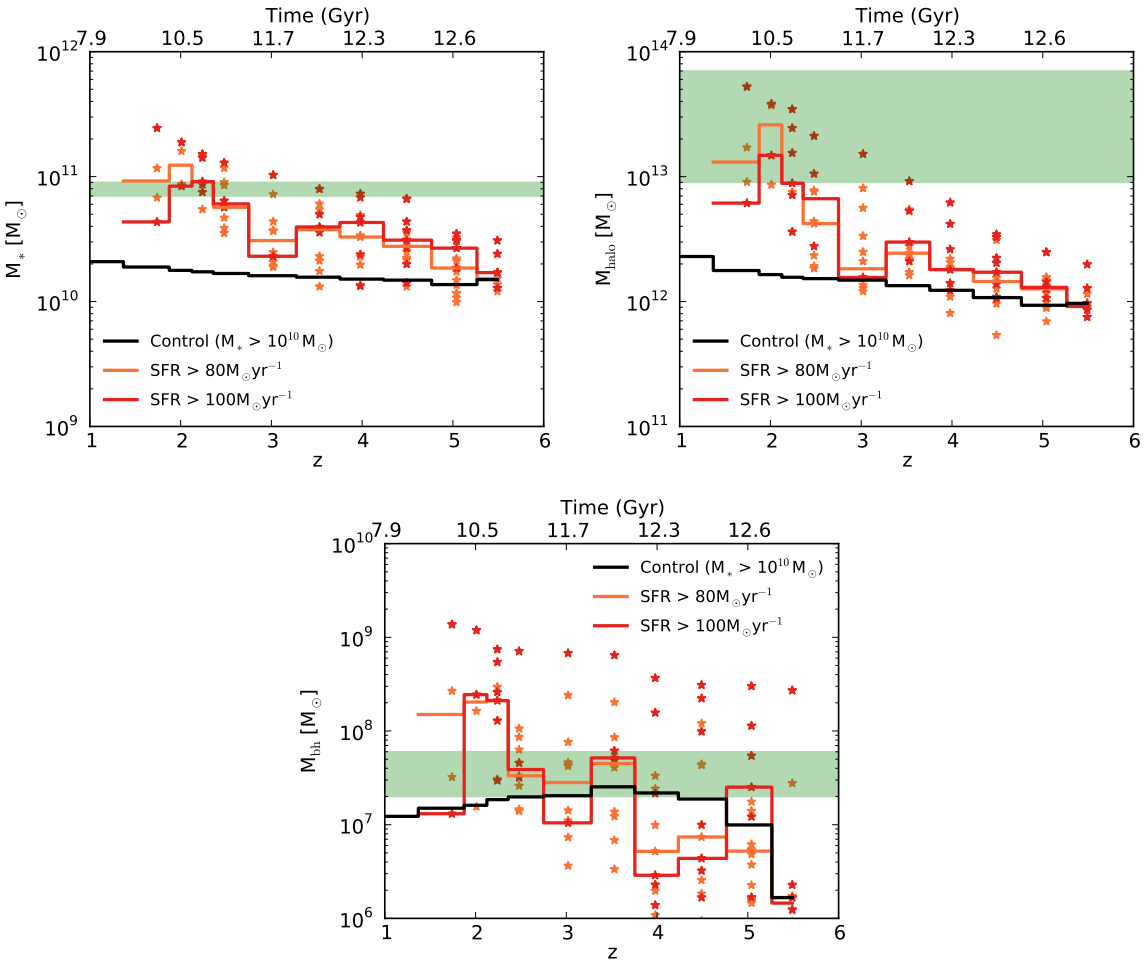

Figure 7.3: The stellar masses, halo masses and black hole masses of highly star forming galaxies in the simulation are shown in the three panels. Colour coding is as is Fig. 7.2. The green band shows the range of observed galaxy properties estimate in Swinbank et al. (2014) (for stellar masses) and Hickox et al. (2012) (for halo and black hole masses).

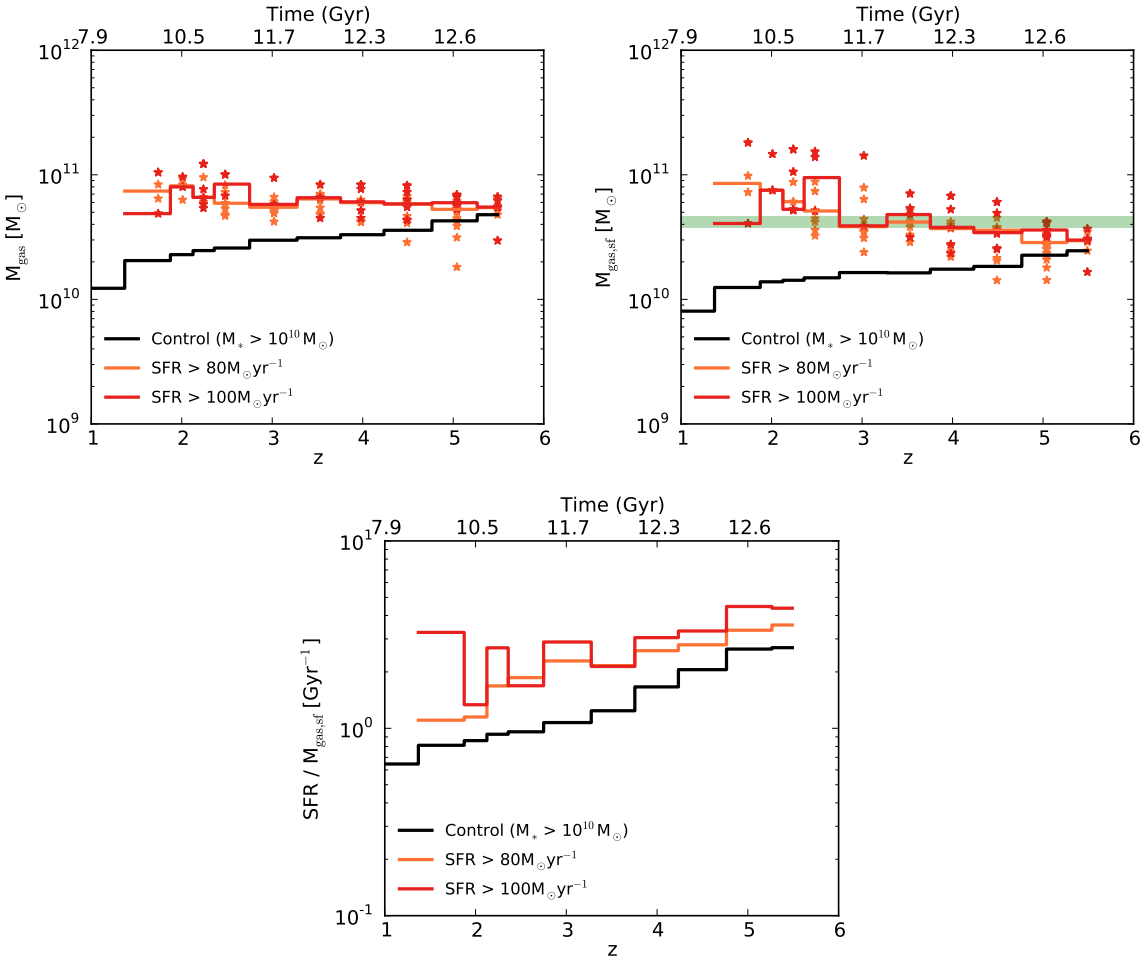

Figure 7.4: The total and star forming gas masses of highly star forming galaxies in the simulation are shown in the top panels, the efficiency of star formation is in the bottom panel. Colour coding is as is Fig. 7.2. The green band shows the range of observed molecular gas estimated for sub-mm galaxies from Swinbank et al. (2014).

the observed stellar masses, above redshift 3 the simulated galaxy masses are low by up to 0.8 dex. Higher stellar mass galaxies at these high redshifts are not probed in a simulation box size of (100 cMpc)^3.

The halo mass of the highly star forming galaxies also grows by an order of magnitude over the redshift range 5.5 to 1.5, from 10^{12} to 10^{13} $M_\odot$. Comparing to the control sample, above redshift 3 the halos of the selected sample is ~ 0.2 dex higher than the control sample, below redshift 3 the selected sample has much higher halo masses of more than 0.5 dex. Again these trends can be understood due to the mass selected nature of the control sample, as there is a tight correlation between stellar and halo mass. The green band shows the estimated halo masses of galaxies from Simpson et al. (2014). As for the stellar masses, below redshift 3 the halo masses of the highly star forming simulated galaxies are consistent with those observed.

The black hole masses of highly star forming galaxies in the simulation show a general increase of up to two orders of magnitude from redshift 5.5 to 1.5, with a very large scatter. The black hole masses show no significant difference to those of the control sample. The values are also broadly consistent with those estimated from observations by Hickox et al. (2012).

In Fig. 7.4 the total and star forming gas masses are shown as a function of redshift. The gas mass of highly star forming simulated galaxies shows very little variation with redshift and is in the region of $10^{10.5}$ $M_\odot$. At high redshifts this gas mass is similar to that of the control sample, but it deviates to lower redshift, with less gas mass in the control sample.

In contrast to the total gas mass, the star forming gas in the simulation shows an increase of 0.5 dex over the redshift range 5.5 to 1.5. This is a probe of the gas on the equation of state in the simulation, which is cold and dense. Surprisingly, the galaxies with star formation rates greater than 100 $M_\odot$ yr^{-1} do not have systematically higher star forming gas masses than those forming stars at 80 $M_\odot$ yr^{-1}. The star forming gas mass of the control sample decreases with redshift, showing the opposite trend to the highly star forming galaxies. The quantity of highly star

forming gas is within 0.3 dex of the molecular gas observed in sub-mm galaxies (Swinbank et al., 2014).

In the final panel of Fig. 7.4 the efficiency of the highest star forming galaxies is shown. This relates the rate at which stars can form to the rate at which gas is available for star formation. At high redshift gas can form stars more efficiently than at lower redshift (both for the highly star forming galaxies and the control sample). This trend will be relevant when looking at the redshift distributions of the galaxies.

7.3.2 Redshift Distributions

Many observational studies of the redshift distribution of sub-mm galaxies have been carried out, e.g. Weiß et al. (2013), Smolčić et al. (2012), Simpson et al. (2014). Studies typically find the median redshift for sub-mm galaxies to be between 2 and 3. The shape of the distribution is found to have a peak around the median, with a tail to higher redshifts, although the shape of the tail differs between the studies. It is interesting to consider the redshift distribution of the highest star forming galaxies in the simulation, to see how it qualitatively compares to the observations.

The top left panel of Fig. 7.5 shows the redshift distribution for galaxies with SFRs above 60, 80 and 100 $M_\odot$ yr^{-1} in increasingly red colours. All cuts result in a reasonably flat distribution from redshift 5 to 2, with decreasing galaxy numbers for increasing SFR cut. There are few highly star forming galaxies outside the redshift range of 2 to 5. There is some evidence that the 80 and 100 $M_\odot$ yr^{-1} galaxies in fact peak between redshift 4.5 and 5. The overall reasonably flat trend with redshift differs significantly from the peaked distribution in observations. We will return to the question of the detectability of simulated galaxies in Section 7.3.2.

The bottom panels of Fig. 7.5 show the redshift distribution of galaxies with total gas mass and star forming gas mass of greater than $10^{10.7}$, $10^{10.8}$, $10^{10.9}$ and $10^{11.0}M_\odot$. These masses were chosen to result in similar number densities of objects as those found for the SFRs.

