THE UNIVERSITY *of* EDINBURGH

Mapping galaxy evolution at multiple wavelengths and cosmic epochs

Rachel K. Cochrane

Doctor of Philosophy
The University of Edinburgh
September 2019

Abstract

Some of the most fundamental measurements we can make of the Universe are where and when stars formed in galaxies. In recent years, astronomers have converged on a picture in which the star formation rate density of the Universe peaks at approximately redshift (z) 2, when the Universe is around a quarter of its present age. There, star-forming galaxies harbour large reservoirs of molecular gas, assemble stellar mass rapidly, and typically display disordered morphologies.

In this thesis, I study the evolution of galaxies on large and small scales, with a particular focus on the epoch around the peak of cosmic star formation. My overarching aim is to understand the physical processes that drive and quench star formation in galaxies over cosmic time. In the first half of this thesis, I focus on global measurements of star formation, using the High-z Emission line survey (HiZELS), a deep, near-infrared narrow-band survey, which identifies star-forming galaxies at $z = 0.8 - 2.2$. I characterise the dark matter halo environments of these galaxies via a clustering analysis, along with a Halo Occupation Distribution model fitting procedure, then study the relationships between host dark matter halo environment and galaxy properties.

I show that the clustering strength and the host dark matter halo masses of the HiZELS galaxies increase linearly with Hα luminosity (and, by implication, star formation rate) at all three redshifts. The typical galaxies in our samples are star-forming centrals, residing in dark matter haloes of mass $\sim 10^{12}\,M_{\odot}$. I find a remarkably tight redshift-independent relation between the Hα luminosity scaled by the characteristic luminosity, $L_{\mathrm{H}\alpha}/L^{*}_{\mathrm{H}\alpha}(z)$, and the host dark matter halo mass of central galaxies. Simple analytic modelling suggests that this is consistent with a model in which the dark matter halo environment is a strong driver of galaxy star formation rate and therefore of the evolution of the star formation rate density in the Universe. I investigate this further by distinguishing the stellar mass and star formation rate dependencies of the clustering of HiZELS galaxies. I compare my observational results to the predictions of a pioneering cosmological hydrodynamical simulation, the Virgo Consortium's Evolution and Assembly of

GaLaxies and their Environments project, known as EAGLE.

In the subsequent chapters of this thesis, I focus more heavily on simulations of galaxy formation, which are important tools for constraining and understanding the physics at play in galaxies. I use EAGLE to investigate the quenching of star formation in simulated galaxies via a novel application of Principal Component Analysis. I show that the key relations between halo mass, stellar mass and star formation rate are in good agreement with observed low-redshift galaxies.

Having studied the global properties of star-forming galaxies, I then turn to smaller scales, investigating what we can learn from spatially-resolved imaging. I present a detailed study of the spatially-resolved dust continuum emission of realistic simulated high-redshift galaxies. These galaxies, drawn from the FIRE-2 simulations, reach Milky Way masses by $z \sim 2$. Post-processing them using radiative transfer techniques, I obtain predictions for the full rest-frame far-ultraviolet to far-infrared Spectral Energy Distributions of these simulated galaxies, as well as maps of their emission across the wavelength spectrum. As has been observed in distant galaxies, the rest-frame far-infrared emission of the simulated galaxies is compact, spanning half-light radii of $\sim 0.5 - 4\,\mathrm{kpc}$. The derived morphologies of simulated galaxies are notably different in different wavebands; a galaxy can appear clumpy and extended in the far-ultraviolet yet compact at far-infrared wavelengths.

Finally, I perform a multi-wavelength study of a single observed galaxy, SHiZELS-14 ($z = 2.24$), drawn from the HiZELS survey and subsequently imaged at $0.15''$ resolution at multiple wavelengths. The data comprise kpc-resolution imaging in three different widely used tracers of star formation: the Hα emission line (from SINFONI/VLT), rest-frame far-ultraviolet continuum (from HST F606W imaging), and the rest-frame far-infrared (from ALMA), as well as the rest-frame optical (from HST F140W imaging). SHiZELS-14 displays a compact, dusty centre, as well as extended emission in both Hα and the rest-frame FIR. The ultraviolet emission is spatially offset from the extended dust emission, and appears to trace holes in the dust distribution. I find that the dust attenuation varies across the spatial extent of the galaxy, reaching up to ~ 5 magnitudes of extinction at Hα wavelengths in the most dusty regions. Global star formation rates inferred using standard calibrations to the different tracers vary from $\sim 10 - 1000\,M_{\odot}/\mathrm{yr}$, and are particularly discrepant in the galaxy's dusty centre. This galaxy highlights the biased view of galaxy evolution provided by short-wavelength data in the absence of long-wavelength data, and is in line with my simulations.

Lay Summary

The evolution of galaxies in our Universe is one of the most extensively-studied topics in modern astrophysics. This field encapsulates the birth of the first galaxies just a few hundred years after the Big Bang, and their growth and merging into larger structures, all the way to the present day.

Observational astronomers learn by interpreting light from galaxies they observe on the sky. The light that galaxies emit in different wavebands provides clues to the physical processes taking place within them. By combining data with models, we can infer properties such as the mass of stars, gas, dust and heavy elements within a galaxy, and the rate at which it is forming new stars. Large observational samples enable us to build a picture of galaxy evolution, and put constraints on the physics that goes into theoretical models.

In this thesis, I study star formation within very distant galaxies. The bulk of this work focuses on galaxies that were forming stars when the Universe was $\sim 4-7$ billion years old, around a third to half of its current age. Astronomers believe that around this time, galaxies formed stars most vigorously, consuming gas and assembling their stellar mass at rapid rates. To perform this research, I use measurements of thousands of galaxies, gathered from telescopes around the world and in space. In the first chapters of this thesis, I characterise the environments in which these galaxies reside, using their spatial distribution on the sky. I show that a galaxy's environment correlates with its star formation rate; more highly star-forming galaxies tend to reside in higher density regions of the Universe. I also make comparisons to a new cosmological simulation, EAGLE, to try to understand the physics behind correlations such as this.

In the second half of this thesis, I turn to look at the detailed properties of individual galaxies on sub-galactic scales. I zoom-in on one galaxy, one of the most distant in my sample, and observe it using different telescopes, probing optical and longer-wavelength emission from ground-based telescopes in Chile, and ultraviolet emission from the Hubble Space Telescope. Each of these telescopes provides data that is used to infer star formation rates of galaxies, though they trace

slightly different physical processes. I show that star formation rates estimated from the different telescopes are actually very different. The shape of the galaxy also varies significantly depending on which of the images we look at. I show that these differences are likely caused by the galaxy's large dust content: dust blocks light at some wavelengths from escaping the galaxy and travelling to our telescopes. I also use simulations to study the spatially resolved emission from a small number of simulated galaxies. The sizes and morphologies of these galaxies are different in different wavebands, just like my observed galaxy. This work highlights the importance of studying galaxies with multiple instruments to gain the most insight into the physical processes taking place within them, and to minimise observational bias.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified. In all chapters, I use 'we' to acknowledge that this work was performed under the guidance of my supervisor and with the help of many collaborators. Nevertheless, all work stems from first author papers and this thesis is my own.

Chapters 2, 3 and 4 are based upon work published in Cochrane et al. (2017) and Cochrane et al. (2018). Chapter 2 and Chapter 5 include work published in Cochrane & Best (2018). Chapter 6 is based upon work published in Cochrane et al. (2019). Chapter 7 will be submitted to Monthly Notices of the Royal Astronomical Society in the near future.

(*Rachel K. Cochrane, September 2019*)

Acknowledgements

My most sincere thanks go to my supervisor, Philip Best, for his constant academic and personal support. Philip has been extremely generous with his time, shown staggering patience, and encouraged me to explore new ideas. His gentle encouragement and reassurance has been critical to the completion of this thesis. Thank you so much, Philip!

Many collaborators have contributed to the papers that went into this thesis. I must thank the HiZELS collaboration, in particular David Sobral for his extensive input and prompt email guidance. My time at the Kavli summer program was one of the most exciting periods of my PhD. I thank my mentors, Chris Hayward, Daniel Anglés-Alcázar and Jennifer Lotz, for their enthusiasm and generosity of time and ideas.

I was fortunate to take part in formative summer and undergraduate projects. My thanks go to my supervisors, Chris Lintott, Brooke Simmons, Yingjie Peng and Roberto Maiolino, for introducing me to astrophysics and motivating me to start this PhD. Long before that, I had the privilege of an unusual education at The Henrietta Barnett School. I thank all my teachers there, especially Mrs Khan, for her spirited maths lessons and memorable stories.

I feel extremely lucky to have spent these years in Edinburgh, and benefited from the expertise of many at the ROE. I thank Ross McLure, Annette Ferguson, Romeel Davé, Jim Dunlop and José Sabater Montes for their academic guidance, and Ken Rice for his leadership of all the graduate students. This PhD would not have been nearly as enjoyable without friends and colleagues, particularly the bumper 2013, 2014 and 2015 cohorts, with whom I had a lot of fun. I thank Clémence Fontanive for saving this thesis after an unfortunate spillage 10 days before its submission and trusting me with a second laptop.

Lastly, I thank my parents, Gillian O'Brien and David Cochrane, and siblings, Rosie and Robert Cochrane, for their support in this endeavour and all others.

Funding Acknowledgements

I have been fortunate to receive funding from a number of sources over the course of this PhD. I thank the UK Science and Technology Facilities Council and the Isaac Newton Group for their support via studentships. The Institute for Astronomy has funded numerous conference trips, and I have also benefited from the University of Edinburgh's Principal's Go Abroad Scheme. I thank the Kavli Foundation and the Simons Foundation for funding the Kavli Summer Program.

Data Acknowledgements

Chapters 3 and 4 are based on observations obtained using the Wide Field CAMera (WFCAM) on the 3.8-m United Kingdom Infrared Telescope (UKIRT), as part of the High-redshift(Z) Emission Line Survey (HiZELS; U/CMP/3 and U/10B/07). The survey also relies on observations conducted with HAWK-I on the ESO Very Large Telescope (VLT), programme 086.7878.A, and observations obtained with Suprime-Cam on the Subaru Telescope (S10B-144S).

In Chapter 5 I draw upon archived data from the Sloan Digital Sky Survey (SDSS; *www.sdss.org*). Funding for the SDSS and SDSS-II was provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS was managed by the Astrophysical Research Consortium for the Participating Institutions.

I acknowledge the Virgo Consortium for making their simulation data available. The EAGLE simulations used in Chapters 4 and 5 were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyères-le-Châtel. The simulations used in Chapter 6 were run using XSEDE, supported by NSF grant ACI-1053575, and Northwestern University's compute cluster 'Quest'.

Chapter 7 is based on observations with the SINFONI spectrograph, collected at the European Southern Observatory's VLT under programme 084.B-0300. This Chapter is also based on observations made with the NASA/ESA Hubble Space Telescope (Program 14719), and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/-NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA) and the Canadian Astronomy Data Centre (CADC/NRC/CSA). This thesis makes use of the following ALMA data: ADS/JAO.ALMA2015.1.00026.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"The end is in the beginning and yet you go on."

- Samuel Beckett, *Endgame*

The evolution of galaxies in our Universe is one of the most extensively-studied topics in modern astrophysics. This field encapsulates the birth of the first galaxies just a few hundred million years after the Big Bang, and their growth and merging into larger structures, all the way to the present day. Here, I present a brief introduction to the foundations of galaxy formation, as understood within a cold dark matter (CDM)-dominated Universe. Our basic ingredients are dark matter and baryonic matter, which account for 26.8% and 4.9% of the mass-energy budget of the Universe respectively, the rest comprising dark energy (Planck Collaboration, 2016).

The complex history of the Universe is shown in Figure 1.1. Around $380,000$ years after the Big Bang (at $z \sim 1000$), the Universe underwent a process known as 'recombination'. During this time, neutral atoms, mainly hydrogen and helium, formed from cooled plasma at $\sim 3000\,\mathrm{K}$. While photons were previously scattered off free electrons, the Universe became transparent during this time, and the photons released during recombination were free to propagate. We can still trace the recombination process via the cosmic microwave background (CMB; Penzias & Wilson 1965), at its characteristic $2.73\,\mathrm{K}$ temperature. The Universe then entered the so-called 'Dark ages', during which it continued to cool and expand. In the paradigm of hierarchical galaxy formation, weak density fluctuations in the early Universe were amplified by gravitational instabilities that acted on cold dark matter (Blumenthal et al., 1984; Davis et al., 1985). This led to the collapse of dark matter into gravitationally-bound structures known as dark matter haloes.

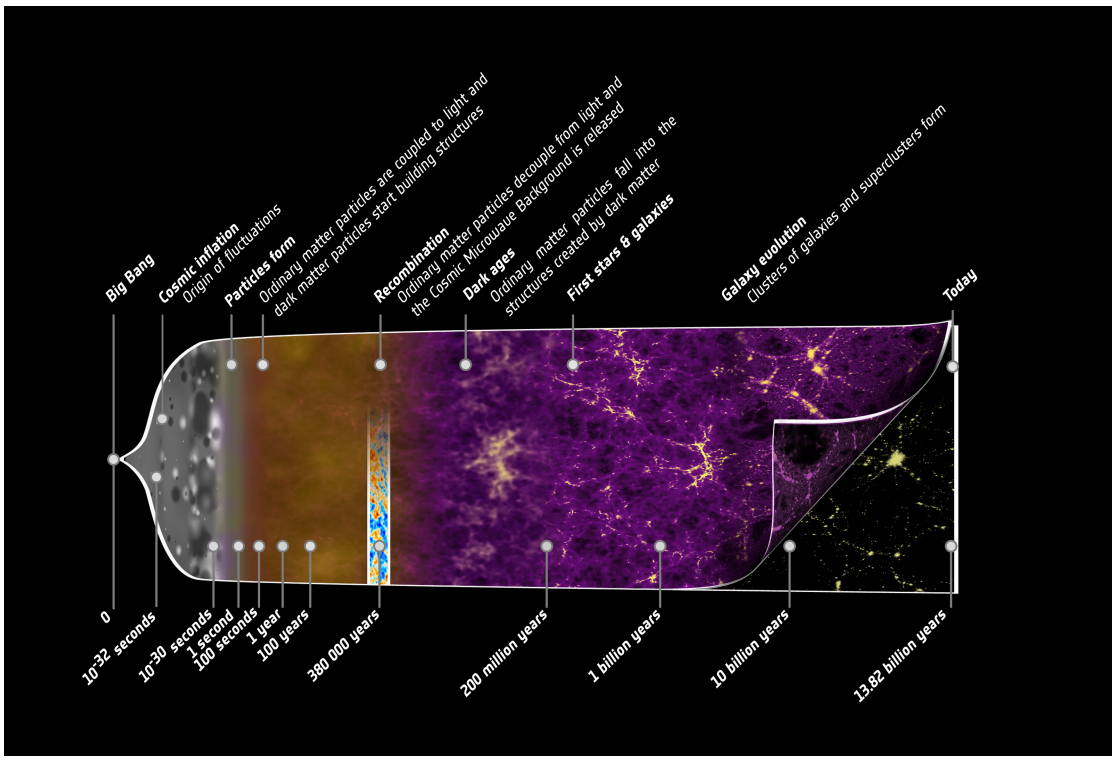

Figure 1.1 *The history of the Universe. The first years of the Universe are radiation-dominated, but by* $47,000$ *years after the Big Bang, the Universe has cooled sufficiently for matter to dominate. At* $380,000$ *years, the Universe becomes transparent upon the formation of neutral atoms, a process known as 'recombination'. Atoms, dominated by hydrogen and helium, form and release photons that form the Cosmic Microwave Background. The Universe then enters the so-called 'Dark ages', in which baryonic structure begins to develop as matter collapses under the gravitational attraction of collapsing dark matter. At around 400 million years, the first stars and galaxies begin to form, tracing the dark matter overdensities. These begin the reionise the Universe. Dark matter haloes and galaxies continue to merge and grow, forming larger structures over the following* ~ 13 *billion years. Credit: ESA - C. Carreau.*

Baryons - in the form of gas - also collapsed under gravity, following the spatial distribution of the collapsing dark matter.

The traditional picture of galaxy formation then proceeds as follows. Baryons falling onto a dark matter halo's potential well were heated by shocks, resulting in a pressure-supported hot gas halo at the virial temperature. Gas cooled radiatively, at high redshifts ($z > 10$) by the inverse scattering of CMB photons by electrons in the hot halo gas. Upon cooling, the pressure of the gas dropped, and it fell towards the centres of the haloes, forming rotating disks (Rees & Ostriker, 1977; White & Rees, 1978; White & Frenk, 1991). Of the order 100 Myr

after the Big Bang, the first, very metal-poor stars formed from this pristine gas. These early stars influenced their environments by a collection of processes we call 'stellar feedback'. Heavy elements synthesised within stars were injected into their surrounding regions via stellar winds and supernovae explosions, increasing the metallicity of the circumgalactic medium (CGM) and the interstellar medium (ISM), and future generations of stars. The highly energetic ultraviolet (UV) photons released by these massive, young stars also ionised pockets of gas. This contributed to the process known as 'reionisation', by which the Universe once again became ionised. Subsequent mergers of smaller dark matter haloes and the galaxies within them gradually led to the assembly of the diverse structures we see in the local Universe (see Figure 1.2).

Over the last few decades, a concerted effort has gone into studying the history of cosmic star formation and understanding the drivers of star formation over time. Wide, deep, spectroscopic surveys began with the Canada-France Redshift Survey (Lilly et al., 1995, 1996), which obtained spectra of 591 galaxies in the redshift range $0.02 < z < 1.2$. Lilly et al. (1996) found an order of magnitude decline in the 2800 Å luminosity function since $z = 1$, suggesting a decline in cosmic star formation. This prompted further Hubble Space Telescope (HST)

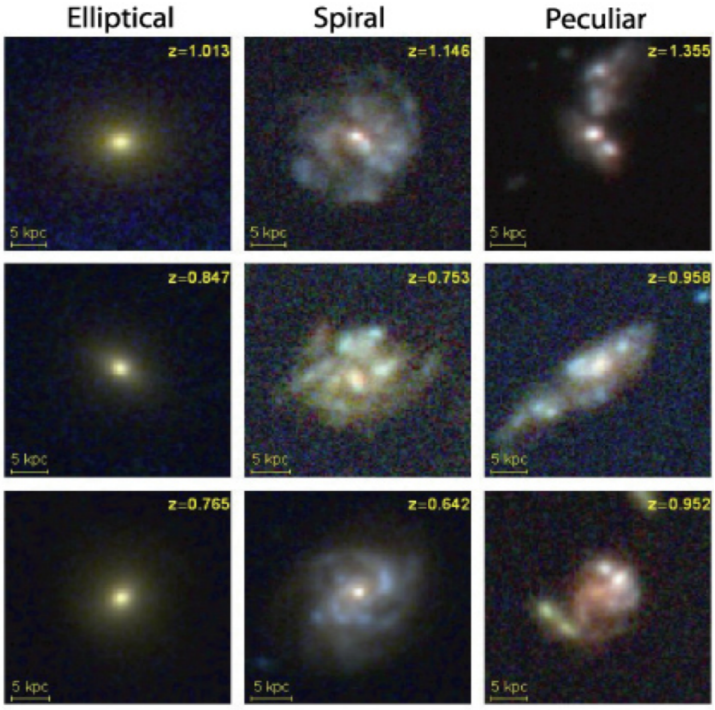

Figure 1.2 *Galaxies display a range of morphologies: shown in this figure are examples of elliptical (left), spiral (centre) and peculiar galaxies (right). Typical galaxy morphologies become more irregular at higher redshift. Figure from Abraham & Van den Bergh (2001).*

optical and ground-based near infrared (NIR) imaging back to $z = 4$ (Madau et al., 1996; Connolly et al., 1997; Pascarelle et al., 1998). Together, these early studies presented a broad view of star formation over cosmic time, with volume-averaged star formation density increasing from high redshift to peak somewhere in the range $1 \lesssim z \lesssim 2$ and then declining towards the present day. By the mid-2000s, the cosmic star formation rate density (SFRD) had been constrained fairly tightly back to $z \sim 1$, using a range of tracers (Hopkins & Beacom, 2006). However, its form at higher redshift and the exact position of the peak remained less well determined.

Recent work has focused on identifying galaxies at higher redshifts ($z > 3$). Surveys tend to select galaxies using either the Lyman break (the characteristic step in UV continuum emission at $\lambda_{\rm rest} = 1216\,\mathring{A}$ due to neutral hydrogen absorption) or the Lyman alpha emission line (see Dunlop, 2013, for a review). The identification of Lyman break galaxies (LBGs) can be performed with broad band filters via the 'dropout method' (e.g. Bouwens et al., 2015), and so is particularly efficient. These methods tend to select galaxies that are both young and fairly dust poor. As will be discussed in detail in Sections 1.1.1 and 1.1.2, dust within galaxies absorbs short-wavelength light and re-emits at infrared (IR) wavelengths. Dust corrections are therefore applied to scale the UV-derived SFRD, bringing it into line with IR-derived values where the two overlap (at $z < 3$) (Madau & Dickinson, 2014).

There is now broad agreement between samples of hundreds of thousands of galaxies back to $z \sim 5$ (see the SFRD compiled by Madau & Dickinson 2014, shown in Figure 1.3, top panel; see also Driver et al. 2018). Madau & Dickinson (2014) also combined surveys to derive the stellar mass density of the Universe as a function of redshift. Encouragingly, they found good agreement between the global stellar mass density and the integral of the past star formation rate history. This is shown in Figure 1.3 (bottom panel).

The physical mechanisms that drive the decline of the volume-averaged star formation density since $z \sim 2$ remain less well-established. It is clear that, at fixed stellar mass, star-forming galaxies at higher redshifts have higher star formation rates (SFRs) than those at lower redshift (see Section 1.2.1). These high SFRs are believed to be driven by large gas supplies; indeed, large molecular gas reservoirs have been observed in galaxies around the peak of cosmic star formation (e.g. Daddi et al., 2010; Tacconi et al., 2010, 2013, 2017; Papovich et al., 2016; Falgarone et al., 2017). Theoretical work suggests that these accumulate due to steady accretion of cold ($\sim 10^4\,$K) gas along filaments of the cosmic web (Kereš

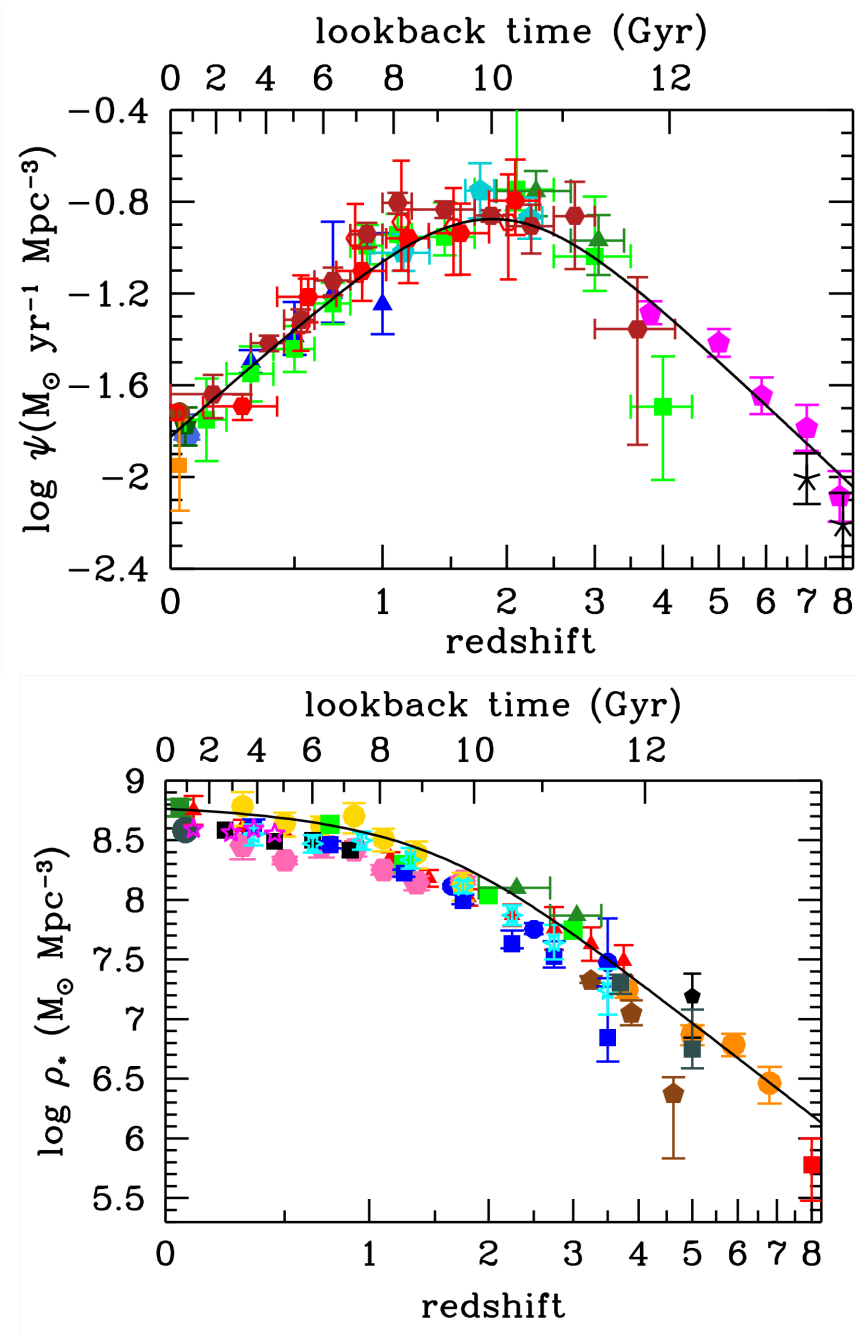

Figure 1.3 *Top: The volume-averaged star formation rate density over cosmic time, derived from a combination of far ultraviolet (corrected for dust attenuation; pink, green, blue data points) and infrared (dust-reprocessed light; red and orange data points) measurements. Star formation rate density increases between $z \sim 8$ and the peak of cosmic star formation around $z = 2$, and then declines towards lower redshift. Bottom: The cosmic stellar mass density history. Observational data (coloured points) agree well with the integrated instantaneous star formation rate density (black line). Taken from Madau & Dickinson (2014).*

et al., 2005; Dekel & Birnboim, 2006; Dekel et al., 2009). Tentative observational work has lended support to this scenario (Cantalupo et al., 2014; Martin et al., 2015, 2016). The streams of cold gas both maintain a galaxy's high gas fraction and set the turbulent conditions that are required for gas to break up into giant clumps. The fraction of galaxies that are star-forming also decreases towards low redshifts (e.g. Peng et al. 2010), with more 'quenched' galaxies contributing substantially to the stellar mass budget but little to the SFR budget at low-z (Renzini & Peng, 2015, see Section 5.4). I return to these issues in more detail in Section 1.2. First, I will provide a more detailed overview of the main methods used to measure star formation in distant galaxies.

1.1 Measuring star formation in galaxies

Understanding the physical processes that drive and quench star formation over cosmic time can only be possible if we are able to measure star formation in galaxies accurately at a range of epochs. There are two main approaches to constructing the star formation histories of galaxies: reconstructing the histories of nearby galaxies via detailed stellar population studies, and measuring the in-situ star formation using samples of galaxies at various redshifts. The former was initially performed using integrated galaxy spectra, with results broadly consistent with in-situ measurements (Panter et al., 2003; Heavens et al., 2004). With the latest integral field unit (IFU) surveys such as MaNGA (Bundy et al., 2015), SAMI (Bryant et al., 2015), and CALIFA (Sánchez et al., 2012; Husemann et al., 2013; García-Benito et al., 2015) providing resolved measurements of local galaxies, this now becoming possible for tens of thousands of galaxies in a spatially-resolved manner (e.g. Li et al., 2015; González Delgado et al., 2017; Goddard et al., 2017; Rowlands et al., 2018).

In this thesis, I adopt the latter approach. I study star-forming galaxies at a range of redshifts, with a particular focus on the period of most vigorous activity, the peak of cosmic star formation. In order to draw statistically significant conclusions, we need large samples of galaxies at a range of redshifts. In general, we must therefore rely on global (rather than spatially resolved) measurements of galaxies, due to the seeing-limited nature of ground-based survey telescopes. Galaxies emit at a range of wavelengths, with emission in different wavebands tracing different physical processes (see Figure 1.4). In this section I will provide an overview of the wide range of tracers of star formation, and review the basic physical processes that motivate their use. A more extensive review of this topic

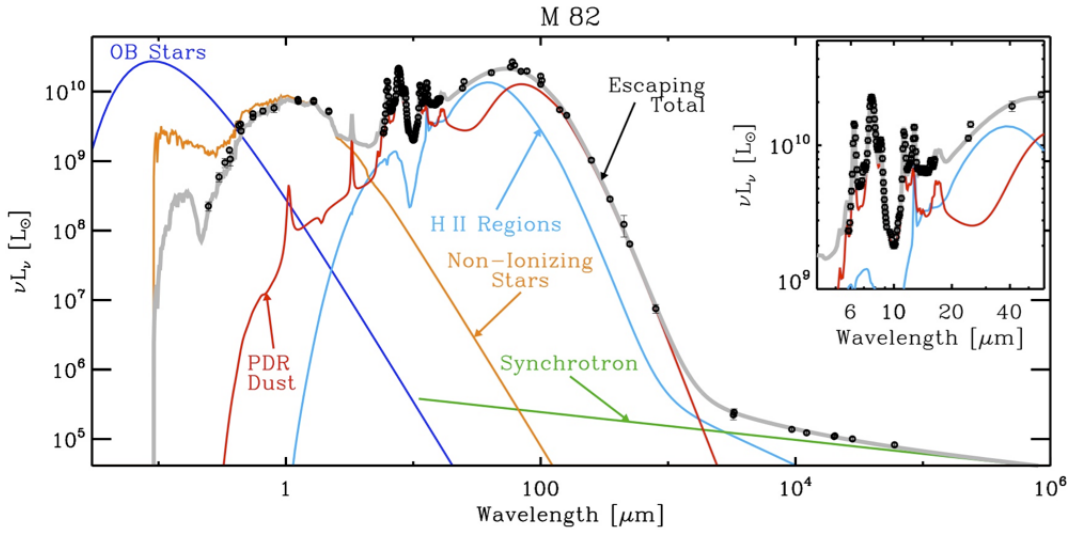

Figure 1.4 *Rest-frame spectral energy distribution of a nearby starburst galaxy, M82, with components modelled by Galliano et al. (2008), fitted to data in black. Emission in different wavebands is dominated by different physical processes. At UV and optical wavelengths, young, short-lived massive stars of spectral types O or early-type B (together, OB) emit radiation that quickly ionises the giant molecular gas clouds in which they formed (dark blue). Towards slightly longer wavelengths, non-ionising radiation from less massive stars (orange) contributes more. At mid-infrared wavelengths, thermal emission from small dust grains in photo-dissociation regions (PDRs; red) and hot dust in polycyclic aromatic hydrocarbon (PAH)-free HII regions (pale blue) dominate the SED. At radio wavelengths, a star-forming galaxy's emission is dominated by synchotron radiation.*

may be found in Madau & Dickinson (2014).

1.1.1 Ultraviolet

Massive, luminous, short-lived stars of spectral types O or early-type B (together, OB) rapidly ionise the giant molecular gas clouds in which they formed, producing HII regions. Over half of the total luminosity produced by a population of stars over 10 Gyr will be emitted within the first 100 Myr, mostly in the ultraviolet (UV) (Figure 1.5). The emission in the wavelength range $1250-2500$ Å is dominated by young massive stars, and we can use UV data to derive an average star formation rate over the last ~ 100 Myr. UV measurements are often used for studies of the star formation rate density (e.g. Arnouts et al. 2005; Bouwens et al. 2007) as they can be obtained for galaxies at a wide range of redshifts; emission at rest-frame far-ultraviolet (FUV) wavelengths, typically at 1500Å, can be measured from the ground using optical and near-infrared telescopes over the whole range

$1.4 < z < 6$, making this a sensitive and efficient probe of star formation. At lower redshifts, we require space-based telescopes such as HST to obtain FUV data.

UV luminosity-SFR calibrations are not trivial, however. The UV luminosity is sensitive to both recent star formation and to metallicity, with metal-rich galaxies emitting less for a given star formation rate than their metal-poor counterparts (Bicker & Fritzev. Alvensleben, 2005). SFRs can thus be overestimated at high redshifts, where star-forming galaxies are much less metal-rich than the local galaxies on which calibrations are usually based (Mehlert et al., 2003). Calibrations are also sensitive to the form of the initial mass function (IMF), which characterises the distribution of initial masses of a population of stars (e.g. Kroupa, 2002; Chabrier, 2003; Salpeter, 1955). UV luminosities trace the population of young massive stars, rather than the lower mass stars which make up a larger proportion of the stellar population, and different IMFs will therefore yield different luminosity-to-mass conversions.

A critical aspect of interpreting UV emission is robustly characterising the

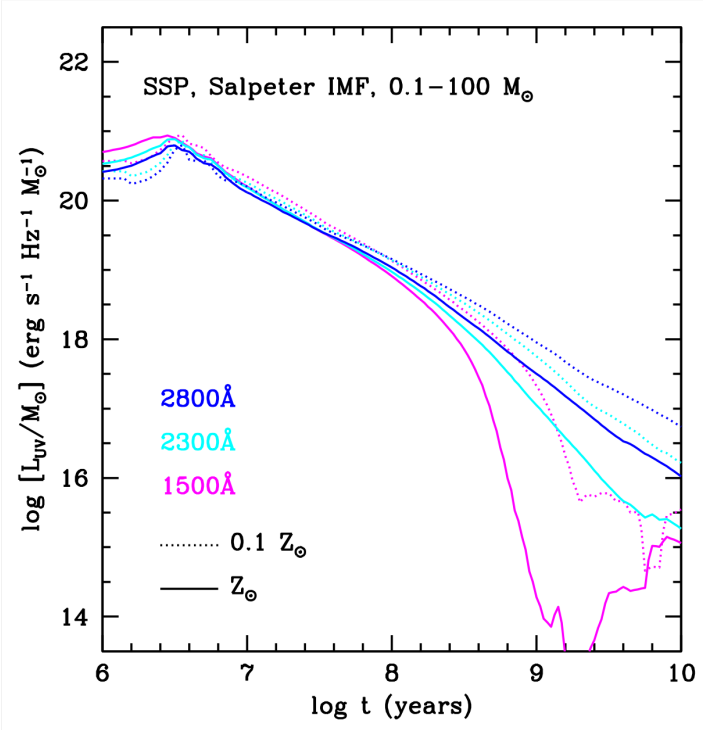

Figure 1.5 *UV luminosity at* 1500 *Å,* 2300 *Å and* 2800 *Å as a function of time for a population of stars of initial mass* $1\,M_{\odot}$*, derived using the stellar population synthesis models of Conroy et al. (2010). The luminosity drops off rapidly after* $\sim$ 100 Myr*. Figure taken from Madau & Dickinson (2014).*

degree of dust attenuation. The efficient coupling of short-wavelength light with small dust grains in the interstellar medium (ISM) leads to the absorption and scattering of UV light. This attenuation is difficult to correct for, given its dependence on the size and shape of individual dust grains and the chemical composition of the dust (Draine & Li, 2001; Weingartner & Draine, 2001), as well as the geometry of the dust and star-forming region (Misselt et al., 2001; Safarzadeh et al., 2017). Typically, wavelength-dependent dust corrections are applied (see Section 1.1.5), following prescriptions such as Calzetti et al. (1994), but these may not hold for the most dusty systems (see Chapter 7). UV-selected samples will tend to be biased towards galaxies with lower dust content, and thus evolution of the UV luminosity density is more easily studied than evolution of the whole star formation density. However, there are good statistical dust corrections established, such as the IRX-β relation (see Section 1.1.5).

1.1.2 Infrared

Short-wavelength light absorbed by dust particles is re-radiated at longer wavelengths, and constitutes thermal dust continuum emission. A combination of IR and UV luminosities (or of IR and Hα luminosities, e.g. Ibar et al. 2013), tracing the dust-obscured and dust unobscured light, is then a reliable measure of the total star formation rate. The total IR luminosity, L_{TIR}, usually measured in the wavelength range $8-1000\,\mu$m, is frequently adopted alone as a star formation rate indicator, particularly for dusty galaxies.

The IR region of a galaxy's spectrum is complex, with different sizes of dust grains at different temperatures all contributing to the bolometric infrared luminosity. The mid infrared (MIR) wavelength range ($\lambda < 30$ μm) is dominated by emission from small, hot dust grains in star-forming regions. Polycyclic aromatic hydrocarbons (PAHs), which are thought to be created in the circumstellar envelopes of carbon-rich AGB stars, emit and silicates absorb in this wavelength range. PAH emission is metallicity and radiation field dependent, which must be taken into account in dusty star-forming galaxies if we wish to estimate L_{IR} from these wavebands. Elbaz et al. (2011) found that the spectral energy distributions of main sequence galaxies (those that lie on the observed tight relation of galaxy mass and star formation rate, see Section 1.2.1) are markedly different from those of starburst galaxies (galaxies which lie above this sequence due to an extreme burst of star formation), with the former displaying strong PAH emission and a broad far infrared (FIR) 'bump' and the latter showing weaker PAH emission and peaking more sharply in the FIR.

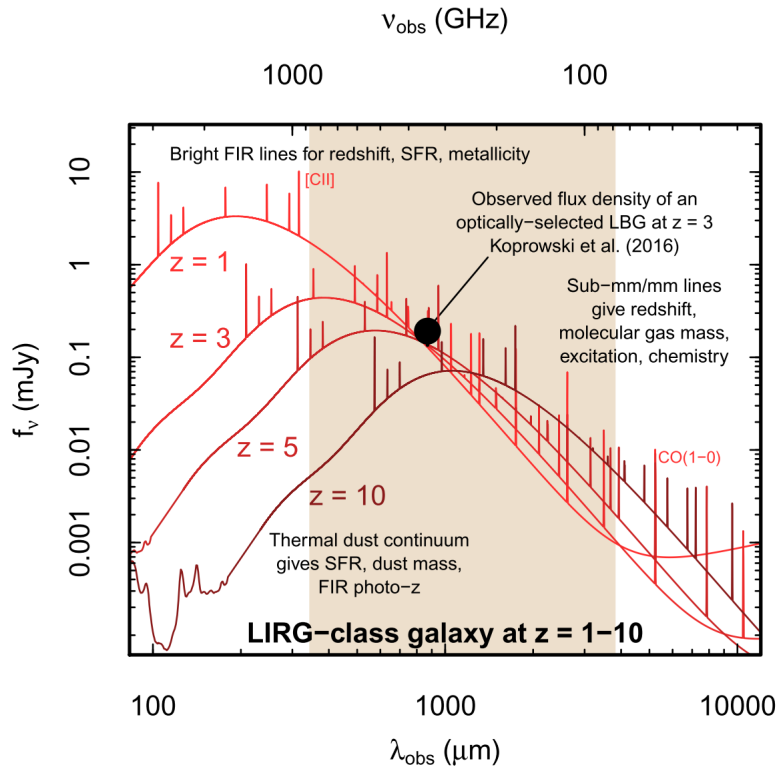

Figure 1.6 *The negative K-correction. $\lambda \sim 850\,\mu$m flux density from galaxies with similar intrinsic spectra remains approximately constant across the redshift range $z \sim 1{-}6$, as we trace further up the Rayleigh-Jeans tail at higher redshift. Figure from Geach (2019).*

Where star formation rates and dust content are high, as is the case at high redshift, IR emission can reliably trace total star formation. The dust continuum emission, at rest-frame FIR wavelengths, is a particularly fruitful way of studying distant galaxies, due to the 'negative k-correction'; observed-frame 850 μm flux remains approximately constant for a simulated galaxy as we move it to higher redshift (see Figure 1.6). However, SFR can be overestimated for galaxies that host older stellar populations with substantial contributions to the FIR flux (Hayward et al., 2014). While these conditions are satisfied at low redshift, this is less of a concern at high redshifts, and there is evidence that IR measurements are a better trace of the star formation rate density evolution at high redshifts than UV data: IR luminosity functions extend to higher luminosities than UV luminosity functions at the same redshift, and display strong luminosity evolution (Magnelli et al., 2013; Gruppioni et al., 2013). There can be substantial contributions to IR flux from AGN emission, however (e.g. Symeonidis et al., 2016, see Section 1.1.7).

1.1.3 Nebular line emission

Massive, short-lived stars ($M > 10\,M_{\odot}$, lifetime $< 20\,$Myr) ionize the gas in the HII regions that surround them, removing electrons from atoms. When a free electron is captured by a positive ion (a process known as 'recombination'), electrons typically end up in high energy states, and subsequently cascade down to the ground state. Radiation is emitted at characteristic wavelengths that correspond to the energy differences between allowed states. Nebular recombination lines are therefore direct, sensitive indicators of the ionising flux and hence of the young massive stellar population. The Lyman series of lines corresponds to electron transitions down to ground state ($n = 1$), and the Balmer series corresponds to transitions to $n = 2$ (see Figure 1.7). Hα and Hβ are the most commonly used Balmer lines.

In approximately half of cases of recombination, Hα is emitted; it is therefore the brightest nebular emission line, and most widely-used to infer SFR on $\sim 10\,$Myr timescales. Bell & Kennicutt (2001) analysed a sample of 50 nearby star-forming galaxies and found that, before applying any extinction corrections, specific star formation rates derived from Hα measurements were systematically higher than those derived from UV measurements, by an average factor of 1.5. Therefore Hα surveys are often said to provide data that are less biased by dust

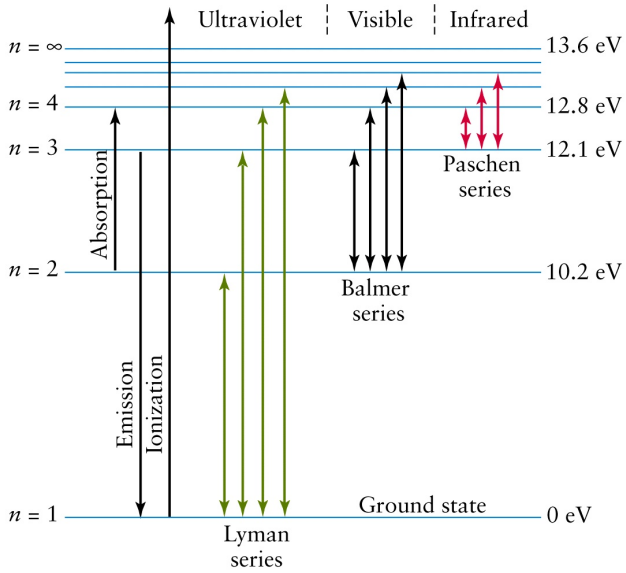

Figure 1.7 *Transitions of the hydrogen atom.* Hα, Hβ *are Balmer series transitions; electrons transition from an energy level at* $n \geq 3$ *to* $n = 2$. *In Chapters 3, 4 and 7, I use* Hα *as a star formation indicator. Credit: https://www.askiitians.com/iit-jee-structure-of-atom-and-nucleus/energy-levels-of-hydrogen-atom/*

extinction. Whilst longer wavelength hydrogen emission lines suffer even less dust extinction, they have much lower fluxes due to fewer transitions; in this respect, Hα presents a compromise between dust extinction and sensitivity. It has been used for decades (e.g. Cohen 1976; Kennicutt 1983) and is therefore well calibrated (e.g. Kennicutt et al., 1994). Although Hα emission is less strongly attenuated by dust than UV emission, extinction is still a substantial source of uncertainty. The conversion of Hα flux to a star formation rate is, as for the UV, affected by extinction, sensitive to the form of the IMF, and dependent on the assumption that ionised gas traces all of the massive star formation. Extinction is of particular concern in circumnuclear starbursts, where HII regions are very dense. Sometimes the equivalent width (EW; the emission-line flux normalised to the adjacent continuum flux density), a measure of the SFR per unit luminosity, is used. This broadly expresses the ratio of new-to-old stars: emission line flux is only from the newly formed stars, whereas older stars also contribute to the continuum emission.

The Hα line is redshifted out of the optical and into the near-IR above $z = 0.5$. Wide-field NIR surveys based on IR mosaic arrays such as UKIRT's WFCAM, VISTA's WIRCAM and the VLT's HAWK-I have enabled detection of large samples of Hα emitters up to $z \sim 2.5$, where the line is redshifted out of the near-IR waveband. Narrow band emission-line surveys enable selection of star-forming galaxies in narrow, well-defined redshift ranges, with identical selection techniques at different redshifts. One such survey is the narrow band High-redshift(Z) Emission Line Survey (HiZELS) (Sobral et al., 2013a), on which Chapters 3 and 4 of this thesis are based.

1.1.4 Radio

Star formation in galaxies can be traced independently of dust content by radio emission. At high radio frequencies ($\gtrsim 5\,\mathrm{GHz}$) thermal Bremsstrahlung dominates the emission. This stems from regions of gas that are ionised by hot, young, massive stars. Acceleration of electrons by interactions with ions results in the emission of radiation with a spectrum following $S_\nu \propto \nu^{-0.1}$. Non-thermal emission from supernovae at centimetre wavelengths is more frequently used as a delayed, indirect tracer of star formation. Relativistic electrons spiralling in weak magnetic fields emit synchotron radiation, characterised by a smooth spectrum ($\nu^{\sim -0.8}$) over a large wavelength range and linear polarization in the plane perpendicular to the magnetic field.