The total gas mass distribution is more peaked than the SFR distribution and

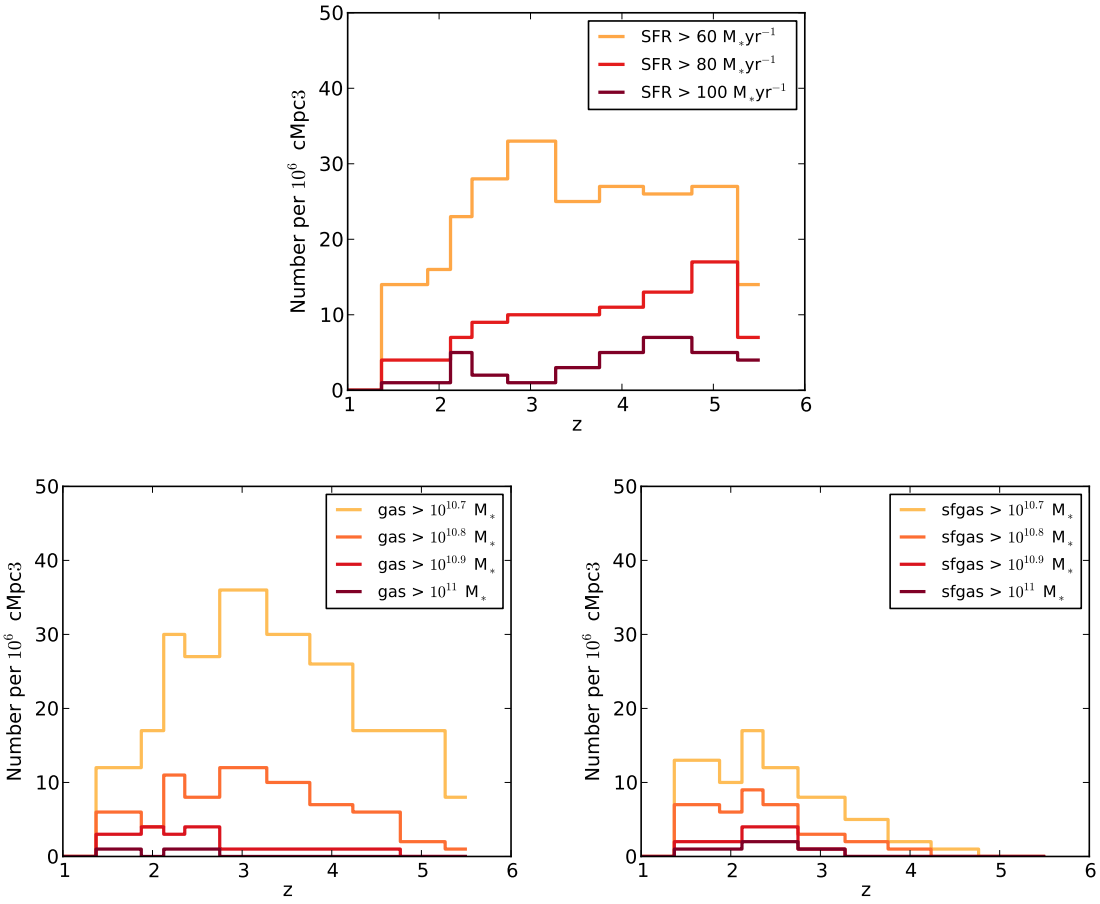

Figure 7.5: The redshift distributions of galaxies selected by SFRs, total gas and star forming gas in the simulation. We find a reasonably flat redshift distribution for the samples selected by SFR, however the gas and star forming gas selected samples are peaked around redshift of 2 with a tail to higher redshifts.

with a peak around redshift 3. The peaked distribution is even more pronounced when considering the star forming gas mass distribution with redshift, peaking around redshift 2.5. The difference in the SFR distribution and gas distribution highlights the evolving efficiency of galaxies at converting gas to stars. At high redshift galaxies are more compact, and so the higher densities can produce more efficient star formation.

The detection of sub-mm galaxies

In Fig. 7.5 we have shown the redshift distribution of highly star forming galaxies in the simulation, which is surprisingly flat relative to the expection from observations. We ask are all the selected simulated galaxies equally detectable in the sub-mm, based on a simple determination of their 850 μm flux. It is well known that one of the primary advantages of the sub-mm rest frame wavebands is that, for the same SED, the negative k-correction results in a minimal variation in the flux density with redshift, between one and six (Casey et al., 2014). However, this lack of sensitivity to redshift relies on the assumption that the galaxy SED remains unchanged.

The SED shape and normalisation depend on the SFR and dust temperature, $T_{\rm d}$. For the highly star forming galaxies in the simulation, the SFRs vary between 80 $\rm M_{\odot}$ $\rm yr^{-1}$ and 260 $\rm M_{\odot}$ $\rm yr^{-1}$, which is not a significant variation. However, $T_{\rm d}$ depends on the radius, as well as the SFR. In the left panel of Fig. 7.6 the half mass radii for the most star forming galaxies in the simulation as a function of redshift are shown. They increase by over an order of magnitude over the redshift range 5.5 to 1.5. Due to the limit of the gravitational softening, we may also over-estimate the size of galaxies at high redshift which could be more compact.

To determine $T_{\rm d}$, we convert the SFR to an infra-red luminosity, $L_{\rm IR}$, by multiplying by 10^{10} (Kennicutt, 1998) and then apply the following equation for optically thick dust,

$$T_{\rm d} = \frac{630(L_{\rm IR}[10^{12}{\rm L}_{\odot}])^{0.25}}{(R[{\rm pc}])^{0.5}}, \tag{7.1}$$

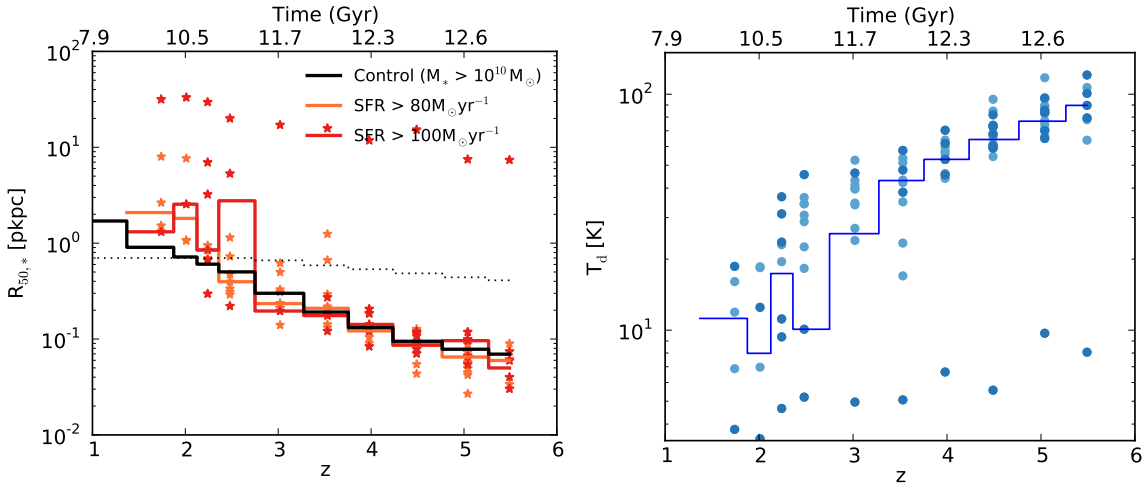

Figure 7.6: The left panel shows the half mass radius in physical units of the stars as a function of redshift from the simulation. Orange stars show galaxies with SFRs between 80 and 100 $M_\odot$ yr^{-1}, red stars show galaxies with SFRs greater than 100 $M_\odot$ yr^{-1}. Solid orange and red lines show the median half mass radius for galaxies with SFRs greater than 80 and 100 $M_\odot$ yr^{-1}. The solid black line shows the median value for the control sample. The dotted black line is the gravitational softening in the simulation, which is comoving above redshift three and fixed at 0.7 pkpc below. There is an increase of an order of magnitude in half mass radius from redshift six to one, where objects are very compact at high redshift. The right panel shows the derived dust temperature, using eq. 7.1. The blue line shows the median dust temperature for these galaxies. The increase in radius results in high dust temperature, of up to 100K at high redshift, with dust temperatures dropping to $\sim$ 10K. A typical dust temperature assumed for sub-mm galaxies at redshift two is 30K.

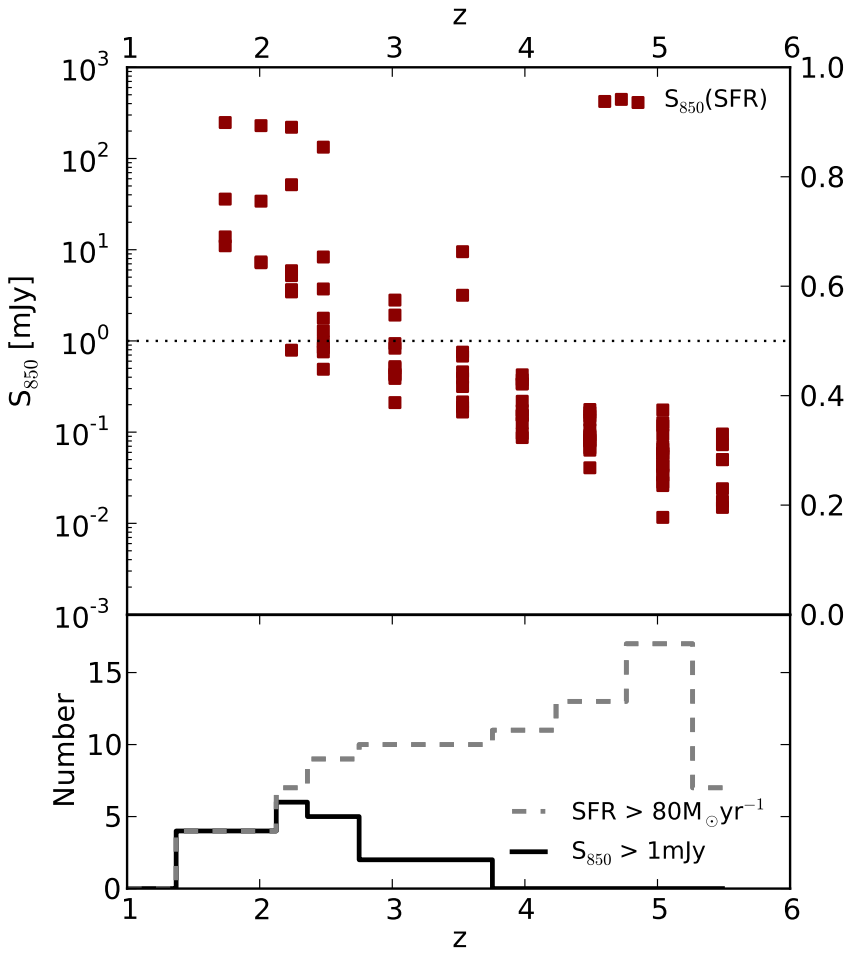

Figure 7.7: The top panel shows the estimated S_{850} mJy flux as a function of redshift for the highly star forming galaxies in the simulation. The flux is normalised to 5mJy for a 500 $M_\odot$ yr^{-1} galaxy with a dust temperature of 30 K. The dotted black line shows where the 1 mJy flux limit is. The bottom panel shows the redshift distribution of highly star forming galaxies in the simulation. The original distribution from Fig. 7.5 is shown as a dashed line, the redshift distribution obtained by imposing a detection limit of 1 mJy is shown as a solid line. The effect of the evolving galaxy radii with redshift is that many high redshift galaxies in the simulation are undetectable with current instruments.

The derived dust temperatures are shown in right panel of Fig. 7.6. The values for $T_{\rm d}$ decrease from around 100 to 10 K over the redshift range shown. Typical sub-mm dust temperatures are ~ 30 K (Swinbank et al., 2014). Over the redshift range 1.5 to 4 the values in the simulation are similar to this, varying between 10 and 50 K.

As the SED, and hence the sub-mm flux, depends on the dust temperature, we use a simple formula from Casey et al. (2014) to see if the changes in galaxy radii with redshift change the detection of galaxies. The 850 μm flux is computed using

$$S_{850} \propto \frac{L_{\rm IR}}{T_{\rm d}^{-3.5}}, \tag{7.2}$$

normalised to 5 mJy for a galaxy with $L_{\rm IR}$ of $10^{12.5}$ $L_{\odot}$ and $T_{\rm d}$ of 30 K.

In the top panel of Fig. 7.7, S_{850} is shown as a function of redshift for the highly star forming galaxies. As with the galaxy sizes, the 850 μm flux increases with redshift. Current observational surveys can detect sub-mm flux down to ~ 1 mJy (Casey et al., 2014), this limit is shown as a dotted line. Only galaxies above this line would be detected.

In the bottom panel of Fig. 7.7 the change in the redshift distribution, when applying a sub-mm flux of 1 mJy, is shown. The distribution becomes peaked around redshift 2.5, with none of the simulated galaxies at redshifts above 4 detected. The new distribution is more consistent with the redshift distribution in observations. This change to the redshift distribution also changes the number density of galaxies, on which the initial selection of galaxies was based. The change in number density will result in a lower SFR limit being selected, which will increase the tension with observed sub-mm galaxy SFRs.

As the properties of the highest star forming galaxies in the simulation vary with redshift, the detection of galaxies is affected (using a simple approximation for the 850 μm flux). This highlights that the selection of galaxies based on a SFR cut is not adequate for comparing to the sub-mm population. A complete dust modelling of the galaxies is necessary to better understand the connection between simulated galaxies and the observed sub-mm population.

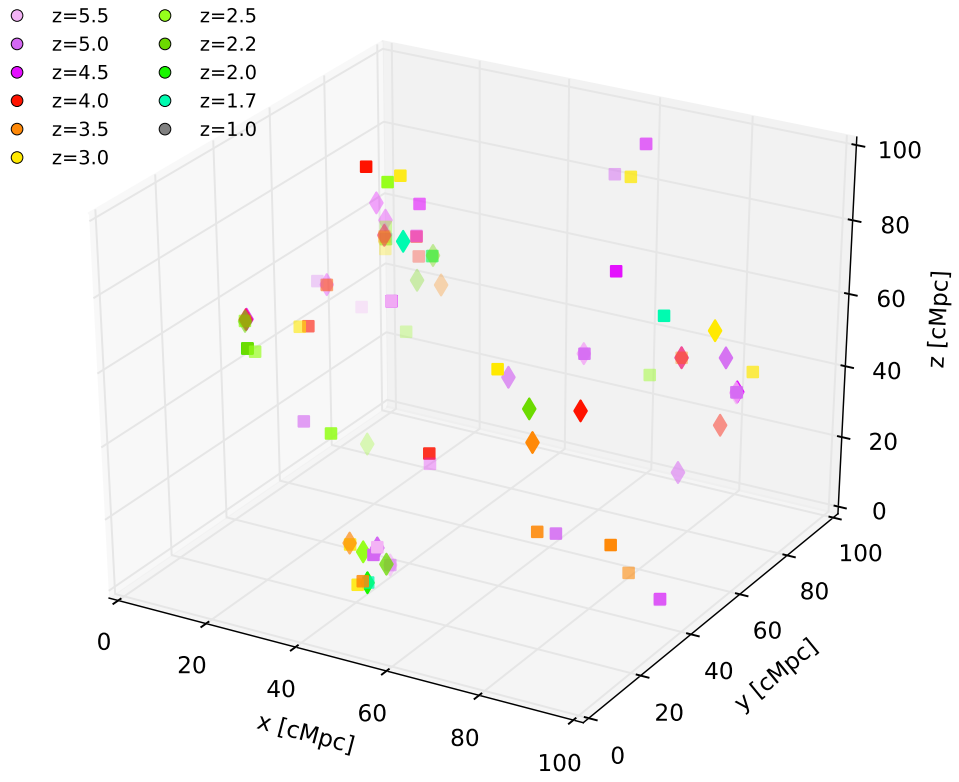

Figure 7.8: A 3D plot showing coordinates of highly star forming galaxies in co-moving Mpc in the simulation. Galaxies with SFRs between 80 and 100 $M_\odot$ yr^{-1} are shown as squares and greater than 100 $M_\odot$ yr^{-1} are shown as diamonds. Points are coloured by redshift, from red to blue going from redshift 6 to 1 redshift. We can see that highly star forming galaxies are clustered in space.

7.3.3 Clustering

One interesting aspect of the sub-mm galaxies to investigate their clustering. While a complete analysis is yet to be carried out, as an initial step we provide a 3D visual impression of the highly star forming galaxy population in the simulation across redshifts 5.5 to 1. Points are coloured by redshift and symbols represent SFR. It is clear that many of the highly star forming galaxies in the simulation are indeed highly clustered. One such example is around (10, 90, 60), while another is at (30,

40, 10). It is possible that the bursts of high star formation in these objects are driven by mergers with other galaxies in their vicinity. However, not all high star formation episodes are clustered, in these cases further investigation is require to determine the environments of these galaxies and the cause of the enhanced star formation.

7.4 Summary and future work

In this initial exploration of sub-mm galaxies in the EAGLE Ref-L100N1504 simulation, a comparison has been carried out between the highly star forming galaxies in the simulation and observations. Galaxies from the simulation are selected based on the observed number density of sub-mm galaxies, determining a cut in SFR. The SFRs of the selected galaxies in the simulation are lower than observed by a factor of more than two, and the extreme SFRs of observed galaxies, of ~ 1000 $M_\odot$ yr^{-1}, are not seen. The previous tension between the SFRs in observations and simulation for the total galaxy population was highlighted in *Chapters 3 and 4*. As this offset for the total galaxy population with SFR is approximately constant with redshift, we adjust for it by increasing the SFRs by a factor of two. This helps to resolve the offset with the median SFR for the sub-mm population. The extreme SFRs observed however present more of a challenge to understand.

In spite of the difference in SFRs, many other galaxy properties of the selected simulated galaxies are consistent with estimates for the observed sub-mm population. In particular, stellar masses, halo masses, black hole masses and cold gas masses were all found to be in general agreement with observations.

The redshift distribution of the simulated sub-mm galaxy population is reasonably flat with redshift, between redshifts 5 and 2. This differs from what is typically found in observations, where the distribution is peaked around redshift 2 to 3. A more detailed investigation into the redshift distribution in the simulation was then carried out. The sub-mm flux was computed using a simple formula, based on the SFRs and radii of the simulated galaxies. This revealed an interesting result, that

many of the high redshift galaxies in the simulation would indeed not be detected based on current detection limits, due to their small radii, resulting in high dust temperatures. Applying a flux limit to the simulation, the recovered distribution is peaked at redshift 2.5, in reasonable agreement with observations. The conclusion of this work is, however, that a more detailed analysis of the simulated galaxy population is required, including radiative transfer, to carry out a comparison with the sub-mm population.

In spite of the requirement for further analysis to do a comparison with observational data, the highest star forming galaxies in the simulation are an interesting population to study in their own right. We have begun to study the triggering and quenching mechanisms of these galaxies in the simulation using the EAGLE merger trees. Connected to the triggering of galaxies, the environments of highly star forming galaxies are of interest. Clustering studies are also underway to look at this in more detail.

Chapter 8 | *Conclusions*

In recent years the abundance of observational data of galaxies has expanded significantly, more specifically, wide field surveys now cover a large fraction of cosmic time with reasonable depths. These observations challenge our understanding of galaxy formation and demand more of galaxy formation theory, in its scope and complexity. In this work we look to cosmological hydrodynamical simulations to determine the ability of models to reproduce realistic virtual universes, in light of the broad redshift coverage of observations. We then use the simulations to understand the growth and evolution of galaxies, from the first galaxies formed to the properties of the redshift zero population.

The analysis in this thesis is based on the EAGLE project (Schaye et al., 2015), a suite of new hydrodynamical simulations. The largest of these simulations, RefL100N1504, took 5 million CPU hours to run, while the code development and calibration required 3 years. The time invested in the project has resulted in a simulation with unprecedented agreement to the observed redshift zero galaxy stellar mass function for a hydrodynamical simulation (Schaye et al., 2015).

As a builder of this simulation suite, I have been involved in development, optimisation and testing of the code, as well as running and calibrating simulations. The EAGLE code, as described in *Chapter 2*, is a GADGET (Springel, 2005) based code, with an improved SPH implementation (Dalla Vecchia, in prep.) and subgrid physics prescriptions based on the OWLS (Schaye et al., 2010) and GIMIC (Crain et al., 2009) projects. The subgrid physics prescriptions in the simulations encompass a range of processes known to be important in the modelling of baryons. These include star formation, radiative cooling, metal enrichment, stellar feedback, black hole growth and AGN feedback. As hydrodynamical simulations trace the baryons, and the processes by which they are affected, concurrently with the dark

matter, their gravitational impact on the dark matter is inherent in the simulations. The baryons in the IGM and CGM are sampled in the simulation and, in particular, the accretion onto, expulsion from and recycling through halos is explicitly tracked. The primary advantage of the hydrodynamical simulations is the more stringent way of implementing subgrid physics processes than is possible in other theoretical models, providing a more self-consistent model.

A particular advantage of the models from the EAGLE project amongst other hydrodynamical simulations is that the free parameters in the feedback prescriptions are calibrated to reproduce the observed galaxy stellar mass function at redshift zero and the galaxy sizes; the redshift zero galaxy populations considered are in dark matter halos of similar masses to those in the real Universe (assuming the cosmology of the simulation is accurate). This implies that the accretion histories of these galaxies (modulo recycling) is representative of galaxies in the real Universe, and hence it is of particular interest to compare to observations and to understand the simulated galaxy population's evolution.

We began in *Chapter 3* by carrying out a multi-epoch comparison between the Reference and Recalibrated EAGLE models and observational data. The focus of the comparison was galaxy stellar masses and star formation rates. Hydrodynamical simulations previously struggled to reproduced the observed evolution in these galaxy properties, however, the Ref-L100N1504 simulation reproduces the observed trends. While the level of agreement is within the systematic uncertainty in the observations, we noted a potential issue with the specific star formation rates, which were found to be 0.2 - 0.3 dex low in the simulation at all redshifts. Overall, however, the evolution of the EAGLE virtual universe was shown to be representative of the observed evolution.

The issue with specific star formation rates has been reported previously, for both hydrodynamical and semi-analytical models (Weinmann et al., 2012; Mitchell et al., 2014; White et al., 2014). In *Chapter 4* we investigated this problem, along with the evolution of the normalisation of the galaxy stellar mass function, another common problem. A number of EAGLE simulations varying the stellar and AGN

feedback schemes were considered. The study revealed some interesting conclusions, (i) any problems with the evolution of the normalisation of the galaxy stellar mass function reported in previous work can be resolved with feedback recipes that preferentially suppress star formation at high redshifts, (ii) the preferential suppression of star formation at high redshifts is also necessary to reproduce results consistent with the observed stellar mass density across cosmic time, (iii) only an extreme model can reproduce the observed specific star formation rates across cosmic time. The extreme model failed to reproduce any other observables considered. The results of this study suggest that there is potential inconsistency with the observed specific star formation rates and the observations of stellar mass.

Following the success of the Reference EAGLE model across cosmic time, in *Chapter 5* we consider how the redshift zero population is formed and what the simulation can reveal about the growth of galaxies. The central and satellite galaxies test the efficiency of environmental quenching in the simulation. The star forming populations in different environments have very similar properties, although satellite galaxies, in higher density environments than centrals of equivalent mass, have a higher quenched fraction. The quenching process of satellites occurs on short timescales. The growth and shape of the galaxy stellar mass function was also considered. It is the star forming galaxies that dominate the galaxy stellar mass function at low masses, the exponential break, however, is controlled by the passive galaxies. The high mass passive galaxies in the simulation are quenched by AGN feedback.

Having compared the galaxy population to observations at redshifts 7 and below, we explore the properties of the first galaxies formed in the simulation in *Chapter 6*. These galaxies are responsible for producing a considerable abundance of ionising photons in the early Universe, and have not yet been studied in a simulation with such good constraints on the evolution of galaxy properties. The agreement of the EAGLE simulation with observations across cosmic time sets a limit on the ionising photons that can be produced by galaxies at high redshifts, while remaining consistent with observations. The dominant source of ionising photons in the

early universe from galaxies is those with masses less than 10^8 $M_\odot$ exhibiting star formation rates significantly above the median value. In the local Universe, the remnants of stars that produced the ionising photons can be found in Milky Way mass galaxies, and higher.

Finally, in *Chapter 7* we look at the most highly star forming galaxies (> 80 $M_\odot$ yr^{-1}) in the simulation. This is a preliminary work, but already reveals some interesting results. Galaxies are selected using a star formation rate cut based on the observed number density of sub-mm galaxies, which are typically high-redshift, dusty, star-bursting galaxies. The simulated galaxies are shown to be consistent with observed sub-mm galaxies in many of their properties. However their star formation rates have a lower median than those observed, and are less extreme - even after accounting for the systematic offset in specific star formation rates reported in *Chapter 3*. A solution to this issue may lie in the physics of star formation included in the simulation, for example, variations in the stellar initial mass function have been shown to alleviate the discrepancy between models and observations. Although, extreme observed star formation rates have been decreasing with improved observational constrains for carrying out spectral energy density fitting. The redshift distribution of the highly star-forming galaxies was found to be flat with redshift, differing from observations. However, a further investigation revealed that the selection criteria of a star formation rate cut was inadequate to select sub-mm like galaxies. The dust temperatures of the galaxies play an important role in the selections. There is much more work to be carried out on this topic, with only an initial investigation presented here.

8.1 Future Work

Many of the topics discussed in this thesis raise further questions and highlight interesting topics for future investigation. Those that we intend to address in the near future are as follows,

- Galaxy quenching in the simulation follows similar trends to observations

and comes about through environmental processes and AGN feedback. We are looking into the contribution of different environmental processes, such as stripping and strangulation, in producing these effects. The triggering mechanisms for AGN feedback and the connection between AGN luminosity and star formation rates are also under investigation. The further work on quenching will also consider the phenomenon of galaxy conformity.

- The study on reionization will be expanded to consider high resolution resimulations of the Local Group (Sawala et al., 2014), employing similar physics to the EAGLE cosmological simulations. We will ask whether the local group produced sufficient ionising photons to reionize itself, or if external photons were required. The range of resolutions available within the Local Group simulations will also allow more complete resolution tests to be performed.

- The triggering and quenching mechanisms of highly star forming galaxies will be considered in the simulation, along with their clustering. From observations it is often difficult to determine if sub-mm galaxies are experiencing a merger, due to the range of merger stages and sometimes limited by resolution. Understanding if, in the simulation, it is mergers or disc instabilities that drive the high star formation rates can help guide observations.

- One of the most interesting findings is the continued inability of models to reproduce the observed extreme star formation rates without appealing to non-standard physics, such as varying stellar initial mass functions. We intend to carry out radiative transfer on the highly star forming objects in the simulation to provide a more realistic comparison to observational data and continue the investigation into the discrepancy.

There are many other projects based on the EAGLE simulations currently underway. For example, a study of all EAGLE variations (Crain et al., 2015) will further motivate the choice of Reference model. Projects on galaxy colours and luminosities (Trayford and et al., in prep.) and galaxy histories using the merger tree data

(Qu and et al., in prep.) are also being carried out. With hydrodynamical simulations that are capable of producing virtual universes, the scope for exploring galaxy formation increases. As this simulation will be publicly available, the benefit will hopefully be experienced by a wide community.

Beyond the EAGLE project, an obvious future goal is to build bigger and better simulation suites with future generation software and hardware. But what does bigger and better mean? Higher resolution simulations, in larger volumes, will expand the dynamic range of galaxy formation that can be probed simultaneously. However, the EAGLE simulation resolution and volume is at the boundary of what can be achieved with the current machinery and increasing both resolution and volume concurrently is a distant future goal.