The radio emission is known to correlate tightly with IR flux, over a range of

redshifts. The first correlation was noted between 21 cm (1415 MHz) radio flux and 10 μm IR flux, initially for AGN and then for spiral galaxies (van der Kruit, 1971, 1973). Dickey & Salpeter (1984) also found a good correlation between 21 cm radio flux and 60 μm IR flux (IRAS) for $\sim$ 20 spiral galaxies (but not ellipticals), holding for both bright and faint optical sources. They noted that this was expected if both trace the population of young massive stars in spiral galaxies, and radio emission from ellipticals stems from AGN activity rather than star formation. This radio-IR correlation, now well-established at low and high redshifts, can be used to distinguish star-forming galaxies from radio-loud AGN, since the latter show a strong radio-excess. This technique will be discussed later in Section 1.1.7.

1.1.5 Dust attenuation

As stressed in Sections 1.1.1 and 1.1.3, galactic emission can be strongly attenuated by dust, particularly at shorter wavelengths (see Figure 1.8 for an example of dust obscuration within our own galaxy). In local galaxies, over a third of the bolometric luminosity is reprocessed by dust and emitted in the infrared. Robust models of the extinction as a function of wavelength are therefore critical for work that relies on data at rest-frame UV and optical wavelengths. Reviews of this topic may be found in Calzetti (2001), Draine (2003) and Casey et al.

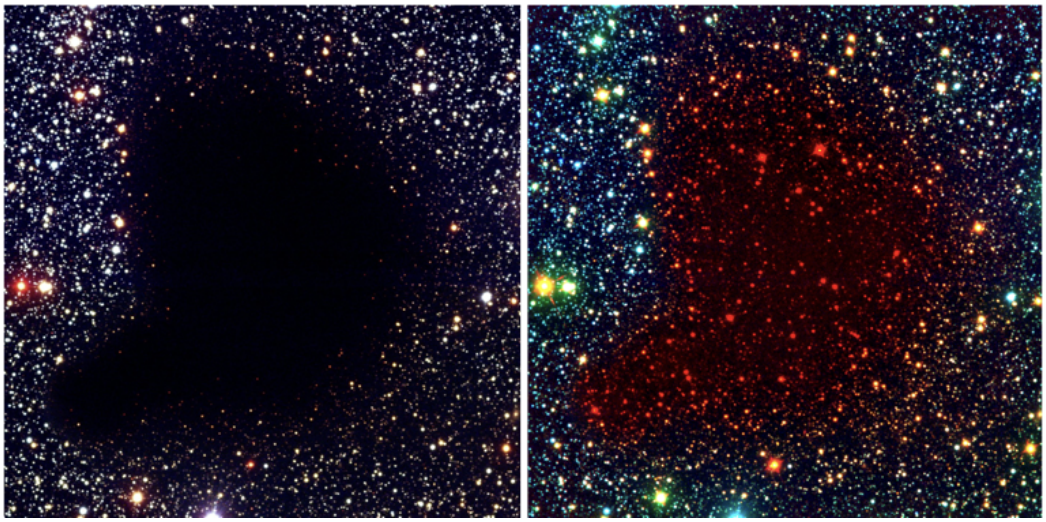

Figure 1.8 *Optical (B, V, and I-band composite; left) and IR (B, I, K_S-band composite; right) images of Barnard 68, a dusty molecular cloud within our own galaxy. $\sim$ 1000 background stars are obscured at the optical wavelengths probed by the VLT. At NIR wavelengths (K_S-band data from the New Technology Telescope, Chile), these stars are visible. Images taken from Alves et al. (2001).*

(2014). In this section, I recap the methods used to limit the biases introduced by dust attenuation.

Where detailed galaxy spectra are available, we can estimate dust attenuation via the ratio of Hα to Hβ flux. In the absence of dust, we expect Hα/H$\beta \approx 2.86$, from known Hydrogen atom transitions. This is known as the 'Balmer decrement', and is calculated for T $= 10^4$ K and electron density $n_e = 10^2\,\mathrm{cm}^{-3}$ gas. Since dust preferentially absorbs bluer light, upon the presence of dust in a galaxy, Hβ is more heavily attenuated by dust than Hα and hence the observed line ratio becomes larger than 2.86. We need good quality emission line measurements for this analysis, as Hβ is intrinsically weaker than Hα.

Given a measured Balmer decrement, we can calculate a dust attenuation analytically. The dust attenuation law relates the attenuation in magnitudes at wavelength λ (A_λ) to the colour excess E(B-V), the change in colour between the B-band and the V-band due to differential attenuation (Calzetti et al., 2000). It is:

$$k_\lambda = \frac{A_\lambda}{\mathrm{E(B-V)}}, \tag{1.1}$$

The attenuation at any wavelength, A_λ can then be calculated from Hα and Hβ fluxes, $S_{\mathrm{H}\alpha}$ and $S_{\mathrm{H}\beta}$, using:

$$A_\lambda = \frac{-2.5k_\lambda}{k_{\mathrm{H}\beta} - k_{\mathrm{H}\alpha}} \log_{10}\left(\frac{2.86}{S_{\mathrm{H}\alpha}/S_{\mathrm{H}\beta}}\right) \tag{1.2}$$

For any adopted dust attenation law (e.g. Calzetti et al., 2000), this then allows the correction of UV and Hα SFR estimates. This method assumes uniform dust attenuation throughout, although in reality dust coverage tends to be more patchy (see Ma et al. 2019, Chapters 6 and 7).

Estimating dust attenuation in galaxies without Hα and Hβ measurements can be challenging. The IRX $-$ β relation (Calzetti et al., 1994; Meurer et al., 1999) is an empirical relation that is frequently employed to derive FIR luminosities from UV data (e.g. Bouwens et al., 2016). It relates the ratio of the FIR and UV luminosity (IRX $= \mathrm{L_{FIR}/L_{1600}}$) to the UV spectral slope (β, where $f_\lambda \propto \lambda^\beta$) evaluated at 1600 Å and is frequently employed to derive FIR luminosities from UV data (e.g. Bouwens et al., 2016). The IRX$-\beta$ relation is of particular interest to studies of high-z galaxies as a method to estimate SFRs when only rest-frame UV luminosities exist. McLure et al. (2018) show that star-forming galaxies at $z = 2.5$ in the stellar mass range $\log_{10}(M_*/M_\odot) > 9.75$ are consistent with this relation, and a Calzetti et al. (2000) dust attenuation law (see also Bourne et al.

2017).

Correlations between global measurements of dust attenuation and galaxy mass (Burgarella et al., 2005; Pannella et al., 2015; Puglisi et al., 2016), star formation rate (Hopkins et al., 2001; Sullivan et al., 2001) and metallicity (Asari et al., 2007) can also be used to estimate dust attenuation and hence derive total SFR. Physical drivers of such correlations include the following: massive galaxies have larger dust reservoirs since they have had more star formation and therefore created more dust; galaxies with higher star formation rates have large, dusty star-forming regions; metal-rich galaxies have higher dust-to-gas ratios. However mass, metallicity and SFR are also correlated (see Section 1.2.4), and so it can be difficult to determine which, if any, of the parameters is the most fundamental in driving dust attenuation.

Garn & Best (2010) sought to disentangle these dependencies, using a large sample of 90,000 star-forming galaxies from the Sloan Digital Sky Survey (SDSS; York & Adelman 2000), which provided measurements of stellar mass, star formation rate and gas-phase metallicity, as well as Hα and Hβ emission line fluxes. They concluded that knowledge of the galaxies' stellar masses is sufficient to predict the dust attenuation of a population statistically at $z = 0$. Ibar et al. (2013) proposed that the relation between dust attenuation and stellar mass holds to higher redshifts, finding a consistent $A_{\mathrm{H}\alpha} \propto M_*$ relation at $z = 1.47$. In contrast, Qin et al. (2019) also suggest a universal relation of dust attenuation across cosmic time, but driven by IR luminosity and galaxy size, instead of stellar mass.

1.1.6 Modelling galaxy spectral energy distributions

As shown in Figure 1.4, a galaxy's spectral energy distribution (SED) reflects a wide array of physical processes happening within it. The galaxy's star formation history, information about stellar populations such as their mass, metallicity, abundance pattern, and the form of the IMF, as well as its dust and gas, are all encoded in the SED. Conroy (2013) provides a comprehensive review of this field, discussing both the process of building mock galaxy SEDs and the challenges of inferring physical properties from those of observed galaxies. In Figure 1.9, we reproduce a key figure from this review. This figure summarises the ingredients that go into building a stellar population synthesis (SPS) model. The basic building block is the simple stellar population (SSP; e.g. Leitherer et al. 1999; Bruzual & Charlot 2003; Maraston et al. 2006), which describes the time evolution of a single population of stars, born at the same time and having the same

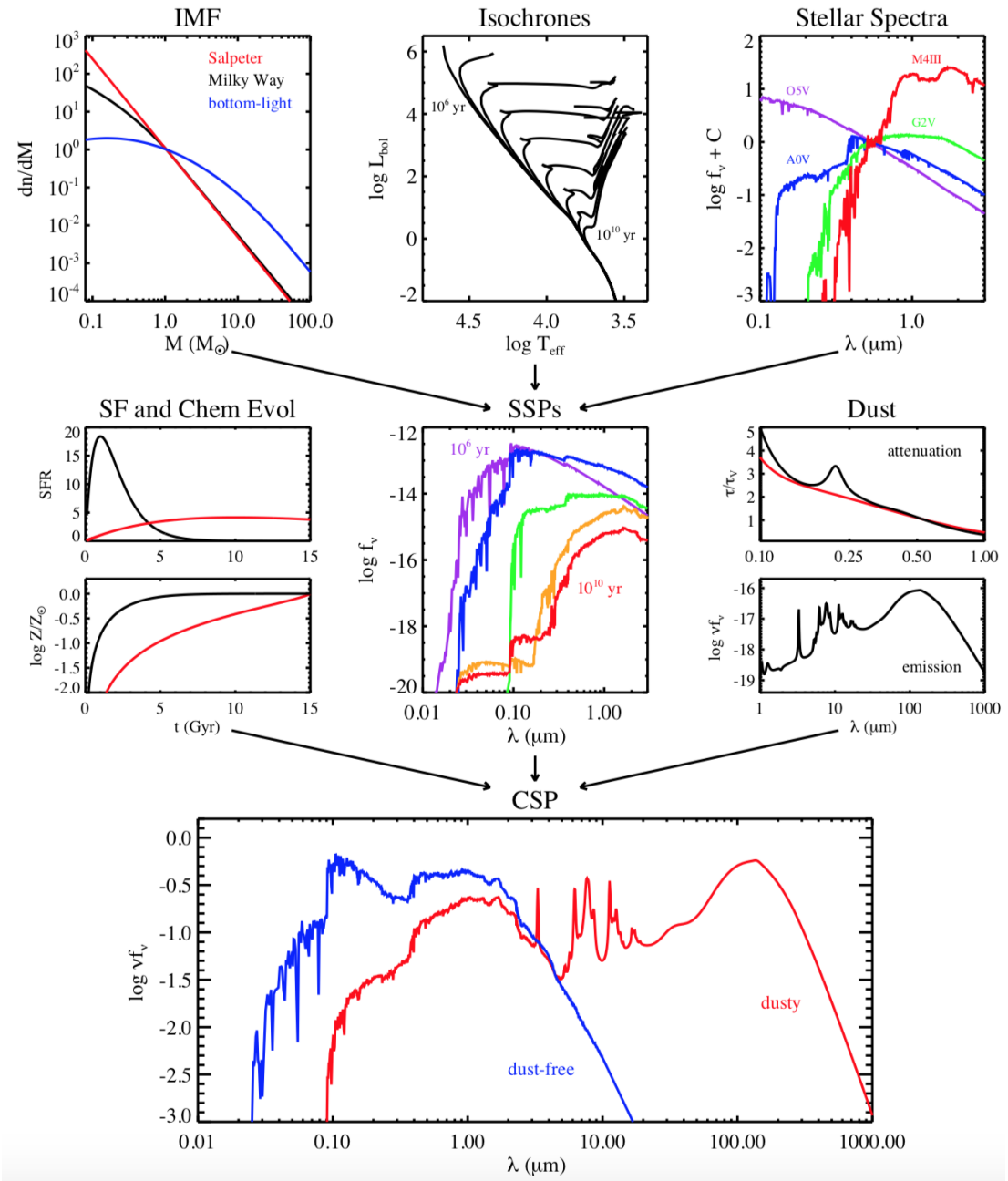

Figure 1.9 *The ingredients of a model galaxy SED, described in Section 1.1.6. From Conroy (2013). Models of the IMF, stellar isochrones and stellar spectra (top row) are combined to form simple stellar populations (SSPs). Summing multiple SSPs of different ages, with models for dust and the star formation and chemical evolution histories (middle row), yields a composite stellar population (bottom), which can be compared to observed galaxy SEDs.*

metallicity and abundance pattern. The IMF (e.g. Salpeter, 1955; Kroupa, 2002; Chabrier, 2003) describes the distribution of stellar masses within this single population, and stellar isochrones (e.g. Marigo et al., 2017; Cordier et al., 2007; Dotter, 2016) describe the position of stars with a given age and metallicity on the Hertzsprung-Russell (HR) diagram (which comes from models of stellar evolution). Stellar spectral libraries transform the output from stellar evolution models to libraries of SSPs. Composite stellar populations (CSPs) are then put together using stars with a range of ages and metallicities (based on the star formation history of the galaxy), with the addition of dust.

This technique can be employed to infer properties such as galaxy stellar mass, star formation rate and metallicity from observed SEDs (e.g. using the MAGPHYS, CIGALE or BAGPIPES codes; Da Cunha et al. 2008, 2011; Buat et al. 2012; Carnall et al. 2018, 2019; see Chapter 7). In principle, high-quality spectra can put strong constraints on these properties, and on star formation histories, since different star formation indicators are sensitive to star formation on different timescales. However, obtaining robust inferences is made more difficult by degeneracies in age, dust content and metallicity, and by biases introduced by the priors imposed on the form of the star formation history and on the dust model (Papovich et al., 2002).

1.1.7 Distinguishing between emission from AGN and star formation

Distinguishing the signatures of AGN from those of star formation is key when deriving star formation rates. In this section, I will provide a very brief introduction to AGN emission, and the ways in which AGN-dominated galaxies are commonly separated from their purely star-forming counterparts.

AGN can be divided into two main categories: 'radiative-mode AGN', and 'jet-mode AGN' (see Heckman & Best 2014 for a comprehensive review). Radiative-mode AGN tend to have bright UV/Hα/IR emission, and often have associated star formation. A straightforward way to identify these types of AGN is via their SEDs. Over half of the bolometric emission from an unobscured AGN is emitted in the region extending from $< 1000\,\mathring{A}$ to the near infrared at $\sim 1\,\mu$m, the so-called 'Big Blue Bump' (see Figure 1.10). This emission is believed to arise from a number of physical processes. Geometrically thin, optically thick accretion disks surrounding the supermassive black hole emit strongly in the UV. The 'broad-line' emission regions (dense gas clouds near the accretion disk) and 'narrow-line' emission regions (lower density clouds at larger radii) are heated

and photoionised by this emission, emitting in the UV, optical and IR. Type 1 (unobscured) radiative-mode AGN are those for which the accretion disk and broad-line regions are directly visible. The accretion disk and broad-line regions in Type 2 (obscured) radiative-mode AGN are blocked from sight by a dusty torus. These AGN are best identified in the IR, and sometimes by hard X-rays or emission lines from the narrow-line region.

Jet-mode AGN occur when accretion rates are much lower. These are radiatively inefficient. The accretion disk for this type of AGN is geometrically thick, with a radiative cooling time that is much longer than the gas inflow time. The short-wavelength emission from jet-mode AGN is therefore much less significant, however inefficient accretion flows launch jets with detectable synchotron emission at radio wavelengths.

Since different types of AGN are best detected in different wavebands, multi-wavelength data are required to obtain the cleanest samples of star-forming galaxies and AGN. Where extensive optical/IR photometry exists, it can be shifted into the rest frame and fitted using template SEDs (fitting software that will incorporate AGN emission includes AGNfitter and CIGALE; Buat et al.

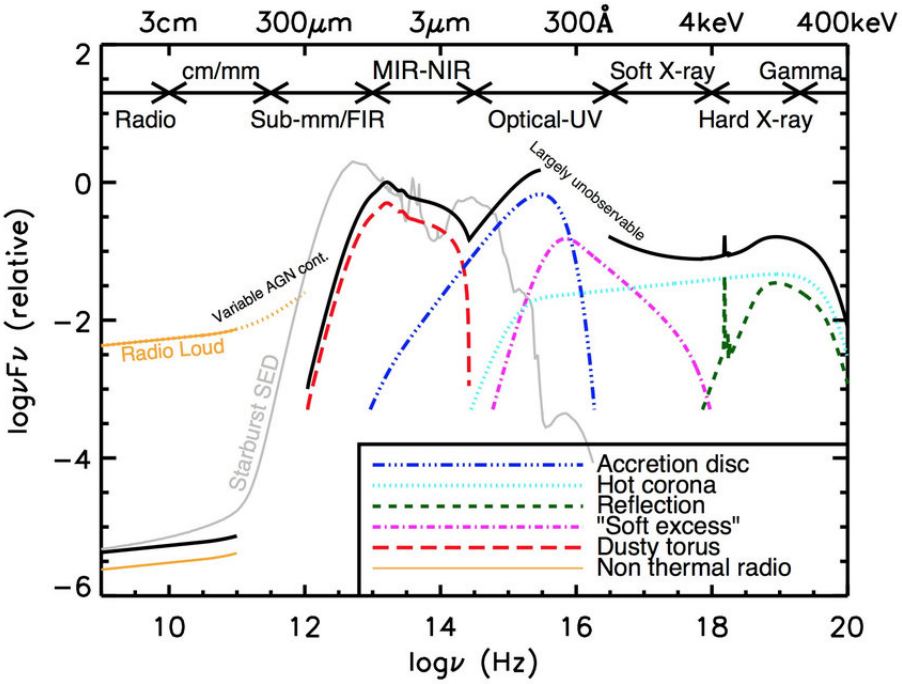

Figure 1.10 *An example AGN SED, decomposed into contributions shown by coloured lines. A typical starburst SED is shown in grey. The AGN SED has a striking optical-UV excess due to emission from the accretion disk, and substantial emission in the MIR-NIR due to the hot torus. Figure from Marin (2018).*

2012; Calistro Rivera et al. 2016). Template SEDs are typically drawn from observations of nearby galaxies with a range of morphologies and ages (spiral, elliptical, starburst), and AGN (Type 1, Type 2, composite, Seyfert). Stellar population synthesis models often don't fold dust effects in very well (Conroy et al., 2010), so using template SEDs that include dust can be an advantage, particularly when fitting at IR wavelengths. This technique can enable the identification of AGN, but is limited by the range of templates available (in particular, low redshift templates may not be fully applicable to high redshift galaxies). When photometric data spanning a range of wavelengths is not available, 2-band colour diagnostics such as the Stern wedge (Stern et al., 2005) can be used instead.

Various optical emission line diagnostics have been developed to distinguish AGN from purely star-forming galaxies. Type 1 AGN display characteristic broadened emission lines, whereas Type 2 AGN have narrow emission lines just like star-forming galaxies. The BPT diagram (Baldwin et al., 1981; Kewley et al., 2001; Kauffmann et al., 2003) allows a quantitative classification via emission line ratios. The ratios most commonly compared are [OIII]5007/Hβ and [NII]6584/Hα. This method is based on detecting increased photoionization and heating due to radiation from the AGN.

Selection of AGN via the radio-IR correlation (see Section 1.1.4) is also possible (Appleton et al., 2004). $q_{\rm IR}$, the ratio of radio and IR luminosities, appears to remain constant to at least $z = 2$ (Ivison et al., 2010; Sargent et al., 2010) and perhaps as far as $z = 3.5$ (Ibar et al., 2008). Galaxies with radio emission in excess of this relation can be identified as radio-loud AGN. Hard X-ray emission from an AGN's hot corona and accretion disk (Brandt & Hasinger, 2004) can similarly be used to identify X-ray bright AGN (Maccacaro et al., 1998). However, neither of these diagnostics provide complete samples of AGN.

1.2 Key galaxy scaling relations

Observations allow us to infer general scalings between galaxy properties, and to put constraints on physical models of galaxy evolution. In this section I will review a small number of key observed relations.

1.2.1 The main sequence of star-forming galaxies

The distribution of low-redshift galaxies on the SFR vs stellar mass plane (Figure 1.11, left-hand panel) shows two very clear peaks, with a linear locus of (blue)

star-forming galaxies separated from their (red) passive counterparts by a ridge, sometimes called the 'green valley' (Renzini & Peng, 2015). The locus of star-forming galaxies is linear to the highest stellar masses. Galaxies that lie in between the peaks are either passive galaxies undergoing a short burst of merger-induced star formation or previously star-forming galaxies currently undergoing quenching.

The vast majority of star-forming galaxies have been found to follow a similar tight relation between stellar mass and SFR, from $z = 0$ (Brinchmann et al., 2004) up to at least $z = 3.5$ (Elbaz et al., 2007; Daddi et al., 2007; Karim et al., 2011; Rodighiero et al., 2011, 2014) and possibly even up to $z = 6$ (Pearson et al., 2018). Large samples have enabled multi-wavelength compilation studies (e.g. Speagle et al., 2014; Lee et al., 2015), which put constraints on the evolution of the shape of the relation, the so-called 'star-forming main sequence'. There remains debate about the slope and normalisation, however, with some studies finding flattening at high stellar masses (see Figure 1.11, right-hand panel). Such discrepancies in the slope and normalisation tend to arise due to selection effects and the definition of a 'star-forming' galaxy. Inclusion of massive galaxies in which star formation is ceasing will naturally pull the mean SFR down at high stellar masses. Similarly, a sample selected by SFR will favour a flatter relation.

Multiplying the number of galaxies by the star formation rate enables us to see where most of the star formation takes place (Renzini & Peng, 2015). This is in the peak of star-forming galaxies, as expected, but there is a small bump in

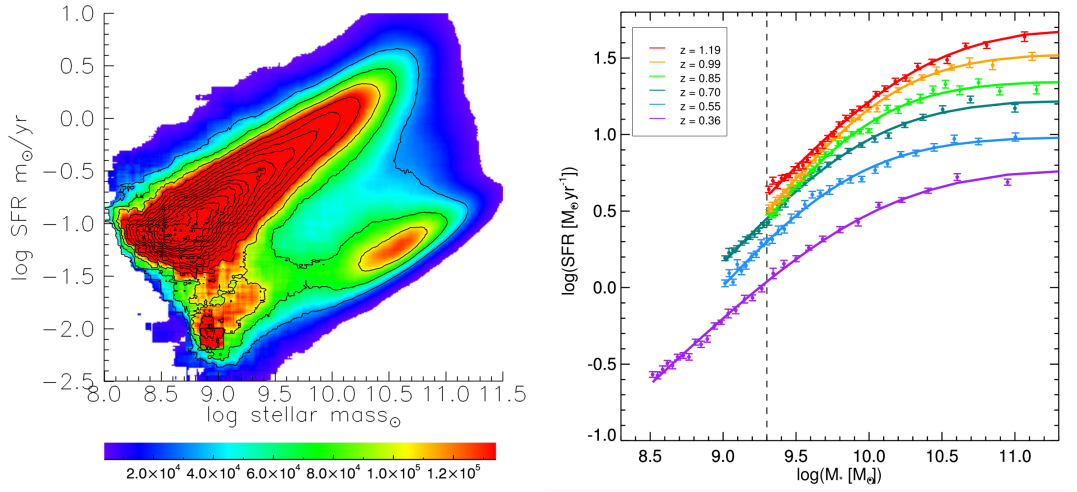

Figure 1.11 *Left: The distribution of galaxies within the SFR vs stellar mass plane, showing the linear sequence of star-forming galaxies and two peaks of quenched galaxies, from Renzini & Peng (2015). Right: the evolution of the main sequence, as derived by Lee et al. (2015).*

the quenched peak, where there is a little residual star formation. Multiplying by the stellar mass rather than star formation shows where most of the stellar mass is contained. As expected from the decline in volume-averaged star formation history, most of the stellar mass in the local universe resides in quenched galaxies.

Galaxies that lie above the main sequence, with particularly high SFR for their stellar mass, are known as 'starbursts'. Such galaxies are rare by definition; Rodighiero et al. (2014) show that starbursts make up just 2% of mass-selected SFGs at $z = 1.5 - 2.5$, though they account for a more substantial amount ($\sim$ 10%) of the cosmic SFR density at this epoch. Many studies have suggested that main sequence and starburst galaxies are forming stars via different mechanisms, with the former in a steady-state with long gas-consumption timescales, and the latter in a rapid, merger-driven mode (e.g. Daddi et al., 2010; Genzel et al., 2010). I will discuss the possible enhanced star formation efficiency in starburst galaxies in Section 1.2.2.

A number of other naming conventions are used for highly star-forming, IR-luminous galaxies. Galaxies with $10^{11} L_{\odot} < L_{\rm IR} < 10^{12} L_{\odot}$ are known as Luminous Infrared Galaxies (LIRGs), and those with $10^{12} L_{\odot} < L_{\rm IR} < 10^{13} L_{\odot}$ are known as Ultra-luminous Infrared Galaxies (ULIRGs). Due to the fixed luminosity nature of this definition and the evolution of the main sequence, LIRGs and ULIRGs are rare in the local Universe but numerous at early times, as the upper end of the main sequence reaches ULIRG luminosities (Pérez-González et al., 2005; Caputi et al., 2006). Indeed, at $z = 2$, ULIRGS are believed to contribute approximately 50% of the total energy density (Caputi et al., 2007). Since these highly star-forming, dusty, IR-luminous galaxies are detected at high redshifts in the sub-mm, they are often named 'sub-mm galaxies' (SMGs).

1.2.2 Fuel for star formation in galaxies: the Kennicutt-Schmidt law

Schmidt (1959) was the first to study the relationship between the star formation rate volume density ($\rho_{\rm SFR}$) and gas volume density ($\rho_{\rm gas}$) of individual galaxies. He obtained measurements of neutral HI gas volume density from 21-cm line observations, and combined the initial luminosity function of nearby stars with estimates of stellar lifetimes to derive star formation rates. He found that $\rho_{\rm SFR} \propto \rho_{\rm gas}^{n}$, with $n \sim 2$ within the Milky Way. Kennicutt (1998) extended this analysis outside our own galaxy, formulating a relation using the global properties of 61 star-forming spiral galaxies and 36 starburst galaxies. By this time, CO measurements were available, which enabled far better estimates of the

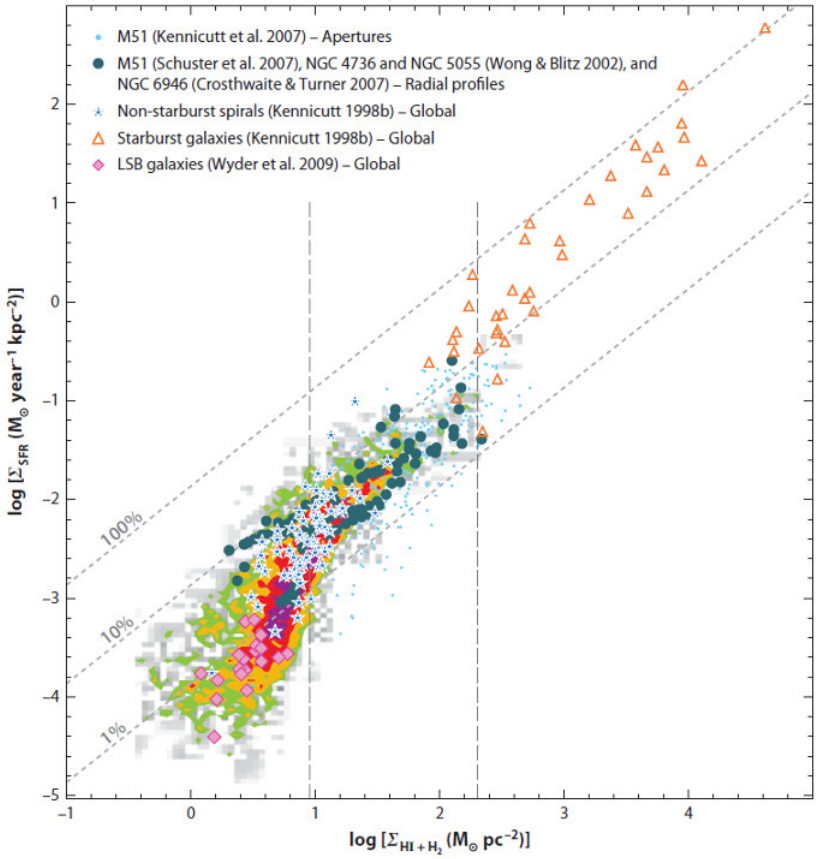

Figure 1.12 *The Kennicutt-Schmidt relation between star formation rate surface density and gas surface density. Grey, green, yellow and red regions are sub-kpc scale regions of galaxies from the SIRTF Nearby Galaxies Survey (SINGS) (Kennicutt et al., 2003). Light blue dots show* $0.5 - 2\,\mathrm{kpc}$ *scale regions of NGC 5194 (M51a) (Kennicutt et al., 2007). Radial profiles of other nearby galaxies are shown as dark green circles. Global measurements of 61 spiral (blue stars) and 36 starburst (orange triangles) galaxies are also shown (Kennicutt, 1998). Pink diamonds show 20 low surface brightness galaxies from Wyder et al. (2009). Figure taken from Kennicutt & Evans (2012).*

H_2 molecular gas from which stars form than HI atomic gas measurements had allowed. Kennicutt (1998) used Hα measurements to derive global star formation rates for the main sequence spiral galaxies, and IR emission as an indicator for the dusty starburst galaxies. Remarkably, the two populations obeyed the same power-law relation (the 'Kennicutt-Schmidt relation', K-S; Figure 1.12), $\Sigma_{\mathrm{SFR}} \propto \Sigma_{\mathrm{gas}}^{n}$, with $n \sim 1.4$, though there was significant variation in n, depending on the SFR tracer used. Two possibilities were proposed for the physical basis of the K-S relation: a model where gravitational forces control star formation and the timescale is the free-fall time ($\rho_{\mathrm{SFR}} \propto \frac{\rho_{\mathrm{gas}}}{t_{\mathrm{ff}}} \implies \rho_{\mathrm{SFR}} \propto \rho_{\mathrm{gas}}^{1.5}$), and one where

the timescale is the local orbital timescale, as in the scenario of spiral arms or bars triggering star formation. In that case, $\Sigma_{\rm SFR} \propto \frac{\Sigma_{\rm gas}}{\tau_{\rm dyn}} \propto \Sigma_{\rm gas}\Omega_{\rm gas}$, where $\Omega_{\rm gas}$ is the angular rotation speed, which determines the orbital timescale $\tau_{\rm dyn}$. The normalisation of this relation shows that star formation on galactic scales is very inefficient; only a few per cent of the total gas mass is converted to stars on the free-fall timescale.

Resolved gas measurements show that a power-law relation also holds on sub-galactic scales at low redshift (e.g. Kennicutt et al., 2007; Bigiel et al., 2008; Azeez et al., 2016) and at high redshift (e.g. Freundlich et al., 2013). However, there is some indication that it may break down on sub-kpc scales due to small spatial offsets between peaks of cold gas emission and star formation, incomplete sampling of the mass function of giant molecular clouds on very small scales, and the subtleties of star formation and gas depletion timescales (Onodera et al., 2010; Calzetti et al., 2012; Feldmann et al., 2012; Diederik Kruijssen & Longmore, 2014). Nevertheless, the relation has been widely recovered on larger scales, and successfully used to infer gas depletion timescales. Bigiel et al. (2011) propose that the broadly linear, 1kpc-scale relation between $\Sigma_{\rm SFR}$ and $\Sigma_{\rm H_2}$ across 30 nearby disk galaxies implies a constant gas depletion timescale of $\sim 2.35\,$Gyr. Work at higher redshifts finds a similarly tight relation, but a steeper slope, with shorter gas depletion timescales of $0.84 \pm 0.07\,$Gyr implied (Freundlich et al., 2019).

Powerful starbursts and SMGs appear to follow a parallel relation above that of typical spirals (e.g. Bouche et al., 2007; Pereira-Santaella et al., 2016, see the

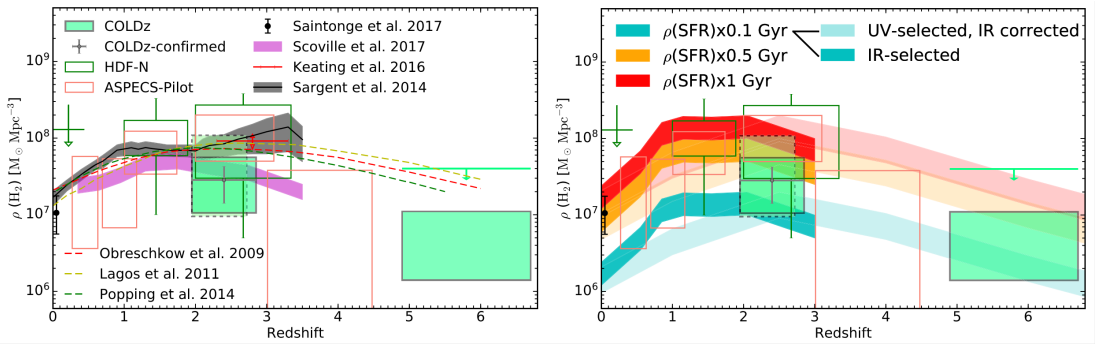

Figure 1.13 *Left: The evolution of the molecular gas density of the Universe, from Riechers et al. (2018). A number of CO surveys are plotted, as labelled in the upper left-hand corner, along with the dust-based relation derived by Scoville et al. (2017). Predictions from simulations, shown with dashed lines, are broadly in agreement with observations. Right: the molecular gas density evolution predicted by scaling the observed SFR density by constant timescales of gas consumption.*

starburst galaxies in Figure 1.12), converting gas to stars much more efficiently (though note the uncertain conversion between CO luminosity and H_2 gas mass).

1.2.3 Molecular gas accretion drives both star formation and AGN activity

The well-known tight correlation between the masses of black holes and the masses of the stellar bulges of their host galaxies (Magorrian et al., 1998) indicates the past joint growth of galaxies and AGN. A significant amount of evidence for this has been collected since this observation. The star formation rate density and supermassive black hole (SMBH) accretion rates have been observed to follow the same rise back to at least $z = 2$ (Heckman et al. 2004, Figure 1.14). The molecular gas density as a function of redshift has a similar form (e.g. Riechers et al., 2018, see Figure 1.13), suggesting that both the evolution of star formation rates and AGN activity are driven by cold gas accretion (Scoville et al., 2017).

On the scales of individual galaxies, specific star formation rates increase with AGN luminosity (Kauffmann et al., 2003), and average star formation rates and specific star formation rates of star-forming galaxies and galaxies with AGN agree well across cosmic time (Mainieri et al., 2011; Santini et al., 2012; Heckman & Best, 2014). This latter relation is weakened when we include quiescent galaxies (Santini et al., 2012; Vito et al., 2014), indicating that star formation and AGN

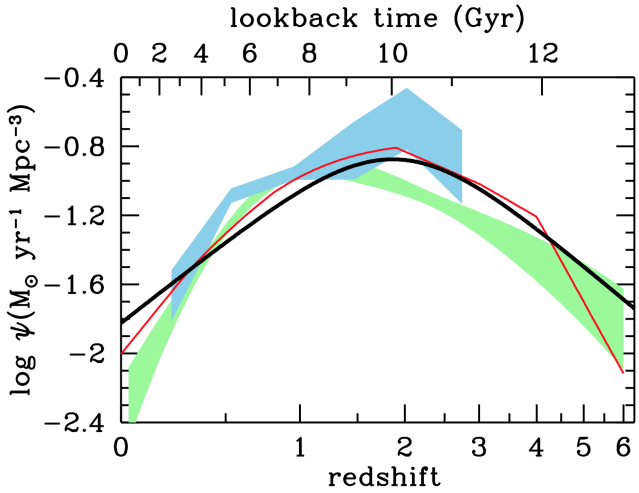

Figure 1.14 *The massive black hole accretion history, from Madau & Dickinson (2014). The red curve and light green shading are derived from X-ray data, the light blue shading is from infrared measurements. The black hole accretion history follows the best-fit star formation history (black line) remarkably well.*

have the same drivers, both being closely related to the supply of cold gas.

1.2.4 Galaxy chemical enrichment: a 'fundamental relation' between star formation, mass and metallicity

Early work on metallicity focussed on the apparently linear relation between galaxy luminosity and metallicity (Lequeux & Peimbert, 1979). As galaxy stellar mass measurements became more reliable, a lower-scatter galaxy stellar mass - gas phase metallicity was derived (Tremonti et al., 2004, see left-hand panel of Figure 1.15). The physical explanation for this, supported by models (e.g. Heckman et al., 2000), is the preferential loss of metals from the small potential wells of low mass galaxies via galactic winds. Tremonti et al. (2004) concluded that only half of the scatter of the mass-metallicity relation could be attributed to measurement errors. Mannucci et al. (2010) argued that the remainder of this scatter is due to variations in SFRs between galaxies of a given mass and metallicity, introducing a third parameter. The new relation they proposed between mass, metallicity and SFR takes the form of a plane with very little (~ 0.5 dex) scatter at low redshift, which is shown to hold to at least $z = 2$ (Erb et al., 2006). They called this the Fundamental Metallicity Relation (FMR, Figure 1.15, right-hand panel). Given that both the mass-metallicity and the mass-SFR relations evolve with redshift whereas the three-dimensional plane does not, Mannucci et al. (2010) claimed that this relation is the more fundamental.

Interpretations of the physical basis of the FMR centre on the relationship between inflows, outflows, and star formation. Infall of 'pristine' (metal poor) gas dilutes the metallicity of a galaxy's gas, and provides fuel for star formation,

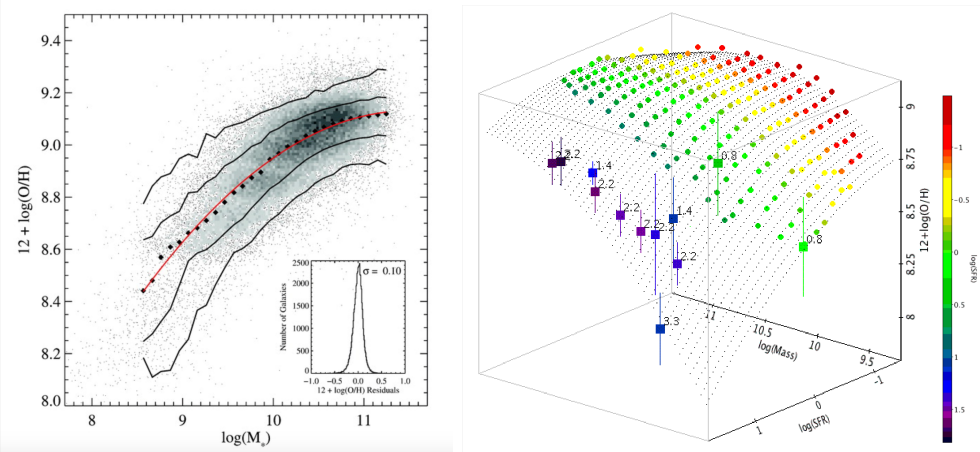

Figure 1.15 *The Mass-Metallicity Relation (Tremonti et al. 2004, left) and the Fundamental Metallicity Relation (Mannucci et al. 2010, right).*

boosting the star formation rate. In this scenario, thought to be particularly important at high redshifts where both inflow rates and star formation rates are high, the gas inflow and star formation rates vary on shorter timescales than the overall chemical enrichment. High star formation rates correlate with high inflow rates, which dilute the gas in the galaxy, lowering its gas-phase metallicity. Models of local galaxies favour outflow-dominant scenarios, where gas inflows and star formation rates vary on longer timescales than the chemical enrichment. Mannucci et al. (2010) modelled an outflow proportional to $\frac{\mathrm{SFR}^{0.65}}{M_*}$, in which outflows are more important for low mass galaxies, explaining their metal depletion.

These equilibrium-style models (see also Lilly et al., 2013) are motivated, in part, by the small scatter in the FMR, which indicates that common secular processes drive galactic metal enrichment, at least in the local universe. In a merger-dominant scenario, we would expect larger scatter in the FMR. In the case of a merger, some of the fuel from star formation is the metal enriched gas within the interacting galaxies, rather than pristine gas from inflows. This is particularly significant in boosting the gas metallicity of the smaller merging galaxy (see Michel-Dansac et al. 2008, who found a mean metallicity excess of 0.2 dex for strongly interacting galaxies with $M* < 10^9\,M_\odot$). The dominance of mergers at high redshift could be the physical cause of the apparent departure from the FMR at redshifts above $z \sim 3$ found by Mannucci et al. (2010).

1.2.5 The quenching of star formation in galaxies

The observed bimodality in the colour, star formation rate and morphologies of galaxies is long established (e.g. Baldry et al. 2004; see Figure 1.11). Galaxies in the local Universe tend to be in one of two camps: blue, star-forming spirals, and red, passive (non star-forming) ellipticals, though outliers exist (e.g. the small samples of red spirals identified by Masters et al. 2010). These classes correlate with environment (as shown originally by Oemler 1977 and Dressler 1980, galaxies in rich clusters are preferentially passive ellipticals, whereas field galaxies tend to be star-forming and disk-like, with increasing star formation rates and star-forming fractions further from cluster centres; Lewis et al. 2002; Gomez et al. 2003) and with galaxy stellar mass (high mass galaxies are far less likely to be star-forming than their low mass counterparts; Baldry et al. 2006). A significant body of work has developed in recent years, focusing on the physical processes which drive galaxy 'quenching', the process by which a previously star-forming galaxy halts star formation and becomes passive. Given the implications of these

possible quenching mechanisms for galaxy colours, gas content and quenching timescales, there are many potential observational avenues. Large galaxy surveys such as the Sloan Digital Sky Survey (SDSS, York & Adelman 2000) have yielded the statistical power to test possible mechanisms.

The power of the SDSS to break degeneracies was shown clearly by Peng et al. (2010), who used data from the SDSS and zCOSMOS survey to study the relationships between mass, star formation rates and environment back to $z \sim 1$. They proposed two primary quenching mechanisms, 'mass quenching' and 'environment quenching', which dominate at different epochs and galaxy masses. 'Mass quenching' refers to the observation that most high-mass galaxies are quenched (e.g. $\sim 50\%$ of $M_* > 10^{10.3}\,M_\odot$ galaxies are passive at $z = 0$, even in the field), and 'environment quenching' to the observation that most galaxies in clusters are passive, regardless of their mass. The empirical modelling of Peng et al. (2010) showed that the two effects are separable: mass quenching efficiency is largely independent of overdensity and vice-versa (see Figure 1.16, left-hand panel). The fraction of galaxies that are satellites is very tightly correlated with environmental overdensity, but very little dependence on mass or redshift exists.

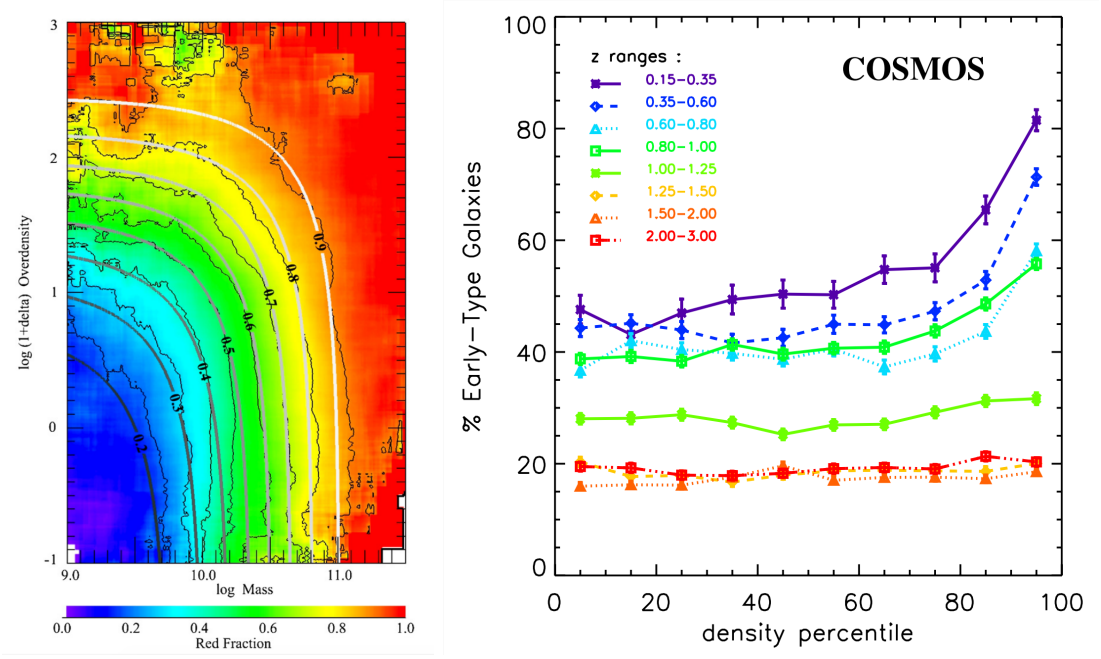

Figure 1.16 *Left: fraction of $z \sim 0$ galaxies that are passive as a function of stellar mass and environmental overdensity, from Peng et al. (2010). Right: observed fractions of early-type (mainly quenched) galaxies as a function of overdensity, at different epochs, from Scoville et al. (2013). The percentage of galaxies that are early type only correlates with overdensity at $z < 1$.*

Thus, quenching of satellites may provide the physical basis for 'environment quenching' (see also Peng et al., 2012). In this scenario, satellite galaxies may be quenched by either mass or environment quenching, whereas central galaxies can only be quenched by mass quenching. Mass quenching is less physically motivated: the 'mass quenching death rate' for a galaxy is proportional to its the star formation rate, and possibly related to AGN feedback.