If we increase resolution there is a concern that the adequacy of the subgrid physics schemes is not sufficient. What is the important physics, such as molecular cooling or magnetic fields, that is not accounted for? Is it possible to improve on the subgrid physics implementations that are already included? Small scale, high resolution simulations can provide more physically motivated prescriptions for cosmological simulations. Before investing significant time in high resolution simulations, such physically motivated models should be included.

Increasing the volume of simulations is less computationally expensive than resolution increases. Certain studies such as those on the sub-mm galaxy population or studies on galaxy cluster properties would benefit from the larger samples in larger volumes. However, the increased volumes would not provide any further physical insight into the intermediate mass galaxies that are well sampled in a volume such as that used in EAGLE. Perhaps of more interest is the possibility to build mock catalogues with larger volume simulations. Such mock catalogues would be a valuable resource for observers.

There are many potential directions for future cosmological hydrodynamical simulations. With the fantastic computing facilities available and the ever-improving codes, the next decade will undoubtedly see an exploration of these possibilities.

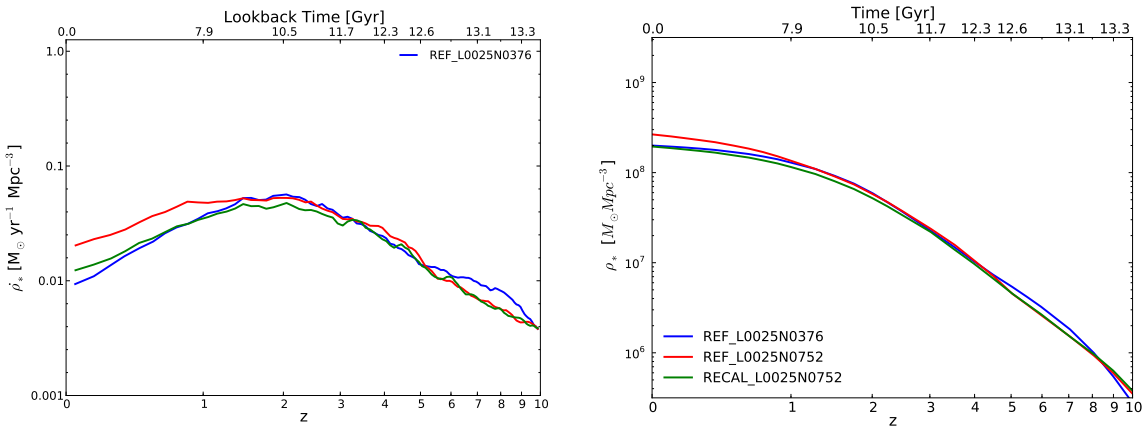

Figure A.1: The star formation rate density and stellar mass density as a function of redshift in the top and bottom panels for Ref-L025N0376, using the same physics as for Ref-L100N1504 shown in all previous plots, Ref-L025N0752, with eight times the resolution and Recal-L025N0752, with eight times higher resolution and recalibrated subgrid parameters in blue, red and green respectively.

Appendix A

Volume average numerical convergence

Here we show resolution tests for the evolution of stellar masses and star formation rates in the EAGLE simulation to provide an overview of the convergence found. Three models are considered, Ref-L025N0376, equivalent in resolution to Ref-L100N1504 except in a 25 cMpc box as opposed to 100 cMpc, Ref-L025N0752, with the same subgrid parameters as Ref-L100N1504 with 8 times higher mass resolution in a 25 cMpc box and Recal-L025N0752, with recalibrated subgrid parameters, 8 times higher mass resolution in a 25 cMpc box. In all three cases we compare 25 cMpc boxes due to the volume averaging required for the stellar mass density and star formation rate density. The small box size results in less extreme regions, for example there are no groups or clusters in the box. As these typically fall in

overdense regions with higher stellar mass density and different star formation histories (as seen in Fig. 3.5) than galaxies outside groups and clusters, the box size affects the ρ_* and $\rho_{\rm SFR}$. To ensure we do not obscure the effects of resolution with other effects such as box size, we carry out this test using the same box size for all models.

In this comparison we use two high resolution models to test strong and weak convergence in the simulation. In S14 it is argued that aiming to produce a simulation with subgrid models that does not require parameter changes when varying the resolution is not necessarily resulting in a 'better' model. For example, with the aim of achieving resolution independence feedback variations often depend on dark matter properties, which are more robust against resolution effects than baryonic properties. However, this results in simulations with subgrid models driven by unphysical parameterisations. As a result, along with the high resolution run without any parameter changes, we have produced a model at high resolution that has been recalibrated to the same criteria as the standard resolution model. We refer to test with the unchanged parameters, Ref-L025N0752, as strong convergence tests and with the recalibrated parameters, Recal-L025N0752, as weak convergence tests.

First we consider both weak and strong convergence tests for the $\rho_{\rm SFR}$ and ρ_* in Fig. A.1. For the $\rho_{\rm SFR}$ between redshifts 6 and 10 we see and excess of star formation in the standard resolution model relative to both higher resolution models. This follows from the higher mass particles, and a coarser minimum star formation rate per particle at the standard resolution. Between redshifts 6 and 1 the two Ref models have similar $\rho_{\rm SFR}$. Below redshift one the higher resolution simulation shows more star formation than the standard resolution. The Ref-L025N0376 model and Recal-L025N0752 model are similar across the redshift range 6 to 0, although the Recal model peak in star formation is below the standard resolution Ref model. From redshift 0.5 Recal-L025N0752 shows a slight upturn relative to Ref-L025N0376. Overall, the star formation rate density for the Ref-L025N0376 and Recal-L025N0752 lie within 0.1 dex of each other, showing good weak convergence

for this property. The Ref-L025N0376 and Ref-L025N0752 have a maximum offset of 0.3 dex, that occurs at redshift zero. The ρ_*, showing the integral of the star formation rate density, shows similar trends across redshift to $\rho_{\rm SFR}$ when comparing the different models. Between redshifts 7 and 4, there is an excess of 20% in the standard resolution model relative to the higher resolution simulations. Below redshift 1 there is an increase in the stellar mass density in the Ref-L025N0752 simulation relative to the standard resolution model consistent with the higher $\rho_{\rm SFR}$, although this amount to only around 0.1 dex.

Overall the level of agreement shown for the global evolution properties for both weak and strong convergence is good.

Bibliography

G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *Astrophys. J. Suppl.*, 208:19, October 2013. doi: 10.1088/0067-0049/208/2/19.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. *ArXiv e-prints, arxiv:1303.5076*, March 2013.

M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, T. Budavari, A. Connolly, D. J. Eisenstein, D. Finkbeiner, J. A. Frieman, J. E. Gunn, A. J. S. Hamilton, L. Hui, B. Jain, D. Johnston, S. Kent, H. Lin, R. Nakajima, R. C. Nichol, J. P. Ostriker, A. Pope, R. Scranton, U. Seljak, R. K. Sheth, A. Stebbins, A. S. Szalay, I. Szapudi, L. Verde, Y. Xu, J. Annis, N. A. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, I. Csabai, J. Loveday, M. Doi, M. Fukugita, J. R. Gott, III, G. Hennessy, D. W. Hogg, Ž. Ivezić, G. R. Knapp, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. A. McKay, P. Kunszt, J. A. Munn, L. O'Connell, J. Peoples, J. R. Pier, M. Richmond, C. Rockosi, D. P. Schneider, C. Stoughton, D. L. Tucker, D. E. Vanden Berk, B. Yanny, D. G. York, and SDSS Collaboration. The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey. *ApJ*, 606:702–740, May 2004. doi: 10.1086/382125.

S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, N. J. G. Cross, G. Dalton, V. R. Eke, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, K. Glazebrook, C. Jackson, A. Jenkins, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, B. A. Peterson, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362:505–534, September 2005. doi: 10.1111/j.1365-2966.2005.09318.x.

B. A. Reid, W. J. Percival, D. J. Eisenstein, L. Verde, D. N. Spergel, R. A. Skibba, N. A. Bahcall, T. Budavari, J. A. Frieman, M. Fukugita, J. R. Gott, J. E. Gunn, Ž. Ivezić, G. R. Knapp, R. G. Kron, R. H. Lupton, T. A. McKay, A. Meiksin, R. C. Nichol, A. C. Pope, D. J. Schlegel, D. P. Schneider, C. Stoughton, M. A. Strauss, A. S. Szalay, M. Tegmark, M. S. Vogeley, D. H. Weinberg, D. G. York, and I. Zehavi. Cosmological constraints from the clustering of the Sloan Digital Sky Survey DR7 luminous red galaxies. *MNRAS*, 404:60–85, May 2010. doi: 10.1111/j.1365-2966.2010.16276.x.

A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, H. Ebeling, W. R. Forman, A. Hornstrup, C. Jones, S. S. Murray, D. Nagai, H. Quintana, and A. Voevodkin. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *ApJ*, 692:1060–1074, February 2009. doi: 10.1088/0004-637X/692/2/1060.

D. J. Reiss, L. M. Germany, B. P. Schmidt, and C. W. Stubbs. The Mount Stromlo Abell cluster supernova search. *AJ*, 115:26, January 1998. doi: 10.1086/300191.

M. Kowalski, D. Rubin, G. Aldering, R. J. Agostinho, A. Amadon, R. Amanullah, C. Balland, K. Barbary, G. Blanc, P. J. Challis, A. Conley, N. V. Connolly, R. Covarrubias, K. S. Dawson, S. E. Deustua, R. Ellis, S. Fabbro, V. Fadeyev, X. Fan, B. Farris, G. Folatelli, B. L. Frye, G. Garavini, E. L. Gates, L. Germany, G. Goldhaber, B. Goldman, A. Goobar, D. E. Groom, J. Haissinski, D. Hardin, I. Hook, S. Kent, A. G. Kim, R. A. Knop, C. Lidman, E. V. Linder, J. Mendez, J. Meyers, G. J. Miller, M. Moniez, A. M. Mourão, H. Newberg, S. Nobili, P. E. Nugent, R. Pain, O. Perdereau, S. Perlmutter, M. M. Phillips, V. Prasad, R. Quimby, N. Regnault, J. Rich, E. P. Rubenstein, P. Ruiz-Lapuente, F. D. Santos, B. E. Schaefer, R. A. Schommer, R. C. Smith, A. M. Soderberg, A. L. Spadafora, L.-G. Strolger, M. Strovink, N. B. Suntzeff, N. Suzuki, R. C. Thomas, N. A. Walton, L. Wang, W. M. Wood-Vasey, J. L. Yun, and Supernova Cosmology Project. Improved Cosmological Constraints from New, Old, and Combined Supernova Data Sets. *ApJ*, 686:749–778, October 2008. doi: 10.1086/589937.

S. Perlmutter, G. Aldering, M. della Valle, S. Deustua, R. S. Ellis, S. Fabbro, A. Fruchter, G. Goldhaber, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, R. A. Knop, C. Lidman, R. G. McMahon, P. Nugent, R. Pain, N. Panagia, C. R. Pennypacker, P. Ruiz-Lapuente, B. Schaefer, and N. Walton. Discovery of a supernova explosion at half the age of the universe. *Nat*, 391:51, January 1998. doi: 10.1038/34124.

T. M. Heckman, M. D. Lehnert, D. K. Strickland, and L. Armus. Absorption-Line Probes of Gas and Dust in Galactic Superwinds. *Astrophys. J. Suppl.*, 129:493–516, August 2000. doi: 10.1086/313421.

C. M. Schwartz, C. L. Martin, R. Chandar, C. Leitherer, T. M. Heckman, and M. S. Oey. Kinematics of Interstellar Gas in Nearby UV-selected Galaxies Measured with HST STIS Spectroscopy. *ApJ*, 646:858–871, August 2006. doi: 10.1086/504961.

K. A. Kornei, A. E. Shapley, C. L. Martin, A. L. Coil, J. M. Lotz, D. Schiminovich, K. Bundy, and K. G. Noeske. The Properties and Prevalence of Galactic Outflows at z ˜ 1 in the Extended Groth Strip. *ApJ*, 758:135, October 2012. doi: 10.1088/0004-637X/758/2/135.

J. Kormendy and L. C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *ARA&A*, 51:511–653, August 2013. doi: 10.1146/annurev-astro-082708-101811.

M. Volonteri and J. Bellovary. Black holes in the early Universe. *Reports on Progress in Physics*, 75 (12):124901, December 2012. doi: 10.1088/0034-4885/75/12/124901.

A. R. King and K. A. Pounds. Black hole winds. *MNRAS*, 345:657–659, October 2003. doi: 10.1046/j.1365-8711.2003.06980.x.

S. S. Murray, A. Kenter, W. R. Forman, C. Jones, P. J. Green, C. S. Kochanek, A. Vikhlinin, D. Fabricant, G. Fazio, K. Brand, M. J. I. Brown, A. Dey, B. T. Jannuzi, J. Najita, B. McNamara, J. Shields, and M. Rieke. XBootes: An X-Ray Survey of the NDWFS Bootes Field. I. Overview and Initial Results. *Astrophys. J. Suppl.*, 161:1–8, November 2005. doi: 10.1086/444378.

T. Di Matteo, V. Springel, and L. Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. *Nat*, 433:604–607, February 2005. doi: 10.1038/nature03335.

A. C. Fabian, A. Celotti, and M. C. Erlund. Radiative pressure feedback by a quasar in a galactic bulge. *MNRAS*, 373:L16–L20, November 2006. doi: 10.1111/j.1745-3933.2006.00234.x.

T. A. Thompson, A. C. Fabian, E. Quataert, and N. Murray. Dynamics of Dusty Radiation Pressure Driven Shells: Fast Outflows from Galaxies, Star Clusters, Massive Stars, and AGN. *ArXiv e-prints*, June 2014.

S. P. Driver, P. Norberg, I. K. Baldry, S. P. Bamford, A. M. Hopkins, J. Liske, J. Loveday, J. A. Peacock, D. T. Hill, L. S. Kelvin, A. S. G. Robotham, N. J. G. Cross, H. R. Parkinson, M. Prescott, C. J. Conselice, L. Dunne, S. Brough, H. Jones, R. G. Sharp, E. van Kampen, S. Oliver, I. G. Roseboom, J. Bland-Hawthorn, S. M. Croom, S. Ellis, E. Cameron, S. Cole, C. S. Frenk, W. J. Couch, A. W. Graham, R. Proctor, R. De Propris, I. F. Doyle, E. M. Edmondson, R. C. Nichol, D. Thomas, S. A. Eales, M. J. Jarvis, K. Kuijken, O. Lahav, B. F. Madore, M. Seibert, M. J. Meyer, L. Staveley-Smith, S. Phillipps, C. C. Popescu, A. E. Sansom, W. J. Sutherland, R. J. Tuffs, and S. J. Warren. GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50(5):12–5, October 2009. doi: 10.1111/j.1468-4004.2009.50512.x.

D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs,

J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120:1579–1587, September 2000. doi: 10.1086/301513.

M. Colless, B. A. Peterson, C. Jackson, J. A. Peacock, S. Cole, P. Norberg, I. K. Baldry, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, W. Sutherland, and K. Taylor. The 2dF Galaxy Redshift Survey: Final Data Release. *ArXiv Astrophysics e-prints*, June 2003.

O. Ilbert, H. J. McCracken, O. Le Fèvre, P. Capak, J. Dunlop, A. Karim, M. A. Renzini, K. Caputi, S. Boissier, S. Arnouts, H. Aussel, J. Comparat, Q. Guo, P. Hudelot, J. Kartaltepe, J. P. Kneib, J. K. Krogager, E. Le Floc'h, S. Lilly, Y. Mellier, B. Milvang-Jensen, T. Moutard, M. Onodera, J. Richard, M. Salvato, D. B. Sanders, N. Scoville, J. D. Silverman, Y. Taniguchi, L. Tasca, R. Thomas, S. Toft, L. Tresse, D. Vergani, M. Wolk, and A. Zirm. Mass assembly in quiescent and star-forming galaxies since z $\sim$ 4 from UltraVISTA. A&A, 556:A55, August 2013. doi: 10.1051/0004-6361/201321100.

A. Muzzin, D. Marchesini, M. Stefanon, M. Franx, H. J. McCracken, B. Milvang-Jensen, J. S. Dunlop, J. P. U. Fynbo, G. Brammer, I. Labbé, and P. G. van Dokkum. The Evolution of the Stellar

Mass Functions of Star-forming and Quiescent Galaxies to z = 4 from the COSMOS/UltraVISTA Survey. *ApJ*, 777:18, November 2013. doi: 10.1088/0004-637X/777/1/18.

R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Stiavelli, P. van Dokkum, M. Trenti, D. Magee, I. Labbé, M. Franx, C. M. Carollo, and V. Gonzalez. Discovery of z ˜ 8 Galaxies in the Hubble Ultra Deep Field from Ultra-Deep WFC3/IR Observations. *ApJ*, 709:L133–L137, February 2010. doi: 10.1088/2041-8205/709/2/L133.

P. A. Oesch, R. J. Bouwens, G. D. Illingworth, I. Labbé, R. Smit, M. Franx, P. G. van Dokkum, I. Momcheva, M. L. N. Ashby, G. G. Fazio, J.-S. Huang, S. P. Willner, V. Gonzalez, D. Magee, M. Trenti, G. B. Brammer, R. E. Skelton, and L. R. Spitler. The Most Luminous z ˜ 9-10 Galaxy Candidates Yet Found: The Luminosity Function, Cosmic Star-formation Rate, and the First Mass Density Estimate at 500 Myr. *ApJ*, 786:108, May 2014. doi: 10.1088/0004-637X/786/2/108.

J. P. Stott, D. Sobral, A. M. Swinbank, I. Smail, R. Bower, P. N. Best, R. M. Sharples, J. E. Geach, and J. Matthee. A relationship between specific star formation rate and metallicity gradient within z 1 galaxies from KMOS-HiZELS. *MNRAS*, 443:2695–2704, September 2014. doi: 10.1093/mnras/stu1343.

M. P. van Daalen, J. Schaye, I. G. McCarthy, C. M. Booth, and C. D. Vecchia. The impact of baryonic processes on the two-point correlation functions of galaxies, subhaloes and matter. *MNRAS*, 440: 2997–3010, June 2014. doi: 10.1093/mnras/stu482.

M. Schaller and et al. Eagle: SPH paper. in prep.

A. P. Cooper, S. Cole, C. S. Frenk, S. D. M. White, J. Helly, A. J. Benson, G. De Lucia, A. Helmi, A. Jenkins, J. F. Navarro, V. Springel, and J. Wang. Galactic stellar haloes in the CDM model. *MNRAS*, 406:744–766, August 2010. doi: 10.1111/j.1365-2966.2010.16740.x.

T. Sawala, C. S. Frenk, R. A. Crain, A. Jenkins, J. Schaye, T. Theuns, and J. Zavala. The abundance of (not just) dark matter haloes. *MNRAS*, 431:1366–1382, May 2013. doi: 10.1093/mnras/stt259.

S. J. Lilly, C. M. Carollo, A. Pipino, A. Renzini, and Y. Peng. Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos. *ApJ*, 772:119, August 2013. doi: 10.1088/0004-637X/772/2/119.

Y.-j. Peng, S. J. Lilly, K. Kovač, M. Bolzonella, L. Pozzetti, A. Renzini, G. Zamorani, O. Ilbert, C. Knobel, A. Iovino, C. Maier, O. Cucciati, L. Tasca, C. M. Carollo, J. Silverman, P. Kampczyk, L. de Ravel, D. Sanders, N. Scoville, T. Contini, V. Mainieri, M. Scodeggio, J.-P. Kneib, O. Le Fèvre,

S. Bardelli, A. Bongiorno, K. Caputi, G. Coppa, S. de la Torre, P. Franzetti, B. Garilli, F. Lamareille, J.-F. Le Borgne, V. Le Brun, M. Mignoli, E. Perez Montero, R. Pello, E. Ricciardelli, M. Tanaka, L. Tresse, D. Vergani, N. Welikala, E. Zucca, P. Oesch, U. Abbas, L. Barnes, R. Bordoloi, D. Bottini, A. Cappi, P. Cassata, A. Cimatti, M. Fumana, G. Hasinger, A. Koekemoer, A. Leauthaud, D. Maccagni, C. Marinoni, H. McCracken, P. Memeo, B. Meneux, P. Nair, C. Porciani, V. Presotto, and R. Scaramella. Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function. *ApJ*, 721:193–221, September 2010. doi: 10.1088/0004-637X/721/1/193.

R. E. Angulo, V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk. Scaling relations for galaxy clusters in the Millennium-XXL simulation. *MNRAS*, 426:2046–2062, November 2012. doi: 10.1111/j.1365-2966.2012.21830.x.

V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White. The Aquarius Project: the subhaloes of galactic haloes. *MNRAS*, 391:1685–1711, December 2008. doi: 10.1111/j.1365-2966.2008.14066.x.

A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole. What Shapes the Luminosity Function of Galaxies? *ApJ*, 599:38–49, December 2003. doi: 10.1086/379160.

C. M. Baugh. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156, December 2006. doi: 10.1088/0034-4885/69/12/R02.

P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. *ApJ*, 770:57, June 2013. doi: 10.1088/0004-637X/770/1/57.

B. P. Moster, T. Naab, and S. D. M. White. Galactic star formation and accretion histories from matching galaxies to dark matter haloes. *MNRAS*, 428:3121–3138, February 2013. doi: 10.1093/mnras/sts261.

S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk. Hierarchical galaxy formation. *MNRAS*, 319: 168–204, November 2000. doi: 10.1046/j.1365-8711.2000.03879.x.

A. J. Benson and R. Bower. Galaxy formation spanning cosmic history. *MNRAS*, 405:1573–1623, July 2010. doi: 10.1111/j.1365-2966.2010.16592.x.

R. G. Bower, I. Vernon, M. Goldstein, A. J. Benson, C. G. Lacey, C. M. Baugh, S. Cole, and C. S. Frenk. The parameter space of galaxy formation. *MNRAS*, 407:2017–2045, October 2010. doi: 10.1111/j.1365-2966.2010.16991.x.

R. G. Bower, A. J. Benson, and R. A. Crain. What shapes the galaxy mass function? Exploring the roles of supernova-driven winds and active galactic nuclei. *MNRAS*, 422:2816–2840, June 2012. doi: 10.1111/j.1365-2966.2012.20516.x.

A. Pontzen and F. Governato. Cold dark matter heats up. *Nat*, 506:171–178, February 2014. doi: 10.1038/nature12953.

I. G. McCarthy, C. S. Frenk, A. S. Font, C. G. Lacey, R. G. Bower, N. L. Mitchell, M. L. Balogh, and T. Theuns. Ram pressure stripping the hot gaseous haloes of galaxies in groups and clusters. *MNRAS*, 383:593–605, January 2008. doi: 10.1111/j.1365-2966.2007.12577.x.

C. Scannapieco, M. Wadepuhl, O. H. Parry, J. F. Navarro, A. Jenkins, V. Springel, R. Teyssier, E. Carlson, H. M. P. Couchman, R. A. Crain, C. Dalla Vecchia, C. S. Frenk, C. Kobayashi, P. Monaco, G. Murante, T. Okamoto, T. Quinn, J. Schaye, G. S. Stinson, T. Theuns, J. Wadsley, S. D. M. White, and R. Woods. The Aquila comparison project: the effects of feedback and numerical methods on simulations of galaxy formation. *MNRAS*, 423:1726–1749, June 2012. doi: 10.1111/j.1365-2966.2012.20993.x.

R. C. Kennicutt, Jr. The Global Schmidt Law in Star-forming Galaxies. *ApJ*, 498:541–552, May 1998. doi: 10.1086/305588.

J. Schaye and C. Dalla Vecchia. On the relation between the Schmidt and Kennicutt-Schmidt star formation laws and its implications for numerical simulations. *MNRAS*, 383:1210–1222, January 2008. doi: 10.1111/j.1365-2966.2007.12639.x.

V. Springel and L. Hernquist. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *MNRAS*, 339:289–311, February 2003. doi: 10.1046/j.1365-8711.2003.06206.x.

J. Schaye, C. Dalla Vecchia, C. M. Booth, R. P. C. Wiersma, T. Theuns, M. R. Haas, S. Bertone, A. R. Duffy, I. G. McCarthy, and F. van de Voort. The physics driving the cosmic star formation history. *MNRAS*, 402:1536–1560, March 2010. doi: 10.1111/j.1365-2966.2009.16029.x.

B. D. Oppenheimer, R. Davé, D. Kereš, M. Fardal, N. Katz, J. A. Kollmeier, and D. H. Weinberg. Feedback and recycled wind accretion: assembling the z = 0 galaxy mass function. *MNRAS*, 406: 2325–2338, August 2010. doi: 10.1111/j.1365-2966.2010.16872.x.

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *MNRAS*, 444:1518–1547, October 2014. doi: 10.1093/mnras/stu1536.

N. Khandai, T. Di Matteo, R. Croft, S. M. Wilkins, Y. Feng, E. Tucker, C. DeGraf, and M.-S. Liu. The MassiveBlack-II Simulation: The Evolution of Halos and Galaxies to z˜0. *ArXiv e-prints*, February 2014.

C. Dalla Vecchia and J. Schaye. Simulating galactic outflows with thermal supernova feedback. *MNRAS*, 426:140–158, October 2012. doi: 10.1111/j.1365-2966.2012.21704.x.

J. Schaye, R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly, A. Jenkins, Y. M. Rosas-Guevara, S. D. M. White, M. Baes, C. M. Booth, P. Camps, J. F. Navarro, Y. Qu, A. Rahmati, T. Sawala, P. A. Thomas, and J. Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *MNRAS*, 446: 521–554, January 2015. doi: 10.1093/mnras/stu2058.

G. S. Stinson, C. Brook, A. V. Macciò, J. Wadsley, T. R. Quinn, and H. M. P. Couchman. Making Galaxies In a Cosmological Context: the need for early stellar feedback. *MNRAS*, 428:129–140, January 2013. doi: 10.1093/mnras/sts028.

C. Dalla Vecchia and J. Schaye. Simulating galactic outflows with kinetic supernova feedback. *MNRAS*, 387:1431–1444, July 2008. doi: 10.1111/j.1365-2966.2008.13322.x.

A. J. Richings, J. Schaye, and B. D. Oppenheimer. Non-equilibrium chemistry and cooling in the diffuse interstellar medium - I. Optically thin regime. *MNRAS*, 440:3349–3369, June 2014a. doi: 10.1093/mnras/stu525.