The physical mechanisms that drive quenching are less well-established than empirical relations. Such mechanisms tend to be based on a lack of cold gas inhibiting further star formation. There are a multitude of environmental processes that could drive quenching in satellite galaxies, including strangulation (when cool gas flows onto the central rather than the satellite; Larson 1980), ram-pressure stripping (the rapid removal of gas from a galaxy as it moves through the intra-cluster medium; Gunn & Gott 1972; Abadi et al. 1999) and tidal stripping (when gas is pulled from a low mass galaxy via gravitational interactions with a more massive galaxy). Stripping effects have in the past been inferred statistically via depleted neutral hydrogen content in cluster galaxies (Giovanelli & Haynes, 1985), and are now observed directly via spectacular 'jellyfish'-like structures (e.g. Poggianti et al., 2017; Bellhouse et al., 2017). Gas stripping is believed to be particularly important close to the centre of galaxy clusters, at higher overdensities and velocities, and for low mass galaxies with small potential wells (Boselli & Gavazzi 2006 provides a thorough review).

For high mass galaxies, star formation is believed to be inhibited either by preventative (i.e. gas is prevented from accreting onto a massive galaxy, perhaps by a central AGN) or ejective (i.e. gas is expelled from the galaxy via stellar or AGN-driven outflows) feedback (see Figure 1.17). Ejective feedback is easier to observe, with high outflow rates observed from small samples of galaxies (e.g. Maiolino et al., 2012; Forster Schreiber et al., 2014; Geach et al., 2018). Preventative feedback has traditionally been based on cooling time arguments: below some characteristic dark matter halo mass ($\sim 10^{12}\,M_{\odot}$), gas cooling times are short compared to the dynamical time of the dark matter halo, and cold gas accretes efficiently. Above this halo mass, cooling times are long, and the gas that accretes onto the galaxy is hot (Dekel & Birnboim, 2006). This is particularly important at low redshift, where haloes of $\sim 10^{12}\,M_{\odot}$ mass are typical. At high redshifts, haloes this massive are rare, existing in nodes of the cosmic web where narrow, dense streams of cold gas can penetrate the hot halo. Slow quenching timescales derived for low-redshift galaxies (e.g. Peng et al., 2015; Trussler et al., 2018) favour such a scenario, as cold gas supplies are gradually used and not

replenished. Heating by a central AGN could also play a role in preventing gas from cooling. In particular, jet-mode AGN, fuelled by slow cooling of the hot gas halo, inject their energy back into the halo via radio jets in a feedback cycle (Bower et al., 2006; Heckman & Best, 2014).

One area of very recent progress is the field of 'post-starburst' galaxies (Wild et al., 2009). These are galaxies identified via their strong Balmer absorption, which indicates that they had a strong starburst that ended within the past $\sim 1\,\mathrm{Gyr}$. They are believed to trace a period of rapid change from star forming

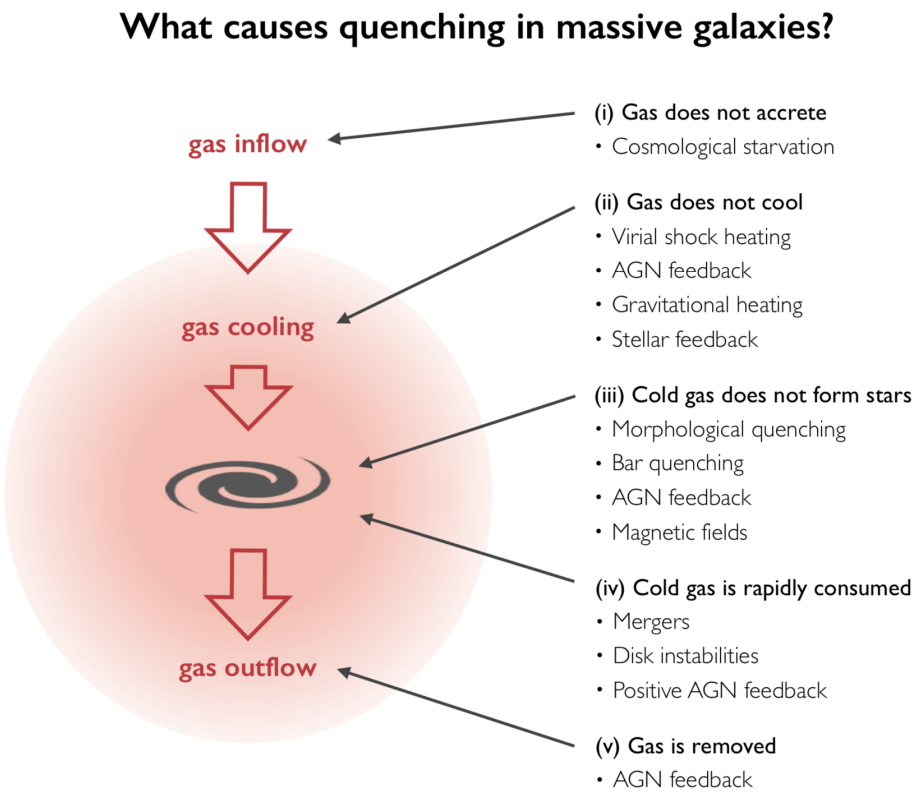

Figure 1.17 *The variety of quenching mechanisms proposed for massive galaxies, from Man & Belli (2018). (i) and (ii) are preventative mechanisms, whereby gas is inhibited from cooling onto the galaxy, either because the galaxy has become detached from the cosmic web, or because the gas that accretes is heated by virial shocks from the halo, or by feedback from stars or an AGN. In scenario (iii), gas is accreted onto the galaxy but is prevented from forming stars for reasons internal to the galaxy, for example, the stabilisation of the gas disk against fragmentation into star-forming clumps by a bulge or bar. In scenario (iv), a process such as a merger or disk instability triggers intense star formation that uses the bulk of the galaxy's reservoir of cold gas, preventing further star formation. In scenario (v), gas is removed from the galaxy via AGN-driven jets.*

.

galaxy to quiescence. Many of these galaxies have substantial reservoirs of molecular gas, as traced by the CO (1-0) line, which is sensitive to densities of $\sim 100\,\mathrm{cm}^{-3}$ (e.g. French et al., 2015; Alatalo et al., 2016). This implies that they are experiencing a period of extremely low star formation efficiency rather than a lack of fuel for star formation. French et al. (2018) probed the dense gas in two CO-luminous, post-starburst galaxies, via HCN $(1-0)$ and HCN^{+} $(1-0)$, which trace densities of $3\times10^{6}\,\mathrm{cm}^{-3}$ and $2\times10^{5}\,\mathrm{cm}^{-3}$, respectively. The low HCN/CO ratios they found imply that some mechanism is preventing molecular gas from collapsing into denser gas. French et al. (2015) showed that all but one of their sample of post-starburst galaxies lie in the low-ionisation nuclear emission-line (LINER) region of the BPT diagram. Together with the rapid decline in SFR during the post-starburst phase, which cannot be explained by stellar feedback, this points towards heating from AGN suppressing star formation.

Simulation work supports a quenching mechanism driven by AGN feedback; in the EAGLE simulations, a hot corona develops around massive galaxies, and star formation-driven outflows are unable to escape haloes (Bower et al., 2017). This leads to a buildup of gas density in the central regions of the galaxy, which, in turn, triggers a period of high rates of black hole accretion. This could also explain the morphological changes observed to go hand-in-hand with the quenching of star formation.

1.2.6 Modelling the relationships between galaxies and their host dark matter haloes

The physical and statistical connection between dark matter haloes and the galaxies that reside within them is a wide-ranging field, described more fully in the recent review of Wechsler & Tinker (2018). It has been recognised for decades that the spatial clustering of haloes is strongly dependent on halo mass (Bardeen et al., 1986; Mo & White, 1996), and hence measuring the clustering of populations of galaxies can tell us about their host haloes. Our understanding of galaxies and their relationships with their host haloes has been accelerated by the development of both cosmological simulations, in particular those that are able to resolve dark matter sub-structures, and large scale galaxy surveys spanning a range of redshifts.

A direct way of relating individual galaxies to their host haloes is via 'abundance matching' (Wechsler et al., 1998; Colin et al., 1999; Kravtsov & Klypin, 1999). Originally this technique was used to relate central galaxies to haloes, but now that substructure can be resolved within N-body simulations,

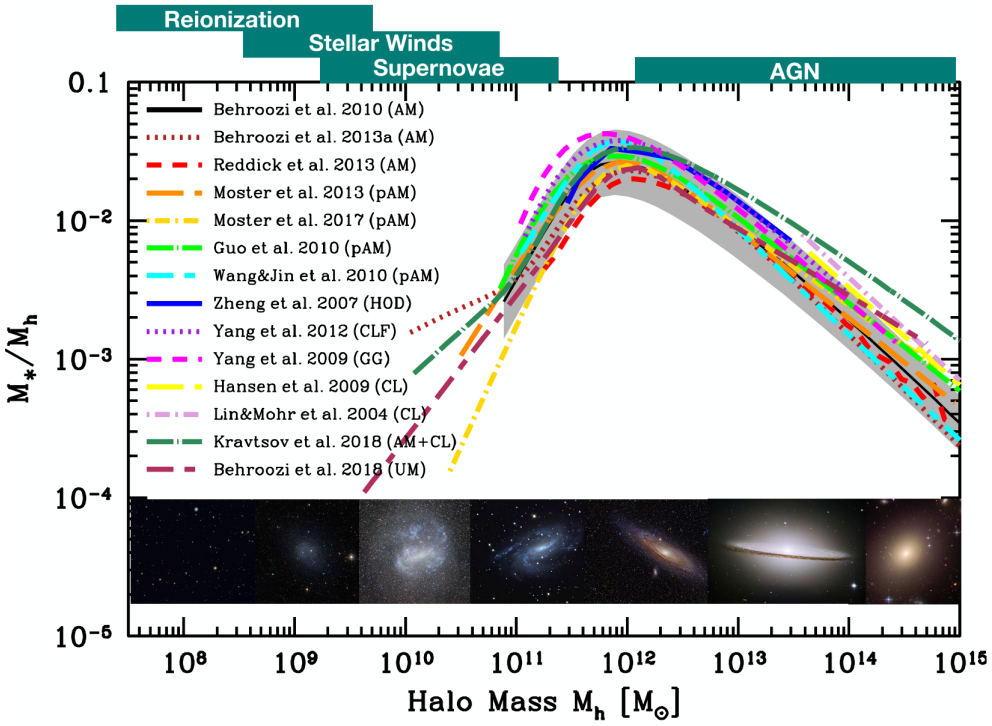

Figure 1.18 *The stellar-to-halo-mass relation (SHMR), adopted from the review of Wechsler & Tinker (2018). Determinations of the relation from abundance matching (AM), parametrised abundance matching (pAM), Halo Occupation Distribution (HOD) modelling, Conditional Luminosity Function (CLF), and from direct measurements of groups and clusters (GG) are shown in the legend. Below the relation, the typical morphologies of galaxies hosted by haloes of increasing masses are shown. Star formation is less efficient at low stellar masses due to stellar feedback, and at high stellar masses due to feedback from the central black hole.*

it has become possible to include satellite galaxies within this formalism. Dark matter haloes within the simulation are ranked by some property, usually mass (sometimes maximum circular velocity). Galaxies, ranked by stellar mass or luminosity, are then assigned to haloes, starting, for example, with the most massive galaxy being placed into the most massive halo. Ideally, galaxy properties are defined as those at the time of accretion onto the halo, since there can be a delay between the stripping of material from subhaloes and the satellite galaxies within them (Conroy et al., 2006). One advantage of abundance matching is that it is non-parametric; no functional form of halo occupation need be assumed. A form of abundance matching was used by Yang et al. (2005) to derive dark matter halo masses for galaxies in the SDSS and produce the group catalogue used in Chapter 5.

A key output of abundance matching has been the galaxy stellar-to-halo mass

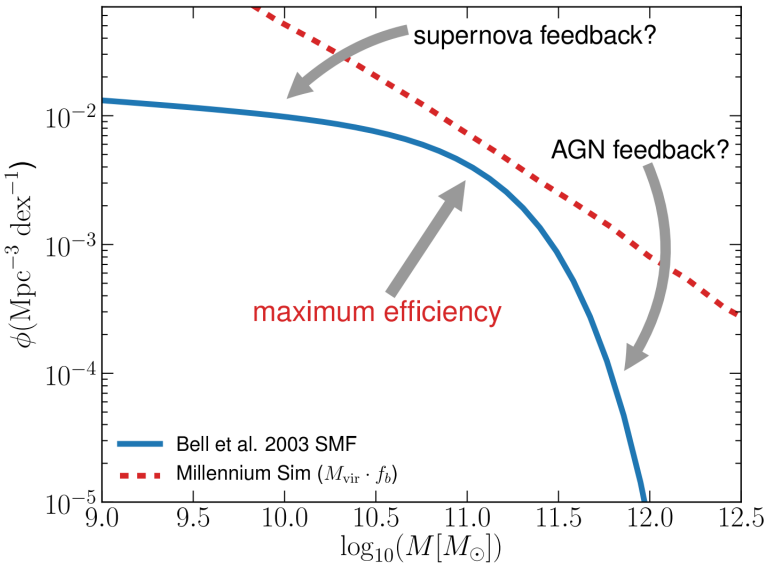

Figure 1.19 *The number density of galaxies and dark matter haloes, as a function of mass, from Mutch et al. (2013). The galaxy stellar mass function at $z = 0$ is taken from Bell et al. (2003) (many more recent determinations of the stellar mass function and its evolution exist; see for example Davidzon et al. 2017). The halo mass function from the Millennium simulation (Springel et al., 2005) is scaled by the Universal baryon fraction, to represent the total number of baryons available for star formation. The efficiency of conversion of baryonic mass to stellar mass within galaxies peaks at $M_* \sim 10^{11}\,M_\odot$.*

relation (SHMR; Figure 1.18). This can also also be inferred via parametrisation, populating an N-body model with galaxies, and then calibrating to well-constrained observational relations such as the stellar mass function (e.g. Moster et al., 2010, 2013; Behroozi et al., 2010, 2013). The SHMR is fundamentally related to the mismatch between the galaxy stellar mass function and the dark matter halo mass function (see Figure 1.19). The galaxy stellar mass function is shallower than the halo mass function at low stellar masses, and steeper at high stellar masses. Between these two regimes lies a region of maximum star formation efficiency; in haloes of $M \sim 10^{12}\,M_\odot$, stellar mass is formed and retained most efficiently (see Figure 1.18, though note that galaxy stellar mass still comprises just a few per cent of the halo mass). Imprinted upon the shape of the SHMR are the feedback processes that act to suppress star formation: stellar feedback, which affects primarily galaxies in haloes below the peak, and AGN feedback, which acts at the highest stellar/halo masses. Both feedback mechanisms are believed to prevent star formation via heating or ejection (see Section 1.2.5).

Hydrodynamical simulations provide the most physical way to model the

galaxy-halo connection. Gravity and hydrodynamics are combined within a cosmological context, with gas, stars and dark matter traced by particles at some mass resolution characteristic to the simulation. Cosmological hydrodynamical simulations such as EAGLE (Crain et al., 2015; Schaye et al., 2015; McAlpine et al., 2016) and Illustris (Genel et al., 2014; Vogelsberger et al., 2014a,b; Sijacki et al., 2015) model key physical processes on a range of scales, including cooling of gas onto haloes and galaxies, the feedback from stellar winds, supernovae and black holes. However, these simulations can be extraordinarily computationally expensive, and fundamentally limited in their predictive power by their resolution. Parametrisation at scales below the resolution of the simulation, known as 'sub-grid physics', can be tuned via calibration to observed relations such as the galaxy stellar mass function at $z = 0$. However, such calibrations limit physical insight gained, and can be restricted by availability of statistical observational samples and the instrumental systematics.

An alternative approach to modelling populations of galaxies is Halo Occupation Distribution (HOD) modelling (Berlind & Weinberg, 2002; Bullock et al., 2002). This is a statistical approach that involves parametrising the probability that a halo of given mass hosts a galaxy (with flexibility on the type of galaxy) and constraining the free parameters using observational data, which I use extensively in this thesis. A detailed discussion of HOD modelling is deferred to Chapter 2.

1.2.7 Our own galaxy

To understand the nature of the galaxies studied in this thesis, it is helpful to put them in context with our own galaxy. The Milky Way (MW) lies at the peak of the SHMR. It resides in a fairly low density environment, as one of the two most massive galaxies in the Local Group. Like the other, Andromeda, the MW is a spiral galaxy with neighbouring satellites and total mass $\sim 10^{12}\,M_{\odot}$. Recent estimates of the stellar mass of the MW consider its main structural components: the bulge and bar contribute $M_* = 0.91\pm0.07\times10^{10}\,M_{\odot}$, while the disk dominates at $M_* = 5.17\pm1.11\times10^{10}\,M_{\odot}$, giving total stellar mass $M_* = 6.08\pm1.14\times10^{10}\,M_{\odot}$ (Licquia & Newman, 2015). Its modest central black hole mass of $\sim 4\times10^{6}\,M_{\odot}$ (Lacy et al., 1980) places it below the M $-\,\sigma$ relation (Kormendy & Ho, 2013).

In line with its low density environment, the MW is not thought to have experienced a major merger within the last 10 Gyr (Stewart et al., 2008). Its star-formation rate has been low ($1 - 3\,M_{\odot}$/yr) for a number of Gyr (Snaith et al., 2014). The galaxies we study with HiZELS in Chapters 3 and 4 are high redshift MW analogues in many ways. The MW is an $L*$ galaxy, like many of

those we identify at higher redshift, with similar stellar mass. However due to the evolution of the luminosities of star-forming galaxies, $L*$ galaxies at high redshift are more highly star-forming than $L*$ galaxies at $z = 0$.

1.3 Structure of this thesis

In the following Chapter, I detail the process of characterising the clustering strength of observed samples of galaxies via the two-point correlation function. I then derive a flexible HOD parametrisation using simulations of galaxies and their host haloes, and show that this parametrisation is suitable for galaxy samples selected via a range of different criteria. In Chapter 3, I describe HiZELS, a narrow-band survey that provides samples of high redshift galaxies. I characterise the clustering and halo occupation of these samples, and of subsamples binned by Hα luminosity, using the clustering formalism laid out in Chapter 2. In Chapter 4, I extend this work to study clustering and halo environment as a function of both galaxy stellar mass and star formation rate. I also make comparisons to simulated galaxies drawn from the EAGLE simulations. In Chapter 5, I extend this work with EAGLE to study the drivers of star formation quenching. I then move from statistical studies of galaxy populations to spatially resolved studies of individual star-forming galaxies. In Chapter 6, I present a study of the multi-wavelength properties of galaxies drawn from the FIRE-2 simulations. I perform radiative transfer in post-processing on selected sub-mm bright snapshots, and study the sizes of the modelled dust continuum emission. In Chapter 7, I study the multi-wavelength emission from an observed galaxy at $z = 2.24$, selected from the HiZELS sample but subsequently targeted at high angular resolution using HST, SINFONI (on the VLT) and ALMA. I show that there are substantial discrepancies between the morphologies and SFRs derived using the different images. In Chapter 8, I summarise the main results of this thesis and outline future work.

I use a $H_0 = 70\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ cosmology throughout this thesis (see Appendix A for more details).

Chapter 2

Quantifying galaxy clustering

As described in the Introduction, galaxy environment has been identified as one of the key drivers of galaxy evolution. There are many methods used to characterise galaxy environments. In the local Universe, nearest-neighbour methods are often implemented. These characterise local overdensity via the projected distance to the n^{th} (often the 5^{th} or 10^{th}) nearest neighbour. Local overdensity measures provide an measure of the richness of the small-scale environment of individual galaxies, but are sensitive to the depth and completeness of the data and do not characterise large-scale environment. In this Chapter, I introduce techniques that characterise typical halo environments of populations of galaxies statistically via their observed clustering. My particular focus will be on the two-point correlation function, and how this can be fitted with HOD models to constrain the host dark matter halo masses and satellite fractions of populations of observed galaxies.

2.1 Angular two-point clustering statistics

Broadly, the two-point correlation function compares the clustering of an observed sample to a uniformly distributed random sample with the same areal coverage. It quantifies overdensities on a large range of scales; unlike nearest-neighbour estimators, it can yield insights into both the local environment within haloes and the large scale environment. When quantifying galaxy clustering, we construct correlation functions based on angular or projected distances between pairs of galaxies on the sky.

The angular two-point correlation function, $w(\theta)$, is a popular estimator of the clustering strength of galaxies. It is defined as the excess probability of finding a pair of galaxies separated by a given angular distance, relative to that probability

for a uniform (unclustered) distribution. The probability $dP(\theta)$ of finding objects in solid angles $d\Omega_1$ and $d\Omega_2$ separated by angular distance θ is:

$$dP(\theta) = N^2(1 + w(\theta))\, d\Omega_1 d\Omega_2, \tag{2.1}$$

where N is the surface density of objects.

Many estimators of $w(\theta)$ have been proposed. We use the minimum variance estimator proposed by Landy & Szalay (1993), which has been shown to be less susceptible to bias from small sample sizes and fields than other estimators:

$$w(\theta) = 1 + \left(\frac{N_R}{N_D}\right)^2 \frac{DD(\theta)}{RR(\theta)} - 2\frac{N_R}{N_D}\frac{DR(\theta)}{RR(\theta)}. \tag{2.2}$$

N_R and N_D are the total number of random and data galaxies in the sample, and $RR(\theta)$, $DD(\theta)$ and $DR(\theta)$ correspond to the number of random-random, data-data, and data-random pairs separated by angle θ. Empirically, it has been shown for a wide range of different galaxy samples that $w(\theta)$ can be fitted with a power law, $w(\theta) = A\theta^{\beta}$, where $\beta = -0.8$. Traditionally, Poissonian errors are used to estimate the uncertainties:

$$\Delta w(\theta) = \frac{1 + w(\theta)}{\sqrt{DD(\theta)}}. \tag{2.3}$$

However, these errors are underestimates (e.g. see Norberg et al., 2009), since they do not account for cosmic variance or correlations between adjacent θ bins. Using these errors also gives unjustifiably large weightings to the largest angular separations, where large DD pair counts result in very low $\Delta w(\theta)$.

Norberg et al. (2009) conclude that while no internal estimator reproduces the error of external estimators faithfully, jackknife and bootstrap resampling methods perform reasonably well, although both overestimate the errors. They note that jackknife resampling estimates the large-scale variance accurately but struggles on smaller scales ($\sim 2 - 3\,\mathrm{h}^{-1}\mathrm{Mpc}$), with the resulting bias strongly dependent on the number of sub volumes. Bootstrap resampling, meanwhile, overestimates the variance by approximately 50% on *all* scales, which may be minimised by oversampling the sub-volumes. In Chapters 3 and 4, we use the bootstrap resampling method with each correlation function constructed from 1000 bootstraps, taking the error on each $w(\theta)$ bin as the diagonal element of the bootstrap covariance matrix.

It can also be necessary to implement the integral constraint, IC, (Groth &

Peebles, 1977), a small correction to account for the underestimation of clustering strength due to the finite area surveyed.

$$\mathrm{IC} = \frac{\sum_\theta \mathrm{A}\theta^\beta \mathrm{RR}(\theta)}{\sum_\theta \mathrm{RR}(\theta)} \tag{2.4}$$

IC is small where fields are large. The HiZELS fields used in this thesis reach square-degree scales, and so IC corrections are largely negligible.

2.1.1 Obtaining a real-space correlation length

In order to compare the clustering strengths of populations of galaxies at different redshifts quantitatively, we convert the angular correlation function to a spatial one. This conversion is often performed using Limber's approximation (Limber, 1953), which assumes that spatial correlations which follow $\xi = (r/r_0)^\gamma$ are projected as angular correlation functions with slopes $\beta = \gamma + 1$. This results in the approximate relation between $\xi(r)$ and $w(\theta)$:

$$w(\theta) = \int_0^{+\infty} p_1(r)p_2(r)dr \int_{-\infty}^{+\infty} \xi(R, r)d\Delta r, \tag{2.5}$$

where $R = \sqrt{r^2\theta^2 + \Delta r^2}$, and $p_1(r), p_2(r)$ are the filter profiles for projected fields 1&2. Substituting $\xi = (r/r_0)^\gamma$ yields:

$$w(\theta) = r_0^\gamma \theta(\mathrm{rad})^{1-\gamma} \times \frac{\Gamma(\gamma/2 - 1/2)\Gamma(1/2)}{\Gamma(\gamma/2)} \times \int_0^{+\infty} p_1(r)p_2(r)r^{1-\gamma}dr, \tag{2.6}$$

where $\Gamma(x)$ is the gamma function. This is a good approximation for small angular scales, and can thus be used to evaluate r_0 from the fitted $w(\theta)$ profile. However, the integral diverges for narrow filters. Simon (2007) shows that in the limiting case of a delta function filter, the observed $w(\theta)$ is no longer a projection, but simply a rescaled $\xi_{\mathrm{gal}}(r)_0$ (thus $\beta = \gamma$ at large separations). Since Limber's approximation is not reliable for our samples of galaxies, which span fields with separations of degrees and use very narrow filters, we perform a numerical integration of the exact equation:

$$w_{\mathrm{model}}(\theta) = \psi^{-1} \int_0^{+\infty} \int_{s\sqrt{2\phi}}^{2s} \frac{2f_s(s - \Delta)f_s(s + \Delta)}{R^{-\gamma-1}r_0^\gamma \Delta} dRds. \tag{2.7}$$

Here, $\psi = 1 + \cos\theta$, $\phi = 1 - \cos\theta$, $\Delta = \sqrt{(R^2 - 2s^2\phi)/2\psi}$, and f_s is the profile of the filter, fitted as a Gaussian profile with μ and σ that depend on the filter being considered. We assume the standard value of $\gamma = -1.8$. χ^2 fitting of observed against modelled $w(\theta)$, generated using different r_0, allows us to estimate r_0 and its error (following Sobral et al., 2010).

Projected-space two-point clustering statistics

The clustering statistic required as input for the halo fitting routine we use in Section 3.4.1 is the projected-space (r_p) two-point correlation function, $w_p(r_p)$. We therefore transform our measured $w(\theta)$ to $w_p(r_p)$. $w_p(r_p)$ is defined by first considering the spatial two-point correlation function along the line of sight (r_l) and perpendicular to the line of sight (r_p):

$$\xi(r_p, r_l) = 1 + \left(\frac{N_R}{N_D}\right)^2 \frac{DD(r_p, r_l)}{RR(r_p, r_l)} - 2\frac{N_R}{N_D}\frac{DR(r_p, r_l)}{RR(r_p, r_l)}. \tag{2.8}$$

$\xi(r_p, r_l)$ is then integrated over r_l to obtain $w_p(r_p)$:

$$w_p(r_p) = 2\int_0^{r_{l,\max}} \xi(r_p, r_l) dr_l. \tag{2.9}$$

This is related to the real-space correlation function by:

$$w_p(r_p) = 2\int_{r_p}^{+\infty} \frac{r\xi(r)}{(r^2 - r_p^2)^{1/2}} dr \tag{2.10}$$

in the limit of a wide filter, and the solution tends to:

$$w_p(r_p) = r_p \left(\frac{r_p}{r_0}\right)^{-\gamma} \frac{\Gamma(\gamma/2 - 1/2)\Gamma(1/2)}{\Gamma(\gamma/2)}. \tag{2.11}$$

In the case of a narrower top-hat filter, we integrate over a finite range of r_l, using $(r_p^2 + r_{l,\max}^2)^{1/2}$ as the upper limit to the integral in Equation 2.10.

Throughout this thesis, we calculate $w_p(r_p)$ from our observed $w(\theta)$. However, our filter profiles are not top-hat (as assumed for the integral in Equation 2.10) but are better approximated by Gaussian profiles (see Table 3.3 for parameters). To account for this difference, we perform numerical integrations to determine

the factor by which $w(\theta)$ differs (for a given $\xi(r)$) if observed over a top-hat of width 2σ as opposed to a Gaussian of width σ (changing f_s in Equation 2.7); we find a required correction of $\sqrt{\pi}$. Using this, and combining Equations 2.2, 2.8, 2.9, with $r_p = \mathrm{D_{ang}}\theta(\mathrm{rad})$, where $\mathrm{D_{ang}}$ is the angular diameter distance, we then obtain:

$$w_p(r_p) \sim 2\sigma\sqrt{\pi}\, w\left(\theta = \frac{r_p}{\mathrm{D_{ang}}}\right)(1+z)^{0.8}. \tag{2.12}$$

2.2 How do galaxies populate dark matter haloes?

2.2.1 Modelling galaxy populations via Halo Occupation Distribution fitting

As described in the Introduction, the HOD formalism extends dark matter halo models to galaxy populations. The HOD describes the bias between galaxies and total mass by quantifying the average number of galaxies per dark matter halo as a function of halo mass. It is an important tool for linking the physics of galaxy evolution to the host halo environment, and frequently used to derive host halo masses and satellite fractions from observed galaxy two-point correlation functions for galaxies of different types.

Given a set of input parameters, an HOD allows us to predict the average number of galaxies of a certain type as a function of dark matter halo mass, $\langle N|M\rangle$. A combination of a cosmological model and an HOD enables us to predict any clustering statistic on any scale; usually observations of galaxy clustering (or weak lensing) are then used to constrain cosmological or galaxy evolution models. In Chapters 3 and 4, HOD modelling enables us to estimate typical host halo masses for HiZELS galaxies. We can also do better than the straight-line r_0 fit to the auto-correlation statistics; HOD fitting takes into account the small dip observed on angular scales of order 10s of arcseconds, below which the clustering is dominated by correlations between galaxies within a single dark matter halo, and above which clustering is dominated by clustering of galaxies in different haloes (see Figure 3.6). This enables us to include the effects of the satellite galaxy population on the observed clustering, no longer assuming that a power-law relationship holds on the smallest scales.

The reliability of this technique is highly dependent on the appropriate choice of an HOD parametrisation. Kravtsov et al. (2004) proposed that the overall HOD can be parametrised by two simple terms. The first describes the probability

that a dark matter halo of mass $M_{\rm halo}$ hosts a central galaxy above some stellar mass limit; this is well-approximated by a step function. Below some minimum halo mass, galaxies will not be found, since energy feedback from supernovae will simply expel baryons from very shallow potential wells, while above $M_{\rm min}$ all haloes host a galaxy. The second term describes the average number of satellite galaxies as a function of halo mass; empirically, this is well-fitted by a power law, for which a slope of unity appears to be appropriate for a wide range of simulated galaxy number densities and redshifts. Parametrisations of this form have been used fairly successfully for many years, for a variety of galaxy types and redshifts (e.g. Zehavi et al., 2005; Zheng et al., 2005, 2007; Tinker et al., 2010a; Zehavi et al., 2011; Wake et al., 2011; Durkalec et al., 2015). While these do well for stellar mass-selected samples (e.g. Wake et al. 2011; Hatfield et al. 2016), they may not be suitable for our samples of Hα-selected galaxies. As noted by Contreras et al. (2013), HODs for stellar mass-selected samples are very different to the HODs of SFR or cold gas mass-selected samples. In particular, HODs for mass-selected samples sensibly assume that above a given halo mass, all haloes contain a central galaxy. However, in not all cases does this central galaxy fall within an SFR or cold gas-selected sample (e.g. due to the suppression of gas cooling in high mass haloes via AGN feedback), so for SFR-limited samples the HOD for central galaxies may be peaked rather than a step function (Contreras et al., 2013).

Given the current availability of large samples of galaxies, increasingly samples can be split into stellar mass or SFR bins too. It is therefore very important to find appropriate HOD parametrisations for limited samples. Geach et al. (2012) argued that a two-component HOD model, composed of a Gaussian distribution at low halo masses and a step function, was more appropriate for centrals in SFR-limited samples, based on the output of *GALFORM* semi-analytic modelling (e.g. Cole et al., 2000a). Contreras et al. (2013) followed this with a detailed study of galaxies drawn from the Durham and Munich semi-analytic models. The halo occupation of galaxies selected above a limiting cold gas mass or SFR were better fitted by an asymmetric peak at low halo masses than by the traditional step function. We make a similar comparison to galaxies drawn from the EAGLE simulation in the following section.

2.2.2 Choosing an HOD parametrisation

Here, we briefly study the typical halo occupations of EAGLE galaxies, as a function of stellar mass and star formation rate. The great advantage of EAGLE

is that properties of the dark matter haloes and the galaxies within them are easily accessible, and so can provide the functional form of the HOD, for use in observational studies.

2.2.3 HOD functional forms from EAGLE

Figure 2.1 shows HODs for samples of EAGLE galaxies at $z = 0$, given different stellar mass and SFR cuts. The blue squares show the total (central & satellite) occupancy, and red circles show only the central occupancy. For mass-limited samples (Figure 2.1, first page, top row), the traditional smoothed step function appears a reasonable choice of parametrisation. Occupancy of centrals flattens at unity, as expected. However, for mass-binned samples (first page of Figure 2.1, bottom row), the red circles, which represent the HOD for central galaxies only, have a Gaussian-like form. This is very different to the canonical step-like function usually assumed in HOD fitting to mass-limited samples. Where samples are mass-incomplete, the central galaxy occupation does not rise and flatten at 1 at high halo masses, as for the mass-limited samples, because not all haloes contain a central galaxy within the chosen stellar mass range.

For star formation rate-limited samples (Figure 2.1, second page, top row), at the lowest SFR limits the HOD is similar to those of the mass-limited samples. For higher SFR cuts, the smoothed step-like function peaks below unity, since such samples do not include all the low star formation rate galaxies that fall into a mass-selected sample. HODs are different again for star formation rate-binned samples (bottom row). Here, although we see a peak in occupation at low halo masses, similar to the mass-binned samples, the HOD does not follow the Gaussian-like form above the peak. Instead, the occupation flattens at high halo masses, but at a value below unity. We thus urge caution in adopting standard forms of the HOD, and suggest that simulations such as EAGLE might be queried for specific classes of galaxies in order to obtain appropriate functional forms that may then be fitted to observed clustering measurements.

Here, we present functional form that appears to do well in fitting EAGLE galaxy HODs. This takes a flexible 6-parameter form, and is based on the parametrisation first presented by Geach et al. (2012)[1]. The numbers of central and satellite galaxies are parametrised as:

[1]In Geach et al. (2012), the factor of $\frac{1}{2}$ in the second term of the central galaxy parametrisation was excluded. We include it here, so that a halo can host a maximum of one (rather than two) central galaxies.

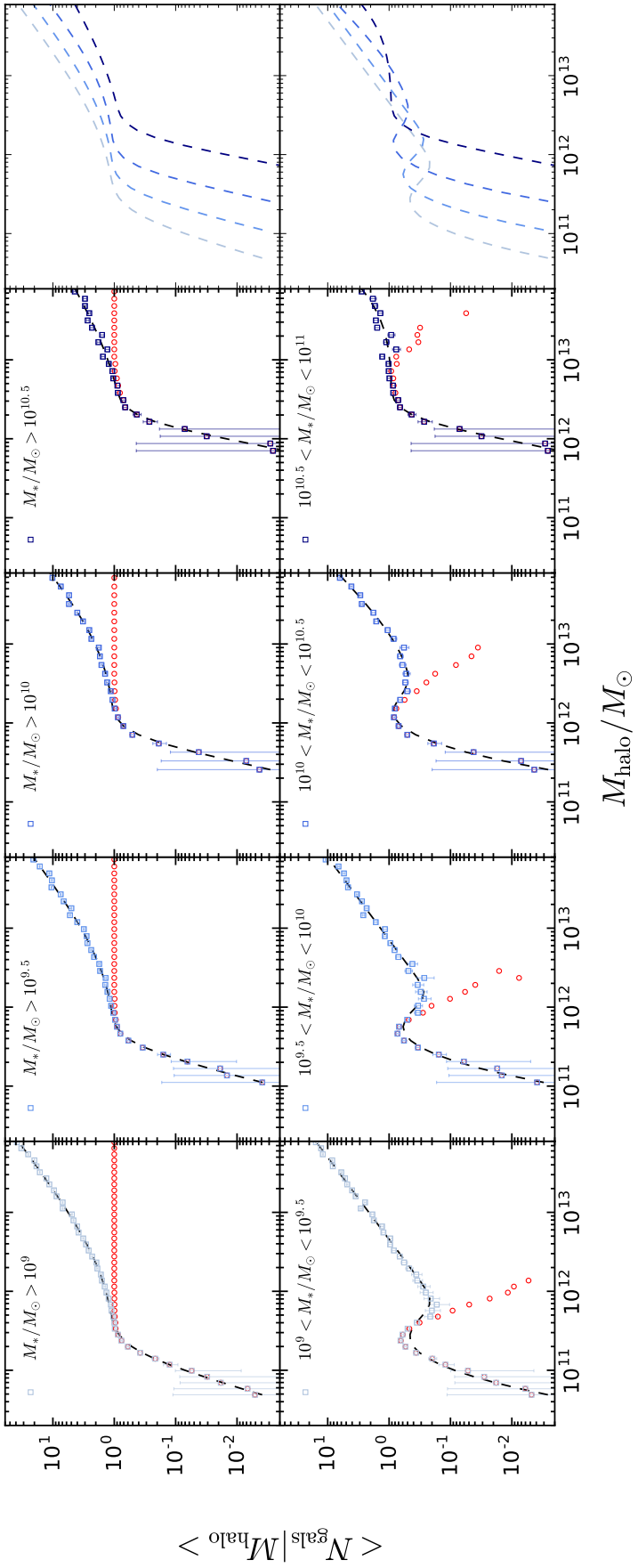

Figure 2.1 *(Figure continued on following page.)*

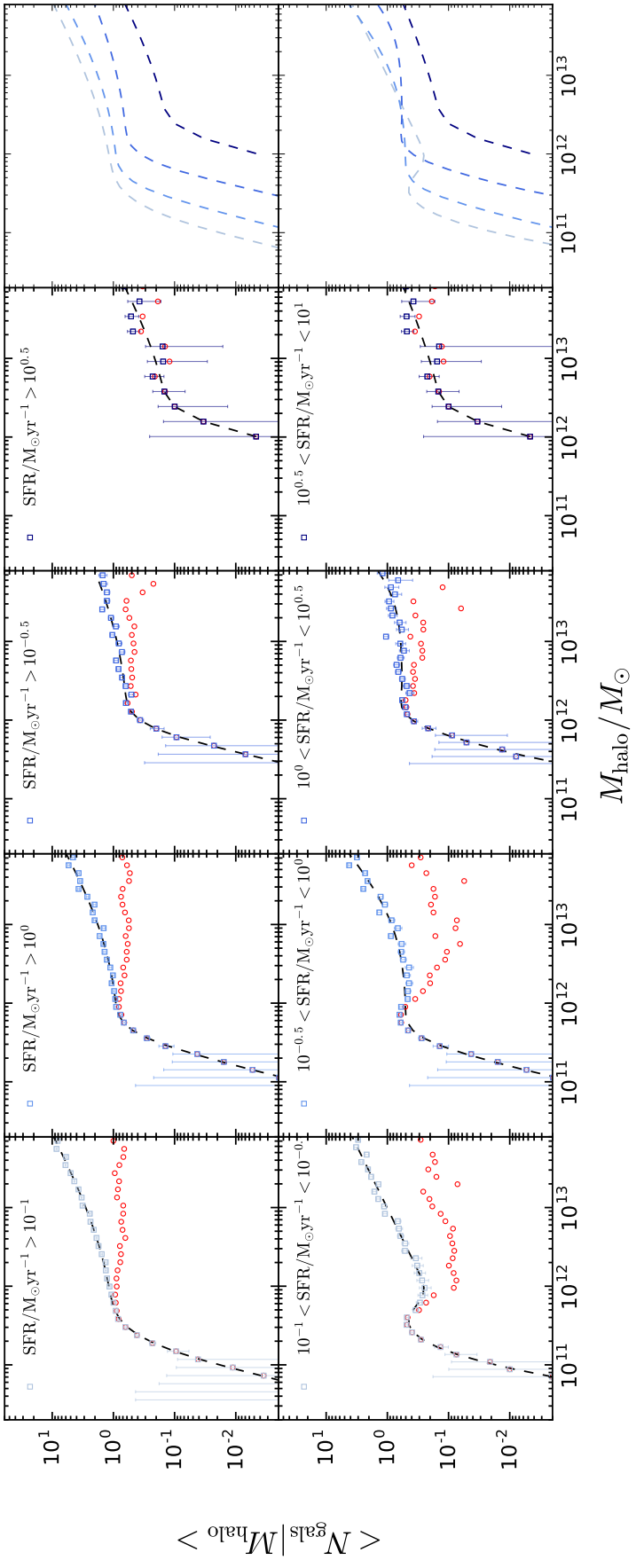

Figure 2.1 *(Caption on following page.)*

Figure 2.1 *Halo Occupation Distributions constructed using EAGLE galaxies at $z = 0.00$, with stellar mass and star formation rate cuts (either limits, or binned ranges) applied. The blue squares show the whole galaxy population (centrals and satellites) and the red circles show only central galaxies. The dashed lines show the best-fitting HOD, given the parametrisation presented in Section 2.2.3. It is encouraging that all samples (SFR and M_* selected; binned and limited) can be well-matched using the same 6-parameter functional form. The best-fitting HODs are shown together in a separate panel (right) to show the differences between the samples more clearly. In general, more massive and more highly star-forming galaxies occupy more massive dark matter haloes. Parameters for all of these fits are provided in Table 2.1.*

$$\langle N_{\rm cen}|{\rm M}\rangle = {\rm F_c^B}(1-{\rm F_c^A}){\rm exp}\left[-\frac{\log({\rm M}/{\rm M_{min}})^2}{2(\sigma_{\log {\rm M}})^2}\right] + \frac{1}{2}F_c^A\left[1+{\rm erf}\left(\frac{\log({\rm M}/{\rm M_{min}})}{\sigma_{\log {\rm M}}}\right)\right], \tag{2.13}$$

$$\langle N_{\rm sat}|{\rm M}\rangle = {\rm F_s}\left[1+{\rm erf}\left(\frac{\log({\rm M}/{\rm M_{min}})}{\sigma_{\log {\rm M}}}\right)\right]\left(\frac{\rm M}{\rm M_{min}}\right)^{\alpha}, \tag{2.14}$$

with the total number of galaxies given by:

$$\langle N|M\rangle = \langle N_{\rm cen}|M\rangle + \langle N_{\rm sat}|M\rangle. \tag{2.15}$$

Some implementations use $\langle N|M\rangle = \langle N_{\rm cen}|M\rangle[1+\langle N_{\rm sat}|M\rangle]$, requiring a central for every satellite galaxy. Given that the observational samples we use in Chapters 3 and 4 sample are essentially star formation rate limited, some of our galaxies could be star-forming satellites around less highly star-forming centrals that do not pass the Hα selection. Therefore we do not impose this condition.

The key parameters are:

- $M_{\rm min}$: the halo mass at which the probability of hosting a central galaxy peaks, and the minimum halo mass that hosts a satellite galaxy.
- $\sigma_{\log M}$: characterises the width of the transition to $\langle N_{\rm sat}|M\rangle = F_s\left(\frac{M}{M_{\rm min}}\right)^{\alpha}$ around $M_{\rm min}$.
- α: the slope of the power-law for $\langle N_{\rm sat}|M\rangle$ in haloes with $M > M_{\rm min}$.

– $F_c^{A,B}$: normalisation factors, in range [0,1].

– F_s: the mean number of satellite galaxies per halo, at $M = M_{\rm min}$

Using this parametrisation, we denote the best-fitting HODs for each subsample in Figure 2.1 by a dashed black line, and provide details of the parameter estimates in Table 2.1. From the successful fits shown, it is clear that this parametrisation is appropriate for a wide range of stellar mass and SFR-selected samples (both binned and limited). Where central galaxies occupy only the lower halo masses, there is a clear Gaussian component to the HOD but no step-function-like occupation at higher halo masses. This is the case for the lowest two stellar mass bins ($10^9 < M_*/M_\odot < 10^{9.5}$ and $10^{9.5} < M_*/M_\odot < 10^{10}$). Here, F_c^A, the step-function normalisation, becomes vanishingly small and F_c^B, which determines the contribution from the low-halo mass Gaussian component, dominates. For the stellar mass-limited samples, the contribution from F_c^A is close to unity, and that of F_c^B generally consistent with zero.

While the slope of the power-law occupancy of satellite galaxies, α, is well-approximated by unity for the mass-limited and mass-binned samples, this appears less suitable for the star formation rate-selected samples. For these, our fits favour a lower α, indicative of satellite quenching in high mass haloes, which removes galaxies from samples selected by star formation rate.