A. J. Richings, J. Schaye, and B. D. Oppenheimer. Non-equilibrium chemistry and cooling in the diffuse interstellar medium - II. Shielded gas. *MNRAS*, 442:2780–2796, August 2014b. doi: 10.1093/mnras/stu1046.

P. F. Hopkins, D. Kereš, N. Murray, E. Quataert, and L. Hernquist. Stellar feedback and bulge formation in clumpy discs. *MNRAS*, 427:968–978, December 2012. doi: 10.1111/j.1365-2966.2012.21981.x.

F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen, and J. Wadsley. Cuspy no more: how outflows affect the central dark matter and baryon distribution in Λ cold dark matter galaxies. *MNRAS*, 422:1231–1240, May 2012. doi: 10.1111/j.1365-2966.2012.20696.x.

S. K. Walch, A. P. Whitworth, T. Bisbas, R. Wünsch, and D. Hubber. Dispersal of molecular clouds by ionizing radiation. *MNRAS*, 427:625–636, November 2012. doi: 10.1111/j.1365-2966.2012.21767.x.

P. Creasey, T. Theuns, and R. G. Bower. How supernova explosions power galactic winds. *MNRAS*, 429:1922–1948, March 2013. doi: 10.1093/mnras/sts439.

M. Aumer, S. D. M. White, T. Naab, and C. Scannapieco. Towards a more realistic population of bright spiral galaxies in cosmological simulations. *MNRAS*, 434:3142–3164, October 2013. doi: 10.1093/mnras/stt1230.

R. A. Crain, J. Schaye, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly, A. Jenkins, Y. M. Rosas-Guevara, S. D. M. White, and J. W. Trayford. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. *ArXiv e-prints*, January 2015.

S. Veilleux, G. Cecil, and J. Bland-Hawthorn. Galactic Winds. *ARA&A*, 43:769–826, September 2005. doi: 10.1146/annurev.astro.43.072103.150610.

C. M. Harrison, D. M. Alexander, J. R. Mullaney, and A. M. Swinbank. Kiloparsec-scale outflows are prevalent among luminous AGN: outflows and feedback in the context of the overall AGN population. *MNRAS*, 441:3306–3347, July 2014. doi: 10.1093/mnras/stu515.

C. Cicone, R. Maiolino, E. Sturm, J. Graciá-Carpio, C. Feruglio, R. Neri, S. Aalto, R. Davies, F. Fiore, J. Fischer, S. García-Burillo, E. González-Alfonso, S. Hailey-Dunsheath, E. Piconcelli, and S. Veilleux. Massive molecular outflows and evidence for AGN feedback from CO observations. A&A, 562:A21, February 2014. doi: 10.1051/0004-6361/201322464.

C. Li and S. D. M. White. The distribution of stellar mass in the low-redshift Universe. *MNRAS*, 398:2177–2187, October 2009. doi: 10.1111/j.1365-2966.2009.15268.x.

I. K. Baldry, S. P. Driver, J. Loveday, E. N. Taylor, L. S. Kelvin, J. Liske, P. Norberg, A. S. G. Robotham, S. Brough, A. M. Hopkins, S. P. Bamford, J. A. Peacock, J. Bland-Hawthorn, C. J. Conselice, S. M. Croom, D. H. Jones, H. R. Parkinson, C. C. Popescu, M. Prescott, R. G. Sharp, and R. J. Tuffs. Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at z ¡ 0.06. *MNRAS*, 421: 621–634, March 2012. doi: 10.1111/j.1365-2966.2012.20340.x.

V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476, August 2000. doi: 10.1086/309179.

A. Jenkins. A new way of setting the phases for cosmological multiscale Gaussian initial conditions. *MNRAS*, 434:2094–2120, September 2013. doi: 10.1093/mnras/stt1154.

A. Jenkins. Second-order Lagrangian perturbation theory initial conditions for resimulations. *MNRAS*, 403:1859–1872, April 2010. doi: 10.1111/j.1365-2966.2010.16259.x.

P. J. E. Peebles. *The large-scale structure of the universe*. 1980.

P. Bode, J. P. Ostriker, and G. Xu. The Tree Particle-Mesh N-Body Gravity Solver. *Astrophys. J. Suppl.*, 128:561–569, June 2000. doi: 10.1086/313398.

J. S. Bagla. TreePM: A Code for Cosmological N-Body Simulations. *Journal of Astrophysics and Astronomy*, 23:185–196, December 2002. doi: 10.1007/BF02702282.

A. A. Klypin and S. F. Shandarin. Three-dimensional numerical model of the formation of large-scale structure in the Universe. *MNRAS*, 204:891–907, September 1983.

S. D. M. White, C. S. Frenk, and M. Davis. Clustering in a neutrino-dominated universe. *ApJ*, 274: L1–L5, November 1983. doi: 10.1086/184139.

V. Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *MNRAS*, 401:791–851, January 2010a. doi: 10.1111/j.1365-2966.2009.15715.x.

J. J. Monaghan. Smoothed particle hydrodynamics. *ARA&A*, 30:543–574, 1992. doi: 10.1146/annurev.aa.30.090192.002551.

V. Springel. Smoothed Particle Hydrodynamics in Astrophysics. *ARA&A*, 48:391–430, September 2010b. doi: 10.1146/annurev-astro-081309-130914.

L. Hernquist and N. Katz. TREESPH - A unification of SPH with the hierarchical tree method. *Astrophys. J. Suppl.*, 70:419–446, June 1989. doi: 10.1086/191344.

V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. 6:79–117, April 2001a. doi: 10.1016/S1384-1076(01)00042-2.

O. Agertz, B. Moore, J. Stadel, D. Potter, F. Miniati, J. Read, L. Mayer, A. Gawryszczak, A. Kravtsov, Å. Nordlund, F. Pearce, V. Quilis, D. Rudd, V. Springel, J. Stone, E. Tasker, R. Teyssier, J. Wadsley, and R. Walder. Fundamental differences between SPH and grid methods. *MNRAS*, 380:963–978, September 2007. doi: 10.1111/j.1365-2966.2007.12183.x.

C. S. Frenk, S. D. M. White, P. Bode, J. R. Bond, G. L. Bryan, R. Cen, H. M. P. Couchman, A. E. Evrard, N. Gnedin, A. Jenkins, A. M. Khokhlov, A. Klypin, J. F. Navarro, M. L. Norman, J. P.

Ostriker, J. M. Owen, F. R. Pearce, U.-L. Pen, M. Steinmetz, P. A. Thomas, J. V. Villumsen, J. W. Wadsley, M. S. Warren, G. Xu, and G. Yepes. The Santa Barbara Cluster Comparison Project: A Comparison of Cosmological Hydrodynamics Solutions. *ApJ*, 525:554–582, November 1999. doi: 10.1086/307908.

C. Power, J. I. Read, and A. Hobbs. The formation of entropy cores in non-radiative galaxy cluster simulations: smoothed particle hydrodynamics versus adaptive mesh refinement. *MNRAS*, 440: 3243–3256, June 2014. doi: 10.1093/mnras/stu418.

N. L. Mitchell, I. G. McCarthy, R. G. Bower, T. Theuns, and R. A. Crain. On the origin of cores in simulated galaxy clusters. *MNRAS*, 395:180–196, May 2009. doi: 10.1111/j.1365-2966.2009.14550.x.

C. Dalla Vecchia. Anarchy. in prep.

P. F. Hopkins. A general class of Lagrangian smoothed particle hydrodynamics methods and implications for fluid mixing problems. *MNRAS*, 428:2840–2856, February 2013. doi: 10.1093/mnras/sts210.

L. Cullen and W. Dehnen. Inviscid smoothed particle hydrodynamics. *MNRAS*, 408:669–683, October 2010. doi: 10.1111/j.1365-2966.2010.17158.x.

D. J. Price. Modelling discontinuities and Kelvin Helmholtz instabilities in SPH. *Journal of Computational Physics*, 227:10040–10057, December 2008. doi: 10.1016/j.jcp.2008.08.011.

H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree . *Advances Comput. Math.*, 4:389, 1995.

W. Dehnen and H. Aly. Improving convergence in smoothed particle hydrodynamics simulations without pairing instability. *MNRAS*, 425:1068–1082, September 2012. doi: 10.1111/j.1365-2966.2012.21439.x.

F. Durier and C. Dalla Vecchia. Implementation of feedback in smoothed particle hydrodynamics: towards concordance of methods. *MNRAS*, 419:465–478, January 2012. doi: 10.1111/j.1365-2966.2011.19712.x.

R. A. Crain, T. Theuns, C. Dalla Vecchia, V. R. Eke, C. S. Frenk, A. Jenkins, S. T. Kay, J. A. Peacock, F. R. Pearce, J. Schaye, V. Springel, P. A. Thomas, S. D. M. White, and R. P. C. Wiersma. Galaxies-intergalactic medium interaction calculation - I. Galaxy formation as a function of large-scale environment. *MNRAS*, 399:1773–1794, November 2009. doi: 10.1111/j.1365-2966.2009.15402.x.

R. P. C. Wiersma, J. Schaye, and B. D. Smith. The effect of photoionization on the cooling rates of enriched, astrophysical plasmas. *MNRAS*, 393:99–107, February 2009a. doi: 10.1111/j.1365-2966.2008.14191.x.

F. Haardt and P. Madau. Modelling the UV/X-ray cosmic background with CUBA. In D. M. Neumann and J. T. V. Tran, editors, *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, 2001.

G. J. Ferland, K. T. Korista, D. A. Verner, J. W. Ferguson, J. B. Kingdon, and E. M. Verner. CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. PASP, 110:761–778, July 1998. doi: 10.1086/316190.

J. Schaye. Star Formation Thresholds and Galaxy Edges: Why and Where. *ApJ*, 609:667–682, July 2004. doi: 10.1086/421232.

G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. PASP, 115:763–795, July 2003. doi: 10.1086/376392.

R. P. C. Wiersma, J. Schaye, T. Theuns, C. Dalla Vecchia, and L. Tornatore. Chemical enrichment in cosmological, smoothed particle hydrodynamics simulations. *MNRAS*, 399:574–600, October 2009b. doi: 10.1111/j.1365-2966.2009.15331.x.

P. Marigo. Chemical yields from low- and intermediate-mass stars: Model predictions and basic observational constraints. A&A, 370:194–217, April 2001. doi: 10.1051/0004-6361:20000247.

L. Portinari, C. Chiosi, and A. Bressan. Galactic chemical enrichment with new metallicity dependent stellar yields. A&A, 334:505–539, June 1998.

Y. M. Rosas-Guevara, R. G. Bower, J. Schaye, M. Furlong, C. S. Frenk, C. M. Booth, R. Crain, C. Dalla Vecchia, M. Schaller, and T. Theuns. The impact of angular momentum on black hole accretion rates in simulations of galaxy formation. *ArXiv e-prints, arXiv:1312.0598*, December 2013.

C. M. Booth and J. Schaye. Cosmological simulations of the growth of supermassive black holes and feedback from active galactic nuclei: method and tests. *MNRAS*, 398:53–74, September 2009. doi: 10.1111/j.1365-2966.2009.15043.x.

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394, May 1985. doi: 10.1086/163168.

V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at [formmu2]z=0. *MNRAS*, 328:726–750, December 2001b. doi: 10.1046/j.1365-8711.2001.04912.x.

K. Dolag, S. Borgani, G. Murante, and V. Springel. Substructures in hydrodynamical cluster simulations. *MNRAS*, 399:497–514, October 2009. doi: 10.1111/j.1365-2966.2009.15034.x.

Y. Qu and et al. Eagle: Galaxy Histories. in prep.

J. Moustakas, A. L. Coil, J. Aird, M. R. Blanton, R. J. Cool, D. J. Eisenstein, A. J. Mendez, K. C. Wong, G. Zhu, and S. Arnouts. PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from z = 0-1. *ApJ*, 767:50, April 2013. doi: 10.1088/0004-637X/767/1/50.

A. R. Tomczak, R. F. Quadri, K.-V. H. Tran, I. Labbé, C. M. S. Straatman, C. Papovich, K. Glazebrook, R. Allen, G. B. Brammer, G. G. Kacprzak, L. Kawinwanichakij, D. D. Kelson, P. J. McCarthy, N. Mehrtens, A. J. Monson, S. E. Persson, L. R. Spitler, V. Tilvi, and P. van Dokkum. Galaxy Stellar Mass Functions from ZFOURGE/CANDELS: An Excess of Low-mass Galaxies since z = 2 and the Rapid Buildup of Quiescent Galaxies. *ApJ*, 783:85, March 2014. doi: 10.1088/0004-637X/783/2/85.

V. González, I. Labbé, R. J. Bouwens, G. Illingworth, M. Franx, and M. Kriek. Evolution of Galaxy Stellar Mass Functions, Mass Densities, and Mass-to-light Ratios from z ˜ 7 to z ˜ 4. *ApJ*, 735:L34, July 2011. doi: 10.1088/2041-8205/735/2/L34.

K. Duncan, C. J. Conselice, A. Mortlock, W. G. Hartley, Y. Guo, H. C. Ferguson, R. Davé, Y. Lu, J. Ownsworth, M. L. N. Ashby, A. Dekel, M. Dickinson, S. Faber, M. Giavalisco, N. Grogin, D. Kocevski, A. Koekemoer, R. S. Somerville, and C. E. White. The mass evolution of the first galaxies: stellar mass functions and star formation rates at 4 ¡ z ¡ 7 in the CANDELS GOODS-South field. *MNRAS*, 444:2960–2984, November 2014. doi: 10.1093/mnras/stu1622.

D. G. Gilbank, I. K. Baldry, M. L. Balogh, K. Glazebrook, and R. G. Bower. The local star formation rate density: assessing calibrations using [OII], H and UV luminosities. *MNRAS*, 405:2594–2614, July 2010a. doi: 10.1111/j.1365-2966.2010.16640.x.

R. Smit, R. J. Bouwens, I. Labbé, W. Zheng, L. Bradley, M. Donahue, D. Lemze, J. Moustakas, K. Umetsu, A. Zitrin, D. Coe, M. Postman, V. Gonzalez, M. Bartelmann, N. Benítez, T. Broadhurst, H. Ford, C. Grillo, L. Infante, Y. Jimenez-Teja, S. Jouvel, D. D. Kelson, O. Lahav, D. Maoz, E. Medezinski, P. Melchior, M. Meneghetti, J. Merten, A. Molino, L. A. Moustakas, M. Nonino, P. Rosati, and S. Seitz. Evidence for Ubiquitous High-equivalent-width Nebular Emission in z ˜ 7 Galaxies: Toward a Clean Measurement of the Specific Star-formation Rate Using a Sample of Bright, Magnified Galaxies. *ApJ*, 784:58, March 2014. doi: 10.1088/0004-637X/784/1/58.

A. Muzzin, D. Marchesini, P. G. van Dokkum, I. Labbé, M. Kriek, and M. Franx. A Near-Infrared Spectroscopic Survey of K-Selected Galaxies at z˜ 2.3: Comparison of Stellar Population Synthesis Codes and Constraints from the Rest-Frame NIR. *ApJ*, 701:1839–1864, August 2009. doi: 10.1088/0004-637X/701/2/1839.

C. Conroy, J. E. Gunn, and M. White. The Propagation of Uncertainties in Stellar Population Synthesis Modeling. I. The Relevance of Uncertain Aspects of Stellar Evolution and the Initial Mass Function to the Derived Physical Properties of Galaxies. *ApJ*, 699:486–506, July 2009. doi: 10.1088/0004-637X/699/1/486.

P. Schechter. An analytic expression for the luminosity function for galaxies. *ApJ*, 203:297–306, January 1976. doi: 10.1086/154079.

P. D. Mitchell, C. G. Lacey, C. M. Baugh, and S. Cole. How well can we really estimate the stellar masses of galaxies from broad-band photometry? *MNRAS*, 435:87–114, October 2013. doi: 10.1093/mnras/stt1280.

E. N. Taylor, A. M. Hopkins, I. K. Baldry, M. J. I. Brown, S. P. Driver, L. S. Kelvin, D. T. Hill, A. S. G. Robotham, J. Bland-Hawthorn, D. H. Jones, R. G. Sharp, D. Thomas, J. Liske, J. Loveday, P. Norberg, J. A. Peacock, S. P. Bamford, S. Brough, M. Colless, E. Cameron, C. J. Conselice, S. M. Croom, C. S. Frenk, M. Gunawardhana, K. Kuijken, R. C. Nichol, H. R. Parkinson, S. Phillipps, K. A. Pimbblet, C. C. Popescu, M. Prescott, W. J. Sutherland, R. J. Tuffs, E. van Kampen, and D. Wijesinghe. Galaxy And Mass Assembly (GAMA): stellar mass estimates. *MNRAS*, 418:1587–1620, December 2011. doi: 10.1111/j.1365-2966.2011.19536.x.

R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Franx, I. Labbé, M. Trenti, P. van Dokkum, C. M. Carollo, V. González, R. Smit, and D. Magee. UV-continuum Slopes at z ˜ 4-7 from the HUDF09+ERS+CANDELS Observations: Discovery of a Well-defined UV Color-Magnitude Relationship for z >= 4 Star-forming Galaxies. *ApJ*, 754:83, August 2012. doi: 10.1088/0004-637X/754/2/83.

D. G. Gilbank, M. L. Balogh, K. Glazebrook, R. G. Bower, I. K. Baldry, G. T. Davies, G. K. T. Hau, I. H. Li, and P. McCarthy. The Redshift One LDSS-3 Emission line Survey (ROLES): survey method and z ˜ 1 mass-dependent star formation rate density. *MNRAS*, 405:2419–2438, July 2010b. doi: 10.1111/j.1365-2966.2010.16607.x.

A. Karim, E. Schinnerer, A. Martínez-Sansigre, M. T. Sargent, A. van der Wel, H.-W. Rix, O. Ilbert, V. Smolčić, C. Carilli, M. Pannella, A. M. Koekemoer, E. F. Bell, and M. Salvato. The Star Formation History of Mass-selected Galaxies in the COSMOS Field. *ApJ*, 730:61, April 2011. doi: 10.1088/0004-637X/730/2/61.

G. Rodighiero, A. Cimatti, C. Gruppioni, P. Popesso, P. Andreani, B. Altieri, H. Aussel, S. Berta, A. Bongiovanni, D. Brisbin, A. Cava, J. Cepa, E. Daddi, H. Dominguez-Sanchez, D. Elbaz, A. Fontana, N. Förster Schreiber, A. Franceschini, R. Genzel, A. Grazian, D. Lutz, G. Magdis, M. Magliocchetti, B. Magnelli, R. Maiolino, C. Mancini, R. Nordon, A. M. Perez Garcia, A. Poglitsch, P. Santini, M. Sanchez-Portal, F. Pozzi, L. Riguccini, A. Saintonge, L. Shao, E. Sturm, L. Tacconi, I. Valtchanov, M. Wetzstein, and E. Wieprecht. The first Herschel view of the mass-SFR link in high-z galaxies. A&A, 518:L25, July 2010. doi: 10.1051/0004-6361/201014624.

O. Cucciati, L. Tresse, O. Ilbert, O. Le Fèvre, B. Garilli, V. Le Brun, P. Cassata, P. Franzetti, D. Maccagni, M. Scodeggio, E. Zucca, G. Zamorani, S. Bardelli, M. Bolzonella, R. M. Bielby, H. J. McCracken, A. Zanichelli, and D. Vergani. The star formation rate density and dust attenuation evolution over 12 Gyr with the VVDS surveys. A&A, 539:A31, March 2012. doi: 10.1051/0004-6361/201118010.

B. E. Robertson, S. R. Furlanetto, E. Schneider, S. Charlot, R. S. Ellis, D. P. Stark, R. J. McLure, J. S. Dunlop, A. Koekemoer, M. A. Schenker, M. Ouchi, Y. Ono, E. Curtis-Lake, A. B. Rogers, R. A. A. Bowler, and M. Cirasuolo. New Constraints on Cosmic Reionization from the 2012 Hubble Ultra Deep Field Campaign. *ApJ*, 768:71, May 2013. doi: 10.1088/0004-637X/768/1/71.

D. Burgarella, V. Buat, C. Gruppioni, O. Cucciati, S. Heinis, S. Berta, M. Béthermin, J. Bock, A. Cooray, J. S. Dunlop, D. Farrah, A. Franceschini, E. Le Floc'h, D. Lutz, B. Magnelli, R. Nordon, S. J. Oliver, M. J. Page, P. Popesso, F. Pozzi, L. Riguccini, M. Vaccari, and M. Viero. Herschel PEP/HerMES: the redshift evolution ($0 \leq z \leq 4$) of dust attenuation and of the total (UV+IR) star formation rate density. A&A, 554:A70, June 2013. doi: 10.1051/0004-6361/201321651.

D. G. Gilbank, R. G. Bower, K. Glazebrook, M. L. Balogh, I. K. Baldry, G. T. Davies, G. K. T. Hau, I. H. Li, P. McCarthy, and M. Sawicki. A spectroscopic measurement of galaxy formation time-scales with the Redshift One LDSS3 Emission line Survey. *MNRAS*, 414:304–320, June 2011. doi: 10.1111/j.1365-2966.2011.18391.x.

A. E. Bauer, A. M. Hopkins, M. Gunawardhana, E. N. Taylor, I. Baldry, S. P. Bamford, J. Bland-Hawthorn, S. Brough, M. J. I. Brown, M. E. Cluver, M. Colless, C. J. Conselice, S. Croom, S. Driver, C. Foster, D. H. Jones, M. A. Lara-Lopez, J. Liske, Á. R. López-Sánchez, J. Loveday, P. Norberg, M. S. Owers, K. Pimbblet, A. Robotham, A. E. Sansom, and R. Sharp. Galaxy And Mass Assembly (GAMA): linking star formation histories and stellar mass growth. *MNRAS*, 434:209–221, September 2013. doi: 10.1093/mnras/stt1011.

K. G. Noeske, B. J. Weiner, S. M. Faber, C. Papovich, D. C. Koo, R. S. Somerville, K. Bundy, C. J. Conselice, J. A. Newman, D. Schiminovich, E. Le Floc'h, A. L. Coil, G. H. Rieke, J. M. Lotz, J. R.

Primack, P. Barmby, M. C. Cooper, M. Davis, R. S. Ellis, G. G. Fazio, P. Guhathakurta, J. Huang, S. A. Kassin, D. C. Martin, A. C. Phillips, R. M. Rich, T. A. Small, C. N. A. Willmer, and G. Wilson. Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies. *ApJ*, 660:L43–L46, May 2007. doi: 10.1086/517926.

V. González, R. J. Bouwens, I. Labbé, G. Illingworth, P. Oesch, M. Franx, and D. Magee. The Rest-frame UV-to-optical Colors and Spectral Energy Distributions of z ˜ 4-7 Galaxies. *ApJ*, 755:148, August 2012. doi: 10.1088/0004-637X/755/2/148.

D. P. Stark, M. A. Schenker, R. Ellis, B. Robertson, R. McLure, and J. Dunlop. Keck Spectroscopy of 3 ¡ z ¡ 7 Faint Lyman Break Galaxies: The Importance of Nebular Emission in Understanding the Specific Star Formation Rate and Stellar Mass Density. *ApJ*, 763:129, February 2013. doi: 10.1088/0004-637X/763/2/129.

D. Utomo, M. Kriek, I. Labbé, C. Conroy, and M. Fumagalli. Simultaneous Modeling of the Stellar and Dust Emission in Distant Galaxies: Implications for Star Formation Rate Measurements. *ApJ*, 783:L30, March 2014. doi: 10.1088/2041-8205/783/2/L30.

P. G. Pérez-González, G. H. Rieke, V. Villar, G. Barro, M. Blaylock, E. Egami, J. Gallego, A. Gil de Paz, S. Pascual, J. Zamorano, and J. L. Donley. The Stellar Mass Assembly of Galaxies from z = 0 to z = 4: Analysis of a Sample Selected in the Rest-Frame Near-Infrared with Spitzer. *ApJ*, 675: 234–261, March 2008. doi: 10.1086/523690.

S. M. Weinmann, A. Pasquali, B. D. Oppenheimer, K. Finlator, J. T. Mendel, R. A. Crain, and A. V. Macciò. A fundamental problem in our understanding of low-mass galaxy evolution. *MNRAS*, 426:2797–2812, November 2012. doi: 10.1111/j.1365-2966.2012.21931.x.

P. D. Mitchell, C. G. Lacey, S. Cole, and C. M. Baugh. The evolution of the star-forming sequence in hierarchical galaxy formation models. *MNRAS*, 444:2637–2664, November 2014. doi: 10.1093/mnras/stu1639.

C. E. White, R. S. Somerville, and H. C. Ferguson. Empirical solutions to the high-redshift overproduction of stars in modeled dwarf galaxies. *ArXiv e-prints*, *arxiv:1407.1850*, July 2014.

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

S. M. Rao, D. A. Turnshek, and D. B. Nestor. Damped Lyα Systems at z ¡ 1.65: The Expanded Sloan Digital Sky Survey Hubble Space Telescope Sample. *ApJ*, 636:610–630, January 2006. doi: 10.1086/498132.

J. X. Prochaska and A. M. Wolfe. On the (Non)Evolution of H I Gas in Galaxies Over Cosmic Time. *ApJ*, 696:1543–1547, May 2009. doi: 10.1088/0004-637X/696/2/1543.

J. Delhaize, M. J. Meyer, L. Staveley-Smith, and B. J. Boyle. Detection of H I in distant galaxies using spectral stacking. *MNRAS*, 433:1398–1410, August 2013. doi: 10.1093/mnras/stt810.

J. Rhee, M. A. Zwaan, F. H. Briggs, J. N. Chengalur, P. Lah, T. Oosterloo, and T. van der Hulst. Neutral atomic hydrogen (H I) gas evolution in field galaxies at z 0.1 and 0.2. *MNRAS*, 435: 2693–2706, November 2013. doi: 10.1093/mnras/stt1481.

T. Zafar, A. Popping, and C. Péroux. The ESO UVES advanced data products quasar sample. I. Dataset and new N_{HI} measurements of damped absorbers. A&A, 556:A140, August 2013. doi: 10.1051/0004-6361/201321153.

A. Rahmati and et al. The distribution of neutral hydrogen around high-redshift galaxies and quasars in the EAGLE simulations. in prep.