When fitting the models to data, we use the observed number density of galaxies as a constraint. For a given $\langle N|M\rangle$ output from the halo model, the predicted number density of galaxies is:

$$n_g = \int n(M)\langle N|M\rangle dM \tag{2.16}$$

where n(M) is the halo mass function. In this thesis, we use n(M) from Tinker et al. (2010b), and number densities from the luminosity functions of our samples, with a 10% error. For each set of HOD parameters, we may then derive a number of parameters of interest for galaxy evolution. The satellite fraction is:

$$f_{\rm sat} = \frac{1}{n_g}\int n(M)\langle N_{\rm sat}|M\rangle dM, \tag{2.17}$$

with the corresponding central fraction $f_{\rm cen} = 1 - f_{\rm sat}$.
The effective halo mass, the typical mass of galaxy host halo is:

$$M_{\rm eff} = \frac{1}{n_g}\int Mn(M)\langle N|M\rangle dM. \tag{2.18}$$

Stellar mass range	$\log_{10} M_{\rm min}$	$\sigma_{\log M}$	α	F_c^A	F_c^B	F_s
$M_*/M_\odot > 10^9$	11.32 ± 0.02	0.26 ± 0.02	0.97 ± 0.03	0.85 ± 0.07	0.5 ± 0.3	0.047 ± 0.008
$M_*/M_\odot > 10^{9.5}$	11.59 ± 0.03	0.22 ± 0.03	0.98 ± 0.05	0.84 ± 0.09	0.5 ± 0.5	0.06 ± 0.01
$M_*/M_\odot > 10^{10}$	11.90 ± 0.03	0.20 ± 0.05	1.05 ± 0.14	0.90 ± 0.18	0.5 ± 0.9	0.04 ± 0.02
$M_*/M_\odot > 10^{10.5}$	12.30 ± 0.05	0.18 ± 0.06	0.8 ± 0.3	0.5 ± 0.4	0.06 ± 0.11	0.10 ± 0.11
$10^9 < M_*/M_\odot < 10^{9.5}$	11.40 ± 0.03	0.22 ± 0.01	0.96 ± 0.04	0.00 ± 0.00	0.43 ± 0.04	0.03 ± 0.01
$10^{9.5} < M_*/M_\odot < 10^{10}$	11.72 ± 0.03	0.21 ± 0.01	0.95 ± 0.09	0.00 ± 0.05	0.55 ± 0.05	0.04 ± 0.02
$10^{10} < M_*/M_\odot < 10^{10.5}$	12.07 ± 0.05	0.20 ± 0.01	1.27 ± 0.23	0.32 ± 0.15	0.97 ± 0.16	0.016 ± 0.015
$10^{10.5} < M_*/M_\odot < 10^{11}$	12.33 ± 0.03	0.19 ± 0.05	1.5 ± 0.6	0.90 ± 0.15	0.28 ± 0.52	0.004 ± 0.009
SFR range	$\log_{10} M_{\rm min}$	$\sigma_{\log M}$	α	F_c^A	F_c^B	F_s
$\mathrm{SFR/M_\odot yr^{-1}} > 10^{-1}$	11.43 ± 0.04	0.28 ± 0.03	0.70 ± 0.08	0.78 ± 0.18	0.06 ± 0.05	0.08 ± 0.03
$\mathrm{SFR/M_\odot yr^{-1}} > 10^{-0.5}$	11.65 ± 0.02	0.22 ± 0.01	0.76 ± 0.07	0.72 ± 0.08	0.21 ± 0.11	0.05 ± 0.02
$\mathrm{SFR/M_\odot yr^{-1}} > 10^{0}$	11.96 ± 0.03	0.21 ± 0.04	0.67 ± 0.22	0.50 ± 0.15	0.06 ± 0.07	0.04 ± 0.04
$\mathrm{SFR/M_\odot yr^{-1}} > 10^{0.5}$	12.31 ± 0.15	0.21 ± 0.23	0.74 ± 0.81	0.09 ± 0.21	0.01 ± 0.07	0.018 ± 0.06
$10^{-1} < \mathrm{SFR/M_\odot yr^{-1}} < 10^{-0.5}$	11.50 ± 0.02	0.20 ± 0.01	0.63 ± 0.06	0.00 ± 0.00	0.39 ± 0.03	0.057 ± 0.018
$10^{-0.5} < \mathrm{SFR/M_\odot yr^{-1}} < 10^{0}$	11.58 ± 0.04	0.20 ± 0.01	1.11 ± 0.14	0.47 ± 0.05	0.09 ± 0.08	0.005 ± 0.003
$10^{0} < \mathrm{SFR/M_\odot yr^{-1}} < 10^{0.5}$	12.00 ± 0.09	0.19 ± 0.01	1.5 ± 0.5	0.57 ± 0.06	0.25 ± 0.29	0.0006 ± 0.0016
$10^{0.5} < \mathrm{SFR/M_\odot yr^{-1}} < 10^{1}$	12.30 ± 0.16	0.20 ± 0.25	0.60 ± 0.13	0.07 ± 0.3	0.01 ± 0.07	0.03 ± 0.09

Table 2.1 *Parameters of the halo occupation distribution model detailed in Section 2.2.3, fitted to EAGLE galaxies selected by different stellar mass and star formation rate criteria. Figure 2.1 shows that our chosen parametrisation is flexible enough to provide good fits to HODs of very differently defined samples.*

The average effective bias factor, which characterises the clustering of galaxies relative to dark matter, is:

$$b_{\rm eff} = \frac{1}{n_g} \int n(M) b(M) \langle N|M \rangle dM, \tag{2.19}$$

where $b(M)$ is the halo mass dependent halo bias (we use Tinker et al. 2010b).

Chapter 3

The $\mathrm{H}\alpha$ luminosity-dependent clustering of star-forming galaxies from $\mathrm{z} \sim 0.8$ to $\mathrm{z} \sim 2.2$ with HiZELS

This chapter contains work that has been published as Cochrane R. K., Best P. N., Sobral D., Smail I., Wake D. A., Stott J. P., Geach E., 2017, Monthly Notices of the Royal Astronomical Society, Volume 469, Issue 3, p.2913-2932.

In this Chapter, we present clustering analyses of identically-selected star-forming galaxies in 3 narrow redshift slices (at $z = 0.8$, $z = 1.47$ and $z = 2.23$), from HiZELS, a deep, near-infrared narrow-band survey. The HiZELS samples span the peak in the cosmic star formation rate density, identifying typical star-forming galaxies at each epoch through their Hα line emission. We quantify the clustering of samples at the three redshifts, and of Hα luminosity-selected subsamples, initially using simple power law fits to the two-point correlation function. We extend this work to link the evolution of star-forming galaxies and their host dark matter haloes over cosmic time using sophisticated dark matter halo models. We find that the clustering strength, r_0, and the bias of galaxy populations relative to the clustering of dark matter increase linearly with Hα luminosity (and, by implication, star formation rate) at all three redshifts, as do the host dark matter halo masses of the HiZELS galaxies. The typical galaxies in our samples are star-forming centrals, residing in haloes of mass $M_{\rm halo} \sim$ a few times $10^{12}\,M_{\odot}$. We find a remarkably tight redshift-independent relation between the Hα luminosity

scaled by the characteristic luminosity, $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}(z)$, and the minimum host dark matter halo mass of central galaxies. This reveals that the dark matter halo environment is a strong driver of galaxy star formation rate and therefore of the evolution of the star formation rate density in the Universe.

3.1 Introduction

The galaxies we observe exist in a wide range of environments, from rich clusters to underdense void regions. They are thought to trace an underlying distribution of dark matter, with more highly clustered galaxies occupying massive dark matter overdensities (Zwicky, 1933; Peebles, 1982). As outlined in Chapter 1, this is commonly explained via the paradigm of hierarchical growth: weak density fluctuations in an expanding, homogeneous Universe are amplified by gravitational instabilities, with smaller structures forming first. Galaxies form due to the collapse of baryonic matter under the gravity of dark matter haloes (White & Frenk, 1991), with the progenitors of the most massive clusters starting to form earliest. Dark matter haloes assemble via successive mergers and accretion of small haloes, which naturally leads to the formation of galaxy groups and clusters, with a single dark matter halo capable of hosting many galaxies.

While the observed 'cosmic web' spatial distribution of dark matter in the Lambda Cold Dark Matter (ΛCDM) paradigm can be successfully modelled using N-body simulations (Davis et al., 1985) as advanced as the Millennium Simulation (Springel et al., 2005), resolution is limited and the evolution of galaxies within this web is harder to model. This complexity reflects the additional baryonic processes present: we must consider not only the underlying distribution of dark matter but also the non-linear physics of galaxy formation and evolution. Key processes such as gas cooling, star formation, and the physics of feedback due to star formation and black hole accretion all act on different timescales with different galaxy mass and environment dependencies. A wealth of observational data is required to fine-tune parameters in galaxy simulations.

The latest galaxy surveys at both low and high redshifts has provoked a flurry of recent work aiming to understand the relationships between stellar mass, star formation rate and environment (e.g. Peng et al., 2010; Sobral et al., 2011; Scoville et al., 2013; Darvish et al., 2015). Both mass and environment are associated with transformations in colour, star formation rate and morphology, popularly known as 'quenching'. As described in Chapter 1, Peng et al. (2010) proposed that two primary quenching mechanisms, 'mass quenching' and 'environment quenching',

act independently and dominate at different epochs and galaxy masses. There is evidence that the trends observed with galaxy mass and environment at low redshift hold to at least $z \sim 1$. At $z \sim 1$, Sobral et al. (2011) and Muzzin et al. (2012) both find that the fraction of galaxies that are star-forming decreases once we reach group densities and at high galaxy masses. However, things become less clear at even higher redshifts. Scoville et al. (2013) find a flattening in the relationship between environmental overdensity and both star-forming fraction and star formation rate above $z \sim 1.2$ for galaxies in the COSMOS field, and note that this flattening holds out to their highest redshift galaxies at $z \sim 3$. Other studies have found an apparent reversal of the low-z star formation rate (or morphology)-density relation at higher redshifts (Butcher & Oemler, 1978). Both Sobral et al. (2011) and Elbaz et al. (2007) find that at $z \sim 1$, median galaxy star formation rates increase with overdensity until cluster densities are reached, at which point star formation rates decrease with overdensity, as in the local universe. Attempting to explain these opposing trends, McGee et al. (2009) propose that the pressure of the intra-cluster medium on infalling galaxies in the outskirts of galaxy clusters actually compresses gas and enhances star formation prior to stripping in the denser environment of the cluster core. Increased galaxy-galaxy interactions may also trigger intense star formation via the disruption of gas disks. At high redshifts, high gas fractions (e.g. Tacconi et al., 2010) permit more efficient starburst responses. Thus, at high redshifts, the richest environments may provide the combination of large gas reservoirs and ICM pressures which fuel high star formation rates and later lead to quenching via gas exhaustion and stripping (Smail et al., 2014).

Quantifying the environmental dependence of star formation activity at high redshift directly is inherently challenging. One approach to studying this is through auto-correlation functions of star-forming galaxies, as discussed in Chapter 2. The dark matter correlation function is the inverse Fourier transform of the dark matter power spectrum. Observing the projected real-space galaxy correlation function, which is a linear scaling of the dark matter correlation function, provides a natural connection between galaxies and the underlying matter distribution which determines their large-scale environments. As described in Chapter 2, the HOD framework provides a powerful technique for characterising dark matter halo masses. A consistent picture has emerged in which more luminous and more massive star-forming galaxies tend to be more strongly clustered, as a result of lying preferentially in high mass dark matter haloes. This holds at both low redshifts (e.g. Norberg et al., 2001; Zehavi et al.,

2011) and at high redshifts (e.g. Sobral et al., 2010; Wake et al., 2011; Geach et al., 2012; Hatfield et al., 2016).

In this Chapter we build upon the work presented in Sobral et al. (2010), which studied the clustering of $\sim$ 700 Hα emitters at $z = 0.84$ from HiZELS. Narrow band Hα surveys such as HiZELS select only those galaxies with emission lines within a very narrow redshift range ($\Delta z \sim 0.02$), and with a well-defined redshift distribution. For clustering measurements, these types of survey are therefore superior to photometric ones, which are often hampered by systematic uncertainties and require a more complex treatment of the spatial distribution in the clustering analysis. Furthermore, unlike many spectroscopic surveys, the narrow band approach provides a clean selection function down to a known flux (star formation rate) limit. Sobral et al. (2010) found evidence for a strong Hα luminosity dependence of the clustering strength of Hα emitters at $z = 0.84$, along with evidence for a single relation with $L_{\rm H\alpha}/L^{*}_{\rm H\alpha}$ from $z \sim 0.2$ to $z \sim 2.2$. Geach et al. (2008, 2012) supplemented this work with the first analyses of the clustering of HiZELS galaxies at $z = 2.23$, though the sample was not sufficiently large to permit binning by luminosity.

Here we analyse a larger sample of $\sim$ 3000 emitters at $z = 0.8$ spanning three fields: COSMOS, UDS and SA22. Crucially, we also use larger samples of Hα emitters at $z = 2.23$, and include new data at $z = 1.47$ (Sobral et al., 2012, 2013a). Our samples, which span large ranges in Hα luminosity and redshift, provide optimal data for revealing the drivers of galaxy evolution over cosmic time. We provide details of the HiZELS sample selection in Section 3.2. In Section 3.3, we present the results of power-law fits to two-point correlation functions. Given the high quality of the correlation functions obtained, we extend these analyses to incorporate a sophisticated HOD modelling treatment. In Section 3.4, we present derived halo properties for our HiZELS galaxies, in particular typical halo masses and galaxy central/satellite fractions. We discuss the implications of these results in Section 3.5.

3.2 The HiZELS survey and sample selection

3.2.1 Sample of Hα emitters

HiZELS (Geach et al. 2008; Sobral et al. 2009, 2012, 2013a) used the United Kingdom Infra-Red Telescope (UKIRT)'s Wide Field CAMera (WFCAM), the Subaru Telescope's Suprime-Cam with the NB921 filter, the Very Large Telescope

Field	$\overline{z}_{\mathrm{H}\alpha\,\mathrm{emitters}}$	No. emitters	Area (deg^2)
NBJ COSMOS & UDS	0.845 ± 0.011	503	1.6
LOW0H2 SA22	0.81 ± 0.011	2332	7.6
NBH COSMOS & UDS	1.47 ± 0.016	451	2.3
NBK COSMOS & UDS	2.23 ± 0.016	727	2.3

Table 3.1 *Numbers and mean redshifts of Hα emitters identified by the HiZELS survey and selected for this analysis. HiZELS uses standard and custom-made narrow-band filters, complemented by broad-band imaging, over well-studied fields. Only emitters which exceed the limiting flux, f_{50}, of their frames are included.*

(VLT)'s HAWK-I camera and the Canada France Hawaii Telescope (CFHT) with MegaCam (CFHiZELS; Sobral et al. 2015) to detect line emitters over large areas within well-studied fields. We present only a brief overview of the survey here, deferring to Sobral et al. (2013a) for a full description of the HiZELS COSMOS and UDS data, and to Sobral et al. (2015) for details of the SA22 CFHiZELS campaign.

HiZELS uses standard and custom-made narrow-band (NB) filters, complemented by broad-band (BB) imaging. Sources identified by the narrow-band filters are matched to those in the broad-band images by using the same aperture size and a search radius of 0.9". True emitters are selected based on their NB-BB colour excess, with a signal-to-noise cut of $S/N > 3$ and an equivalent width selection corresponding to EW > 25 Å for Hα. High quality photometric redshifts derived from data spanning from optical to mid-IR wavelengths (e.g. Cirasuolo et al. 2010; Ilbert et al. 2009; Lawrence et al. 2007) were used to identify which emission line is being selected for each emitter, and thus select a clean sample of Hα emitters. This technique enables the identical selection of Hα emitting galaxies at $z = 0.81, 0.84$ (NBJ: COSMOS, UDS, SA22), $z = 1.47$ (NBH: COSMOS, UDS) and $z = 2.23$ (NBK: COSMOS, UDS); see Table 3.1 for details. Spectroscopic redshifts confirmed that the large sample of galaxies we obtain lies within well-defined redshift ranges (see also Sobral et al. 2016b; Stott et al. 2016).

Hα fluxes are corrected for contamination by the adjacent [NII]λ6548, 6584 lines within the NB filter using the relationship between log([NII]/Hα) and EW$_0$([NII]+Hα) derived by Sobral et al. (2013a) and confirmed spectroscopically in Sobral et al. (2015). They are also corrected for dust attenuation assuming $\mathrm{A}_{\mathrm{H}\alpha} = 1.0$ mag (Garn et al., 2010; Ibar et al., 2013). The median combined correction is 0.307 dex at $z = 0.8$, 0.325 dex at $z = 1.47$ and 0.335 dex at $z = 2.23$.

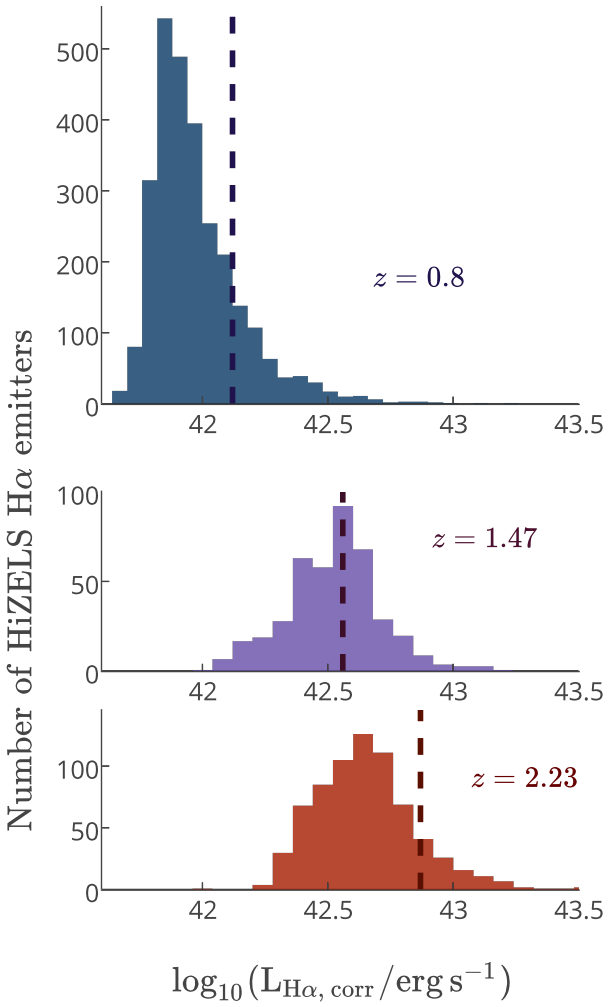

Figure 3.1 *Distribution of dust-corrected* Hα *luminosities of HiZELS emission line-selected galaxies in our samples at the three epochs. Vertical dashed lines show the characteristic luminosity,* $L^*_{\mathrm{H}\alpha}$*, at each redshift. HiZELS galaxies span a large luminosity range at each epoch, probing well below* $L^*_{\mathrm{H}\alpha}$*.*

The distribution of dust-corrected Hα luminosities is shown in Figure 3.1.

3.2.2 Generating random samples

We generated unclustered random samples in order to quantify the clustering of the observed Hα emitters. Variations in coverage and observing conditions have resulted in individual HiZELS frames having different depths, meaning that robustly-constructed random samples are essential to differentiate between true clustering and that introduced by the observing strategy. In this section we describe the construction of random samples which reflect these depths.

Most simply, random sources may be generated by calculating a limiting flux at which each frame is essentially 100% complete, drawing sources from the

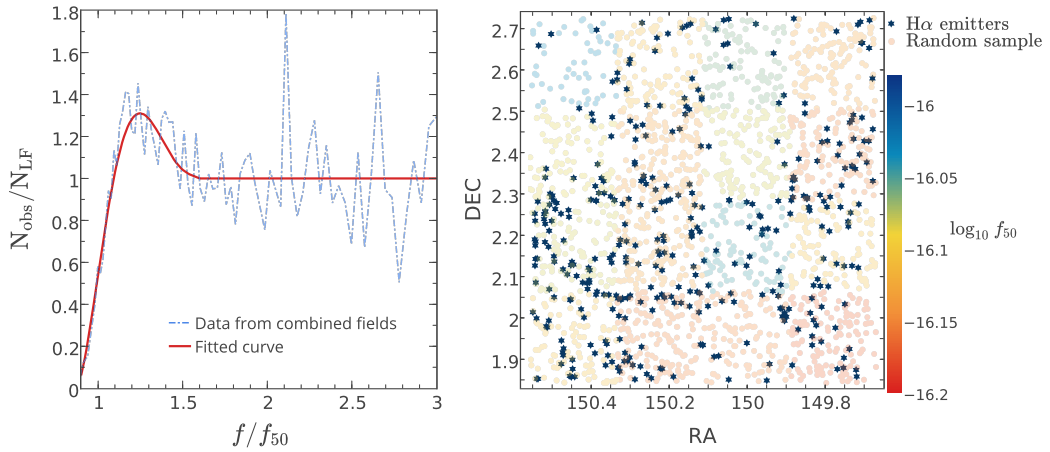

Figure 3.2 *Left: the completeness curve used to place sources in frames with flux limit f_{50}. We account for a small number of excess sources due to flux boosting around the detection limit. Right: example of random sources in the COSMOS field, colour coded by the limiting flux of their frame, with real sources shown by stars overlaid. Fluxes are given in units of* $\mathrm{erg\,s^{-1}cm^{-2}}$.

luminosity function down to this flux, and distributing these randomly across the frame. For this analysis we aim to push further in flux, so as to include as many sources as possible. We include sources down to luminosities corresponding to the 50% completeness flux, f_{50}, as calculated by Sobral et al. (2013a, 2015) for each frame using Monte Carlo simulations. To study source detection as a function of the limiting flux (taking account of both incompleteness and flux boosting biases), we have calculated the ratio of the number of sources recovered, $N_{\rm obs}$, to the number of sources expected from the luminosity function, $N_{\rm LF}$, as a function of f_{50} in each frame. We found a small boost in the number of sources with recovered fluxes around the flux limit, suggesting that flux-boosting effects dominate over incompleteness. We tested different filters, and both deep and shallow fields separately, and found that all show the same general form. We have therefore fitted a single empirically-derived effective completeness curve (Figure 3.2, left) and taken this into account when generating the random catalogues. Numerous tests have confirmed that our results are qualitatively unchanged if the random sources are simply drawn from the luminosity function down to f_{50} or constructed using a slightly different completeness curve.

In this Chapter, we use luminosity functions of the standard Schechter (1976) form:

$$\phi(L)dL = \phi^* \left(\frac{L}{L^*}\right)^{\alpha} e^{-(L/L^*)} d\left(\frac{L}{L^*}\right). \tag{3.1}$$

z	$L^*_{\mathrm{H}\alpha}(\mathrm{erg\,s^{-1}})$	$\phi^*(\mathrm{Mpc^{-3}})$	α
0.810 & 0.845	$42.12^{+0.03}_{-0.02}$	$-2.31^{+0.04}_{-0.05}$	$-1.6^{+0.2}_{-0.2}$
1.466	$42.56^{+0.06}_{-0.05}$	$-2.61^{+0.08}_{-0.09}$	$-1.62^{+0.25}_{-0.29}$
2.231	$42.87^{+0.08}_{-0.06}$	$-2.78^{+0.08}_{-0.09}$	$-1.59^{+0.12}_{-0.13}$

Table 3.2 *LF parameters used in this Chapter, derived in Sobral et al. (2013a, 2015). At $z \sim 0.8$, we use the Schechter function fit to the much larger $z = 0.81$ sample by Sobral et al. (2015), which is more accurate than that presented by Sobral et al. (2013a) and is also a good fit for the $z = 0.84$ data.*

Here, L^* represents the characteristic luminosity 'break' of the LF, ϕ^* is the corresponding characteristic comoving space density, and α is the 'faint-end' slope of the power law, dominant at low luminosities. The parameters we adopt, given in Table 3.2, were derived using the samples of Hα emitters from Sobral et al. (2013a, 2015). We generated a random position for each random source, carefully taking into account the boundaries of each frame and the masked regions due to bright stars and artefacts. The final number of sources generated within a frame depends on both its unmasked chip area and its depth. All random samples are substantially larger (e.g. 1000×) than the real samples. When constructing correlation functions for samples binned by flux, we also require knowledge of the fluxes of the random sources, to account for faint sources being preferentially detected in the deepest frames. The fluxes of random sources are drawn from the luminosity functions given in Table 3.2, scaled by the fitted completeness curve (Figure 3.2) for a given f_{50}. We have also incorporated average corrections for dust and [NII] emission line contamination. We did not include any real or random sources with flux $f < f_{50}$ in this analysis.

3.2.3 Effects of potential contaminants

Here we discuss three classes of possible contaminants: sources that are not true emitters; true emitters that are different lines misclassified as Hα; and AGN interlopers. As discussed in Section 3.2.1, HiZELS emitters are selected based on their NB-BB colour excess, with a signal-to-noise cut of $S/N > 3$. To check the possibility of including false emitters, we have repeated the clustering measurements using a more conservative cut of $S/N > 4$ for various luminosity bins. We find no significant differences in the clustering strengths. We also note that the exclusion of sources with fluxes below their frame's f_{50} serves to remove some potential low-flux contaminants. Contamination from misclassified lines is

also estimated to be small, at $\sim 5\%$, as estimated by Sobral et al. (2013a). Such contaminants will generally have the effect of a small decrease in $w(\theta)$, with much smaller effects than our observed trends.

Our sample could suffer from contamination from AGN, for which Hα emission is not a reliable tracer of star formation rate. Using extensive multi-wavelength data to identify AGN candidates within HiZELS samples in the COSMOS and UDS fields, Garn & Best (2010) estimate an AGN fraction of $\sim 10\%$, but Sobral et al. (2016a) find that this can be much higher at very high Hα luminosities. We expect that the effect of AGN contamination may only be very important in the highest luminosity bins. However, these bins show no evidence of deviation from the linear trend of the low-luminosity regime (see Section 3.3.2). Given that it is difficult to exclude these individual sources from our analyses, we present all results using Hα luminosity rather than converting to star formation rate explicitly. We invoke star formation rate only in our gas-regulator model interpretation in Section 3.5.2.

3.3 Quantifying galaxy clustering using the two-point correlation function

We use the two-point angular correlation function, the formalism for which is laid out in detail in Chapter 2.

3.3.1 Power-law fits to angular correlation functions

We have derived angular correlation functions for large samples of Hα emitters at each redshift and fitted these with power-law models (see Figure 3.3), using the filter profiles detailed in Table 3.3. The exact luminosity ranges of these samples, given in Table 3.4, are chosen to compare similar samples at each redshift, and span the same range in $L_{\rm H\alpha}/L^*_{\rm H\alpha}$: $-0.4 < \log_{10}(L_{\rm H\alpha}/L^*_{\rm H\alpha}) < 0.3$ (albeit with non-matched distributions within this range). We use a power-law of

Redshift	$\mu(h^{-1}\mathrm{Mpc})$	$\sigma(h^{-1}\mathrm{Mpc})$
0.81 ± 0.011	1970	14
1.47 ± 0.016	3010	18
2.23 ± 0.016	3847	18

Table 3.3 *Parameters of Gaussian filter profile fits for the three HiZELS redshifts studied.*

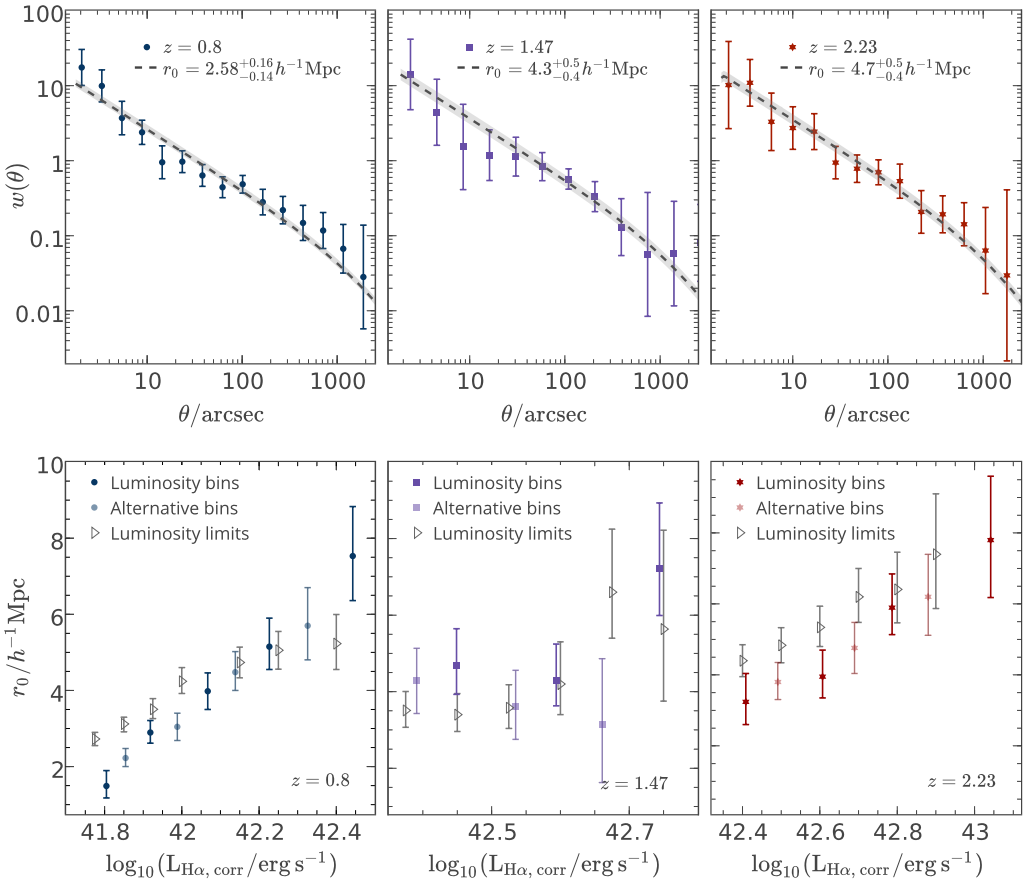

Figure 3.3 *Top: power law fits (with the correction to Limber's approximation at large scales) to the measured angular correlation functions at three redshifts, each over the range* $-0.4 < \log_{10}(L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}) < 0.3$. *Bottom: derived clustering strength,* r_0, *for* Hα *luminosity-binned and luminosity-limited samples. We also show alternative binning. The plotted luminosity value is the mean value of* $\log_{10}(L_{\mathrm{H}\alpha})$ *for the luminosity-binned samples, and the lower limit for the luminosity-limited samples. The clustering strength increases with* $\log_{10} L_{\mathrm{H}\alpha}$ *for all three redshifts surveyed in a broadly linear manner.*

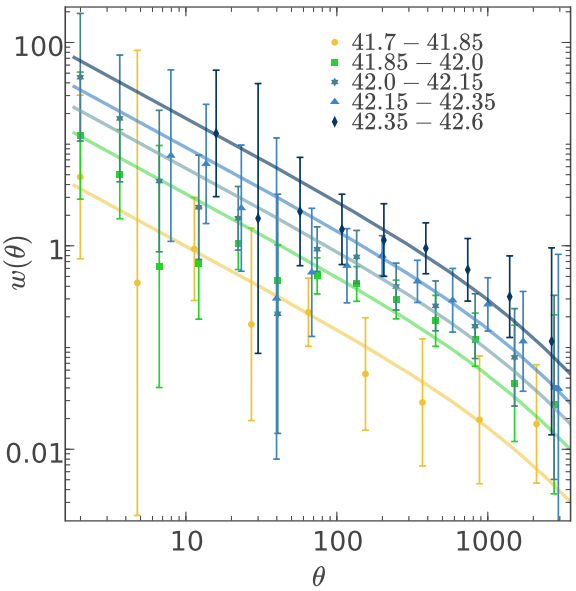

Figure 3.4 *Examples of five correlation functions of luminosity-binned samples at $z = 0.8$, with $\log_{10} L_{\mathrm{H}\alpha}$ selection shown in the key. Although the correlation functions are not as high quality as those of the whole samples shown in Figure 3.3), it is clear that the clustering strength (obtained from the amplitude of the correlation function) increases with* Hα *luminosity.*

fixed gradient -1.8 for the spatial correlation function, leading to a slope of -0.8 in the angular correlation function on small scales and the correction to Limber's approximation at large scales where the angular separation is much greater than the separation along the line of sight. This parametrisation is sufficient to derive indicative clustering strengths. However, the correlation functions of all three samples do show clear departures from the traditional power-law relation fitted here. At angular scales of order 10s of arcseconds the power-law fit consistently overestimates the observed $w(\theta)$, indicative of a dominant contribution from a separate 1-halo term at small angular separations. We explore this further in Section 3.4.

3.3.2 Clustering strength as a function of galaxy Hα luminosity

We have fitted both luminosity-binned data and luminosity-limited data with the same power-law models (see Table 3.4). As shown in the lower panels of Figure 3.3 and in Figure 3.4, the clustering strength, r_0, increases roughly linearly with galaxy Hα luminosity for the luminosity-binned samples. More highly star-forming galaxies are more strongly clustered, and hence may live in more massive

Redshift	$\log_{10}(L_{\rm H\alpha}$ range/erg s^{-1})	Mean $\log_{10}(L_{\rm H\alpha})$	r$_0$/h^{-1}Mpc	b$_{\rm eff}$	$\log_{10}(M_{\rm eff}/M_\odot)$	$\log_{10}(M_{\rm min}/M_\odot)$	$f_{\rm sat}$
'Full' samples: $-0.4 < \log_{10}(L_{\rm H\alpha}/L^*_{\rm H\alpha}) < 0.3$							
0.8	41.72-42.42	41.96	$2.6^{+0.2}_{-0.1}$	$1.12^{+0.06}_{-0.05}$	$12.13^{+0.10}_{-0.09}$	$11.12^{+0.11}_{-0.15}$	$0.05^{+0.01}_{-0.01}$
1.47	42.16-42.86	42.52	$4.3^{+0.5}_{-0.4}$	$1.78^{+0.06}_{-0.08}$	$12.16^{+0.07}_{-0.09}$	$11.45^{+0.06}_{-0.08}$	$0.05^{+0.02}_{-0.02}$
2.23	42.47-43.17	42.71	$4.7^{+0.5}_{-0.4}$	$2.52^{+0.07}_{-0.09}$	$11.96^{+0.05}_{-0.07}$	$11.41^{+0.06}_{-0.06}$	$0.05^{+0.02}_{-0.02}$
Hα **Luminosity-selected subsamples**							
	Bins						
0.8	41.7-41.85	41.80	$1.5^{+0.4}_{-0.3}$	$1.07^{+0.09}_{-0.03}$	$11.92^{+0.14}_{-0.04}$	$11.26^{+0.13}_{-0.08}$	$0.03^{+0.01}_{-0.01}$
0.8	41.775-41.925	41.85	$2.2^{+0.3}_{-0.2}$	$1.16^{+0.06}_{-0.03}$	$12.01^{+0.14}_{-0.06}$	$11.46^{+0.08}_{-0.07}$	$0.02^{+0.01}_{-0.01}$
0.8	41.85-42.0	41.92	$2.9^{+0.3}_{-0.3}$	$1.33^{+0.07}_{-0.06}$	$12.31^{+0.14}_{-0.11}$	$11.69^{+0.07}_{-0.08}$	$0.02^{+0.01}_{-0.01}$
0.8	41.925-42.075	41.99	$3.1^{+0.4}_{-0.4}$	$1.32^{+0.09}_{-0.07}$	$12.30^{+0.16}_{-0.12}$	$11.71^{+0.10}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
0.8	42.0-42.15	42.07	$4.0^{+0.5}_{-0.5}$	$1.49^{+0.08}_{-0.09}$	$12.55^{+0.11}_{-0.14}$	$11.91^{+0.08}_{-0.12}$	$0.02^{+0.01}_{-0.01}$
0.8	42.075-42.25	42.14	$4.5^{+0.5}_{-0.5}$	$1.56^{+0.07}_{-0.09}$	$12.63^{+0.08}_{-0.12}$	$12.01^{+0.07}_{-0.09}$	$0.02^{+0.01}_{-0.01}$
0.8	42.15-42.35	42.23	$5.2^{+0.8}_{-0.6}$	$1.63^{+0.08}_{-0.11}$	$12.71^{+0.09}_{-0.13}$	$12.09^{+0.07}_{-0.12}$	$0.03^{+0.02}_{-0.01}$
0.8	42.25-42.475	42.33	$5.7^{+1.0}_{-0.9}$	$1.79^{+0.08}_{-0.12}$	$12.86^{+0.07}_{-0.12}$	$12.28^{+0.07}_{-0.11}$	$0.03^{+0.02}_{-0.01}$
0.8	42.35-42.6	42.44	$7.5^{+1.3}_{-1.2}$	$2.02^{+0.08}_{-0.13}$	$13.05^{+0.06}_{-0.10}$	$12.50^{+0.06}_{-0.09}$	$0.05^{+0.03}_{-0.02}$
	Limits						
0.8	> 41.775	41.99	$2.7^{+0.2}_{-0.2}$	$1.18^{+0.06}_{-0.06}$	$12.18^{+0.10}_{-0.11}$	$11.24^{+0.10}_{-0.14}$	$0.04^{+0.01}_{-0.01}$
0.8	> 41.85	42.03	$3.1^{+0.2}_{-0.2}$	$1.26^{+0.05}_{-0.06}$	$12.28^{+0.08}_{-0.11}$	$11.40^{+0.08}_{-0.11}$	$0.04^{+0.01}_{-0.01}$
0.8	> 41.925	42.10	$3.5^{+0.3}_{-0.3}$	$1.32^{+0.05}_{-0.07}$	$12.38^{+0.08}_{-0.11}$	$11.54^{+0.07}_{-0.10}$	$0.04^{+0.01}_{-0.01}$
0.8	> 42.0	42.18	$4.2^{+0.4}_{-0.3}$	$1.41^{+0.05}_{-0.07}$	$12.52^{+0.07}_{-0.10}$	$11.69^{+0.07}_{-0.09}$	$0.05^{+0.02}_{-0.02}$
0.8	> 42.075	42.25	$4.7^{+0.4}_{-0.4}$	$1.48^{+0.05}_{-0.07}$	$12.57^{+0.06}_{-0.09}$	$11.84^{+0.07}_{-0.09}$	$0.03^{+0.02}_{-0.01}$
0.8	> 42.15	42.33	$5.1^{+0.5}_{-0.5}$	$1.57^{+0.06}_{-0.08}$	$12.66^{+0.07}_{-0.10}$	$11.98^{+0.07}_{-0.09}$	$0.03^{+0.02}_{-0.01}$
0.8	> 42.25	42.44	$5.2^{+0.8}_{-0.7}$	$1.71^{+0.07}_{-0.11}$	$12.79^{+0.07}_{-0.12}$	$12.18^{+0.07}_{-0.10}$	$0.03^{+0.02}_{-0.01}$
0.8	> 42.4	42.57	$5.8^{+1.5}_{-1.3}$	$2.00^{+0.11}_{-0.17}$	$13.03^{+0.09}_{-0.14}$	$12.53^{+0.07}_{-0.13}$	$0.03^{+0.02}_{-0.01}$
	Bins						
1.47	42.3-42.45	42.39	$4.3^{+0.9}_{-0.8}$	$2.12^{+0.09}_{-0.13}$	$12.4^{+0.07}_{-0.12}$	$11.88^{+0.06}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
1.47	42.375-42.525	42.45	$4.7^{+1.0}_{-0.9}$	$2.21^{+0.09}_{-0.13}$	$12.48^{+0.06}_{-0.11}$	$11.97^{+0.06}_{-0.09}$	$0.02^{+0.01}_{-0.01}$
1.47	42.45-42.6	42.53	$3.6^{+1.0}_{-0.7}$	$2.12^{+0.16}_{-0.15}$	$12.39^{+0.14}_{-0.14}$	$11.97^{+0.09}_{-0.11}$	$0.01^{+0.0}_{-0.0}$
1.47	42.525-42.675	42.59	$4.3^{+1.0}_{-0.9}$	$2.21^{+0.20}_{-0.18}$	$12.46^{+0.15}_{-0.15}$	$12.04^{+0.10}_{-0.16}$	$0.03^{+0.01}_{-0.01}$
1.47	42.6-42.75	42.66	$3.2^{+1.7}_{-1.2}$	$2.28^{+0.22}_{-0.18}$	$12.50^{+0.16}_{-0.15}$	$12.12^{+0.11}_{-0.13}$	$0.02^{+0.01}_{-0.01}$
1.47	42.675-42.85	42.74	$7.2^{+1.7}_{-1.5}$	$2.72^{+0.13}_{-0.20}$	$12.79^{+0.07}_{-0.11}$	$12.34^{+0.07}_{-0.12}$	$0.07^{+0.06}_{-0.04}$
1.47	42.75-43.3	42.87	$6.8^{+2.6}_{-2.2}$	$2.67^{+0.16}_{-0.26}$	$12.76^{+0.09}_{-0.16}$	$12.33^{+0.07}_{-0.15}$	$0.04^{+0.03}_{-0.02}$

Redshift	$\log_{10}(L_{\mathrm{H}\alpha}$ range/erg s^{-1})	Mean $\log_{10}(L_{\mathrm{H}\alpha})$	r_0/h^{-1}Mpc	b_{eff}	$\log_{10}(M_{\mathrm{eff}}/M_\odot)$	$\log_{10}(M_{\mathrm{min}}/M_\odot)$	f_{sat}
	Limits						
1.47	> 42.2	42.55	$3.7^{+0.5}_{-0.4}$	$1.85^{+0.07}_{-0.11}$	$12.18^{+0.08}_{-0.12}$	$11.58^{+0.06}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
1.47	> 42.375	42.59	$3.5^{+0.5}_{-0.4}$	$1.90^{+0.09}_{-0.12}$	$12.22^{+0.09}_{-0.13}$	$11.67^{+0.07}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
1.47	> 42.45	42.63	$3.4^{+0.6}_{-0.4}$	$1.90^{+0.16}_{-0.13}$	$12.21^{+0.16}_{-0.15}$	$11.71^{+0.09}_{-0.13}$	$0.02^{+0.01}_{-0.01}$
1.47	> 42.525	42.66	$3.6^{+0.6}_{-0.5}$	$2.07^{+0.12}_{-0.14}$	$12.36^{+0.11}_{-0.13}$	$11.88^{+0.08}_{-0.10}$	$0.02^{+0.01}_{-0.01}$
1.47	> 42.6	42.73	$4.2^{+1.1}_{-0.8}$	$2.22^{+0.14}_{-0.17}$	$12.47^{+0.11}_{-0.14}$	$12.02^{+0.08}_{-0.12}$	$0.02^{+0.01}_{-0.01}$
1.47	> 42.675	42.80	$6.6^{+1.7}_{-1.2}$	$2.56^{+0.12}_{-0.18}$	$12.71^{+0.07}_{-0.11}$	$12.22^{+0.07}_{-0.11}$	$0.07^{+0.06}_{-0.04}$
1.47	> 42.75	42.87	$5.6^{+2.6}_{-1.9}$	$2.69^{+0.14}_{-0.25}$	$12.77^{+0.08}_{-0.15}$	$12.34^{+0.07}_{-0.14}$	$0.03^{+0.03}_{-0.02}$
	Bins						
2.23	42.2-42.5	42.41	$3.2^{+0.8}_{-0.6}$	$2.30^{+0.11}_{-0.17}$	$11.79^{+0.09}_{-0.15}$	$11.25^{+0.07}_{-0.13}$	$0.03^{+0.02}_{-0.01}$
2.23	42.35-42.6	42.49	$3.8^{+0.6}_{-0.5}$	$2.50^{+0.10}_{-0.15}$	$11.93^{+0.07}_{-0.12}$	$11.43^{+0.06}_{-0.09}$	$0.02^{+0.01}_{-0.01}$
2.23	42.5-42.7	42.61	$4.0^{+0.8}_{-0.6}$	$2.67^{+0.13}_{-0.20}$	$12.03^{+0.09}_{-0.14}$	$11.58^{+0.07}_{-0.12}$	$0.02^{+0.01}_{-0.01}$
2.23	42.6-42.8	42.69	$4.8^{+0.7}_{-0.7}$	$2.87^{+0.09}_{-0.16}$	$12.14^{+0.06}_{-0.10}$	$11.70^{+0.05}_{-0.09}$	$0.02^{+0.01}_{-0.01}$
2.23	42.7-42.9	42.79	$5.9^{+1.0}_{-0.8}$	$3.05^{+0.09}_{-0.14}$	$12.24^{+0.05}_{-0.08}$	$11.81^{+0.05}_{-0.07}$	$0.02^{+0.01}_{-0.01}$
2.23	42.8-43.0	42.88	$6.2^{+1.2}_{-1.1}$	$3.23^{+0.10}_{-0.16}$	$12.33^{+0.05}_{-0.08}$	$11.92^{+0.05}_{-0.07}$	$0.03^{+0.02}_{-0.01}$
2.23	42.9-43.6	43.04	$7.8^{+1.8}_{-1.6}$	$3.23^{+0.11}_{-0.19}$	$12.35^{+0.05}_{-0.09}$	$11.93^{+0.05}_{-0.08}$	$0.05^{+0.03}_{-0.03}$
	Limits						
2.23	> 42.2	42.66	$4.3^{+0.5}_{-0.4}$	$2.23^{+0.07}_{-0.09}$	$11.77^{+0.06}_{-0.08}$	$11.12^{+0.06}_{-0.07}$	$0.07^{+0.02}_{-0.02}$
2.23	> 42.3	42.66	$4.3^{+0.5}_{-0.4}$	$2.29^{+0.07}_{-0.09}$	$11.80^{+0.05}_{-0.08}$	$11.21^{+0.06}_{-0.07}$	$0.04^{+0.02}_{-0.01}$
2.23	> 42.4	42.69	$4.4^{+0.5}_{-0.5}$	$2.39^{+0.07}_{-0.11}$	$11.86^{+0.06}_{-0.10}$	$11.30^{+0.06}_{-0.07}$	$0.04^{+0.01}_{-0.01}$
2.23	> 42.5	42.73	$4.8^{+0.5}_{-0.5}$	$2.52^{+0.07}_{-0.10}$	$11.95^{+0.05}_{-0.08}$	$11.42^{+0.05}_{-0.07}$	$0.04^{+0.02}_{-0.01}$
2.23	> 42.6	42.78	$5.3^{+0.6}_{-0.5}$	$2.65^{+0.08}_{-0.12}$	$12.03^{+0.05}_{-0.09}$	$11.53^{+0.05}_{-0.07}$	$0.04^{+0.01}_{-0.01}$
2.23	> 42.7	42.87	$6.2^{+0.8}_{-0.7}$	$2.83^{+0.08}_{-0.12}$	$12.14^{+0.05}_{-0.07}$	$11.66^{+0.05}_{-0.07}$	$0.04^{+0.02}_{-0.02}$
2.23	> 42.8	42.95	$6.4^{+1.1}_{-1.0}$	$3.03^{+0.09}_{-0.15}$	$12.24^{+0.05}_{-0.09}$	$11.79^{+0.05}_{-0.07}$	$0.04^{+0.02}_{-0.02}$
2.23	> 42.9	43.06	$7.4^{+1.7}_{-1.5}$	$3.26^{+0.12}_{-0.19}$	$12.35^{+0.06}_{-0.10}$	$11.93^{+0.05}_{-0.08}$	$0.05^{+0.04}_{-0.03}$

Table 3.4 *r_0 values and key parameters derived from HOD fitting, for samples of* Hα *emitters at different redshifts and luminosities. We find a clear trend towards increasing r_0, b_{eff}, M_{min} and M_{eff} for samples of galaxies with higher* Hα *luminosities at all redshifts, but little evidence for changing satellite fractions for these SFR-selected samples.*

dark matter halo environments. The trends are similar for the luminosity-limited samples: these also show an increase in clustering strength with galaxy luminosity. The results for the two sample types do not agree exactly because luminosity-limited samples of galaxies with faint limits have their clustering increased by the inclusion of a small number of bright sources, and therefore have a greater clustering strength than that of galaxies entirely within a faint luminosity bin.

Although the absolute values of r_0 agree (within errors) with the previous HiZELS study of a smaller sample of Hα emitters at $z = 0.8$, the apparently linear relationship is at odds with the results of Sobral et al. (2010), who found tentative hints of a more step-like behaviour around the characteristic luminosity. With our much larger sample of ~ 3000 emitters, there is no longer evidence for a break in the r_0 vs $\log_{10}(L_{\mathrm{H}\alpha})$ relationship, and a linear relation provides a far better fit. The trends at $z = 1.47$ and $z = 2.23$ also show no clear departure from a simple linear trend, albeit that the $z \sim 1.47$ results are noisier. These results are also broadly consistent with previous studies. We find $r_0 = 4.3^{+0.5}_{-0.4}\,h^{-1}$Mpc for our sample at $z = 1.47$, while Kashino et al. (2017) obtain $r_0 = 5.2 \pm 0.7\,h^{-1}$Mpc for H$\alpha$ emitters at $1.43 \leq z \leq 1.74$. We find $r_0 = 4.7^{+0.5}_{-0.4}\,h^{-1}$Mpc for the full sample at $z = 2.23$, which is slightly higher than Geach et al. (2012) found using a smaller sample at the same redshift ($r_0 = 3.7 \pm 0.3\,h^{-1}$Mpc), but this depends critically on the luminosity range studied.

In Figure 3.5, we show the Hα luminosity-dependent clustering of $z = 0.8$ HiZELS emitters split into two observed K-band magnitude bins. Observed K-band magnitude is believed to be a rough proxy for galaxy stellar mass. We find that the clustering strength increases broadly linearly with $\log_{10}(L_{\mathrm{H}\alpha})$ within each of the broad K-band magnitude bins, and that this trend is much larger than any differences between the two K-band magnitude bins. We will explore the stellar mass-dependence of the clustering of star-forming galaxies more thoroughly in Chapter 4, but we stress here that the strong trends of clustering strength with Hα luminosity presented in this Chapter are not driven primarily by galaxy stellar mass.

3.4 Physical parameters from HOD modelling

3.4.1 Fitting HOD models to HiZELS Hα-emitting galaxies

We use the HMF (Murray et al., 2013) and HALOMOD codes (Murray, 2015) to fit HOD models to the correlation functions using the HOD parametrisation

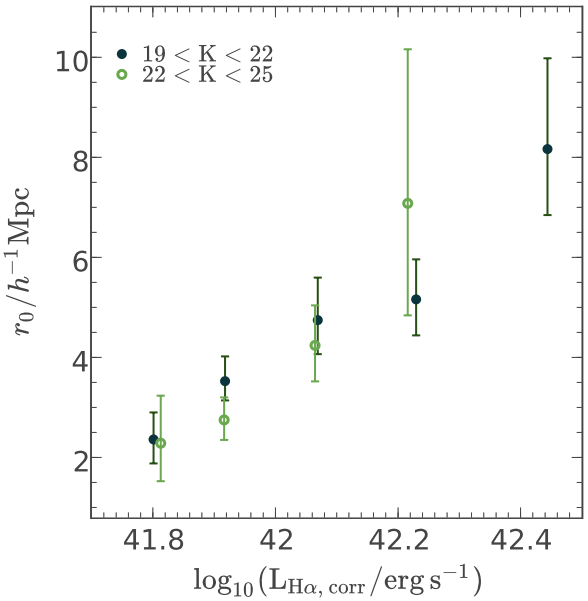

Figure 3.5 *To investigate whether trends with $L_{H\alpha}$ are driven by stellar mass, we plot* r_0 *against $L_{H\alpha}$ for observed K-band magnitude-selected subsamples of the $z = 0.8$ HiZELS emitters. We find that the strong trends of clustering strength with* Hα *luminosity hold for these subsamples. This indicates that trends with $L_{H\alpha}$ are not driven primarily by stellar mass.*

developed in Chapter 2. These take an HOD parametrisation and construct real-space correlation functions for a range of parameter inputs. For each set of parameter inputs, we compare the projection of the modelled real-space correlation function with that observed, and calculate the log likelihood. We use *emcee* (Foreman-Mackey et al., 2013), a fast python implementation of an affine-invariant Markov Chain Monte Carlo (MCMC) ensemble sampler, to sample the parameter space and optimise the fit to the correlation function. Like many studies, we fix $\alpha = 1$, since the low satellite fractions prevent strong constraints. As discussed in Chapter 2, we fit the number density of galaxies in the log-likelihood fitting as a further constraint. We use 500 walkers, each with 1000 steps.

We present examples of the best-fit modelled correlation function and its HOD occupation, decomposed into the central and satellite galaxy terms, in Figure 3.6. The parametrisation, shown here for a correlation function constructed using the full sample of galaxies at $z = 0.8$, provides a good fit to the data, and clearly shows the separate contributions of the clustering within a single halo and between dark matter haloes. In Figure 3.7 we show an example of the MCMC output of this fit. While the five fitted HOD parameters are highly correlated, we can still constrain the derived parameters and obtain good fits to the correlation functions.

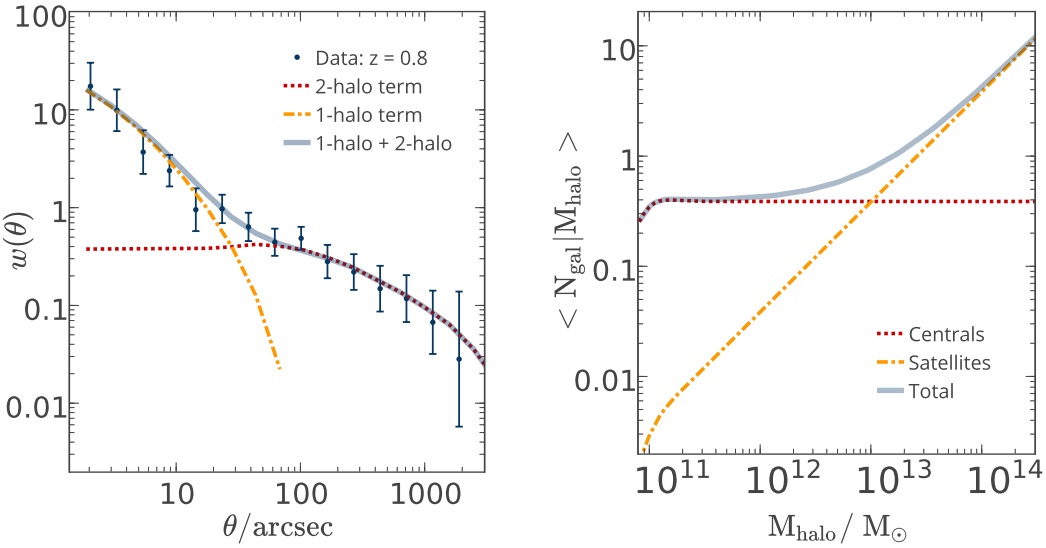

Figure 3.6 *Left: Halo Occupation Model fit to the correlation function of the whole $z = 0.8$ sample using HALOMOD. This multi-parameter model provides a better fit to data than the single power law model and shows the separate contributions of satellite and central galaxies. Right: the best-fitting halo occupation distribution model. The contribution from satellite galaxies becomes significant only in haloes more massive than $\sim 10^{13}\,M_{\odot}$.*

For each correlation function to which an HOD model is fitted, we estimate $M_{\rm min}$ as well as the following derived parameters: $f_{\rm sat}$, $M_{\rm eff}$, $b_{\rm eff}$. We take the 50th, 16th and 84th percentiles of the posterior distribution of each of these derived parameters, to obtain an estimate of the median and associated 1σ errors. The individual HOD input parameters $\sigma_{\log M}$, $F_c^{A,B}$ and F_s, tend to be individually less well constrained due to correlations between them. In Table 3.5, we present the five HOD parameters fitted to the correlation functions of large samples of galaxies within a fixed $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ range at each redshift.

The selection of galaxies within a fixed $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ range, as in Section 3.3.1 (see Table 3.4), allows the comparison of similar galaxies across cosmic time. Interestingly, the derived galaxy occupations as a function of halo mass are similar, consistent within their errors (see Figure 3.8). Although the $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ distributions are not exactly the same across the different redshift ranges, we deduce from this that samples of galaxies selected from HiZELS at similar $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ trace similar dark matter haloes across redshift. Intrigued by this, we compare galaxies within narrower $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ bins in Section 3.4.2.

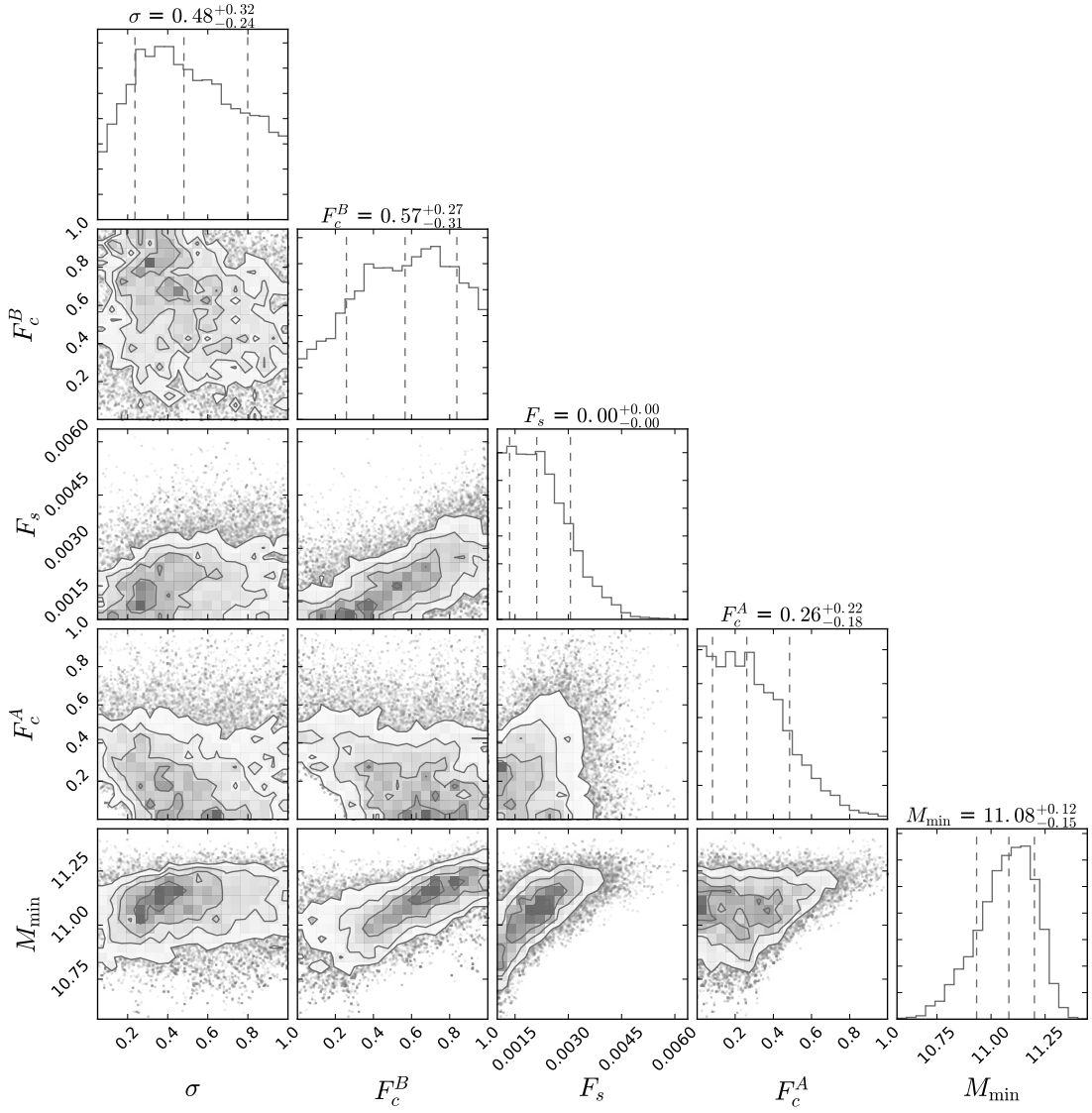

Figure 3.7 *An example of the output from the MCMC HOD fit to the two-point correlation function (Foreman-Mackey, 2016), constructed using the 'full' sample of galaxies at* $z = 0.8$. *The five fitted parameters are highly correlated, but we obtain good constraints on the derived parameters,* $\mathrm{b_{eff}}$, $M_{\rm eff}$, $M_{\rm min}$ *and* $\mathrm{f_{sat}}$.

Redshift	$\log_{10}(M_{\rm min}/M_{\odot})$	F_s	F_c^A	F_c^B	σ
	[10, 13.0, 11.5]	[0.001, 1.0, 0.01]	[0.001, 1.0, 0.9]	[0.001, 1.0, 0.4]	[0.05, 1.0, 0.5]
0.8	$11.08^{+0.12}_{-0.15}$	$0.002^{+0.001}_{-0.001}$	$0.3^{+0.2}_{-0.3}$	$0.6^{+0.3}_{-0.3}$	$0.5^{+0.3}_{-0.2}$
1.47	$11.45^{+0.06}_{-0.08}$	$0.005^{+0.003}_{-0.002}$	$0.7^{+0.2}_{-0.4}$	$0.7^{+0.3}_{-0.4}$	$0.6^{+0.3}_{-0.3}$
2.23	$11.40^{+0.06}_{-0.07}$	$0.007^{+0.003}_{-0.003}$	$0.7^{+0.2}_{-0.4}$	$0.6^{+0.3}_{-0.4}$	$0.6^{+0.3}_{-0.4}$

Table 3.5 *Fitted HOD parameters, with MCMC priors used (minimum, maximum, starting point). All priors are 'uniform', apart from σ, which is 'log' We show here the derived parameters for the large samples of galaxies within a fixed $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ range at each redshift. $M_{\rm min}$ is the minimum mass halo hosting a galaxy, F_s determines the number of satellite galaxies per halo, $F_c^{A,B}$ are normalisation factors, and σ is the width of the Gaussian distribution of centrals around its peak, $M_{\rm min}$.*

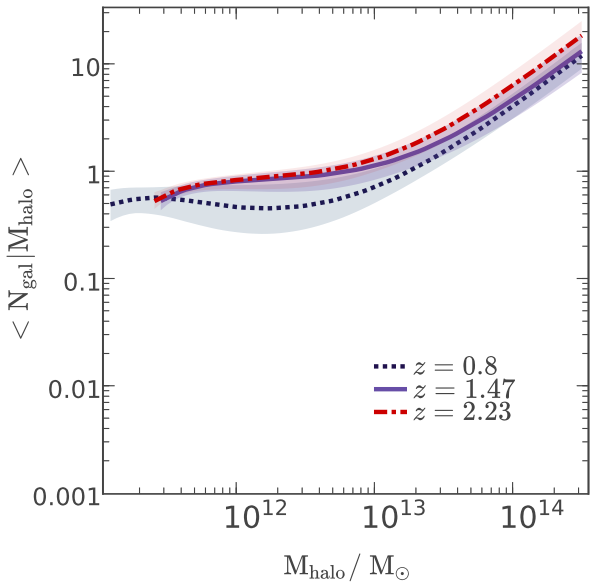

Figure 3.8 *HOD parametrisations of samples of galaxies at $z = 0.8$, $z = 1.47$ and $z = 2.23$, within fixed ranges of $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}$ line up closely. Although the $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}$ distributions are not exactly the same across the different redshift ranges, galaxies selected at similar $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}$ seem to trace similar dark matter haloes across redshift.*

3.4.2 Luminosity dependence of HOD models

Before extending the HOD analysis to Hα luminosity-binned data at all three redshifts, we show fits to luminosity-binned and luminosity-limited data at a single epoch, $z = 0.8$, where we have the largest and most robust sample (Figure 3.9). For the highest luminosity (SFR) bins (e.g. dark blue line), there is a clear shift towards the right, indicating that galaxies typically occupy higher mass dark matter haloes with increasing Hα luminosity. The lowest luminosity bin (yellow line) is also interesting: the central galaxy distribution is strongly peaked around $M_{\mathrm{halo}} \sim 10^{11}\,M_{\odot}$. Therefore high mass haloes do not tend to host central galaxies with these low star formation rates.

The luminosity-binned and luminosity-limited results are largely self-consistent, though there is some discrepancy between the sum of the HODs of independent luminosity bins (black line) and the HOD of the sum of the luminosity bins (grey). This is particularly evident at halo masses in the range $10^{12}\,M_{\odot} < M_{\mathrm{halo}} < 10^{13}\,M_{\odot}$, where the bins sum to more than one central galaxy per halo. We attribute this to the limitations of our parametrisation, and to the uncertainties inherent in fitting HODs to correlation functions constructed using limited numbers of galaxies.

The luminosity-limited HODs broadly agree with the halo occupation of

simulated Hα emitters from the semi-analytic model GALFORM. Geach et al. (2012) show the HOD of GALFORM emitters with $L_{\mathrm{H}\alpha} > 10^{42}\,\mathrm{erg\,s^{-1}}$, which is in excellent agreement with our derived HOD for the same luminosity cut (Figure 3.9, right, green line). Both HODs show the occupation of central galaxies peaking at $M_{\mathrm{halo}} \sim 10^{12}\,M_{\odot}$, with satellites becoming dominant at $M_{\mathrm{halo}} \sim 10^{13}\,M_{\odot}$. HODs derived from the highest luminosity GALFORM sources display a dip in the occupation of haloes around $10^{13}\,M_{\odot}$, with high mass haloes in GALFORM preferentially hosting low luminosity galaxies. We see no evidence for this, but do not reach the high luminosities of $L_{\mathrm{H}\alpha} > 10^{43}\,\mathrm{erg\,s^{-1}}$ where this is most clear in the simulated galaxies. We now explore these trends in greater detail using binned samples at all three redshifts.

At all three redshifts, we observe strong trends in the derived HOD parameters with galaxy Hα luminosity (left-hand panels of Figure 3.10, see also Table 3.4). The effective bias, which characterises the increased clustering of galaxies compared to dark matter, increases roughly linearly with Hα luminosity: more highly star-forming galaxies are therefore more strongly clustered with respect to the underlying dark matter distribution. The effective bias also increases towards higher redshifts. This reflects the growth of the dark matter correlation function

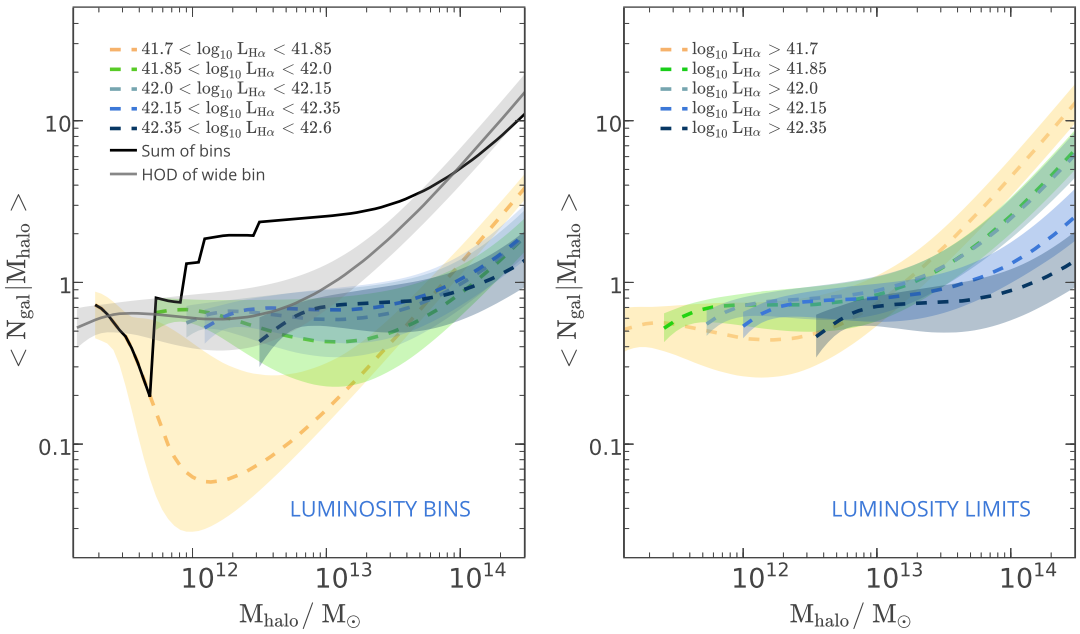

Figure 3.9 *Fitted halo occupation distributions for* Hα *luminosity-binned (left) and luminosity-limited (right) samples at* $z \sim 0.8$. *Higher luminosity* Hα *emitters occupy higher mass dark matter haloes. Our results are qualitatively consistent between the luminosity-binned and luminosity-limited samples, but trends are cleaner for the luminosity-limited samples, which are larger.*

with time (Weinberg et al., 2004). The first galaxies to form - those at high redshift - are more biased relative to the underlying mass distribution which itself is less strongly clustered.

The effective mass ($M_{\rm eff}$) is the average mass of the dark matter halo inhabited by the star-forming galaxies in our samples. The relationship between effective mass of the host dark matter haloes and Hα luminosity is similar to that of the bias: galaxies with higher star formation rates lie, on average, in more massive dark matter haloes. At fixed Hα luminosity, the dark matter halo mass increases steeply towards low redshifts. The minimum mass of dark matter halo that hosts star-forming galaxies scales with Hα luminosity in a similar way: more luminous satellite galaxies are hosted by more massive dark matter haloes.

To compare similar populations of galaxies at the three different redshifts, we scale by the characteristic luminosity once again (see right-hand panels of Figure 3.10). Values of $M_{\rm min}$ from samples at all 3 redshifts form a tight sequence when plotted against $\log_{10}(L_{\rm H\alpha}/L^{*}_{\rm H\alpha})$. This is key: if we select galaxies at a given luminosity relative to the characteristic luminosity at any redshift, they reside in dark matter haloes of the same minimum mass. $M_{\rm eff}$ shows a similar, broadly redshift-independent trend, though there is tentative evidence of evolution to slightly higher masses towards lower redshifts, as the mass of typical dark matter haloes grows with cosmic time. We obtain the following best-fit relations, by fitting to one set of bins at each redshift:

$$\log_{10}(M_{\rm min}/M_{\odot}) = (1.64 \pm 0.11)\ \log_{10}(L_{\rm H\alpha}/L^{*}_{\rm H\alpha}) + (11.94 \pm 0.02) \tag{3.2}$$

$$\log_{10}(M_{\rm eff}/M_{\odot}) = (1.40 \pm 0.12)\ \log_{10}(L_{\rm H\alpha}/L^{*}_{\rm H\alpha}) + (12.46 \pm 0.02) \tag{3.3}$$

We test for evolution in the normalisation of these lines by fixing their gradients to those fitted above (1.64 and 1.40) and fitting the intercept at each redshift individually. We find intercepts of 11.92 ± 0.05 at $z = 0.8$, 11.96 ± 0.06 at $z = 1.47$, and 11.94 ± 0.08 at $z = 2.23$ for the $M_{\rm min}$ fit. Similarly, we obtain 12.54 ± 0.04 at $z = 0.8$, 12.41 ± 0.06 at $z = 1.47$, and 12.36 ± 0.06 at $z = 2.23$ for the $M_{\rm eff}$ fit. The fits are consistent to within 0.04 dex for $M_{\rm min}$ and 0.2 dex for $M_{\rm eff}$.

The satellite fraction for the HiZELS samples is the least well constrained derived parameter. This is because when a halo contains only one star-forming galaxy, the two-point correlation function cannot distinguish whether this is a central galaxy or a satellite of a central quenched galaxy. Satellite galaxies are only constrained by the one-halo term in the most massive haloes, and thus the

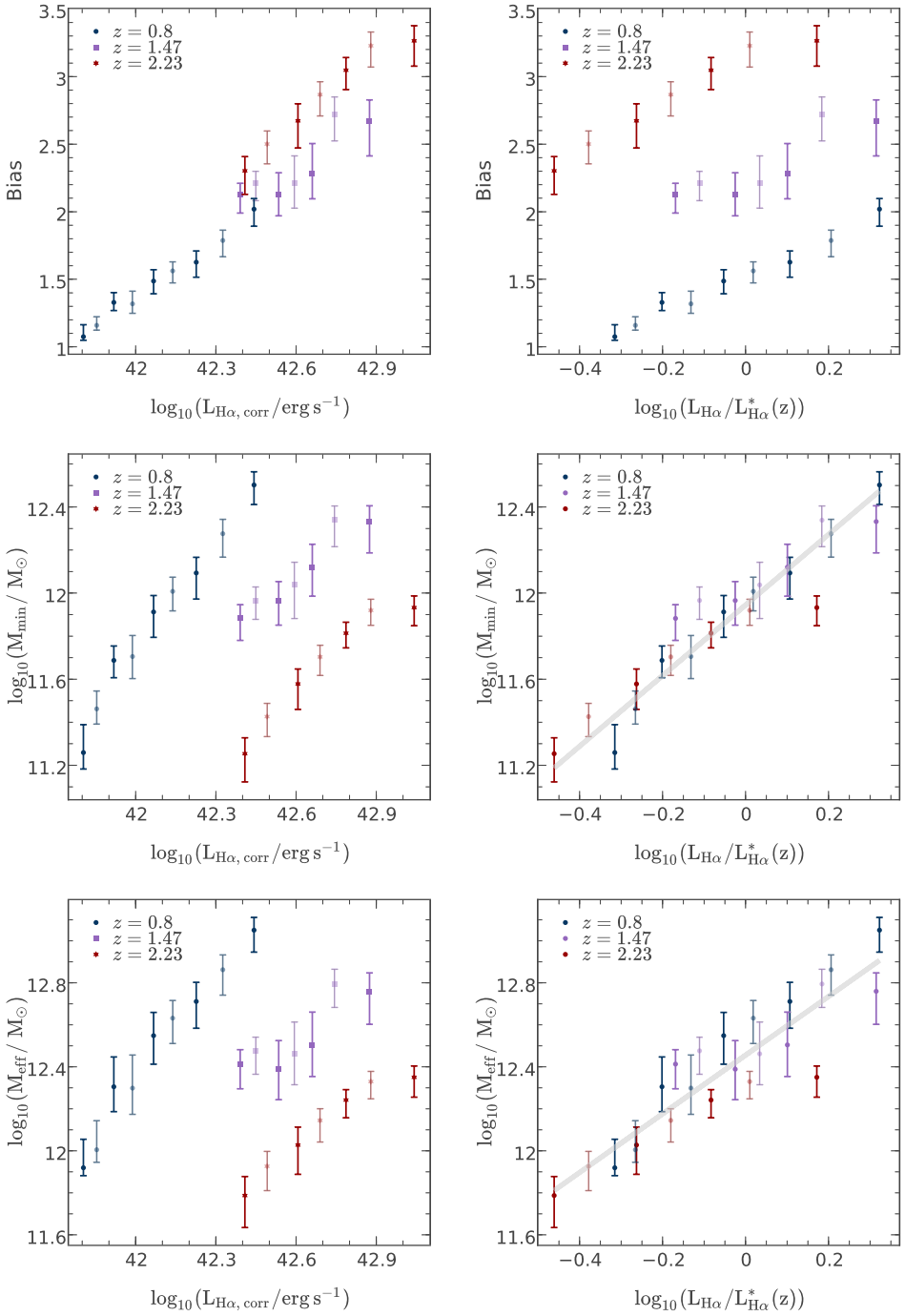

Figure 3.10 *Derived properties of galaxy populations of HiZELS galaxies binned by* Hα *luminosity. We find a linear, broadly redshift-independent relationship between halo mass and* Hα *luminosity. As in Figure 3.3, the paler colours denote alternative binning. The lines of best fit derived in Section 3.4.2 are overplotted:* $\log_{10}(M_{\rm min}/M_\odot) = (1.64 \pm 0.11)\log_{10}(L_{\rm H\alpha}/L^*_{\rm H\alpha}) + (11.94 \pm 0.02)$, $\log_{10}(M_{\rm eff}/M_\odot) = (1.40 \pm 0.12)\ \log_{10}(L_{\rm H\alpha}/L^*_{\rm H\alpha}) + (12.46 \pm 0.02)$.

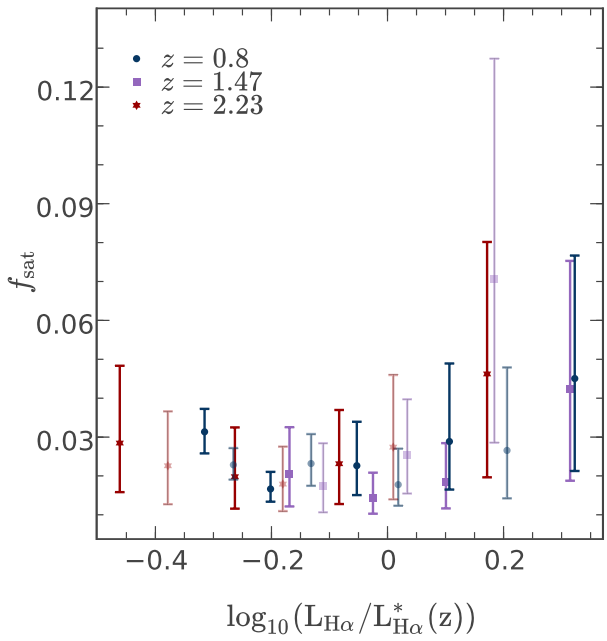

Figure 3.11 *The derived satellite fraction is low for all redshifts and luminosity bins, indicating that HiZELS galaxies are primarily centrals. However, the satellite fraction is the least well constrained of the HOD output parameters. Again, the paler colours denote alternative binning.*

determination of $f_{\rm sat}$ is sensitive to the form of the HOD parametrisation (which extrapolates this to lower masses). Nevertheless, we find no evidence of a change in satellite fraction with luminosity or with redshift (Figure 3.11). As noted previously, this satellite fraction is only the fraction of *star-forming satellites in the sample*, and may be higher if passive populations were included. There is a slight indication of an upturn in satellite fractions at the highest luminosities, but at low significance. Figure 3.9 had shown that the sum of the HODs of luminosity-binned samples clearly exceeds the HOD of the whole sample at moderate halo masses of $M_{\rm halo} \sim 10^{12} - 10^{13}\,M_{\odot}$ by a factor of ~ 2. This suggests that the HOD fits to luminosity-binned samples are overestimating the number of central galaxies in the sample. This would decrease the satellite fraction and explain the discrepancy between the $\sim 5\%$ satellite fractions derived for the whole samples (see Table 3.4) compared to those of luminosity-binned samples, which stand at $\sim 3\%$. The $\sim 5\%$ satellite fraction is likely to be closer to the 'true' satellite fractions of our samples. Nevertheless, the main result of Figure 3.11 is that there is no evidence that $f_{\rm sat}$ changes dramatically with either $L_{\rm H\alpha}$ or redshift.

Finally, we note that when scaled by $L^*_{\rm H\alpha}$, the luminosity-bias relations show strong redshift dependence. This is due to the growth of the dark matter correlation function with time. The different redshifts align better in the $\log_{10} L_{\rm H\alpha}$

vs bias plot, but this is likely to be simply because at fixed $L_{\mathrm{H}\alpha}$, selection of brighter (relative to $L^*_{\mathrm{H}\alpha}$) galaxies at low redshift goes some way towards compensating the dark matter halo growth.

3.5 Discussion

Having studied the halo environments of galaxies at three different redshifts, we draw together the main findings here. The Hα-selected galaxies detected by the HiZELS survey are typical star-forming galaxies residing in dark matter haloes of masses $\sim 10^{12}\,M_{\odot}$. Our typical HiZELS limiting Hα fluxes correspond to star formation rates of $\sim 4\,M_{\odot}$/yr at $z = 0.8$, $\sim 8\,M_{\odot}$/yr at $z = 1.47$ and $\sim 13\,M_{\odot}$/yr at $z = 2.23$, according to the Hα-SFR conversion of Kennicutt (1998). At all redshifts, in all luminosity bins, we find low satellite fractions of $\sim 5\%$, with fitted HODs only reaching above one star-forming satellite per halo in haloes of $\geq 10^{13}\,M_{\odot}$. Whilst there are some uncertainties introduced by the limitations of our HOD parametrisation, the satellite fractions derived are consistently low for both luminosity-limited and luminosity-binned samples of Hα emitters. We conclude that the majority of the star-forming galaxies in our samples are centrals.

The star-forming galaxies detected at lower redshifts ($z = 0.8$ and $z = 1.47$) have lower Hα luminosities than the high-redshift $z = 2.23$ galaxies that reside in equally massive haloes. This reflects the general trend of decreasing star formation rates towards low redshift (see Daddi et al., 2007; Elbaz et al., 2007; Karim et al., 2011; Sobral et al., 2014; Lee et al., 2015). At all three redshifts, we find an increase in estimated average host dark matter halo mass with Hα luminosity of galaxies studied. More highly star-forming galaxies are hosted by more massive dark matter haloes.

We emphasise here that we have performed the analysis on a sample of galaxies selected cleanly by Hα emission line strength. These galaxies are predominantly star-forming, with luminosities close to the characteristic luminosity at each redshift, and are therefore representative of the star-forming population (Oteo et al., 2015). If we were to probe down to much lower star formation rates (including the passive galaxy population), trends in halo mass versus Hα luminosity may eventually reverse. Hartley et al. (2010), for example, found passive galaxies to be significantly more strongly clustered than their star-forming counterparts back to $z \sim 2$ (see also Wilkinson et al., 2017). This fits easily into our interpretation: the passive, massive galaxies at a given redshift formed their mass early (downsizing; Cowie et al. 1996), and hence quickly. Indeed, we find

that the most highly star-forming galaxies at all redshifts are the most strongly clustered.

3.5.1 The halo mass - characteristic luminosity relation

Scaling by the characteristic luminosity at each redshift enables us to compare similar populations of galaxies. The $\log_{10}(L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha})$ versus halo mass relations line up very tightly, and as shown in Figure 3.5, the relationship between $L_{\mathrm{H}\alpha}$ and halo mass is a genuine trend, not driven by stellar mass. This indicates that the mass of the host dark matter halo is driving the typical luminosity of its star-forming galaxies. The minimum halo mass at $L_{\mathrm{H}\alpha} = L^*_{\mathrm{H}\alpha}$ is $\sim 10^{12}\,M_\odot$ for all three redshifts. This exactly coincides with the peak of the stellar mass-halo mass relation (SHMR), the halo mass at which the star formation efficiency peaks, within this redshift range (Behroozi et al., 2010, 2013, see Section 1.2.6). As noted by Behroozi et al. (2013), the halo mass at which the SHMR is at its maximum is also that at which the baryon conversion efficiency (the ratio of the SFR to the halo's baryon accretion rate) is highest. Models predict that this holds across a large redshift range, until at least $z = 4$. Our results support the conclusion that haloes of mass $\sim 10^{12}\,M_\odot$ are the most efficient at forming stars at every epoch. The SHMR decreases at higher halo masses, which are less efficient at forming stars. We obtain $M_{\mathrm{min}} \sim 10^{12-12.4}\,M_\odot$ for our most luminous galaxies, in line with this. This is consistent with the models of Dekel & Birnboim (2006), which posit a roughly redshift-independent limiting halo mass of $M_{\mathrm{shock}} \sim 10^{12}\,M_\odot$, above which efficient gas cooling is prevented by shock heating. Sobral et al. (2016a) find that those HiZELS galaxies with $L_{\mathrm{H}\alpha} > L^*_{\mathrm{H}\alpha}$ have increasing AGN fractions, while Sobral et al. (2009) find that that these very luminous galaxies are much more likely to be mergers than their low-luminosity counterparts (the fraction of $z = 0.84$ HiZELS galaxies with irregular morphologies increases from $< 20\%$ below $L_{\mathrm{H}\alpha} = L^*_{\mathrm{H}\alpha}$ to $\sim 100\%$ at $L_{\mathrm{H}\alpha} > L^*_{\mathrm{H}\alpha}$). This supports the argument that $L^*_{\mathrm{H}\alpha}$ is the luminosity where 'normal', non-merger-driven star formation peaks.

3.5.2 Interpretation via an equilibrium gas regulator model

In this section, we use a few simple ideas from models of the evolution of galaxies and dark matter haloes to link the luminosities of the star-forming galaxies in our sample to the growth of dark matter haloes over cosmic time.

Fakhouri et al. (2010) derive the mean halo mass growth as a function of mass

and redshift, using the Millennium simulation:

$$\left\langle \frac{dm_{\rm halo}}{dt} \right\rangle = 46.1 \left(\frac{m_{\rm halo}}{10^{12}} \right)^{1.1} (1 + 1.11z)\sqrt{\Omega_M (1+z)^3 + \Omega_\Lambda} \tag{3.4}$$

We gather the terms $(1+1.11z)\sqrt{\Omega_M(1+z)^3+\Omega_\Lambda}$ and call them $f(z)$ from here onward.

We define the halo specific mass inflow rate, $sMIR_{\rm DM}$, as:

$$sMIR_{\rm DM} = \frac{1}{m_{\rm halo}} \frac{dm_{\rm halo}}{dt}. \tag{3.5}$$

We now attempt to relate this to star formation in galaxies. Equilibrium models, in which star formation in a galaxy is regulated by the instantaneous mass of gas in its reservoir and mass loss is similarly regulated by the star formation rate, have been successful in reproducing many observed galaxy properties including gas fractions and metallicities to $z \sim 2$ (e.g. Davé et al., 2012; Lilly et al., 2013; Saintonge et al., 2013). In the gas-regulated model of galaxy evolution proposed by Lilly et al. (2013), the specific star formation rate of a central galaxy is related to the average specific mass accretion rate of its dark matter halo via:

$$sSFR = \frac{1}{(1-\eta)(1-R)} sMIR_{\rm DM}, \tag{3.6}$$

where η (the slope of the mass-metallicity relation) and R (which determines the fraction of stars which are long-lived) are observationally-determined constants.

Substituting $sSFR = SFR/m_{\rm star}$, and using $SFR = 7.9 \times 10^{-42} L_{\rm H\alpha}$, from Kennicutt (1998), then combining Equations 3.4 & 3.5 yields:

$$L_{\rm H\alpha} = k\, m_{\rm halo}^{0.1}\, m_{\rm star} f(z), \tag{3.7}$$

where k is a numerical factor.

We found in Section 3.4.2 that the scaled mean Hα luminosity, $L_{\rm H\alpha}/L^*_{\rm H\alpha}$, of a sample of our star-forming galaxies is related to halo mass in a redshift-independent manner:

$$\frac{L_{\rm H\alpha}}{L^*_{\rm H\alpha}(z)} \approx \left(\frac{m_{\rm halo}}{10^{12}} \right)^{1/1.6}. \tag{3.8}$$

Dividing Equation 3.7 by $L^*_{\mathrm{H}\alpha}$, we obtain:

$$\frac{L_{\mathrm{H}\alpha}}{L^*_{\mathrm{H}\alpha}(z)} \approx k\, m_{\mathrm{halo}}^{0.1}\, m_{\mathrm{star}}\, \frac{f(z)}{L^*_{\mathrm{H}\alpha}(z)}, \tag{3.9}$$

which, from our observed relation (Equation 3.8) must remain constant with redshift for a given m_{halo}.

The average galaxy stellar mass, m_{star}, is also related to m_{halo} broadly independently of redshift within our range of redshifts (the SHMR; Behroozi et al., 2013; Birrer et al., 2014; Hatfield et al., 2016). Therefore, to maintain Equation 3.8 across cosmic time in the context of the gas regulator model,

$$\frac{f(z)}{L^*_{\mathrm{H}\alpha}(z)} = const. \tag{3.10}$$

must hold.

To test this, we calculate $\frac{f(z)}{L^*_{\mathrm{H}\alpha}(z)}$ for the HiZELS samples at the three different redshifts. We find that this is, indeed, fairly constant compared to the strong evolution in $L^*_{\mathrm{H}\alpha}$ (see Figure 3.12). Whereas $L^*_{\mathrm{H}\alpha}$ changes by an order of magnitude, $\frac{f(z)}{L^*_{\mathrm{H}\alpha}(z)}$ changes by less than 0.2 dex. Our results therefore support a model in which the evolution in $L^*_{\mathrm{H}\alpha}$ is driven solely by the halo mass growth, in line with a gas regulator model. We thus conclude that our HiZELS galaxies are dominated by typical star-forming galaxies in equilibrium, rather than extreme, merger-driven starburst systems, even at high redshifts. The halo mass accretion rate is the dominant driver of star formation in these galaxies across the large redshift range $0.8 < z < 2.23$.

3.5.3 Satellite fractions and environmental quenching

We have found low satellite fractions ($\sim 5\%$) at all three of the redshifts studied, and in all luminosity bins, using this HOD parametrisation. The gas regulator model, shown in Section 3.5.2 to fit our observations well, does not include any satellite-specific mechanisms like ram pressure stripping. This supports the conclusion that the majority of HiZELS galaxies are centrals.

As discussed earlier, the exact values we derive for the satellite fraction may be significantly dependent on the HOD parametrisation adopted, as two-halo clustering cannot discriminate between centrals and satellites. Nevertheless, it is possible to demonstrate that the satellite fraction must be low. Many HOD models of mass-selected samples of galaxies (including at these redshifts, e.g.

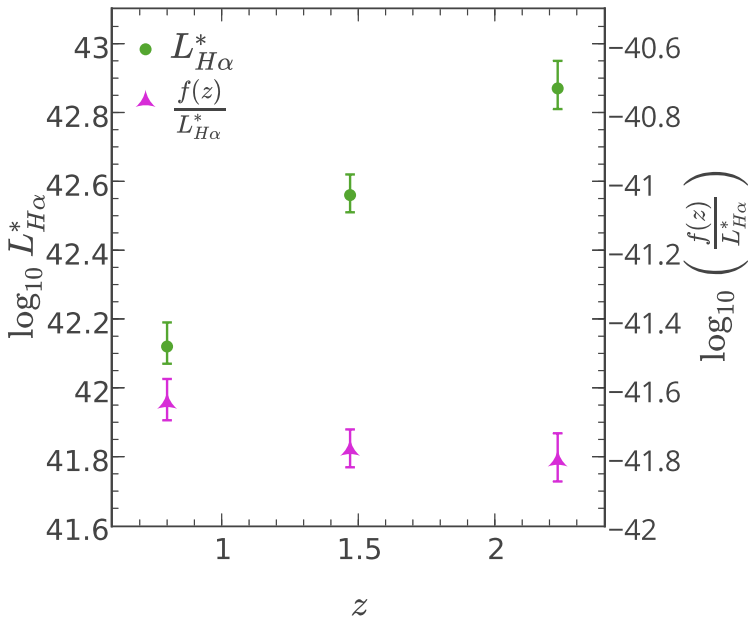

Figure 3.12 *The characteristic* Hα *luminosity,* $L^*_{\mathrm{H}\alpha}$ *displays a striking increase with redshift. Once scaled by the halo mass growth factor,* $f(z)$*, from Fakhouri et al. (2010), we observe little evolution. This indicates that the evolution in* $L^*_{\mathrm{H}\alpha}$*, and therefore in the star formation history of the universe, is driven by dark matter halo mass accretion.*

Wake et al., 2011; Hatfield et al., 2016) use a power-law satellite occupancy model with $\alpha \approx 1$, with a low-mass cut-off below halo mass $\sim 10^{13}\,M_\odot$. As shown in our HOD modelling, obtaining a good fit to our (relatively low-amplitude) correlation functions requires a substantial contribution from low mass haloes, down to $< 10^{12}\,M_\odot$. The scarcity of satellites in these low mass haloes, coupled with the increase in the halo mass function at low halo masses, thus mandates a fairly low overall satellite fraction. To quantify this, we consider a conservative model in which the satellite occupancy of haloes follows a power-law with $\alpha = 1$ down to the lowest masses (i.e. no cut-off), normalised to unity at $M_{\mathrm{halo}} = 2 \times 10^{13}\,M_\odot$ (cf. Wake et al., 2011; Hatfield et al., 2016). Even if all potential satellite galaxies were to be star-forming, our total HOD model for the 'full' sample at $z = 0.8$ then permits a maximum satellite fraction of $\sim 8\%$ (this increases to $\sim 14\%$ for a normalisation of $< N_{\mathrm{sat}}|M >= 1$ at $M_{\mathrm{halo}} = 10^{13}\,M_\odot$). We can thus safely conclude that satellite fractions must be low.