A. Rahmati, J. Schaye, A. H. Pawlik, and M. Raičević. The impact of local stellar radiation on the H I column density distribution. *MNRAS*, 431:2261–2277, May 2013. doi: 10.1093/mnras/stt324.

T. Theuns and S. J. Warren. Intergalactic stars in the Fornax cluster. *MNRAS*, 284:L11–L14, January 1997.

V. Presotto, M. Girardi, M. Nonino, A. Mercurio, C. Grillo, P. Rosati, A. Biviano, M. Annunziatella, I. Balestra, W. Cui, B. Sartoris, D. Lemze, B. Ascaso, J. Moustakas, H. Ford, A. Fritz, O. Czoske, S. Ettori, U. Kuchner, M. Lombardi, C. Maier, E. Medezinski, A. Molino, M. Scodeggio, V. Strazzullo, P. Tozzi, B. Ziegler, M. Bartelmann, N. Benitez, L. Bradley, M. Brescia, T. Broadhurst, D. Coe, M. Donahue, R. Gobat, G. Graves, D. Kelson, A. Koekemoer, P. Melchior, M. Meneghetti, J. Merten, L. A. Moustakas, E. Munari, M. Postman, E. Regős, S. Seitz, K. Umetsu, W. Zheng, and A. Zitrin. Intracluster light properties in the CLASH-VLT cluster MACS J1206.2-0847. A&A, 565: A126, May 2014. doi: 10.1051/0004-6361/201323251.

D. R. Patton, P. Torrey, S. L. Ellison, J. T. Mendel, and J. M. Scudder. Galaxy pairs in the Sloan Digital Sky Survey - VI. The orbital extent of enhanced star formation in interacting galaxies. *MNRAS*, 433:L59–L63, June 2013. doi: 10.1093/mnrasl/slt058.

A. R. Wetzel, J. L. Tinker, and C. Conroy. Galaxy evolution in groups and clusters: star formation rates, red sequence fractions and the persistent bimodality. *MNRAS*, 424:232–243, July 2012. doi: 10.1111/j.1365-2966.2012.21188.x.

A. S. Font, R. G. Bower, I. G. McCarthy, A. J. Benson, C. S. Frenk, J. C. Helly, C. G. Lacey, C. M. Baugh, and S. Cole. The colours of satellite galaxies in groups and clusters. *MNRAS*, 389:1619–1629, October 2008. doi: 10.1111/j.1365-2966.2008.13698.x.

M. Hirschmann, G. De Lucia, D. Wilman, S. Weinmann, A. Iovino, O. Cucciati, S. Zibetti, and Á. Villalobos. The influence of the environmental history on quenching star formation in a Λ cold dark matter universe. *MNRAS*, 444:2938–2959, November 2014. doi: 10.1093/mnras/stu1609.

Q. Guo and et al. Eagle vs SAMs. in prep.

R. G. Bower and et al. Eagle: What triggers black hole feedback? in prep.

R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey. Breaking the hierarchy of galaxy formation. *MNRAS*, 370:645–655, August 2006. doi: 10.1111/j.1365-2966.2006.10519.x.

A. Gallazzi, S. Charlot, J. Brinchmann, S. D. M. White, and C. A. Tremonti. The ages and metallicities of galaxies in the local universe. *MNRAS*, 362:41–58, September 2005. doi: 10.1111/j.1365-2966.2005.09321.x.

G. Kauffmann. Quantitative constraints on starburst cycles in galaxies with stellar masses in the range 10^8-10^{10} $M_\odot$. *MNRAS*, 441:2717–2724, July 2014. doi: 10.1093/mnras/stu752.

G. Kauffmann, C. Li, W. Zhang, and S. Weinmann. A re-examination of galactic conformity and a comparison with semi-analytic models of galaxy formation. *MNRAS*, 430:1447–1456, April 2013. doi: 10.1093/mnras/stt007.

D. J. Mortlock, S. J. Warren, B. P. Venemans, M. Patel, P. C. Hewett, R. G. McMahon, C. Simpson, T. Theuns, E. A. Gonzáles-Solares, A. Adamson, S. Dye, N. C. Hambly, P. Hirst, M. J. Irwin, E. Kuiper, A. Lawrence, and H. J. A. Röttgering. A luminous quasar at a redshift of z = 7.085. *Nat*, 474:616–619, June 2011. doi: 10.1038/nature10159.

P. R. Shapiro. Cosmological H II regions and the photoionization of the intergalactic medium. PASP, 98:1014–1017, October 1986. doi: 10.1086/131865.

R. L. Grissom, D. R. Ballantyne, and J. H. Wise. On the contribution of active galactic nuclei to reionization. A&A, 561:A90, January 2014. doi: 10.1051/0004-6361/201322637.

D. Schaerer. On the properties of massive Population III stars and metal-free stellar populations. A&A, 382:28–42, January 2002. doi: 10.1051/0004-6361:20011619.

J. S. Bolton and M. G. Haehnelt. The observed ionization rate of the intergalactic medium and the ionizing emissivity at z >= 5: evidence for a photon-starved and extended epoch of reionization. *MNRAS*, 382:325–341, November 2007. doi: 10.1111/j.1365-2966.2007.12372.x.

J. H. Wise, V. G. Demchenko, M. T. Halicek, M. L. Norman, M. J. Turk, T. Abel, and B. D. Smith. The birth of a galaxy - III. Propelling reionization with the faintest galaxies. *MNRAS*, 442:2560–2579, August 2014. doi: 10.1093/mnras/stu979.

G. C. So, M. L. Norman, D. R. Reynolds, and J. H. Wise. Fully Coupled Simulation of Cosmic Reionization. II. Recombinations, Clumping Factors, and the Photon Budget for Reionization. *ApJ*, 789:149, July 2014. doi: 10.1088/0004-637X/789/2/149.

M. Raičević, T. Theuns, and C. Lacey. The galaxies that reionized the Universe. *MNRAS*, 410: 775–787, January 2011.

N. Murray and M. Rahman. Star Formation in Massive Clusters Via the Wilkinson Microwave Anisotropy Probe and the Spitzer Glimpse Survey. *ApJ*, 709:424–435, January 2010. doi: 10.1088/0004-637X/709/1/424.

M. A. Alvarez, K. Finlator, and M. Trenti. Constraints on the Ionizing Efficiency of the First Galaxies. *ApJ*, 759:L38, November 2012. doi: 10.1088/2041-8205/759/2/L38.

A. H. Pawlik, J. Schaye, and E. van Scherpenzeel. Keeping the Universe ionized: photoheating and the clumping factor of the high-redshift intergalactic medium. *MNRAS*, 394:1812–1824, April 2009. doi: 10.1111/j.1365-2966.2009.14486.x.

I. T. Iliev, G. Mellema, U.-L. Pen, H. Merz, P. R. Shapiro, and M. A. Alvarez. Simulating cosmic reionization at large scales - I. The geometry of reionization. *MNRAS*, 369:1625–1638, July 2006. doi: 10.1111/j.1365-2966.2006.10502.x.

M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, and M. Zaldarriaga. The morphology of HII regions during reionization. *MNRAS*, 377:1043–1063, May 2007. doi: 10.1111/j.1365-2966.2007.11489.x.

H. Trac and R. Cen. Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and Initial Results. *ApJ*, 671:1–13, December 2007. doi: 10.1086/522566.

S. C. Chapman, A. W. Blain, I. Smail, and R. J. Ivison. A Redshift Survey of the Submillimeter Galaxy Population. *ApJ*, 622:772–796, April 2005. doi: 10.1086/428082.

C. M. Casey, D. Narayanan, and A. Cooray. Dusty star-forming galaxies at high redshift. *PhysRep*, 541:45–161, August 2014. doi: 10.1016/j.physrep.2014.02.009.

J. M. Simpson, A. M. Swinbank, I. Smail, D. M. Alexander, W. N. Brandt, F. Bertoldi, C. de Breuck, S. C. Chapman, K. E. K. Coppin, E. da Cunha, A. L. R. Danielson, H. Dannerbauer, T. R. Greve, J. A. Hodge, R. J. Ivison, A. Karim, K. K. Knudsen, B. M. Poggianti, E. Schinnerer, A. P. Thomson, F. Walter, J. L. Wardlow, A. Weiß, and P. P. van der Werf. An ALMA Survey of Submillimeter Galaxies in the Extended Chandra Deep Field South: The Redshift Distribution and Evolution of Submillimeter Galaxies. *ApJ*, 788:125, June 2014. doi: 10.1088/0004-637X/788/2/125.

C. M. Baugh, C. G. Lacey, C. S. Frenk, G. L. Granato, L. Silva, A. Bressan, A. J. Benson, and S. Cole. Can the faint submillimetre galaxies be explained in the Λ cold dark matter model? *MNRAS*, 356:1191–1200, January 2005. doi: 10.1111/j.1365-2966.2004.08553.x.

R. Davé, K. Finlator, B. D. Oppenheimer, M. Fardal, N. Katz, D. Kereš, and D. H. Weinberg. The nature of submillimetre galaxies in cosmological hydrodynamic simulations. *MNRAS*, 404:1355–1368, May 2010. doi: 10.1111/j.1365-2966.2010.16395.x.

W. I. Cowley, C. G. Lacey, C. M. Baugh, and S. Cole. Simulated observations of sub-millimetre galaxies: the impact of single-dish resolution and field variance. *MNRAS*, 446:1784–1798, January 2015. doi: 10.1093/mnras/stu2179.

A. M. Swinbank, J. M. Simpson, I. Smail, C. M. Harrison, J. A. Hodge, A. Karim, F. Walter, D. M. Alexander, W. N. Brandt, C. de Breuck, E. da Cunha, S. C. Chapman, K. E. K. Coppin, A. L. R. Danielson, H. Dannerbauer, R. Decarli, T. R. Greve, R. J. Ivison, K. K. Knudsen, C. D. P. Lagos, E. Schinnerer, A. P. Thomson, J. L. Wardlow, A. Weiß, and P. van der Werf. An ALMA survey of sub-millimetre Galaxies in the Extended Chandra Deep Field South: the far-infrared properties of SMGs. *MNRAS*, 438:1267–1287, February 2014. doi: 10.1093/mnras/stt2273.

R. C. Hickox, J. L. Wardlow, I. Smail, A. D. Myers, D. M. Alexander, A. M. Swinbank, A. L. R. Danielson, J. P. Stott, S. C. Chapman, K. E. K. Coppin, J. S. Dunlop, E. Gawiser, D. Lutz, P. van der Werf, and A. Weiß. The LABOCA survey of the Extended Chandra Deep Field-South: clustering of submillimetre galaxies. *MNRAS*, 421:284–295, March 2012. doi: 10.1111/j.1365-2966.2011.20303.x.

A. Weiß, C. De Breuck, D. P. Marrone, J. D. Vieira, J. E. Aguirre, K. A. Aird, M. Aravena, M. L. N. Ashby, M. Bayliss, B. A. Benson, M. Béthermin, A. D. Biggs, L. E. Bleem, J. J. Bock, M. Bothwell, C. M. Bradford, M. Brodwin, J. E. Carlstrom, C. L. Chang, S. C. Chapman, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, T. P. Downes, C. D. Fassnacht, E. M. George, M. D. Gladders, A. H. Gonzalez, T. R. Greve, N. W. Halverson, Y. D. Hezaveh, F. W. High, G. P. Holder, W. L. Holzapfel, S. Hoover, J. D. Hrubes, K. Husband, R. Keisler, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, M. Malkan, V. McIntyre, J. J. McMahon, J. Mehl, K. M. Menten, S. S. Meyer, E. J. Murphy, S. Padin, T. Plagge, C. L. Reichardt, A. Rest, M. Rosenman, J. Ruel, J. E. Ruhl, K. K.

Schaffer, E. Shirokoff, J. S. Spilker, B. Stalder, Z. Staniszewski, A. A. Stark, K. Story, K. Vanderlinde, N. Welikala, and R. Williamson. ALMA Redshifts of Millimeter-selected Galaxies from the SPT Survey: The Redshift Distribution of Dusty Star-forming Galaxies. *ApJ*, 767:88, April 2013. doi: 10.1088/0004-637X/767/1/88.

V. Smolčić, M. Aravena, F. Navarrete, E. Schinnerer, D. A. Riechers, F. Bertoldi, C. Feruglio, A. Finoguenov, M. Salvato, M. Sargent, H. J. McCracken, M. Albrecht, A. Karim, P. Capak, C. L. Carilli, N. Cappelluti, M. Elvis, O. Ilbert, J. Kartaltepe, S. Lilly, D. Sanders, K. Sheth, N. Z. Scoville, and Y. Taniguchi. Millimeter imaging of submillimeter galaxies in the COSMOS field: redshift distribution. A&A, 548:A4, December 2012. doi: 10.1051/0004-6361/201219368.

T. Sawala, C. S. Frenk, A. Fattahi, J. F. Navarro, R. G. Bower, R. A. Crain, C. Dalla Vecchia, M. Furlong, A. Jenkins, I. G. McCarthy, Y. Qu, M. Schaller, J. Schaye, and T. Theuns. Bent by baryons: the low mass galaxy-halo relation. *ArXiv e-prints*, April 2014.

J. Trayford and et al. An Eagle Eye: Luminosities and Colours. in prep.