Detailed comparison of our HOD modelling result with those of mass-selected samples at these same redshifts would require us to match the samples in stellar mass; this is beyond the scope of this Chapter, but will be discussed in Chapter 4. Nevertheless, we can gain some initial insights by comparison with the results

of Hatfield et al. (2016), who studied mass-selected samples in a similar stellar mass range as our Hα emitters, in overlapping redshift ranges, using the same HOD fitting code as ours (thus minimising any systematic errors). Hatfield et al. (2016) find satellite fractions of $\sim 13 \pm 1\%$ at $z \sim 0.8$ and $\sim 6 \pm 1\%$ at $z \sim 1.5$, integrating down to the lowest galaxy stellar masses within their samples. Our redshift-independent satellite fraction of star-forming galaxies, when compared to the increasing satellite fraction amongst mass-selected galaxies towards low redshift, indicates that a significantly larger portion of satellites are star-forming at higher redshifts. These results are consistent with those of Tal et al. (2014), who find that the quiescent fraction for satellites increases towards low redshift, from $\sim 10\%$ at $z \sim 1.5$ to $\sim 30\%$ at $z \sim 0.8$, with onset of satellite quenching taking place several Gyr after the first centrals reach quiescence.

Our results may also provide insights into the quenching mechanisms acting at high redshifts. A number of studies find a strong excess of starbursting sub-mm galaxies in high-redshift cluster environments (Elbaz et al., 2007; Smail et al., 2014; Ma et al., 2015). In some cases these starbursting galaxies reside in the cluster core (e.g. Ma et al., 2015), and in others they lie towards the cluster's outskirts, with passive galaxies dominating the central regions (e.g. Smail et al., 2014). If this intense star formation were driven by an enhanced intracluster gas supply, we would expect to see enhanced star formation throughout these high mass haloes, reflected in high satellite fractions and increased effective halo masses for our HiZELS galaxies at higher redshifts. Instead, we find that both of these properties remain broadly consistent. Combined with the sub-mm view, our results support the scenario put forward by McGee et al. (2009), in which upon infall onto a rich cluster, compression of high gas contents within galaxies may provoke intense, dust-obscured star formation, after which quenching proceeds on fairly long timescales ($> 2\,$Gyr) via gas stripping or exhaustion.

3.6 Conclusions

We have used HiZELS galaxies selected by the strength of their Hα emission to study the clustering of star-forming galaxies at three well-defined epochs: $z = 0.8$, $z = 1.47$, $z = 2.23$. Our samples comprise typical star-forming galaxies on and just above the 'main sequence' at each redshift. We have constructed two-point correlation functions and fitted these with simple power-law fits, finding that the clustering strength, r_0, of HiZELS sources at all redshifts increases linearly with their Hα luminosity, from $r_0 \sim 2 - 3h^{-1}\,$Mpc for the lowest luminosity sources in

our samples to $r_0 \sim 7 - 8h^{-1}$ Mpc for the most luminous. We have demonstrated that this is not driven by galaxy stellar mass.

We then used MCMC techniques to fit the same correlation functions with a more sophisticated Halo Occupation Distribution (HOD) models, deriving each galaxy population's effective bias, satellite fraction, and indicative dark matter halo masses. We summarise the key results here.

- Typical Hα-emitting galaxies in the redshift range $z = 0.8 - 2.2$ are star-forming centrals, residing in host haloes of minimum mass $10^{11.2}\,M_\odot - 10^{12.6}\,M_\odot$ and effective mass $10^{11.6}\,M_\odot - 10^{13}\,M_\odot$. At all three redshifts, $L^*_{\mathrm{H}\alpha}$ galaxies typically reside in haloes of mass $\sim 10^{12}\,M_\odot$. This coincides with the halo mass predicted by theory to be maximally efficient at converting baryons into stars.

- The effective bias of the galaxy populations (their clustering relative to the underlying dark matter) decreases towards lower redshifts, reflecting the increase of the clustering of dark matter with time. Similarly, typical masses of host haloes increase with time at fixed $L_{\mathrm{H}\alpha}$.

- Bias increases linearly with Hα luminosity at all redshifts, indicating that the most highly star-forming galaxies thrive in higher dark matter overdensities, where a plentiful gas supply fuels high star formation rates in the central galaxies.

- Samples selected within the same $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}$ range inhabit similar populations of dark matter haloes. Although the dark matter halo mass at fixed $L_{\mathrm{H}\alpha}$ varies by more than an order of magnitude across the three different redshifts, the relationship between scaled galaxy luminosity $L_{\mathrm{H}\alpha}/L^*_{\mathrm{H}\alpha}$ and dark matter halo mass is independent of redshift to within 0.04 dex in M_{min} and 0.2 dex in M_{eff}.

- Comparing our results to models of galaxy evolution based on gas-regulation, we find that $L^*_{\mathrm{H}\alpha}$ evolves in line with average mass growth of the host dark matter haloes.

Together, these results reveal halo environment as a strong driver of galaxy star formation rate and the evolution of the luminosity function over cosmic time. The central galaxies which dominate our samples evolve in equilibrium with their growing dark matter haloes, with typical specific star formation rate directly proportional to the specific mass accretion rate of the host dark matter halo.

Satellite fractions remain low ($\sim$ 5% with the HOD parametrisation we have adopted) for all samples, regardless of redshift or luminosity. This may indicate that their star formation is suppressed, particularly towards low redshifts and in high mass dark matter haloes. This is in line with models of satellite quenching upon accretion onto a massive cluster. In the following Chapter we will extend this study to incorporate stellar mass, exploring the clustering of HiZELS galaxies as a function of Hα luminosity, stellar mass and redshift.

Chapter 4

The dependence of galaxy clustering on stellar mass, star formation rate and redshift at $z = 0.8 - 2.2$, with HiZELS

This chapter contains work that has been published as Cochrane R. K., Best P. N., Sobral D., Smail I., Geach J. E., Stott J. P., Wake D. A., 2018, Monthly Notices of the Royal Astronomical Society, Volume 475, Issue 3, p.3730-3745.

In this Chapter, we distinguish the stellar mass and star formation rate (SFR) dependence of the clustering of HiZELS galaxies at $z = 0.8$, $z = 1.47$ and $z = 2.23$. At high stellar masses ($M_*/M_\odot \gtrsim 2 \times 10^{10}$), where HiZELS selects galaxies close to the so-called star-forming main sequence, the clustering strength is observed to increase strongly with stellar mass (in line with the results of previous studies of mass-selected galaxy samples) and also with SFR. These two dependencies are shown to hold independently. At lower stellar masses, however, where HiZELS probes high specific SFR galaxies, there is little or no dependence of the clustering strength on stellar mass, but the dependence on SFR remains: high-SFR, low-mass galaxies are found in more massive dark matter haloes than their lower SFR counterparts. We argue that this is due to environmentally-driven star formation in these systems. We apply the same selection criteria to the EAGLE cosmological hydrodynamical simulations. We find that, in EAGLE, the high-SFR, low-mass galaxies are central galaxies in more massive dark matter haloes, in which the high SFRs are driven by a (halo-driven) increased gas content.

4.1 Introduction

As discussed in Chapter 3, a rich array of work reveals that key observable galaxy properties including stellar mass, colour, star formation rate, and morphology correlate with galaxy environments (Butcher & Oemler, 1978; Dressler, 1980; Baldry et al., 2006; Peng et al., 2010; Koyama et al., 2013; Scoville et al., 2013; Darvish et al., 2016), with massive, red, quiescent spheroids residing in the densest environments. Studies of galaxy environments can help constrain galaxy formation and evolution processes (e.g. Peng et al., 2010). Yet quantifying galaxy environments on a galaxy-by-galaxy basis can be difficult, particularly at high redshifts, because the accuracy of such measurements is highly dependent on the depth and uniformity of the observations and the quality of the redshifts (e.g. Cooper et al., 2005).

Galaxy clustering measures provide a statistical description for a population of galaxies rather than quantifying environments on a galaxy-by-galaxy basis. Strong trends in clustering strength have been observed with galaxy morphological type (Davis & Geller, 1976), colour (Zehavi et al., 2005; Coil et al., 2008; Simon et al., 2009; Hartley et al., 2010; Zehavi et al., 2011), star formation rate (Williams et al., 2009; Dolley et al., 2014; Wilkinson et al., 2017) and stellar mass (Wake et al., 2011; McCracken et al., 2015; Coupon et al., 2015; Hatfield et al., 2016), with the more recent studies reaching back to $z \sim 2-3$. A limited number of studies of Lyman break galaxies have probed even further, back to $z \sim 6-7$ (e.g. Harikane et al., 2016, 2017; Hatfield et al., 2018). The largest samples have permitted the splitting of galaxy populations by more than one observed property. For example, Norberg et al. (2002), using low-redshift ($z < 0.15$) data from the 2dF survey (Cole et al., 2000b), found that both early- and late-type galaxies display higher r_0 values and therefore stronger clustering at brighter B-band absolute magnitudes (M_B). Coil et al. (2008) found broadly consistent results at $z \sim 1$ using the DEEP2 galaxy redshift survey (Newman et al., 2012), also confirming that at fixed M_B, red galaxies are more strongly clustered than blue galaxies.

Splitting by multiple variables in this manner is important for galaxy evolution studies. A natural consequence of the tight ($\sim$ 0.4 dex scatter) correlation between stellar mass and star formation rate of star-forming galaxies (the 'main sequence', see Chapter 1) is that fundamental trends in one of these properties manifest as trends in the other. Galaxies with star formation rates below the main sequence can also complicate observed trends: the fraction of galaxies that

are passive increases towards higher stellar masses (Peng et al., 2010; Sobral et al., 2011), and this can give rise to trends with stellar mass which might not exist for the star-forming population only (e.g. the bending of the main sequence, Lee et al. 2015). Therefore, in this Chapter, we aim to investigate the dependence of galaxy clustering on galaxy stellar mass and star formation rate separately, using the HiZELS samples studied in Chapter 3.

Sobral et al. (2010) presented the first study of Hα luminosity-binned HiZELS galaxies and found evidence of higher clustering strengths for the strongest emitters at $z = 0.84$. Geach et al. (2008) and Geach et al. (2012) performed the first clustering studies of $L_{\rm H\alpha}$-selected galaxies at $z = 2.23$, though the sample size was insufficient to split by luminosity. In Chapter 3, we confirmed that the trends found by Sobral et al. (2010) hold to higher redshifts, using larger HiZELS samples at $z = 0.8$, $z = 1.47$ and $z = 2.23$, and used HOD fitting to study trends with halo mass. Sobral et al. (2010) used the K-band luminosities of HiZELS galaxies as a proxy for their stellar mass, finding an increase in galaxy clustering with increasing K-band luminosity, though the trend was significantly shallower than was observed for Hα luminosities. Preliminary investigations in Chapter 3 involved splitting our larger sample of galaxies at $z = 0.8$ into two bins by observed K-band magnitude. Intriguingly, we found that the strong, roughly linear relationship between $\log_{10} L_{\rm H\alpha}$ and r_0 held for our two samples, with any differences between the two K-band magnitude bins being much smaller than the trend with Hα luminosity. Khostovan et al. (2018) present consistent results in their study of Hβ + [OII] and [OIII] emitters from HiZELS: clustering strength increases more significantly with emission line strength than with galaxy stellar mass.

In this Chapter, we extend our previous work to study the clustering of HiZELS star-forming galaxies as a function of both Hα luminosity and stellar mass in more detail. Rather than using K-band observed magnitude as a proxy for stellar mass, we use a full SED-fitting approach to estimate stellar masses. We then compare our observational results to the output of the state-of-the-art cosmological hydrodynamical simulation EAGLE (Crain et al., 2015; McAlpine et al., 2016; Schaye et al., 2015). The structure of this Chapter is as follows. In Section 4.2 we discuss our stellar mass estimates in some depth. In Section 4.3 we present our conversion between clustering strength and halo mass, which is based on the techniques used in Chapter 3. In Section 4.4 we present our results, and in Section 4.5 we compare these to the output of the EAGLE simulation. Conclusions are drawn in Section 4.6.

4.2 Deriving stellar masses and SFRs of HiZELS galaxies

We use the samples of Hα emitters as presented in Chapter 3. star formation rates are derived directly from dust-corrected Hα luminosities, $L_{H\alpha}$ using:

$$SFR_{H\alpha}(M_{\odot}year^{-1}) = 4.6 \times 10^{-42} L_{H\alpha}(erg\,s^{-1}), \quad (4.1)$$

adopting the calibration of Kennicutt (1998) and scaling by a factor 1.7 (Speagle et al., 2014) to convert from a Salpeter (1955) IMF to a Chabrier (2003) IMF.

In order to estimate stellar mass, we model each galaxy's stellar populations and dust content via spectral energy distribution (SED) fitting using a similar method to that described in Sobral et al. (2011) and Sobral et al. (2014). The observed photometry is first shifted into the rest-frame. Model galaxy SEDs are then convolved with the detector's spectral response function to compare modelled and observed flux, and fitted via χ^2 minimization.

Our modelling draws upon the stellar population synthesis package of Bruzual & Charlot (2003), using the updated models commonly referred to as CB07. These models assume a Chabrier (2003) IMF and an exponentially declining star formation history of the form $e^{-t/\tau}$, where τ is in the range $0.1-10$ Gyr. Although this is not a realistic description of the star formation histories of individual galaxies, which are likely to be characterized by shorter bursts, triggered by stochastic accretion, τ is a reasonable estimate of the mean age of a galaxy (see also Sobral et al. 2014, who show that using single exponential star formation models does not introduce any significant bias into the stellar mass estimates of HiZELS galaxies). We use a grid of ages from 30 Myr to the age of the Universe at each redshift, with a grid of dust extinctions from Calzetti et al. (2000) up to $E(B-V) = 0.5$, and three metallicities ($0.2 - 1.0\,Z_{\odot}$).

For the COSMOS field, up to 36 wide, medium and narrow bands are used, from *GALEX*'s far-UV band to *Spitzer*'s four IRAC bands. In the UDS field there are only 16 available bands, but J, H and K data from UKIRT/UKIDSS DR5 are very deep. Seven bands ($ugrizJK$) are used in SA22 (see Sobral et al., 2013b). All HiZELS sources are assumed to lie at the central wavelength of the redshift distribution, which is a reasonable approximation since the filter profile is extremely narrow (see Table 3.3). The resultant stellar masses are fairly well constrained, with typical statistical uncertainties of 0.23, 0.24 and 0.26 dex at $z = 0.8$, 1.47 and 2.23, which vary a little from source to source.

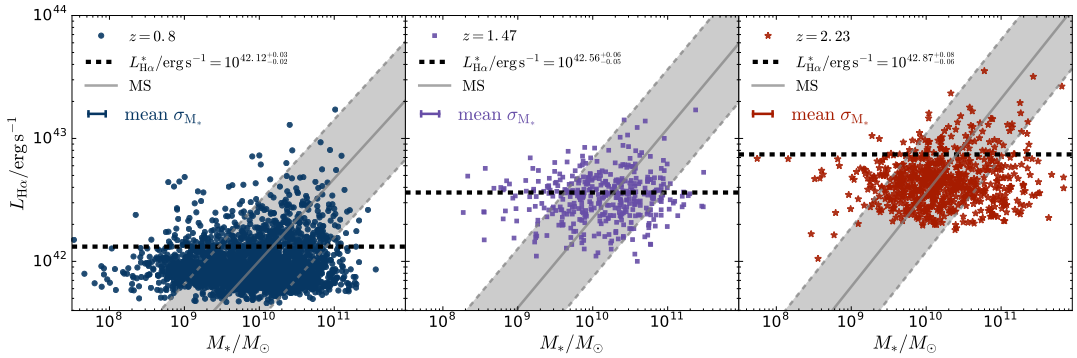

Figure 4.1 *Distributions of SED-estimated stellar masses and dust-corrected* Hα *luminosities for the three samples of HiZELS galaxies, at* $z = 0.8$, $z = 1.47$ *and* $z = 2.23$. *The dashed lines show* $L^*_{\rm H\alpha}$ *at each redshift, derived by Sobral et al. (2013a) and Chapter 3. Overplotted are indicative regions of the 'main sequence' at each redshift with* 2σ *contours, derived by Speagle et al. (2014).*

SED masses are plotted against Hα luminosities for the HiZELS samples in Figure 4.1. At each redshift, our samples cover a very wide range in stellar mass ($10^8 < M_*/M_\odot < 10^{11}$) and also around 1 dex in Hα luminosity, spanning the break of the luminosity function.

As a test of our stellar masses, especially in SA22, where fewer bands are available, we compare our stellar mass estimates to apparent K-band luminosities, which broadly trace the older stellar population (e.g. Kauffmann, 1998; Longhetti & Saracco, 2009). Figure 4.2 shows SED-derived stellar mass versus observed K-band magnitude for HiZELS galaxies in the SA22 field at $z = 0.8$. These galaxies occupy a clear locus in this plane, close to the line expected from direct proportionality between K-band flux (rest-frame 1.2 μm) and stellar mass. At fixed K-band magnitude, redder galaxies (see colour coding) have higher SED masses than would be expected from a naive extrapolation of K-band flux, and bluer galaxies have lower derived SED masses. This is exactly as expected, since the red fraction is higher for higher luminosity sources. These galaxies are dominated by old stars and have high mass-to-light ratios. In contrast, the bluer (typically less luminous) galaxies in our HiZELS samples have younger stellar populations, and are thus particularly luminous for their mass. We conclude that our SED masses are reasonable, and fold in important colour information. Therefore, we use the SED-derived stellar masses for the remainder of this Chapter, with confidence. We note, nevertheless, that our results are qualitatively unchanged whether we use K-band-derived or SED-derived masses.

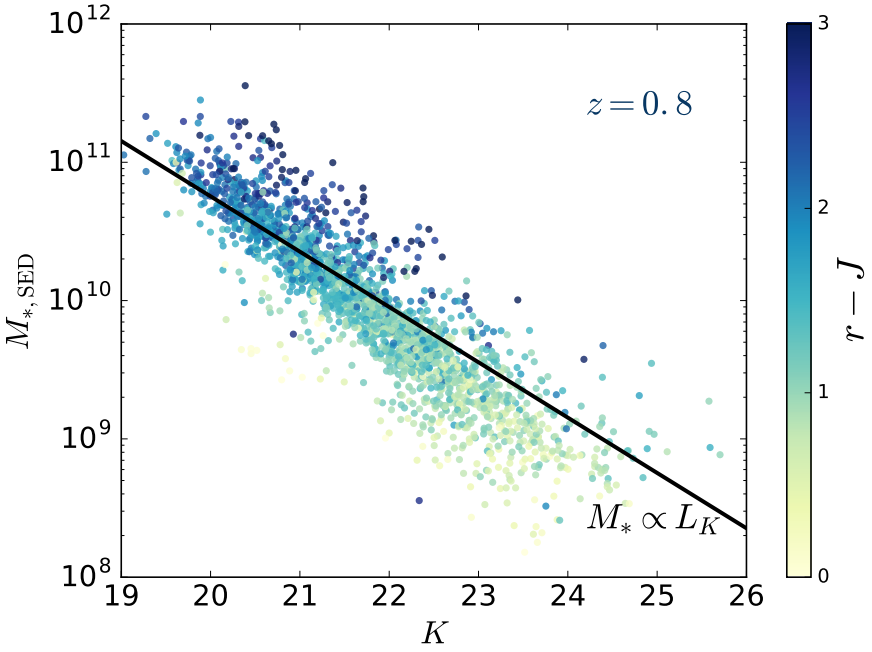

Figure 4.2 *SED-derived stellar mass versus observed K-band magnitude for SA22 galaxies, colour-coded by $r-J$ colour. The black line shows the direct proportionality between K-band flux (rest-frame 1.2 μm) and stellar mass (i.e. gradient fixed at -0.4). The stellar mass is clearly well correlated with K-band flux, but at fixed K-band magnitude, redder galaxies have higher SED-derived stellar masses, as would be expected. This colour dependence appears to drive the scatter in the relation and the deviation of the points from the straight line shown.*

4.3 Quantifying galaxy clustering using the two-point correlation function

We use the two point angular correlation function described in Chapter 2 to probe the clustering strength of HiZELS galaxies binned by stellar mass. In this Chapter, we do not perform HOD fitting as in Chapter 3. Instead, we calibrate the clustering strength to halo mass based on our results in that Chapter.

4.3.1 Calibrating $\mathrm{r_0}$ to $\mathrm{M_{halo}}$ using HOD models

For samples of galaxies with large satellite fractions, there will be a substantial one-halo term in the correlation function at small separations. In such cases, HOD modelling offers a better fit than a simple power-law. In Chapter 3, we found that HiZELS samples at $z = 0.8$, $z = 1.47$ and $z = 2.23$ have low satellite fractions ($\sim 5\%$), and HOD fitting offers only marginal gains in goodness of fit at small scales (see Figure 4.3, left-hand panel). Instead, the main benefit of HOD

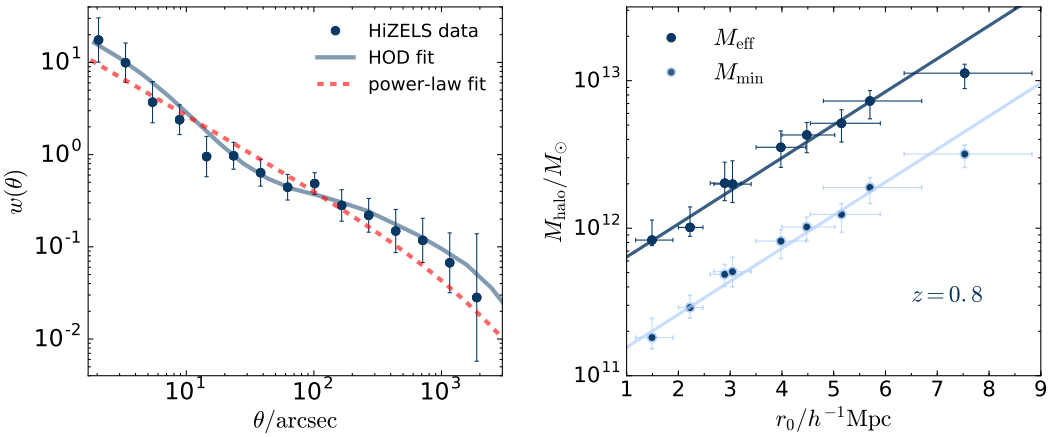

Figure 4.3 *Left: The two-point angular correlation function constructed for the whole sample at $z = 0.8$, fitted with a power-law ($r_0 = 2.58^{+0.16}_{-0.14}\,h^{-1}$Mpc) and HOD model ($M_{\rm eff} = 12.13^{+0.10}_{-0.09}\,M_\odot$). Right: $r_0 - M_{\rm halo}$ calibration from Chapter 3. Overplotted are the best-fitting relations $\log_{10} M_{\rm eff}/M_\odot = 11.7 \pm 0.7 + r_0/(4.5 \pm 0.3)$ and $\log_{10} M_{\rm min}/M_\odot = 10.9 \pm 0.7 + r_0/(4.5 \pm 0.3)$. We find excellent linear fits, so use r_0 as a proxy for halo mass in this Chapter.*

fitting is to allow the conversion of clustering strengths into typical halo masses. Comparing measured r_0 to derived halo masses (Figure 4.3, right-hand panel), we find that these are tightly correlated, and can be reasonably approximated as simple linear fits. At $z = 0.8$, these are given by:

$$\log_{10} M_{\rm eff}/M_\odot = 11.7 \pm 0.7 + r_0/(4.5 \pm 0.3) \tag{4.2}$$

$$\log_{10} M_{\rm min}/M_\odot = 10.9 \pm 0.7 + r_0/(4.5 \pm 0.3). \tag{4.3}$$

Therefore, in some parts of this Chapter (Section 4.4.1 - 4.4.4), we simply derive and quote r_0 values, as these are sufficient to indicate trends of clustering with stellar mass or star formation rate. When we require robust halo masses, as in Sections 4.4.5 and 4.5, we perform the full HOD fitting.

4.4 Clustering of HiZELS galaxies as a function of stellar mass and SFR

4.4.1 Clustering as a function of Hα luminosity

In Chpater 3, we studied the clustering of HiZELS galaxies as a function of their Hα luminosity. We found strong relationships between $L_{\rm H\alpha}$ and r_0. The

clustering strength increases monotonically with Hα luminosity at all redshifts, indicating that the most highly star-forming galaxies thrive in higher dark matter overdensities (see Figure 4.4). We speculated that this is where a plentiful gas supply fuels high star formation rates.

HOD fitting revealed that typical Hα-emitting galaxies are star-forming centrals, residing in host haloes with minimum mass increasing with Hα luminosity from $\sim 10^{11.2}\,M_{\odot}$ to $\sim 10^{12.6}\,M_{\odot}$ and corresponding effective halo masses $\sim 10^{11.6}\,M_{\odot} - 10^{13}\,M_{\odot}$. At all three redshifts, $L^{*}_{\mathrm{H}\alpha}$ galaxies typically reside in haloes of effective mass $\sim 10^{12}\,M_{\odot}$. This coincides with the halo mass predicted by theory to be maximally efficient at converting baryons into stars. Samples selected within the same $L_{\mathrm{H}\alpha}/L^{*}_{\mathrm{H}\alpha}$ range inhabit similar populations of dark matter haloes. The relationship between scaled galaxy luminosity $L_{\mathrm{H}\alpha}/L^{*}_{\mathrm{H}\alpha}$ and dark matter halo mass is largely independent of redshift.

4.4.2 Clustering as a function of stellar mass

Chapter 3 briefly looked at K-band observed luminosities. We found that the trends in clustering strength with $L_{\mathrm{H}\alpha}$ do not differ between two large K-band bins, concluding that they are unlikely to be driven by stellar mass. Here, we extend that study to provide a more definitive answer to the role stellar mass plays.

Initially we bin our sample of $z \sim 0.8$ HiZELS galaxies by stellar mass, construct correlation functions and fit these as described in Section 2.1, obtaining a clustering strength r_0 for each subsample. We use the broad bins in Hα luminosity as defined by Chapter 3 ($-0.4 < \log_{10}(L_{\mathrm{H}\alpha}/L^{*}_{\mathrm{H}\alpha}) < 0.3$) for consistency, but find no significant differences when we re-run the analysis with no luminosity cuts except for the HiZELS selection. We find that the clustering strength is broadly constant with stellar mass at low galaxy masses (see Figure 4.4). This is particularly clear at $z = 0.8$, where our samples are largest and probe lowest in stellar mass, but all three redshifts are consistent with this result. The clustering strength only increases when we reach stellar mass bins that contain a significant number of galaxies below the main sequence: at all three HiZELS redshifts, clustering strength increases significantly above a mass $2 - 3 \times 10^{10}\,M_{\odot}$ and the most massive galaxies are very strongly clustered (see Figure 4.4 and Table 4.1). For our Hα-selected samples, the $M_{*} - r_0$ relationship appears substantially weaker than the $L_{\mathrm{H}\alpha} - r_0$ relation obtained by Chapter 3, and shown in Figure 4.4 for comparison, which continues to decrease at low Hα luminosities.

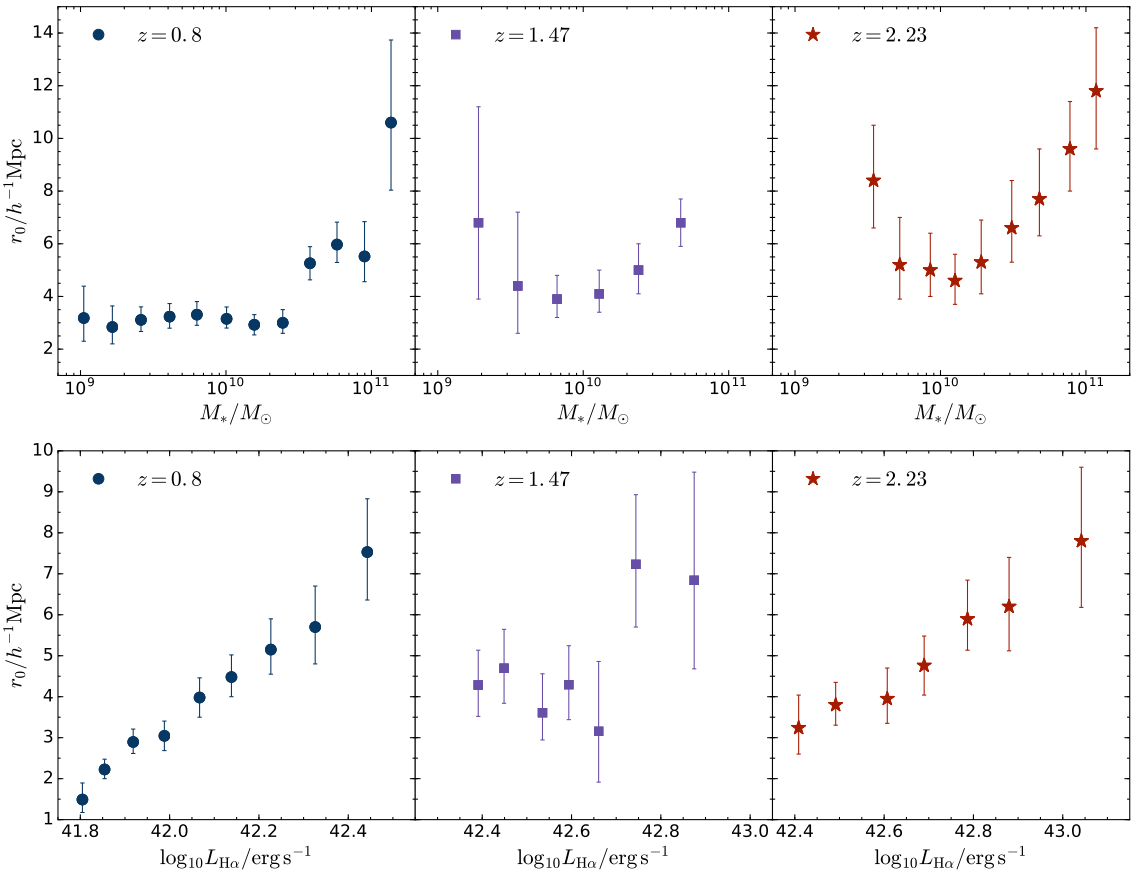

Figure 4.4 *Top: clustering strength, r_0, as a function of stellar mass. At all three redshifts, the clustering strength is broadly flat at low stellar masses, with evidence for an increase for the most massive galaxies (above $\sim 2-3\times10^{10}\,M_\odot$). Bottom: r_0 versus $L_{\mathrm{H}\alpha}$ from Chapter 3, replotted for comparison. Here, a strong monotonic trend is seen between r_0 and $L_{\mathrm{H}\alpha}$ at $z = 0.8$ and $z = 2.2$; as shown in Chapter 3, the $z = 1.47$ data are consistent with the same trend (albeit noisier due to the smaller sample).*

Whilst the gradient of the stellar mass - halo mass relation of mass-selected galaxies does decrease below $M_* \sim 10^{10}\,M_\odot$ (see Section 4.4.5; Moster et al. 2010, 2013; Behroozi et al. 2013 and many others), the flattening we observe for these Hα-selected galaxies is very pronounced. This indicates that low-mass HiZELS galaxies reside in more massive dark matter haloes than would be expected for star-forming central galaxies of these stellar masses. Although this might be surprising, given that in Chapter 3 we found low satellite fractions for these samples, it is important to remember that, at these masses, HiZELS Hα-selected galaxies lie well above the 'main sequence'. We explore the joint dependence of clustering on both stellar mass and $L_{\mathrm{H}\alpha}$ in the following subsection.

4.4.3 Splitting by both stellar mass and Hα luminosity

Within the star-forming population, higher mass galaxies tend to have higher star formation rates (and therefore higher Hα luminosities), so trends in mass can manifest as apparent trends in star formation rate, and vice-versa. Here, r_0 increases significantly at both high $L_{\mathrm{H}\alpha}$ and high stellar masses, and it is hard to tell the extent to which mass and luminosity are each independently correlated with halo mass. Our large samples of HiZELS galaxies allow us to break this degeneracy, and study trends in stellar mass and $L_{\mathrm{H}\alpha}$ luminosity independently.

At $z = 0.8$, where our sample is largest, we split the stellar mass - $L_{\mathrm{H}\alpha}$ plane into $\sim$ 500 overlapping subsamples, constructing and fitting two-point correlation functions for each. In Figure 4.5, we present a 2D plot of stellar mass versus $L_{\mathrm{H}\alpha}$. Each region is colour-coded by its r_0 value, obtained via a smoothed grid using x and y values of each subsample's mean stellar mass and star formation rate, respectively. Note that these r_0 measurements are not independent, due to the overlapping samples. With around 100 galaxies per bin, there are approximately 30 independent subsamples. We find that clustering strength increases broadly monotonically with $L_{\mathrm{H}\alpha}$ at all stellar masses. At high stellar masses $M_* \geq 10^{10}\,M_\odot$, r_0 also increases with stellar mass, as has been found by many clustering studies of mass-selected samples. At low stellar masses, the stellar mass-r_0 relationship breaks down, as had been seen in Figure 4.4. There is little change in r_0 with stellar mass at fixed $L_{\mathrm{H}\alpha}$ (if anything, r_0 increases slightly as we probe to lower stellar mass at higher $L_{\mathrm{H}\alpha}$, where we are probing star formation rates well above the main sequence).

Next, we show projections of this plot for the $z = 0.8$ data, and for the smaller samples at $z = 1.47$ and $z = 2.23$. We divide our galaxies at each redshift slice into two stellar mass bins, and bin further by $L_{\mathrm{H}\alpha}$. We construct two-point correlation

$\log_{10}(M_*/M_\odot)$	Mean $\log_{10}(M_*/M_\odot)$	r_0/h^{-1}Mpc
$z = 0.8$, $41.72 < \log_{10}(L_{H\alpha}/\mathrm{erg\,s^{-1}}) < 42.42$		
$8.8 - 9.2$	9.02	$3.2^{+1.2}_{-0.9}$
$9.0 - 9.4$	9.22	$2.8^{+0.8}_{-0.6}$
$9.2 - 9.6$	9.42	$3.1^{+0.5}_{-0.4}$
$9.4 - 9.8$	9.61	$3.2^{+0.5}_{-0.4}$
$9.6 - 10.0$	9.80	$3.3^{+0.5}_{-0.4}$
$9.8 - 10.2$	10.00	$3.2^{+0.5}_{-0.4}$
$10.0 - 10.4$	10.19	$2.9^{+0.4}_{-0.4}$
$10.2 - 10.6$	10.39	$3.0^{+0.5}_{-0.4}$
$10.4 - 10.8$	10.58	$5.3^{+0.6}_{-0.6}$
$10.6 - 11.0$	10.76	$6.0^{+0.9}_{-0.7}$
$10.8 - 11.2$	10.95	$5.5^{+1.3}_{-1.0}$
$11.0 - 11.4$	11.13	$10.6^{+3.1}_{-2.6}$
$z = 1.47$, $42.16 < \log_{10}(L_{H\alpha}/\mathrm{erg\,s^{-1}}) < 42.86$		
$8.9 - 9.5$	9.28	$6.8^{+4.4}_{-2.9}$
$9.2 - 9.8$	9.55	$4.4^{+2.8}_{-1.8}$
$9.5 - 10.1$	9.82	$3.9^{+0.9}_{-0.7}$
$9.8 - 10.4$	10.11	$4.1^{+0.9}_{-0.7}$
$10.1 - 10.7$	10.38	$5.0^{+1.0}_{-0.9}$
$10.4 - 11.0$	10.67	$6.8^{+1.1}_{-0.9}$
$z = 2.23$, $42.47 < \log_{10}(L_{H\alpha}/\mathrm{erg\,s^{-1}}) < 43.17$		
$9.3 - 9.7$	9.54	$8.4^{+2.1}_{-1.8}$
$9.5 - 9.9$	9.72	$5.2^{+1.8}_{-1.3}$
$9.7 - 10.1$	9.93	$5.0^{+1.4}_{-1.0}$
$9.9 - 10.3$	10.10	$4.6^{+1.0}_{-0.9}$
$10.1 - 10.5$	10.28	$5.3^{+1.6}_{-1.2}$
$10.3 - 10.7$	10.49	$6.6^{+1.8}_{-1.3}$
$10.5 - 10.9$	10.68	$7.7^{+1.9}_{-1.4}$
$10.7 - 11.1$	10.89	$9.6^{+1.8}_{-1.6}$
$10.9 - 11.3$	11.07	$11.8^{+2.4}_{-2.2}$

Table 4.1 *Clustering strength, r_0, for stellar mass-binned samples of HiZELS galaxies at $z = 0.8$, 1.47, and 2.23.*

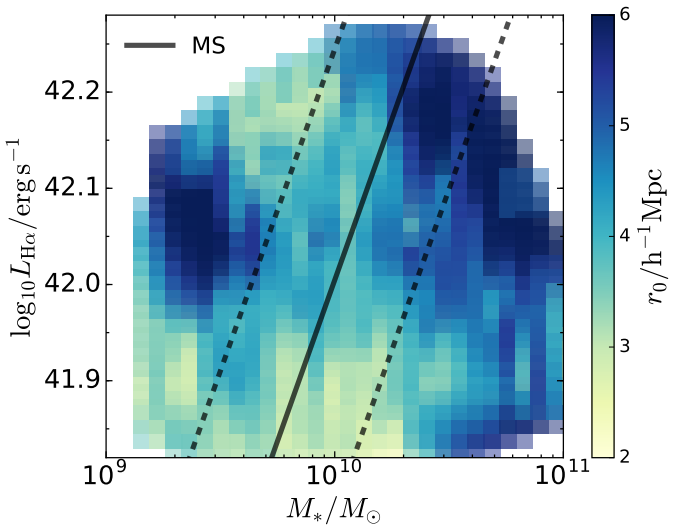

Figure 4.5 *r_0 in the stellar mass - $L_{\mathrm{H}\alpha}$ plane at $z = 0.8$, constructed using ~ 500 overlapping (non-independent) subsamples and plotted using a smoothed linear interpolation. We overplot the main sequence derived by Speagle et al. (2014) at this redshift as a solid line, with the dashed lines showing the standard deviation. Clustering strength increases broadly monotonically with $L_{\mathrm{H}\alpha}$ at all stellar masses. At high stellar masses $M_* \gtrsim 2{\times}10^{10}\,M_\odot$, r_0 increases with stellar mass. We also find large r_0 values for highly star-forming low stellar mass galaxies that are located well above the main sequence.*

functions and obtain correlation strengths for these subsamples. The results are shown in Figure 4.6. We find that the increase in clustering strength with Hα luminosity holds for both stellar mass bins. The trends of the two stellar mass bins are almost indistinguishable. Only the most extremely luminous galaxies at $z = 0.8$ ($L_{\mathrm{H}\alpha} > 10^{42.2}\,\mathrm{erg\,s^{-1}}$) show any departure from this, and, as found by Sobral et al. (2016a), HiZELS samples at these luminosities suffer from significant AGN contamination.

We also divide our galaxies at each redshift slice into two $L_{\mathrm{H}\alpha}$ bins, and bin further by stellar mass. The results are shown in Figure 4.7. Given the size of the sample, our results are clearest at $z = 0.8$. Here, we find that at all stellar masses, the higher luminosity galaxies are more strongly clustered than low luminosity galaxies at the same stellar mass, but this difference is most significant at low stellar masses. The data at $z = 0.8$ (top panel of Figure 4.4) clearly shows that below stellar masses of $M_* \sim 10^{10}\,M_\odot$, HiZELS galaxies have a fairly flat r_0-M$_*$ relation. At these stellar masses, the higher luminosity subsample displays much stronger clustering than the lower luminosity subsample, with $r_0 \sim 6-7\,h^{-1}$Mpc ($M_{\mathrm{eff}} \sim 10^{13}\,M_\odot$), compared to $r_0 \sim 3-4\,h^{-1}$Mpc ($M_{\mathrm{eff}} \sim 10^{12.4}\,M_\odot$). There is even a slight increase in clustering strength towards low masses for the higher

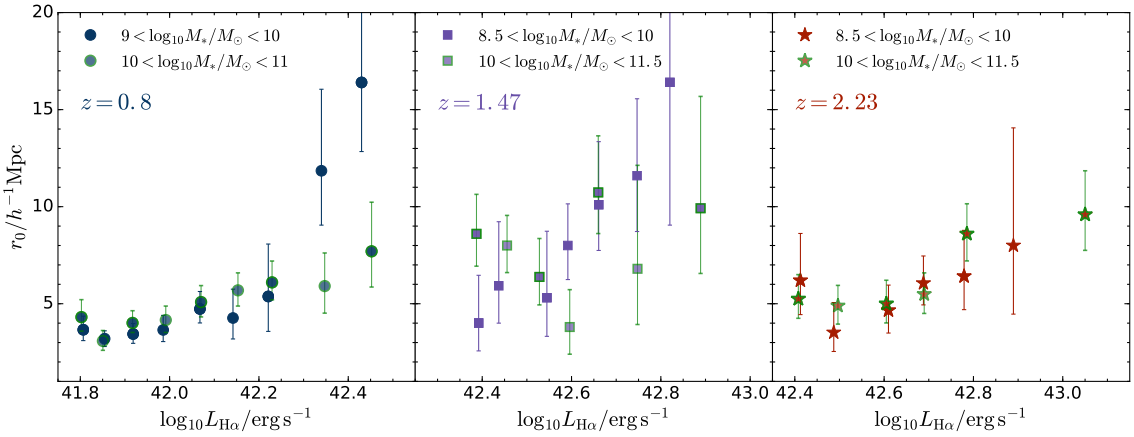

Figure 4.6 *Clustering strength as a function of $L_{\mathrm{H}\alpha}$ for HiZELS galaxies split into two stellar mass bins at each redshift. The calculated r_0 values of the two mass-binned samples are consistent at fixed mass, with the possible exception of the very highest luminosities at $z = 0.8$. This implies that the* Hα *luminosity is the physical property most strongly correlated with clustering strength for our HiZELS galaxies.*

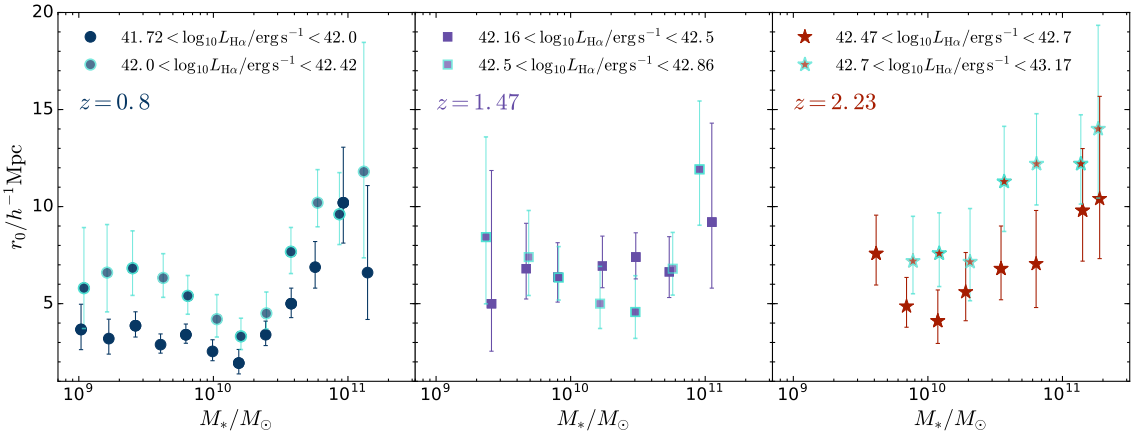

Figure 4.7 *Clustering strength as a function of stellar mass for HiZELS galaxies split into two* Hα *luminosity bins at each redshift. Both high- and low-luminosity massive galaxies are more strongly clustered than their less massive counterparts. Higher* Hα *luminosity galaxies tend to be more strongly clustered than less luminous galaxies at fixed mass. This is clear for the two largest samples, at $z = 0.8$ and $z = 2.23$. The offset in r_0 between the two luminosity bins is particularly large at low stellar masses, suggesting that low-mass galaxies with high luminosities have environmentally triggered star formation.*

luminosity subsample. We find similar trends for our second largest sample, at $z = 2.23$.

Together, our results present clear evidence for a dependence of star formation activity of low-mass galaxies on environment. For these galaxies, Hα luminosity is a better predictor of clustering strength than stellar mass. The key difference between this work and many studies of galaxy clustering that use mass-selected samples is the clean, $L_{\mathrm{H}\alpha}$-selected sample of star-forming galaxies yielded by our survey. In order to satisfy the HiZELS Hα flux limit, low stellar mass galaxies must lie significantly above the main sequence. One physical interpretation of this result is that these galaxies are highly star-forming centrals, which will soon form more stellar mass to put them on the main stellar mass - halo mass relation. Alternatively, we could be observing an increasing contribution of starbursting satellite galaxies (or galaxies that are infalling on to a massive halo and will soon become satellites) at low stellar masses.

4.4.4 Comparison of star-forming galaxies to mass-selected samples

Here, we compare the clustering of our Hα-selected samples to mass-limited samples. Hatfield et al. (2016) measure the clustering of mass-limited galaxy samples from the VIDEO survey at a very similar redshift to our $z = 0.8$ sample, at $0.75 < z < 1.00$ with median redshift $z = 0.88$.[1] Their selection is based on an apparent AB magnitude limit $K_S < 23.5$. Our observations probe slightly deeper, reaching down to $K \sim 25$, but the majority of our sources also satisfy $K < 23.5$. The important difference between our samples is the Hα flux limit of our sample. Whereas we are probing mainly the star-forming population, a substantial proportion of the Hatfield et al. (2016) sample will comprise less highly star-forming and passive galaxies. We characterize the clustering of HiZELS emitters down to the same stellar mass limits as Hatfield et al. (2016), using no luminosity cuts other than the source selection criteria described in Section 3.2. The results, shown in the upper panel of Figure 4.8, are strikingly different. At identical stellar mass limits, HiZELS r_0 values are approximately half of the VIDEO mass-selected sample r_0 values, with this difference only decreasing at the highest stellar masses. This shows that, at fixed stellar mass, star-forming galaxies are markedly less strongly clustered than the galaxy population as a

[1]Note that in Hatfield et al. (2016), r_0 is not derived from a power-law fit as in this work. Instead, r_0 is defined as the radius at which the best-fitting spatial correlation function equals unity.

whole. Note that for the lowest two stellar mass bins of Hatfield et al. (2016), the $K_S < 23.5$ selection may mean that only the reddest (and most passive, thus often most clustered) galaxies are included in the analysis, possibly biasing the points upwards relative to a fully mass-selected sample.

We now compare the clustering of our large samples of star-forming galaxies at the three HiZELS redshifts, $z = 0.8$, $z = 1.47$, $z = 2.23$, to other clustering measurements in the literature, to see whether these stark differences between differently selected samples persist at other redshifts. The lower panel of Figure 4.8 shows the results. We find that samples of passive galaxies and mass-selected samples tend to be more highly clustered than samples of star-forming galaxies at the same redshift, to at least $z \sim 2$.

Those results form a parallel story to that already presented here. While we have studied the clustering of star-forming galaxies and shown that more highly star-forming galaxies are more strongly clustered than their less star-forming counterparts at fixed stellar mass, we show here that passive galaxies are more strongly clustered than star-forming galaxies at fixed mass. How do these two apparently contradictory results fit together? Sobral et al. (2011) show that, at fixed stellar mass for $M_* < 10^{10.6}\,M_\odot$, the mean star formation rate of HiZELS galaxies increases strongly with environmental overdensity (Σ_c) across almost the full range of overdensities probed ($2 < \Sigma_c < 30$), which included field galaxies and small groups. This is consistent with the main part of our study: the clustering strength of the most highly star-forming galaxies is largest. Janowiecki et al. (2017) study the atomic hydrogen gas fraction of field and small group galaxies, finding that low-mass ($M_* \leq 10^{10.2}\,M_\odot$) galaxies in the centres of groups have gas fractions ~ 0.3 dex higher than those in the field at fixed stellar mass. They conclude that the higher star formation activity of these galaxies is driven by their higher gas availability. Sobral et al. (2011) also use the underlying photometric sample to estimate the star-forming fraction for HiZELS galaxies as function of overdensity. Here, the trends are different. The star-forming fraction increases slowly in the range $2 < \Sigma_c < 10$, but displays a sharp fall above these densities, falling to below 15% in the richest clusters. This is entirely consistent with our results: the mass-selected samples of Hatfield et al. (2016) display higher clustering strengths because they are dominated by passive galaxies in richer environments, which are not detected by the HiZELS survey due to its Hα flux selection. This interpretation, driven by the exclusion of environmentally quenched satellites from our HiZELS samples, is in line with both the low satellite fractions found in Chapter 3, and the low $M_{\rm eff}$ values for

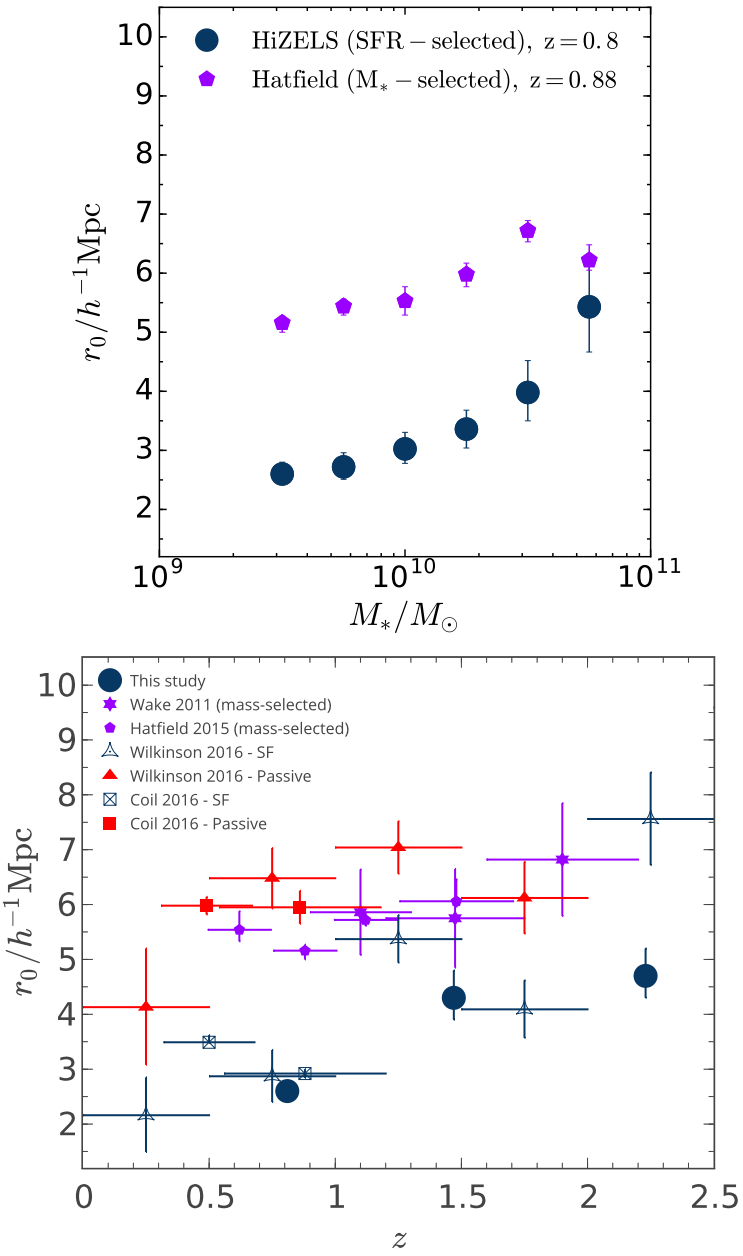

Figure 4.8 *Top: r_0 as a function of stellar mass lower limit, for HiZELS* Hα*-selected galaxies and mass-selected galaxies from Hatfield et al. (2016). At fixed stellar mass limit, the star-forming galaxies display significantly lower r_0 values, with the difference only decreasing at the highest stellar mass limits. Bottom: Comparison of whole-sample r_0 values at different redshifts. There are clear differences in derived r_0 due to sample selection. In general, samples of passive galaxies (red points) and mass-selected samples (purple points) tend to be more highly clustered than samples of star-forming galaxies at the same redshift (blue points).*

HiZELS galaxies in general.

4.4.5 The stellar mass-halo mass relation

The stellar mass to halo mass ratio (SHMR; see Section 1.2.6) is defined as the total stellar mass within a halo divided by the dark matter halo mass. It reflects the relative star formation and satellite galaxy accretion of a halo, compared to its dark matter accretion history, and is effectively a measure of the efficiency of the conversion of baryons into stars. The least massive dark matter haloes build stellar mass inefficiently due to supernova feedback, resulting in low $M_*/M_{\rm halo}$ fractions. Efficiency appears to increase towards higher halo mass, up to $M_{\rm halo} \sim 10^{12}\,M_\odot$. A consensus has emerged that haloes of this mass are most efficient at forming stars, with substantial decrease in efficiency above this halo mass (e.g. Behroozi et al., 2013; Moster et al., 2013), which is associated with AGN feedback. Birrer et al. (2014) find that the reduced stellar-to-halo mass ratio can be accounted for at high halo masses by the quenching of massive galaxies at around M^*, the knee of the stellar mass function. There is little evidence for redshift evolution in the peak of the SHMR. Here, we review one approach to modelling the SHMR, and compare our measurements to predictions.

Moster et al. (2013) follow Moster et al. (2010) in adopting a double power-law parametrisation for the SHMR. The four free parameters are fitted using populations of dark matter haloes and galaxies at redshifts from $z = 0$ to $z = 4$, specifically dark matter halo populations drawn from the Millennium and Millennium-II Simulations (Springel et al., 2005; Boylan-Kolchin et al., 2009) and galaxy populations from Li & White (2009) at low redshifts and Pérez-González et al. (2008) and Santini et al. (2012) at high redshifts. At each redshift, Moster et al. (2013) initiate an SHMR with a given set of parameters, and use this to simulate the stellar masses of galaxies within the dark matter haloes they draw from the N-body simulation at the same redshift. They then compare the stellar masses of their simulated galaxies to the observed stellar mass function, and assign the modelled SHMR a likelihood. They thus optimize the parameters of the SHMR at each redshift. By including observational errors on high-redshift stellar masses, they are able to derive models that agree well with observed stellar mass functions.

Behroozi et al. (2010) show (using another stellar mass-limited approach) that there is little difference between the SHMRs at low halo masses ($M_{\rm halo} < 10^{12}\,M_\odot$) derived when considering the total stellar mass within the halo or just that of the central galaxy. Given that we argued in Chapter 3 that the HiZELS samples

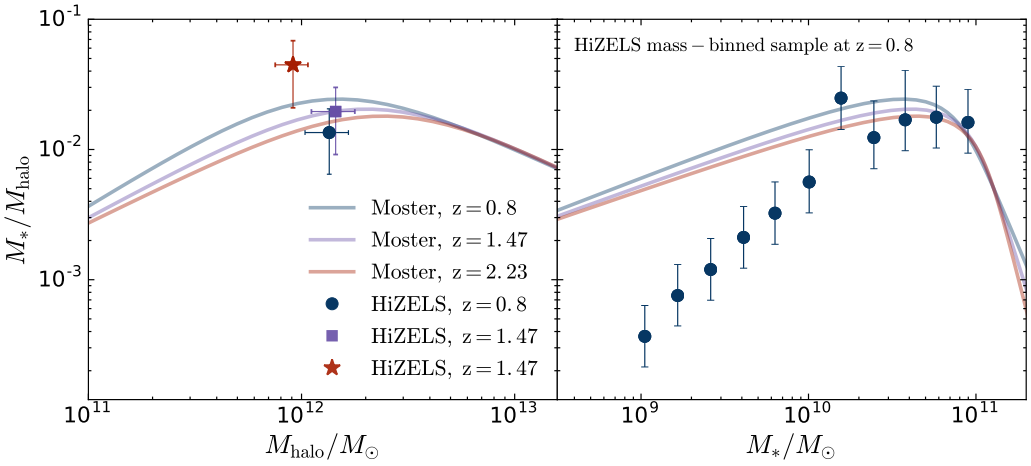

Figure 4.9 *Left: The stellar mass - halo mass relation from Moster et al. (2013), with whole HiZELS samples at each redshift overplotted. We use the effective halo mass estimated via the HOD fitting to the whole HiZELS samples at each redshift (see Chapter 3). Error bars on the y-axis represent the* 1σ *uncertainty derived from the MCMC posterior distribution, combined in quadrature with the typical errors on the stellar mass measurements (*0.23, 0.24, *and* 0.26 dex *for* $z = 0.8$, 1.47 *and* 2.23 *respectively). At all three redshifts, HiZELS galaxies occupy a region at the peak of the SHMR, where conversion of baryons into stellar mass is at a maximum. Right: The stellar mass - halo mass relation from Moster et al. (2013) as a function of stellar mass, with mass-binned HiZELS data from the* $z = 0.8$ *sample within the range* $41.72 < \log_{10} L_{\rm H\alpha}/{\rm erg\,s^{-1}} < 42.42$ *overplotted. While high-mass emitters lie on the relation predicted by Moster et al. (2013), the lowest mass* Hα *emitters lie significantly below it, which indicates that these galaxies are living in more massive haloes than would be expected for central galaxies of their stellar masses.*

are dominated by central galaxies, we use the stellar mass of HiZELS galaxies as a proxy for total stellar mass in the halo. We then compare our estimates of dark matter halo mass for HiZELS galaxies to the predictions of Moster et al. (2013). We take the same samples of galaxies within large $L_{\rm H\alpha}/L^*_{\rm H\alpha}$ bins at each of the three redshifts, as in Chapter 3. We use the average SED masses and the effective halo masses derived from HOD fitting (see Section 2.2.1) to place these samples on to the SHMR. The left-hand panel of Figure 4.9 shows that our data are in excellent agreement with the predictions of Moster et al. (2013). At all three redshifts, HiZELS galaxies occupy a region at the peak of the SHMR. They reside in haloes that are able to support maximum conversion of baryons into stellar mass.

Nevertheless, these global averages include galaxies spread over > 2 dex in stellar mass, so are not necessarily representative of all HiZELS galaxies. To

investigate this, in the right-hand panel of Figure 4.9 we place mass-selected subsamples of our $z = 0.8$ data on to the same relation. When we calculate the SHMR from the mean stellar mass and derived effective halo mass for each subsample, samples of galaxies with $M_* > 10^{10}\,M_\odot$ lie approximately on the Moster et al. (2013) relation. However, at low stellar masses, our samples lie significantly below this modelled relation. As discussed in Section 4.4.3, our low-mass galaxies reside in particularly high-mass haloes for central galaxies of their stellar mass. One possible interpretation of this is that it could be indicative of a substantial amount of stellar mass contained in galaxies that are undetectable by HiZELS within the same halo (i.e. our assumption that the halo's total stellar mass is broadly given by the HiZELS stellar mass is wrong). This points towards some of our low-mass galaxies being satellites. In that case, our low-mass galaxies would be highly star-forming satellites of a (more massive) passive central. However, this would go against the conclusion of the HOD modelling in Chapter 3 that the majority of HiZELS galaxies are centrals. Alternatively, we could be picking out starbursting low-mass centrals that will soon gain sufficient stellar mass to place them on to the main SHMR. Given only the current HiZELS observational data, it is difficult to distinguish between these scenarios. We will return to this issue in Section 4.5.5, where we compare against the EAGLE simulations.

4.5 Comparing our results to simulations

4.5.1 Overview of the EAGLE simulation

Historically, cosmological hydrodynamical simulations have struggled to reproduce observed properties of galaxy populations simultaneously with the same success as semi-analytic models. Observed statistics of galaxy populations such as stellar mass functions, luminosity functions and the detailed properties of individual galaxies such as sizes, bulge/disc masses and star formation histories were poorly matched (see Somerville & Davé, 2015, for a review). This is partly an issue of resolution: to maintain the broadest view of galaxies within the large-scale dark matter structure of the Universe, key processes that determine the detailed evolutionary path of individual galaxies such as star formation and feedback are left unresolved.

The latest generation of hydrodynamical simulations has made notable strides by attempting to improve the calibration of sub-grid models to observed

properties of galaxy populations. The Virgo Consortium's Evolution and Assembly of GaLaxies and their Environments project, EAGLE, comprises a suite of ΛCDM simulations based on SPH code GADGET 3 (Springel et al., 2005). EAGLE represents a significant improvement on previous hydrodynamical simulations due to its simple implementation of energy feedback from both massive stars and AGN. Subgrid models for these processes are calibrated using two main relations at $z = 0.1$: the galaxy stellar mass function, and the galaxy-black hole mass relation. EAGLE's success lies in its reproduction of various other observed relations (e.g. galaxy specific star formation rate distributions, passive fractions and the Tully-Fisher relation; Schaye et al. 2015) that are not explicitly used in the calibration. Artale et al. (2017) also find good agreement between the clustering of blue galaxies in EAGLE and those in the GAMA survey, concluding that these simulated and observed galaxies with similar properties occupy dark matter haloes of similar masses.

A number of EAGLE simulations are publicly available (McAlpine et al., 2016). Here, we use version Ref-L100N1504, due to its large volume (box of side length 100 Mpc, comoving) and particle number (7 billion). We select galaxies at $z = 0.87$, close to the $z = 0.8$ HiZELS redshift slice.

4.5.2 Halo environments of EAGLE galaxies

Rather than calculating halo mass via the two-point correlation function as we have done for HiZELS galaxies, we identify the halo masses of EAGLE galaxies directly. We use the total friends-of-friends (FOF) mass of the galaxy's halo, labelled as GroupMass in the EAGLE FOF table, as opposed to the subhalo mass. We identify central galaxies as those galaxies for which SubGroupNumber = 0, and satellite galaxies as galaxies with SubGroupNumber > 0. In Figure 4.10, we show the typical halo masses of subsamples of EAGLE central and satellite galaxies at $z = 0.87$. The stellar mass and star formation rates used are those within a 30 pkpc (proper, as opposed to comoving, kpc) aperture, taken from the EAGLE Aperture table. We see that the halo masses of central galaxies are strongly correlated with their positions on the SFR-stellar mass plane, with high-stellar mass galaxies residing in massive dark matter haloes. We also see hints of higher halo masses for higher luminosity low-mass central galaxies at fixed stellar mass. We quantify this in more detail in Section 4.5.3. For satellite galaxies, halo masses are less strongly correlated with stellar mass or star formation rate. This reflects the fact that much of a satellite's mass is built up at earlier times, when it is the central of its own subhalo, before this subhalo is accreted on to the larger

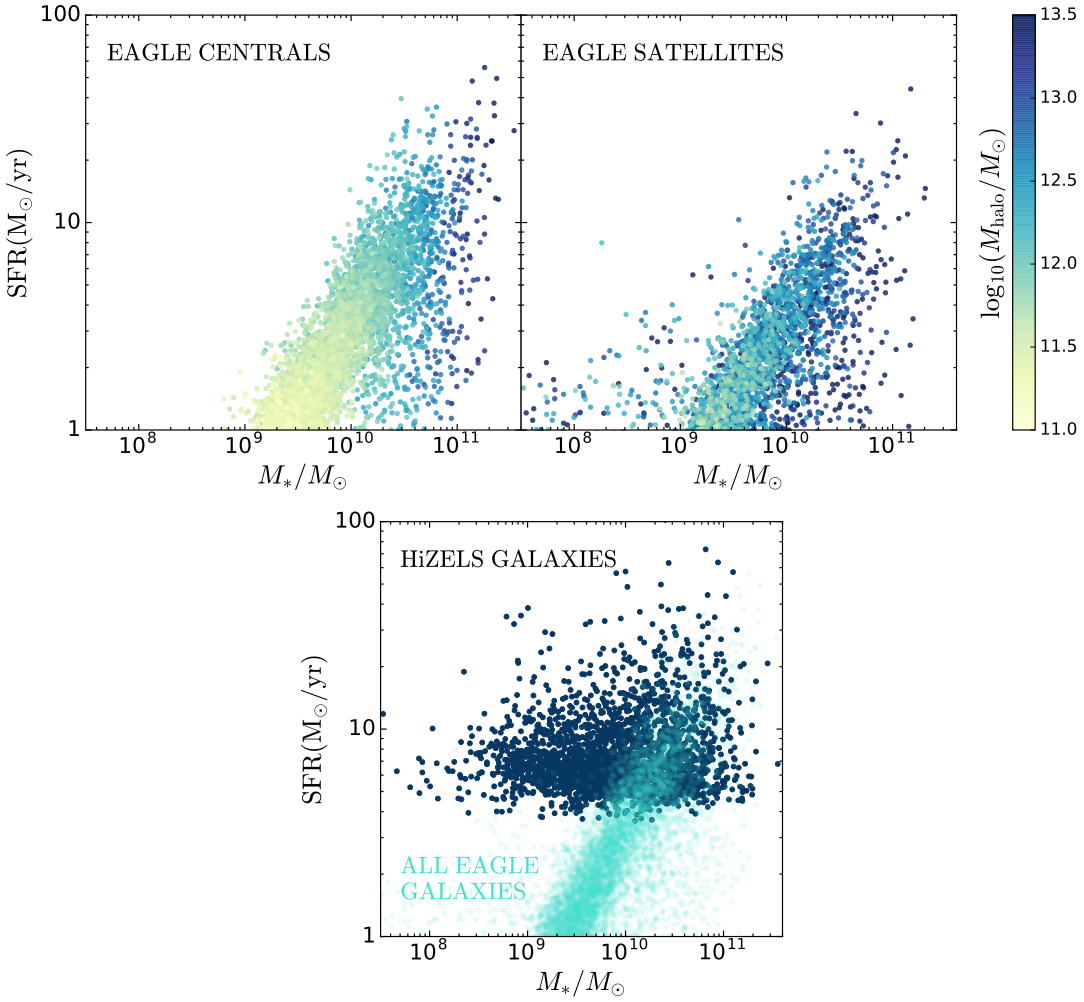

Figure 4.10 *$z = 0.87$ galaxies from EAGLE, plotted on the stellar mass - star formation rate plane using a 30 kpc (proper) aperture, colour-coded by their group halo mass. The halo masses of central galaxies (top, left-hand panel) are strongly correlated with their positions on this plane, with high stellar mass galaxies residing in massive dark matter haloes. The satellite galaxies (top, right-hand panel) have greater variance in halo mass at fixed stellar mass, due to the formation of their stellar mass in a smaller halo, before accretion on to more massive haloes. We also show the positions of $z = 0.8$ HiZELS galaxies (not colour-coded by halo mass) on the same plane (bottom panel). HiZELS star formation rates tend to be slightly higher than those of EAGLE galaxies at low stellar masses.*

halo.

4.5.3 Mass and star formation rate dependencies of halo mass from EAGLE

In Section 4.4.3, we showed that at fixed stellar mass, more highly star-forming low-mass galaxies appear more strongly clustered than their less highly star-forming counterparts. Here, we mimic these stellar mass and star formation rate selections and quantify the average halo masses of EAGLE central galaxies binned in the same way. We convert EAGLE star formation rates to rough Hα luminosities, for comparison with HiZELS, using the Kennicutt (1998) $L_{\mathrm{H}\alpha}-\mathrm{SFR}$ conversion given in Section 4.2 and assuming the same Chabrier (2003) IMF as used by EAGLE.

Our results are presented in Figure 4.11. We see a strong $M_* - M_{\mathrm{halo}}$ correlation at high stellar masses, which flattens at low stellar masses, just like we found for the HiZELS samples. At low stellar masses ($M_* \lesssim 10^{10}\,M_\odot$), average halo mass increases with star formation rate at fixed stellar mass. At high stellar masses ($M_* \gtrsim 10^{10}\,M_\odot$), average halo mass is roughly independent of star formation rate for central galaxies. This is broadly consistent with our HiZELS observational results. However, there appears to be a lack of very highly star-forming, low-mass galaxies in EAGLE (cf. Figure 4.10). EAGLE galaxies do not reach the high luminosities of HiZELS galaxies, perhaps because of insufficiently bursty star formation in the simulations, or the inability to resolve bursts on small time-scales. There are well-known tensions between EAGLE star formation rates and observations. The specific star formation rates of EAGLE star-forming galaxies are $0.2 - 0.5\,\mathrm{dex}$ below those inferred from observations, across all redshifts (Furlong et al., 2015). Despite the offset in global star formation rate density, applying the required $0.3\,\mathrm{dex}$ star formation rate offset to all star formation rates would break the agreement between simulated and observed stellar mass densities. Nevertheless, the broad trends of our observational results are supported by EAGLE: for low stellar mass central galaxies, galaxy dark matter halo mass is not a simple function of stellar mass, but also depends on the galaxy's star formation rate.

4.5.4 Physical interpretation using EAGLE

Here, we use EAGLE to investigate why our most highly star-forming HiZELS galaxies tend to reside in the most massive dark matter haloes. We study the

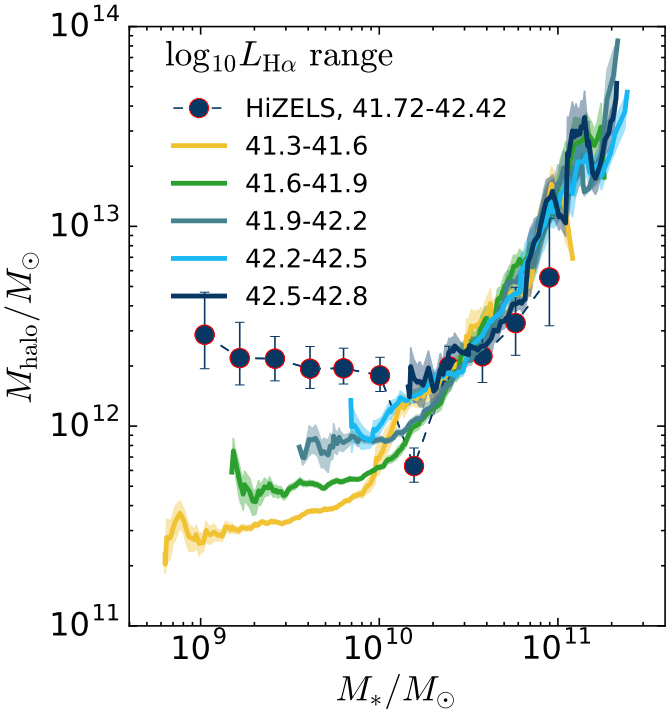

Figure 4.11 *Halo mass as a function of stellar mass for EAGLE central galaxies at* $z = 0.87$*, using moving average bins of size* 0.15 dex*. The errors plotted are the standard error on the mean. We select by EAGLE star formation rate within an aperture of 30kpc (proper), and convert to a rough* $L_{\rm H\alpha}$ *using the Kennicutt (1998) conversion, with correction to a Chabrier IMF. At low stellar masses, the most highly star-forming galaxies lie in more massive haloes than galaxies of the same mass but lower star formation rates, in line with our HiZELS observations. Low-mass HiZELS galaxies tend to reside in higher mass haloes than even the most highly star-forming EAGLE galaxies. As discussed in Section 4.5.2, this could be related to the known* $0.2-0.5$ dex *global offset between the EAGLE star formation rate density and observational measurements.*

average gas content, $M_{\rm gas}$, star formation rate, SFR, and star formation efficiency, SFE $= \frac{\rm SFR}{\rm M_{gas}}$ (the inverse of the gas depletion time-scale), as a function of halo mass and stellar mass. We include only galaxies with SFR > 0 in this analysis. Figure 4.12 shows our results. The $\log_{10} M_{\rm halo} - \log_{10} M_{\rm gas}$ relation for central galaxies is linear, and independent of galaxy stellar mass. At all stellar masses, the most massive haloes supply the most gas to their centrals. The same relation is strikingly different for satellite galaxies: the average gas mass of a satellite galaxy appears broadly independent of its halo mass, but varies significantly with stellar mass. At fixed halo mass, more massive satellite galaxies have larger gas reservoirs. This is likely due to the gas content being established earlier, prior to accretion on to a more massive halo, when the satellite galaxy's gas

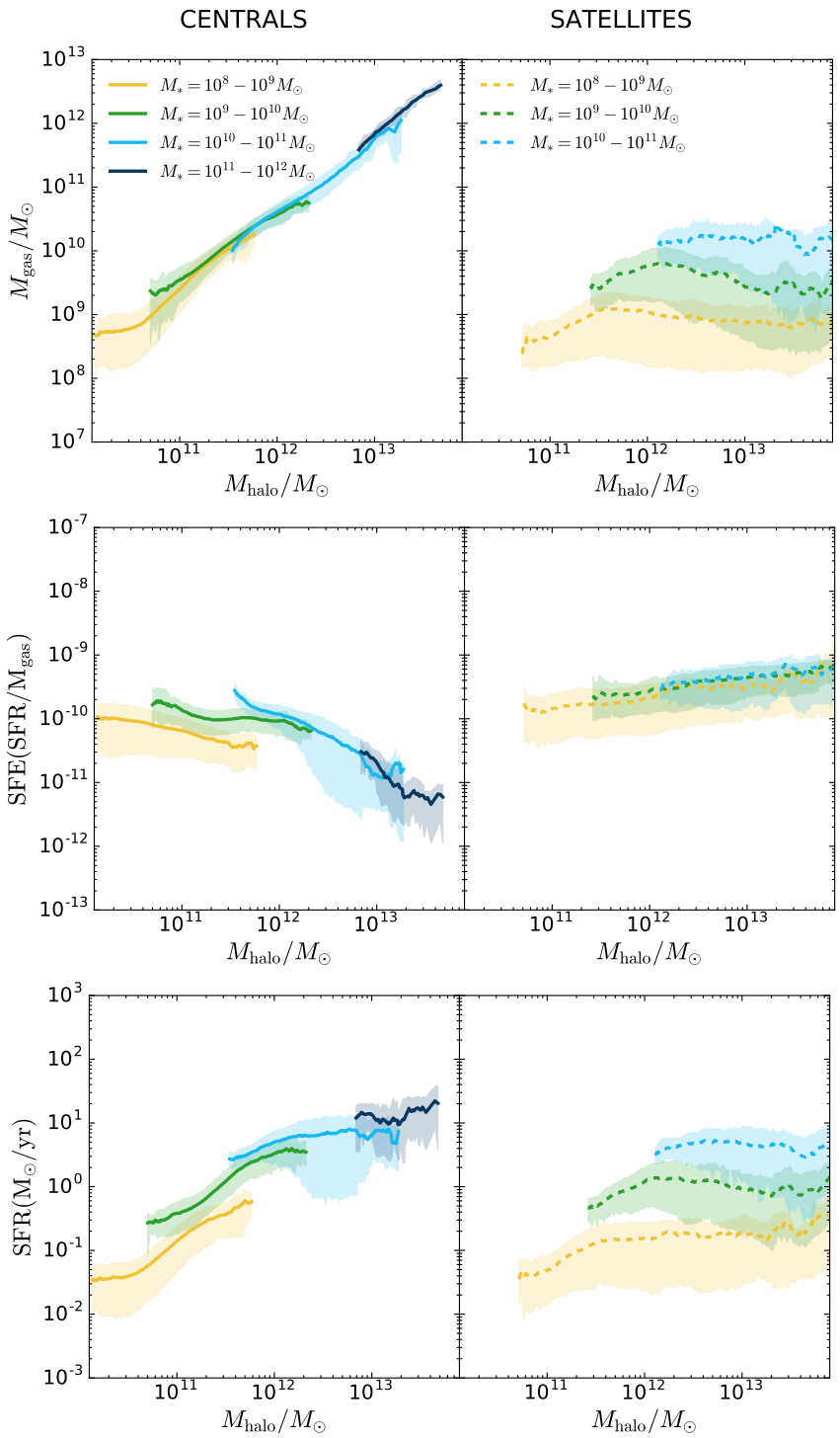

Figure 4.12 *Mean gas mass, star formation efficiency and star formation rate as a function of halo mass for satellite and central EAGLE galaxies at $z = 0.8$, with 1σ error contours. For central galaxies at all stellar masses, galaxy gas mass correlates tightly with host halo mass. Although star formation efficiency decreases with increasing halo mass, mean star formation rate increases with halo mass, for central galaxies in haloes with $M_{\rm halo} < 10^{12}\,M_\odot$. Dependencies on stellar mass are weak by comparison. In contrast, for satellites, star formation rate does not depend strongly on $M_{\rm halo}$, but more on M_*.*

mass would have correlated with the mass of its subhalo (using the mass of the EAGLE subhalo places centrals and satellites on to the same sequence), which in turn correlates more closely with stellar mass. Wetzel et al. (2013) argue that satellite galaxies retain their cold gas reservoirs upon infall and continue to form stars on long time-scales. This is broadly supported by EAGLE, where the gas mass of satellites of fixed stellar mass varies little with halo mass. The role of gas stripping in these galaxies' evolution appears to be sub-dominant.

The star formation efficiencies of central and satellite galaxies are also markedly different. SFE falls with increasing halo mass for central galaxies at all stellar masses, with a particularly steep decrease above $M_{\rm halo} \sim 10^{12}\,M_{\odot}$. Higher stellar mass centrals also have slightly higher star formation efficiencies, particularly in the lowest mass haloes. Satellite galaxies display a weak increase in SFE with halo mass (~ 1 dex over ~ 3 dex in $M_{\rm halo}$), independently of stellar mass, perhaps due to increased intracluster medium pressure in higher mass haloes (e.g. Bekki 2014).

The bottom row of Figure 4.12 shows the combination of the gas content and star formation efficiency: the mean star formation rate as a function of halo mass. Below $M_{\rm halo} \sim 10^{12}\,M_{\odot}$, mean SFR increases with $M_{\rm halo}$ for central galaxies of all stellar masses. This increase appears to be driven by gas content: gas cooling from the halo fuels star formation in central galaxies, with higher cooling rates in more massive haloes and little variation in star formation efficiency. At fixed halo mass, the more massive galaxies have higher SFRs due to increasing efficiency of gas conversion. Above $M_{\rm halo} \sim 10^{12}\,M_{\odot}$, the SFR $-$ M$_{\rm halo}$ relation appears to flatten due to decreasing star formation efficiency; there are also few star-forming galaxies at these high halo masses. Satellite galaxies display a very weak increase in SFR with halo mass at the lowest halo masses, and subsequent flattening at high halo masses. This appears to be driven by a combination of increasing star formation efficiency and decreasing gas content with increasing halo mass. At fixed halo mass, more massive satellites are more highly star-forming due to their higher gas content.

EAGLE thus provides insights into the drivers of the trends we observe with HiZELS. Simulated low-mass, highly star-forming galaxies also reside in higher mass haloes than their less highly star-forming counterparts. EAGLE shows that these trends are likely driven by gas supply rather than increased star formation efficiencies in high-mass haloes. One remaining tension is the paucity of very highly star-forming galaxies in EAGLE compared to those observed. Those EAGLE galaxies that are highly star-forming tend to be satellites (see Figure

4.10). Given the difficulties in an auto-correlation analysis of distinguishing star-forming satellites of passive centrals from star-forming centrals given only a star formation rate-selected sample, there are significant uncertainties in our satellite fraction determination discussed in Chapter 3. Nevertheless, the scarcity of highly star-forming centrals in EAGLE may well be due to star formation in the high redshift Universe being more bursty and stochastic than is simulated or recorded in the timestep-smoothed EAGLE output.

4.5.5 Insights into the SHMR from EAGLE

In Section 4.4.5, we placed our HiZELS samples on to the SHMR, considering the typical halo mass derived from clustering measurements for galaxies in different stellar mass bins. We found that mass-selected subsamples of HiZELS galaxies tend to lie below the SHMR at the lowest stellar masses. We suggested that this could be due to significant additional stellar mass within the same haloes, indicating that some of our low-mass galaxies are satellites of central galaxies which lie below the HiZELS Hα detection limits. Alternatively, these galaxies could be very highly star-forming centrals which will soon gain enough mass to place them on to the main SHMR. Here, we investigate these scenarios, to ascertain whether either star formation at HiZELS observed rates or unaccounted stellar mass within the same halo (as estimated using the EAGLE simulations) can account for the additional stellar mass needed.

We begin by calculating the increase in stellar mass required to move our HiZELS measurements diagonally on to the Moster et al. (2013) SHMR, assuming little change in halo mass. For moderate to high-mass galaxies ($M_* = 10^{10} - 10^{11}\,M_\odot$), the SHMR offsets are very small, but we find higher offsets (factors of tens) for galaxies at lower stellar masses. The required growth factors are shown as a function of stellar mass in Figure 4.13.

Next, we use the average $L_{\mathrm{H}\alpha}$ within each stellar mass bin to calculate a typical stellar mass increase over 1 Gyr of star formation if either the current star formation rate or the current specific star formation rate is maintained.

Finally, we select a sample of galaxies in EAGLE with comparable SFRs to those observed by HiZELS to evaluate the mass contribution of other galaxies in the halo. We do this in two ways. The first selects only star-forming central galaxies. This is motivated by Chapter 3, which estimated low satellite fractions for these samples. The second allows our star-forming EAGLE comparison galaxies to be either centrals or satellites. For each EAGLE comparison sample, we identify other EAGLE galaxies within the same dark matter haloes, and

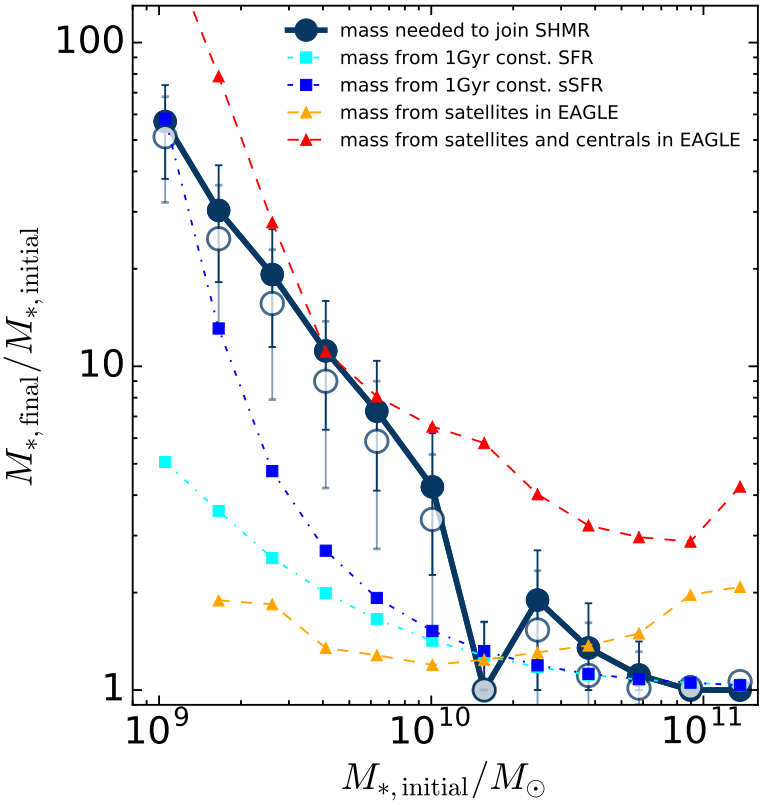

Figure 4.13 *The growth factor required, as a function of stellar mass, to bring the $z = 0.8$ HiZELS galaxies on to the SHMR (thick blue line). Closed circles use the SHMR relation from Moster et al. (2013), and open circles use the SHMR constructed using EAGLE. This indicates the approximate uncertainty on the SHMR itself. We model corrections to the mass of the HiZELS galaxy obtained under the assumption of* 1 Gyr *star formation at the measured star formation rate and specific star formation rate. For comparison, the other lines show the simulated corrections to the mass contained in the dark matter host haloes of HiZELS galaxies using EAGLE. High-mass HiZELS galaxies already lie on the SHMR. Low-mass EAGLE galaxies ($M_* < 10^{10}\,M_\odot$) with comparable star formation rates reside in dark matter haloes with significant stellar mass contributions from companion galaxies. A correction from these places HiZELS galaxies on or above the main SHMR.*

calculate a stellar mass correction, the difference between the stellar mass in the detected star-forming galaxy and the total stellar mass in the halo. These correction factors are shown in Figure 4.13.

Figure 4.13 shows that for the high-mass galaxies, which already lie on the SHMR, stellar mass is little affected by ~ 1 Gyr of star formation at either fixed SFR or fixed sSFR, and that similarly accounting for satellite galaxies makes little difference to the stellar mass of the haloes. At lower stellar masses, ongoing

star formation at fixed SFR over $\sim 1\,\mathrm{Gyr}$ time scales can produce a significant increase in stellar mass (up to a factor of a few), but falls far short of that required to bring the galaxies on to the SHMR. Likewise, 1 Gyr of star formation at fixed sSFR or considering the contribution of satellite galaxies in the same halo, both appear insufficient. Instead, it appears likely that some contribution from centrals within the same halo is required if our samples are going to move on to the SHMR, indicating that a proportion of our low-mass star-forming galaxies may be satellites of centrals with lower SFRs. Otherwise, we are detecting low-mass central galaxies that lie significantly below the SHMR, and will remain so for more than a Gyr, even if they maintain their current high specific star formation rates.

4.6 Conclusions

We have studied the clustering of intermediate redshift star-forming galaxies and its dependence on star formation rate and stellar mass. Our samples comprise Hα-selected galaxies predominantly on and above the star-forming main sequence at three redshifts, $z = 0.8$, 1.47 and 2.23. We summarize the key results here.

- At all three redshifts, we find clear evidence for a monotonic increase in clustering strength, r_0, with stellar mass above $M_* \sim 2 - 3 \times 10^{10}\,M_\odot$. At lower stellar masses, where star-forming galaxies selected by HiZELS lie significantly above the main sequence, this relation flattens. The $M_* - r_0$ relation is very different from the $\log_{10} L_{\mathrm{H}\alpha} - r_0$ relation studied in Chapter 3, which shows a significant and monotonic increase of r_0 with increasing Hα luminosity, with no flattening at the lowest luminosities.

- At fixed stellar mass, higher Hα luminosity subsamples are more strongly clustered than their less luminous counterparts. This is particularly pronounced at the lowest stellar masses ($M_* < 10^{10}\,M_\odot$). We find consistent results when we mimic our $L_{\mathrm{H}\alpha}$ cuts using the EAGLE simulations. We deduce that these highly star-forming low-mass galaxies are undergoing environmentally driven star formation. Investigating the cause of this using EAGLE reveals that our trends are likely driven by enhanced gas supply in small groups compared to the field.

- We compare our mass-binned clustering measurements of $L_{\mathrm{H}\alpha}$-selected galaxies to those obtained from mass-selected samples, and show that

measurements of galaxy clustering are strongly dependent on the galaxy selection criteria. We find that HiZELS star-forming galaxies are less strongly clustered than mass-selected galaxies at fixed stellar mass. Compilations of literature measurements confirm that passive and mass-selected samples tend to be more strongly clustered than star-forming samples back to at least $z \sim 2$. Mass-selected samples seem to be picking up many more quenched satellites in massive haloes. We argue that our results are in line with average star formation rates increasing towards group densities but decreasing at the highest cluster densities, where environmentally driven quenching plays a stronger role.

- We place HiZELS samples on the SHMR obtained empirically using mass-selected galaxy samples by Moster et al. (2013). We find that, on average, these highly star-forming galaxies lie at its peak, where baryon to stellar mass conversion is most efficient. Extending this to mass-binned subsamples, we show that high-mass HiZELS galaxies ($M_* > 10^{10}\,M_\odot$) lie on the SHMR, but that at lower stellar masses, our samples lie below the relation.

- Finally, we consider the effect of ongoing star formation and show that current star formation rates are insufficient to return low-mass galaxies to the SHMR. Using EAGLE, we find that if a proportion of these are satellites, typical stellar mass corrections from HiZELS-undetected galaxies within the same haloes can easily bring low-mass galaxies up on to the main SHMR.

In conclusion, we use the clustering of carefully selected star-forming galaxies with well-defined redshift distributions to determine their typical halo masses. We present evidence for environmentally driven star formation in low-mass galaxies, some of which lie well above the main sequence. We use the EAGLE simulation to strengthen the physical interpretation, and show that it is likely that these star formation rates are driven by increased gas content in galaxies residing in higher mass haloes.

Chapter 5

Dissecting the roles of mass and environment quenching in galaxy evolution with EAGLE

This chapter contains work that has been published as Cochrane, R. K. & Best, P. N., Monthly Notices of the Royal Astronomical Society, Volume 480, Issue 1, p.864-878.

In this Chapter, we use the EAGLE simulation to study how the connection between halo mass ($M_{\rm halo}$), stellar mass (M_*) and star formation rate (SFR) evolves across redshift. Using Principal Component Analysis we identify the key axes of correlation between these physical quantities, for the full galaxy sample and split by satellite/central and low/high halo mass. The first principal component of the $z = 0$ EAGLE galaxy population is a positive correlation between $M_{\rm halo}$, M_* and SFR. This component is particularly dominant for central galaxies in low mass haloes. The second principal component, most significant in high mass haloes, is a negative correlation between $M_{\rm halo}$ and SFR, indicative of environmental quenching. For galaxies above $M_* \sim 10^{10}\,M_\odot$, however, the SFR is seen to decouple from the $M_{\rm halo}$–M_* correlation; this result is found to be independent of environment, suggesting that mass quenching effects are also in operation. We find extremely good agreement between the EAGLE principal components and those of SDSS galaxies; this lends confidence to our conclusions. Extending our study to EAGLE galaxies in the range $z = 0 - 4$, we find that, although the relative numbers of galaxies in the different subsamples change, their principal components do not change significantly with redshift. This indicates

that the physical processes that govern the evolution of galaxies within their dark matter haloes act similarly throughout cosmic time.

5.1 Introduction

As discussed in previous Chapters, a myriad of recent work in extragalactic astrophysics has focused on revealing the physical processes which drive galaxy 'quenching', the process by which a previously star-forming galaxy halts star formation and becomes passive. Peng et al. (2010) suggest that these can be separated into two separate (and independent) quenching modes: 'mass quenching' and 'environment quenching'. The latter has been proposed to be primarily important for satellite galaxies, however the extent to which these trends are driven by local density as compared to the direct influence of the cosmic web remains unclear (e.g. Eardley et al., 2015). Peng et al. (2012) proposed that the satellite quenching process is more closely linked to local galaxy density than overall halo mass. However, others interpret the same data as indicating a stronger role of halo mass. Woo et al. (2013) show that the passive fraction of central galaxies is more correlated with halo mass at fixed stellar mass than with stellar mass at fixed halo mass. For satellite galaxies, there is a strong dependence on both halo mass and distance to the halo centre. Woo et al. (2013) suggest that local overdensity measurements can be unreliable and dependent on the number of observed group members, and instead argue that the halo mass is the key driver of quenching.

Gabor & Davé (2015) argue that both mass and environment quenching can be attributed to hot gas in massive host dark matter haloes (see also Birnboim & Dekel, 2003; Kereš et al., 2005). Below some characteristic dark matter halo mass (typically $\sim 10^{12}\,M_{\odot}$, the approximate peak of the stellar mass - halo mass relation, Moster et al. 2010), gas cooling times are short compared to the dynamical time of the dark matter halo, and cold gas accretes efficiently and forms stars (Dekel & Birnboim, 2006). Above this halo mass, cooling times are long, and the gas that accretes onto the galaxy is hot, so star formation is inefficient. Bower et al. (2017) explore this in more detail, proposing that the effectiveness of star formation-driven outflows depends on their buoyancy compared to that of the halo. Above some characteristic halo mass scale, these outflows are unable to clear gas from the galaxy, resulting in the buildup of gas in the central regions which then drives a rapid increase in black hole mass. This, in turn, heats the halo, preventing further gas accretion. Galaxies are then not replenished with fuel for

star formation, and star formation in high mass haloes is thus inefficient (see also Peng et al., 2015, for observational evidence for quenching via gas-exhaustion, or 'strangulation'). Similar arguments have been made within radio AGN feedback models, whereby the presence of hot intracluster gas in more massive dark matter haloes provides both a fuel source and an energy repository for recurrent radio AGN activity, which acts as a self-regulating feedback cycle controlling gas cooling rates and hence star formation (e.g. see the review by Heckman & Best, 2014).

Investigating whether two physically distinct quenching mechanisms are really required by the data, Zu & Mandelbaum (2016) study whether quenching is primarily driven by stellar mass or halo mass by modelling the clustering and weak lensing of galaxies in SDSS. They conclude that models in which the quenching of both central and satellite galaxies depends solely on halo mass (but in different ways) provide the best fit to observations, without the need for a second variable such as galaxy stellar mass. Furthermore, they find a critical quenching mass of $M_{\rm halo} \sim 1.5 \times 10^{12}\, h^{-1} M_{\odot}$ for both central and satellite galaxies.

Despite this work, the influence of the dark matter halo on its galaxies is not understood in detail. This is partly due to the inherent difficulties of linking galaxies to their host haloes observationally. As described in Section 1.2.6, this is normally attempted using one of two methods: Halo Occupation Distribution (Ma & Fry 2000; Peacock & Smith 2000, see Chapter 2, or Cooray & Sheth 2002 for a review) modelling, whereby the occupation of haloes as a function of mass is modelled for central and satellite galaxies separately, then fitted to clustering or weak-lensing observations; and Subhalo Abundance Matching, which traditionally assigns galaxies to dark matter haloes by ranking them by stellar mass and subhalo mass (e.g. as measured by circular velocity). This becomes more difficult when we seek to explore different populations of galaxies (i.e. those selected in terms of mass, star formation rate or colour).

In this Chapter, we take a simpler approach. We draw simulated galaxies and their host haloes directly from the EAGLE simulations (Crain et al., 2015; Schaye et al., 2015; McAlpine et al., 2016), described in Chapter 4. In Section 5.2, we introduce the sample and present the relationships between stellar mass, halo mass and star formation rate as seen by EAGLE over cosmic time. In Section 5.3, we quantify the strength of these relations using a statistical technique, Principal Component Analysis, over the redshift range $z = 0 - 4$. We also compare the $z = 0$ results to observational data from SDSS using an equivalent analysis. We discuss the implications of our results for the quenching of star formation in Section 5.4, and draw conclusions in Section 5.5.

5.2 EAGLE galaxies across cosmic time

5.2.1 Sample selection and galaxy properties

There are a number of EAGLE simulations available (McAlpine et al., 2016). We draw our galaxy samples from version Ref-L100N1504, due to its large volume (box of side length 100 Mpc, comoving) and particle number (7 billion). We select EAGLE galaxies with $M_* > 10^9\,M_\odot$. Large numbers of particles are required to sample the formation history of each galaxy, and EAGLE galaxy properties become unreliable below this stellar mass (McAlpine et al., 2016; Schaye et al., 2015). Imposing this stellar mass limit also makes comparison to observational data easier. The minimum SFR resolved by EAGLE is $\sim 10^{-3}\,M_\odot \mathrm{yr}^{-1}$ due to gas particle resolution, and some galaxies ($< 15\%$ at $z = 0$ and fewer at higher redshifts) are assigned SFR $= 0\,\mathrm{M}_\odot\mathrm{yr}^{-1}$. We exclude these SFR $= 0\,\mathrm{M}_\odot\mathrm{yr}^{-1}$ galaxies from the PCA analysis described in Section 5.3.1, since we use the logarithm of the SFR (note that our results are largely unchanged if we instead assign these galaxies a low 'limit' star formation rate).

As in Chapter 4, we use the total friends-of-friends (FOF) mass of the galaxy's halo (Davis et al., 1985), labelled as GroupMass in the EAGLE FOF table, as opposed to the subhalo mass. We identify central galaxies as those galaxies for which SubGroupNumber = 0, and satellite galaxies as galaxies with SubGroupNumber > 0. The stellar mass and star formation rates used are those within a 30 pkpc (proper, as opposed to comoving, kpc) aperture, taken from the EAGLE Aperture table.

5.2.2 Relationships between halo mass, stellar mass and SFR and evolution with redshift

As described in Section 1.2.1, galaxies have long been found to inhabit a particular region in the stellar mass - star formation rate plane, often dubbed the 'star-forming main sequence' (e.g. Noeske et al., 2007; Renzini & Peng, 2015; Lee, 2015). This broadly linear relation appears to persist with redshift (with evolving normalisation; see Figure 1.11), though its absolute normalisation and slope differ from sample-to-sample. The extent to which more passive galaxies occupy a wholly separate region of the plane has also been questioned (Eales et al., 2017).

In Figure 5.1 we plot two commonly studied relations as output by EAGLE. In the upper panels, we present the stellar mass vs star formation rate relation of EAGLE central and satellite galaxies at $z = 0$, and overplot the evolution of the

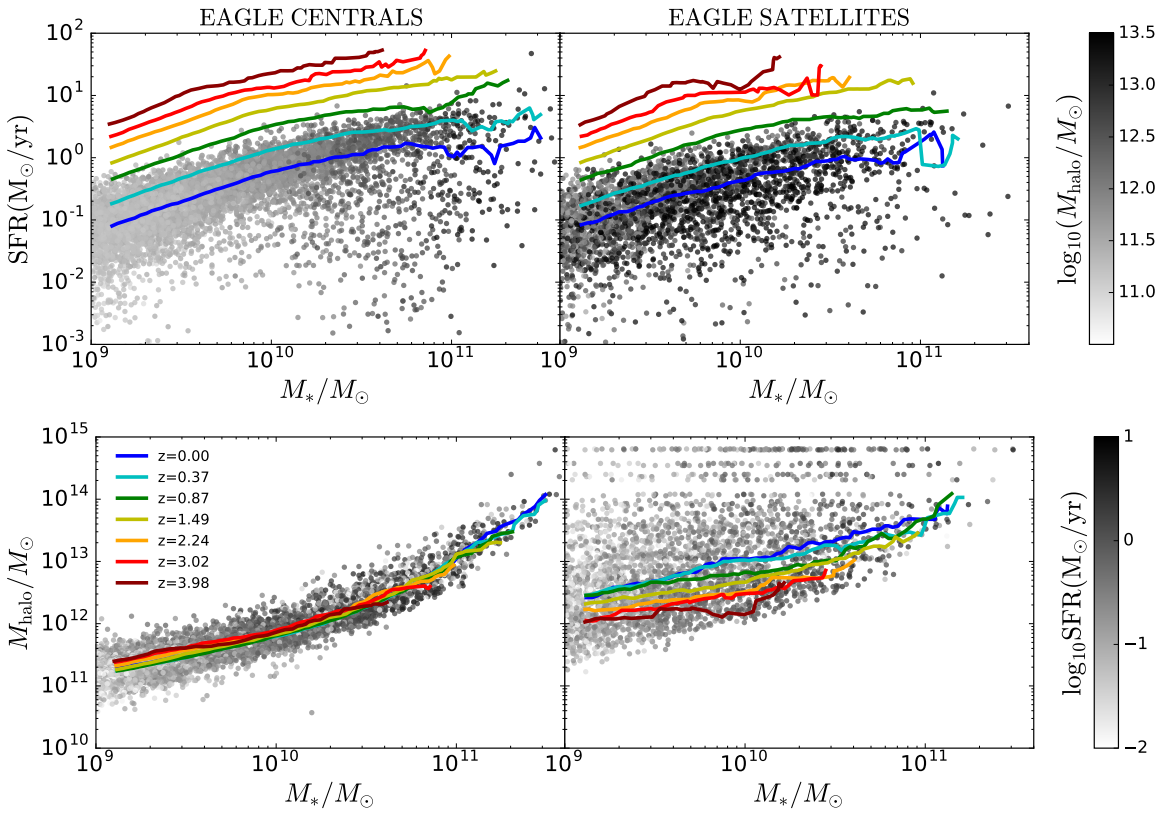

Figure 5.1 *Top: the positions of EAGLE central (left) and satellite (right) galaxies in the stellar mass - star formation rate plane at $z = 0$, colour-coded by halo mass. Bottom: the same galaxies plotted in the stellar mass-halo mass plane, colour-coded by star formation rate. On all panels, we overplot the evolution in the median relation with redshift, using a 0.25 dex running median. At fixed stellar mass, star formation rates increase substantially towards higher redshift for both central and satellite galaxies. However, the typical halo mass of central galaxies at fixed stellar mass is largely invariant with redshift.*

median relation back to $z = 4$. The evolution of this relation is fairly smooth, with both central and satellite galaxies in the simulations forming stars at a faster rate at higher redshift, for fixed stellar mass. Galaxies at $z = 0$ are colour-coded by their group halo mass. For centrals, there is a strong trend that more massive galaxies are hosted by more massive dark matter haloes, as expected. Furthermore, at lower stellar masses ($M_* < 10^{10}\,M_\odot$) there is a weak trend that (at fixed stellar mass) more highly SF galaxies reside in more massive dark matter haloes; this is discussed in more detail, and found to match observational results, in Chapter 4. Satellite galaxies inhabit similar regions of this plane, but their halo mass appears to correlate less strongly with position.

The lower panels of Figure 5.1 show stellar mass versus halo mass at different

redshifts. The relationship between stellar mass and halo mass reflects the time-integrated efficiency of stellar mass growth relative to halo growth. As found in many other studies, the host dark matter halo mass to stellar mass relation does not evolve with redshift for central galaxies. This could be because star formation in galaxies tracks the specific mass accretion rate of the halo, as described in Chapter 4 (see also Rodríguez-Puebla et al., 2016).

Some work has already used EAGLE to study these relations in detail. For example, Matthee et al. (2017) found that the scatter in stellar mass at fixed halo mass decreases with increasing halo mass, from $\sim 0.25\,\mathrm{dex}$ at $M_{\mathrm{halo}} = 10^{11}\,M_{\odot}$ to $\sim 0.12\,\mathrm{dex}$ at $M_{\mathrm{halo}} = 10^{13}\,M_{\odot}$, stressing that this scatter is not, as is often assumed, independent of halo mass. They attributed some of this scatter (up to $\sim 0.04\,\mathrm{dex}$) to the halo formation time, but found no dark matter halo property that can account for the remaining scatter. In this Chapter, we look at the role of star formation rate in driving this scatter.

5.3 Distinguishing the roles of $\mathrm{M_{halo}}$, $\mathrm{M_*}$ and SFR using Principal Component Analysis

Principal Component Analysis (PCA) is a statistical approach used to describe the variance within a dataset. Observed variables - here, halo mass, stellar mass and star formation rate - are converted into a set of uncorrelated variables, the orthogonal principal components. The first component reveals the direction of maximum variance. Successive components contain less of the variance of the population. This way, some latter components may be dominated by noise, leaving the data decomposed into fewer dimensions.

PCA has been used in a number of recent galaxy evolution studies. Bothwell et al. (2016) selected (mostly low redshift) galaxies with cold gas measurements, arguing that the relation between stellar mass, molecular gas mass and gas-phase metallicity is more fundamental than the traditional 'Fundamental Metallicity Relation' (Mannucci et al., 2010) which uses star formation rate rather than molecular gas mass. Lagos et al. (2016) used PCA to show that EAGLE galaxies occupy a nearly flat surface within the neutral gas - stellar mass - star formation rate plane, with little redshift evolution. Neither of these studies look at the role of halo mass, nor is environment studied in great detail in the follow-up work of Hashimoto et al. (2018).

In the following subsections, we identify the principal components within the 3 parameters of halo mass, stellar mass, and star formation rate, for central

and satellite galaxies within the EAGLE simulation. We also investigate the differences between the principal components of galaxies hosted by low mass haloes ($10^{10} - 10^{12}\,M_{\odot}$) and high mass haloes ($10^{12} - 10^{14}\,M_{\odot}$). This roughly splits haloes into those above and below the peak of the stellar mass - halo mass relation (SHMR), which quantifies the efficiency of stellar mass build-up as a function of dark matter halo mass (e.g. Moster et al., 2013; Behroozi et al., 2013).

5.3.1 PCA procedure

Principal Component Analysis describes data in terms of linear combinations of the input variables. Therefore, we take the logarithm of all three quantities, supplying vectors of the form $[\log_{10} M_{\rm halo}/M_{\odot},\ \log_{10} M_{*}/M_{\odot},\ \log_{10} {\rm SFR/M_{\odot}yr^{-1}}]$. We use the PCA python tool *scikit.learn* to perform the PCA analysis. Each variable is normalised to its mean and scaled to unit variance for each galaxy sample input to the PCA.

5.3.2 The whole EAGLE sample at $z = 0$

Initially, we perform PCA on our whole sample of EAGLE galaxies with $M_{*} > 10^{9}\,M_{\odot}$, within the halo mass range $M_{\rm halo} = 10^{10} - 10^{14}\,M_{\odot}$ at $z = 0$. The resulting principal components are provided in Table 5.1. The primary relation is a positive correlation between halo mass, stellar mass and star formation rate. This axis encapsulates the majority ($\sim 63\%$) of the sample variance. The secondary component is a negative correlation between halo mass and star formation rate, with little dependence on stellar mass. This reflects the tendency of galaxies in high mass haloes to have low star formation rates, broadly independent of their stellar mass, and is suggestive of environmental quenching.

Next, we divide the galaxies into four subsamples, splitting by central/satellite galaxy and by halo mass but retaining the $M_{*} > 10^{9}\,M_{\odot}$ stellar mass cut[1]. Our two halo mass bins are $10^{10}\,M_{\odot} < M_{\rm halo} < 10^{12}\,M_{\odot}$ and $10^{12}\,M_{\odot} < M_{\rm halo} < 10^{14}\,M_{\odot}$, but note that, because of the stellar mass cut applied to select only well-resolved galaxies, most of our haloes in the mass range $10^{10}\,M_{\odot} < M_{\rm halo} < 10^{12}\,M_{\odot}$ are actually at $M_{\rm halo} > 10^{11}\,M_{\odot}$. We find that the principal components vary between the four subsamples (see Table 5.1 for full details of the $z = 0$ principal components). We summarise the results here.

[1]We have tested the impact of this stellar mass cut, and find that including galaxies with lower stellar masses (e.g. imposing a lower limit of $M_{*} = 10^{8}\,M_{\odot}$), where host halo masses are typically lower, makes little difference to our results.

- For $z = 0$ central galaxies in low mass haloes ($10^{10}\,M_{\odot} < M_{\rm halo} < 10^{12}\,M_{\odot}$), $\sim$ 79% of the variance of the population is contained in PC1, which represents a positive correlation between halo mass, stellar mass and star formation rate. Note that the star formation rate is a key component in this, i.e. we don't just obtain a halo mass - stellar mass component, nor do we obtain two separate components that encode the halo mass - stellar mass and the stellar mass - SFR correlations. PC2, which contains a comparatively small $\sim 14\%$ of the variance, reflects the secondary negative correlation between star formation rate and the other two parameters. This is significantly smaller for the centrals in low mass haloes than for the $z = 0$ EAGLE sample as a whole, reflecting the low passive galaxy fraction of this subsample.

- For $z = 0$ central galaxies in high mass haloes ($10^{12}\,M_{\odot} < M_{\rm halo} < 10^{14}\,M_{\odot}$), the primary relation is solely between halo mass and stellar mass ($\sim 61\%$), with essentially no component of SFR. PC2 then represents SFR only, containing 33% of the scatter. In high mass haloes, the SFR of the central galaxy thus appears to be decoupled from its stellar mass and halo mass. Note here that since the SFR correlates with neither stellar mass nor halo mass, it is not possible to tell from this alone whether the quenching of star formation for centrals in high mass haloes is driven by stellar or halo mass. We return to this question in Section 5.3.5.

- The first principal component of satellite galaxies in low mass haloes ($10^{10}\,M_{\odot} < M_{\rm halo} < 10^{12}\,M_{\odot}$) is again between halo mass, stellar mass and star formation rate, though less variance is contained in this component than for the central galaxies in haloes of the same mass ($\sim 58\%$ compared to $\sim 79\%$). This is likely to be due to the smaller role of the group halo compared to the subhalo in the growth of the satellite galaxy. Indeed, if the subhalo mass is used instead of halo mass in the analysis, then principal components similar to those of the central galaxies are recovered. PC2 indicates scatter in the halo mass - star formation rate relation ($\sim 24\%$), and PC3 is the scatter in the stellar mass - halo mass relation ($\sim 18\%$).

- For satellite galaxies in high mass haloes ($10^{12}\,M_{\odot} < M_{\rm halo} < 10^{14}\,M_{\odot}$), the primary correlation is between stellar mass and star formation rate (55%). Although halo mass is also positively correlated with these two, it has a much weaker contribution, probably reflecting the history of the satellites, which formed most of their mass prior to accretion onto a more massive

dark matter halo. PC2 (33%) is driven by the negative correlation between halo mass and star formation rate. Stellar mass does not contribute to this component. This clearly reflects the important role of halo environment, rather than stellar mass, in quenching star formation in satellite galaxies.

We have tested changing the halo mass threshold between high and low halo mass samples. The change in principal components is quite gradual with halo mass, and our results are insensitive to the exact threshold selected.

5.3.3 Comparison to SDSS $z \sim 0$ galaxies

To compare our results from the EAGLE simulation with observations, we select galaxies with $M_* > 10^9\,M_\odot$ from the 7th data release (DR7; Abazajian et al. 2009) of the Sloan Digital Sky Survey (SDSS; York & Adelman 2000). We draw stellar masses and star formation rates from the value-added spectroscopic catalogues produced by MPA-JHU[2] (Kauffmann & Heckman, 2003; Brinchmann et al., 2004). We obtain halo mass and central/satellite estimates from the group catalogues of Yang et al. (2007). These primarily ascribe halo masses of $M_{\rm halo} > 10^{12}\,M_\odot$, so we can only reliably compare these observational data with simulated EAGLE galaxies in high mass haloes. Our final sample consists of 319,158 SDSS galaxies at $z < 0.2$.

The populations of EAGLE and SDSS galaxies are not perfectly matched, with EAGLE galaxies having lower masses and star formation rates, on average, than the observed SDSS galaxies. This is in part because the lowest mass (hence, broadly, lowest luminosity) galaxies in SDSS will only be detectable at the lowest redshifts, and hence over a smaller observed volume than is available to higher mass (luminosity) galaxies. It is also well-known that the specific star formation rates of EAGLE star-forming galaxies are $0.2-0.5$ dex below those inferred from observations, across all redshifts (Furlong et al., 2015). Nevertheless, despite these small inconsistencies in the distributions and absolute values of stellar mass and star formation rate, we are still able to make comparisons between the simulations and our data. This is because the PCA approach considers the broad trends between stellar mass, star formation rate and halo mass, and it is therefore not necessary to select a sample of galaxies from EAGLE that matches the observed population exactly. For the same reasons, we find that applying different redshift cuts to the SDSS sample, to generate a sample better matched in stellar mass

[2]http://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/

Halo mass range	PC1	Var1	PC2	Var2	PC3	Var3
Whole EAGLE sample, $M_* > 10^9\,M_\odot$, SFR $> 0\,\mathrm{M_\odot yr^{-1}}$						
$10^{10}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$	[0.52, 0.65, 0.55]	63.2%	[0.75, -0.04, -0.66]	25.8%	[0.41, -0.76, 0.51]	11.0%
EAGLE central galaxies						
$10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$, $M_* > 10^9 M_\odot$	[0.59, 0.60, 0.54]	78.6%	[0.46, 0.31, -0.83]	14.4%	[0.67, -0.74, 0.09]	7.0%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^9\,M_\odot$	[0.71, 0.71, 0.04]	61.1%	[0.07, -0.01, -1.0]	33.4%	[0.71, -0.71, 0.06]	5.5%
$10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$, $M_* > 10^{10}\,M_\odot$	[0.67, 0.62, 0.41]	47.7%	[0.15, 0.43, -0.89]	30.8%	[0.73, -0.65, -0.20]	21.5%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^{10}\,M_\odot$	[0.71, 0.71, -0.04]	58.9%	[0.04, -0.10, -0.99]	33.6%	[0.70, -0.70, 0.10]	7.5%
EAGLE satellite galaxies						
$10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$, $M_* > 10^9\,M_\odot$	[0.53, 0.61, 0.59]	58.1%	[0.84, -0.28, -0.46]	23.7%	[0.11, -0.74, 0.66]	18.2%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^9 M_\odot$	[0.39, 0.70, 0.60]	54.9%	[0.84, -0.02, -0.54]	32.6%	[0.37, -0.71, 0.60]	12.5%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^{10}\,M_\odot$	[0.72, 0.68, -0.09]	43.5%	[0.24, -0.37, -0.90]	36.4%	[0.65, -0.63, 0.44]	20.2%
SDSS central galaxies						
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^9\,M_\odot$	[0.71, 0.71, 0.03]	55.5%	[0.09, -0.05, -0.99]	33.5%	[0.70, -0.70, 0.10]	11.0%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^{10}\,M_\odot$	[0.71, 0.71, 0.03]	55.5%	[0.09, -0.05, -0.99]	33.5%	[0.70, -0.70, 0.10]	11.0%
SDSS satellite galaxies						
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^9\,M_\odot$	[0.48, 0.69, 0.54]	51.4%	[0.76, -0.02, -0.65]	31.7%	[-0.44, 0.72, -0.53]	16.9%
$10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$, $M_* > 10^{10}\,M_\odot$	[0.60, 0.71, 0.37]	47.4%	[0.52, 0.00, -0.85]	33.3%	[-0.60, 0.71, -0.37]	19.3%

Table 5.1 *Principal components of EAGLE galaxies at z=0.00 (snapshot 28), with vectors signifying* $[\log_{10} M_{\rm halo}/M_\odot,\ \log_{10} M_*/M_\odot,\ \log_{10}\mathrm{SFR/M_\odot yr^{-1}}]$. *Central and satellite galaxies with* $M_* > 10^9\,M_\odot$ *and* $\mathrm{SFR} > 0\,\mathrm{M_\odot yr^{-1}}$ *were included in the analysis, though there are no significant differences in the principal components when less massive galaxies are included, or if* $\mathrm{SFR} = 0\,\mathrm{M_\odot yr^{-1}}$ *galaxies are included at some low SFR limit. We also show principal components calculated using solely high mass galaxies (*$M_* > 10^{10}\,M_\odot$*; see Section 5.3.5). There is little cosmic evolution in the variance contained by each component (see Figure 5.2), so we detail only this one redshift, presenting others in Section 5.6. We also present the principal components for SDSS galaxies in high mass haloes, which show strong agreement with the simulated galaxies.*

distribution, does not change the principal components significantly. Thus, given that this would only reduce the sample size, we choose not to apply a further redshift selection to the SDSS data.

We perform exactly the same analysis for SDSS galaxies as for EAGLE and find excellent agreement between the principal components of the observational and simulated data for both satellites and centrals in high mass haloes at $z \sim 0$ (see Table 5.1 and Figure 5.2). For observed central galaxies in high mass haloes, the first principal component embodies the positive correlation between halo mass and stellar mass, with star formation rate decoupled from this as the second principal component. For observed satellite galaxies in high mass haloes, the key relation is between all three variables, but the secondary component, which contains $\sim 32\%$ of the variance, is the negative correlation between halo mass and star formation rate. Both the components and the magnitudes of the variance they contain are very similar to those found in EAGLE, given the same stellar mass, halo mass and central/satellite sample selections. Thus, we are confident in the conclusions that we draw from EAGLE. This strong agreement between SDSS and EAGLE also gives us further confidence in the viability of the EAGLE HOD modelling in Chapter 2, where we established the HOD parametrisation adopted in Chapters 3 and 4.

5.3.4 PCA evolution with redshift

The EAGLE catalogues span a wide range in redshifts. Therefore, it is possible to study the evolution of the principal components over cosmic time. We repeat the principal component analysis at all EAGLE redshifts back to $z \sim 4$ (Section 5.6 provides a full table of results). It is remarkable how consistent both the principal components and the variances are for most of the samples. We show the evolution in the variance contained by each principal component in the top panel of Figure 5.2. There is little evolution in these values, at fixed halo mass within the central and satellite galaxy population. In the middle and bottom panels, we plot the magnitudes of each component of the vectors themselves for the first and second principal components. These, too, show little evolution in most cases. One exception is PC2 of centrals in low-mass haloes, but this is simply noisy due to low variance in that principal component. A second exception is the SFR component of PC1 for central galaxies in high mass haloes. In the higher redshift slices of EAGLE, the star formation rate of central galaxies in high mass haloes is positively correlated with their stellar mass and host halo mass very similarly to that of lower-mass haloes. However, the star formation becomes increasingly

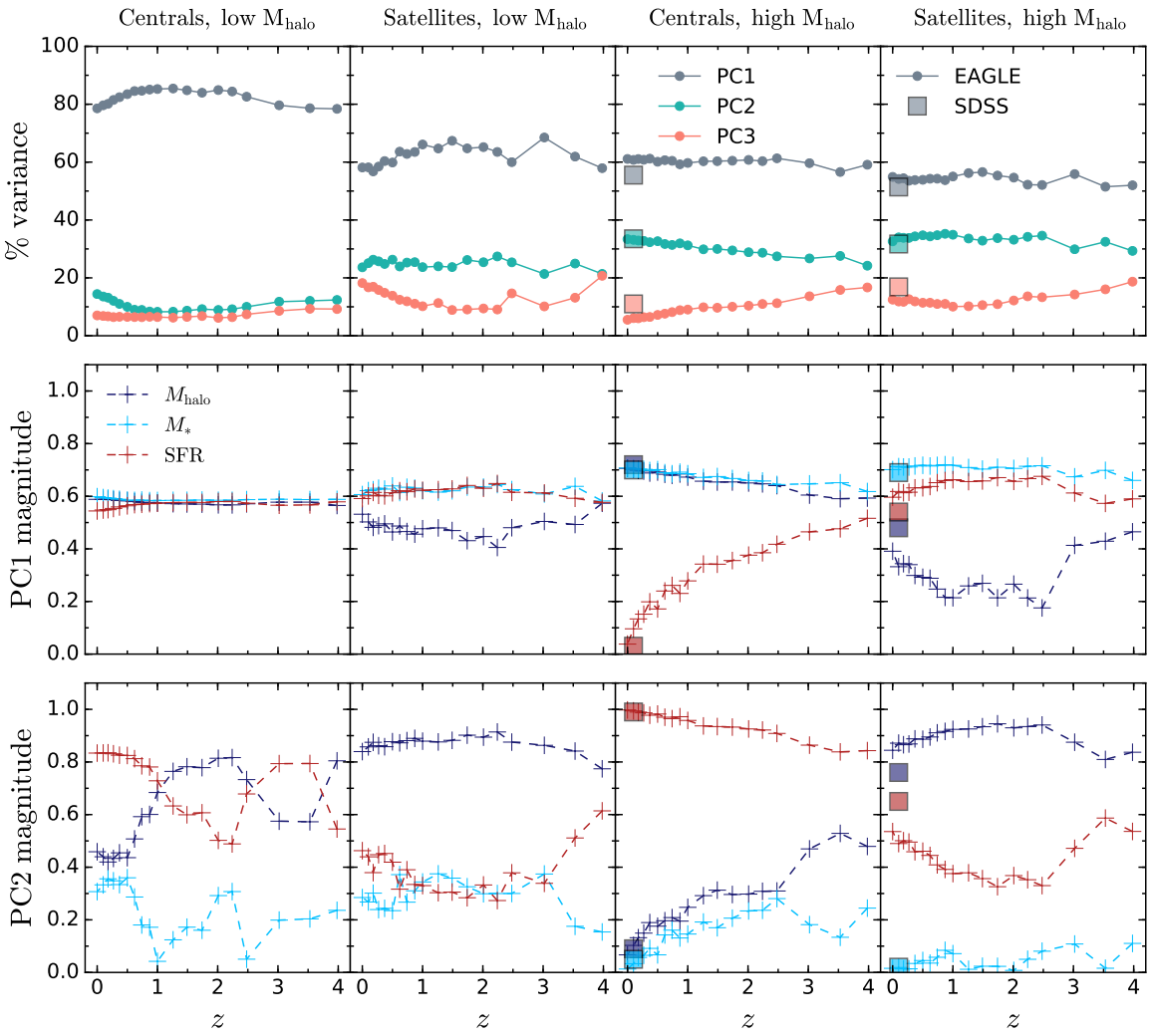

Figure 5.2 *Top: Evolution of the variance contained by the three principal components of EAGLE galaxies, split into centrals and satellites, and into high and low mass haloes. There is remarkably little evolution back to $z = 4$ for either central or satellite galaxies. Middle and bottom: The magnitudes of the vectors of the first two principal components for EAGLE galaxies at each redshift, split as above. Again, there is little evolution in these, apart from the decoupling of star formation rate from the stellar mass and halo mass for central galaxies in high mass haloes at low redshifts (and PC2 for centrals in low mass haloes, which is noisy due to low variance in this component). The square symbols show data points for SDSS galaxies in high mass haloes. These are in very good agreement with the EAGLE results at $z \sim 0$.*

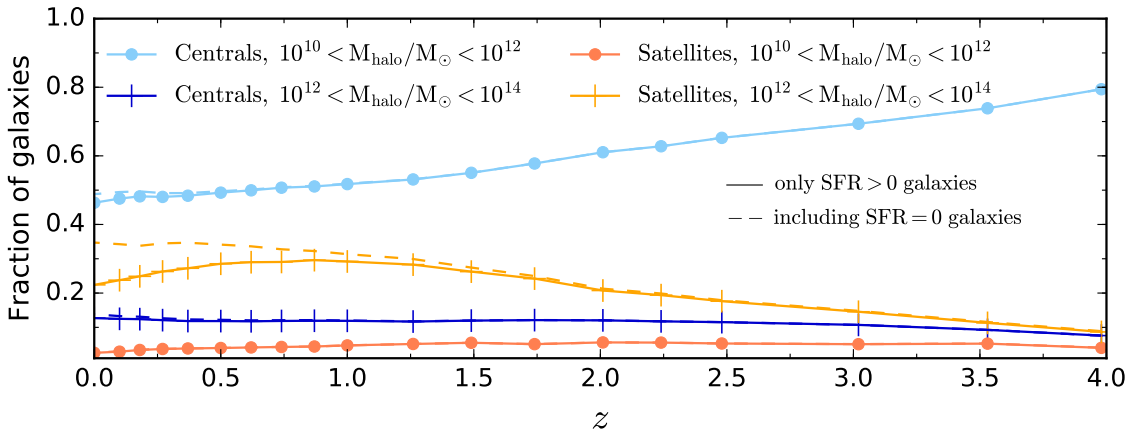

Figure 5.3 *The cosmic evolution of the distribution of $M_* > 10^9\,M_\odot$ galaxies between centrals and satellites and between low ($10^{10} < M_{\rm halo}/M_\odot < 10^{12}$) and high ($10^{12} < M_{\rm halo}/M_\odot < 10^{14}$) mass haloes. Solid lines represent the $\mathrm{SFR} > 0\,\mathrm{M_\odot yr^{-1}}$ population. The vast majority of EAGLE galaxies fall into this category. Dashed lines represent the whole EAGLE population, including $\mathrm{SFR} = 0\,\mathrm{M_\odot yr^{-1}}$ galaxies. The only population with significant numbers of $\mathrm{SFR} = 0\,\mathrm{M_\odot yr^{-1}}$ galaxies is the satellite galaxies in high mass haloes; these grow in number significantly below $z \sim 1$.*

decoupled from the halo and stellar mass towards low redshift. Interestingly, this seems not to occur for central galaxies in lower mass haloes; the positive $M_{\rm halo}$, M_*, SFR relation of PC1 holds to $z = 0$ with little change in the magnitudes of the principal components, and there is only a small decrease ($< 8\%$) in the percentage of variance contained by PC1 since $z = 1$.

The lack of evolution in the PCA view of satellite galaxies is also interesting, given that the percentage of passive galaxies evolves so strongly at low redshift, particularly at low stellar masses (see, for example, the stellar mass functions of Moutard et al., 2016). Our results indicate that the mechanism of environment quenching does not evolve with redshift. This is in line with Peng et al. (2010), who find that the environmental quenching efficiency as a function of overdensity is invariant with redshift back to $z = 1$.

Figures 5.3 and 5.4 present a complementary view of the evolution of these different populations of galaxies. In Figure 5.3, we show the fraction of the total sample that are central and satellite galaxies in haloes of different masses, as a function of redshift. The fraction of galaxies that are satellites in high mass haloes increases significantly, from $< 10\%$ at $z \sim 4$ to $\sim 30\%$ at $z \sim 0$. From $z \sim 4$ to $z \sim 1$, this reflects increasing numbers of star-forming satellites. Below $z \sim 1$, there are a large number of $\mathrm{SFR} = 0\mathrm{M_\odot yr^{-1}}$ satellite galaxies in massive haloes

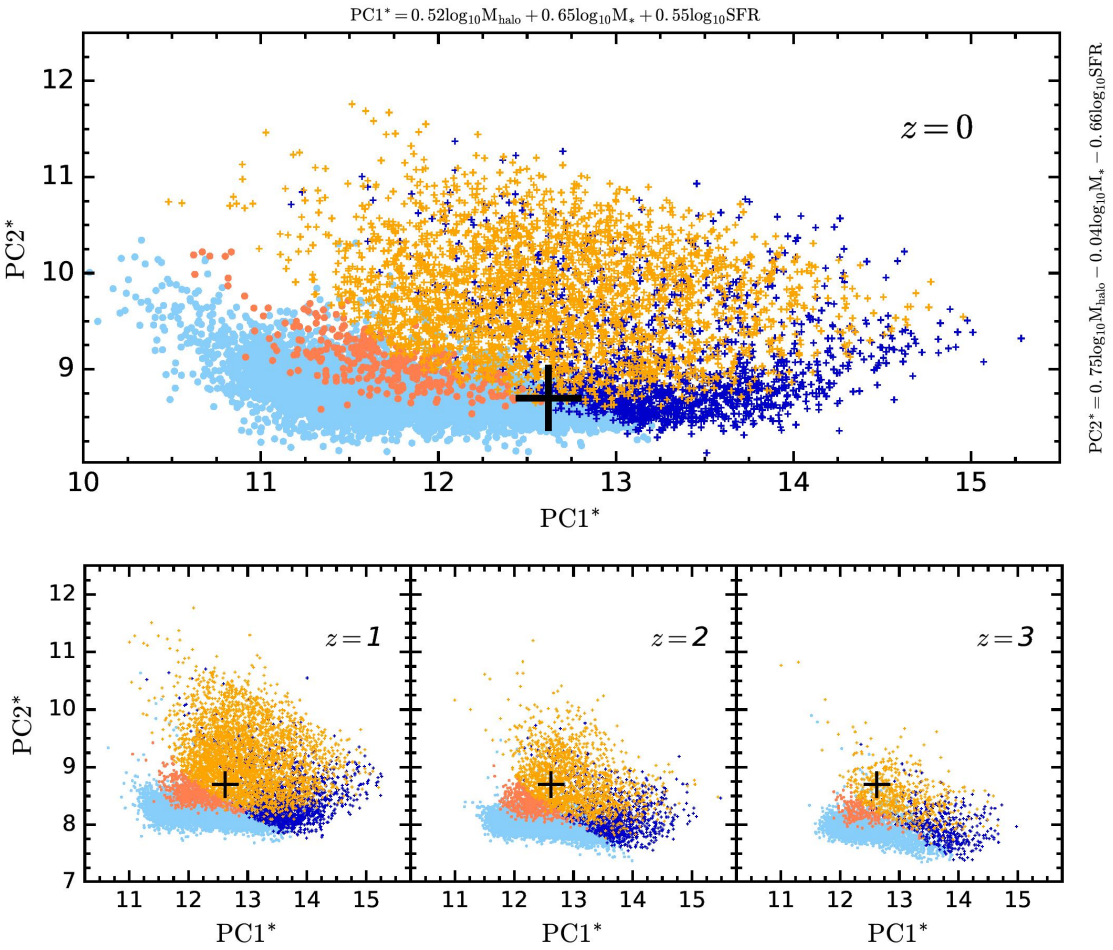

Figure 5.4 *Top: All four $z = 0$ EAGLE samples are plotted on the axes of the first two principal components of the whole $z = 0$ EAGLE sample ($M_* > 10^9\,M_\odot$, $10^{10} < M_{\rm halo}/M_\odot < 10^{14}$). The black cross shows the approximate meeting point of the populations at $z = 0$, to guide the eye. Bottom: The same plot for EAGLE galaxies at different redshifts. The numbers of galaxies in each of the four samples changes, but the typical positions of the four subsamples on the PCA plot do not (apart from moving upwards towards lower redshifts as typical star formation rates decrease, as shown by the relative position of the $z = 0$ black cross). We find that the principal correlations between $M_{\rm halo}$, M_* and* SFR *are fundamental, and independent of cosmic time.*

(around a third of EAGLE satellite galaxies in high mass haloes have unresolved SFRs).

In Figure 5.4 we plot our EAGLE subsamples on the PC1-PC2 plane defined by the whole sample at $z = 0$, as given in Table 5.1. The top panel shows the different regions of the plane that these populations occupy at $z = 0$. Each subsample occupies a fairly well-defined region of the plane. We do not show EAGLE galaxies with low star formation rates that are unresolved by EAGLE and assigned SFR $= 0\,\mathrm{M_\odot yr^{-1}}$, since their exact SFRs are unknown. Depending on the exact SFR adopted, these will lie towards the upper-left corner of the PC1-PC2 plot, naturally extending the plotted distribution of high-mass halo satellites.

In the lower panel, we show examples of the same plot at different redshifts from EAGLE, with the rough meeting point of the four populations at $z = 0$ shown by a black cross. Although the numbers of galaxies within the different classes change significantly, there is little redshift evolution in the regions of the plane occupied by galaxies within the same class, save for an overall shift upwards and to the left towards lower redshifts. This reflects typical star formation rates dropping with cosmic time.

5.3.5 Evidence for stellar mass quenching?

Peng et al. (2010) argues that mass quenching dominates the quenching of massive galaxies at $M_* > 10^{10.2}\,M_\odot$ (with the stellar mass threshold decreasing slightly towards higher redshift). If, at these high stellar masses, the role of environment is minimal, we might expect the principal components of very massive galaxies to be different. However, any such trend will be hidden in the analysis of Section 5.3.2, because the most massive galaxies are greatly outweighed by the lower mass galaxies which dominate the stellar mass function (except for central galaxies in high mass haloes, which are mostly high mass due to the strong $M_{\mathrm{halo}} - M_*$ correlation). Therefore, to probe the role of stellar mass in more detail, we select a 'high stellar mass' subsample of EAGLE galaxies with $M_* > 10^{10}\,M_\odot$, and repeat the analysis on this subsample.

We present the principal components of high mass EAGLE galaxies in Figure 5.5. Note that we do not show high mass satellite galaxies in low mass haloes, due to their scarcity. It is clear that the principal components of the central galaxies in high and low mass haloes and of the satellite galaxies in high mass haloes are extremely similar, once this stellar mass cut is made. For all three subsamples, PC1 is dominated by the correlation between halo mass and stellar

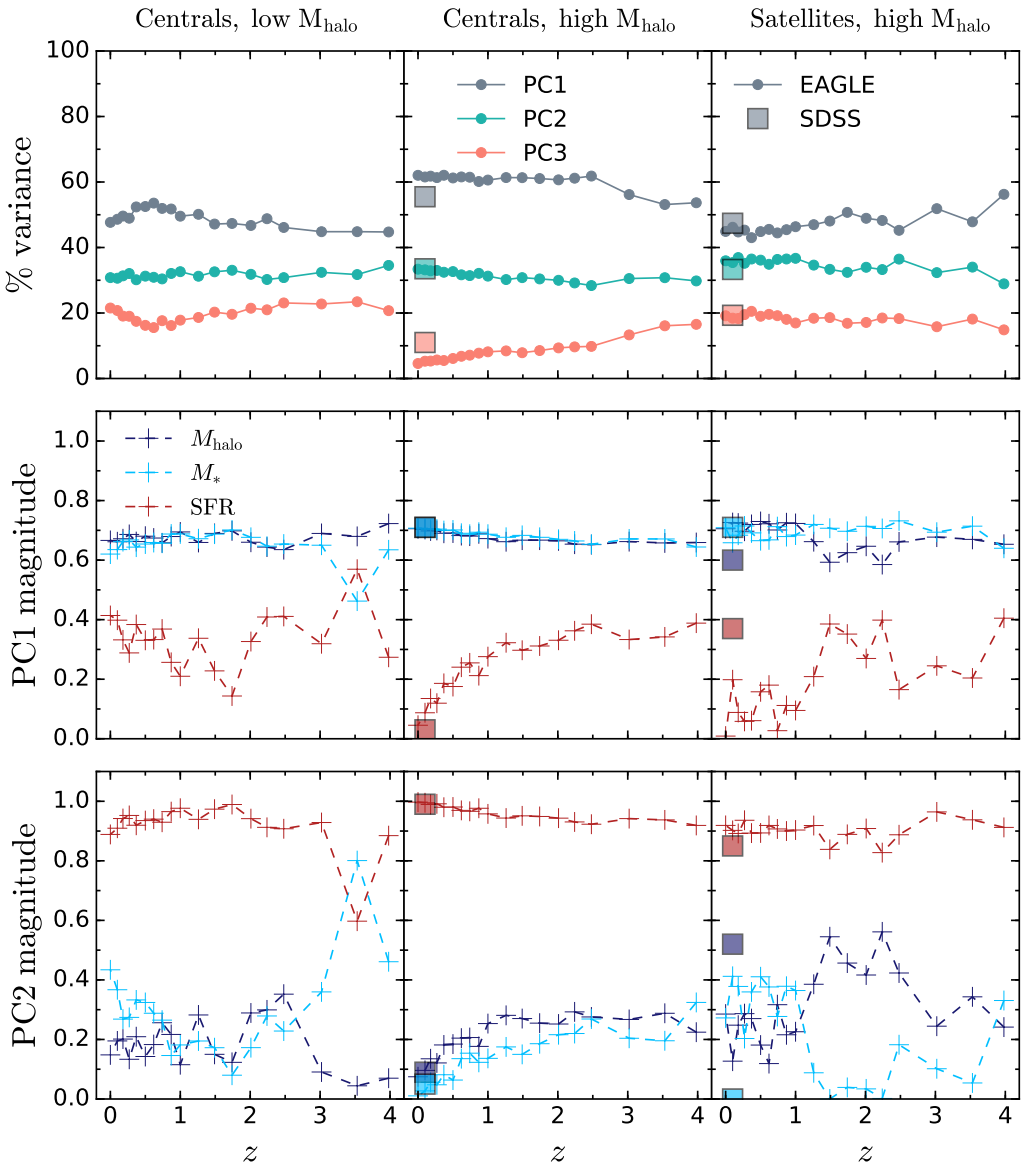

Figure 5.5 *Evolution in the variance contained by the three principal components of high mass EAGLE galaxies ($M_* > 10^{10}\,M_\odot$) in different environments. For these high stellar mass subsamples, the principal components of the central galaxies in high and low mass haloes and of the satellite galaxies in high mass haloes are very similar (there are insufficient high mass satellites in low mass haloes to investigate these). Star formation rate is largely decoupled from the stellar mass - halo mass relation, apparently due to processes related to stellar mass but not halo environment. There is little redshift evolution for the samples in any of the halo environments. Again, the SDSS data points agree fairly well with EAGLE.*

mass. While star formation rate makes a fairly small contribution towards PC1, it completely dominates PC2, reflecting the decoupling of the star formation rate from the coevolving stellar mass and halo mass. This trend is seen across halo environments (indeed, although very noisy, high mass centrals in low mass haloes are also consistent with these principal components), and across cosmic time. Thus, star formation rate decoupling in high stellar mass galaxies appears to be driven by stellar mass rather than halo environment.

Motivated by other studies of stellar mass quenching, we initially chose a 'high stellar mass' threshold of $M_* > 10^{10}\,M_\odot$. To investigate where stellar mass quenching becomes important, we repeat the principal component analysis for samples of galaxies at $z = 0$ selected using different minimum stellar mass thresholds. We present our results in Figure 5.6. We find that the principal components begin to change at $M_{\rm cut} = 10^{9.5}\,M_\odot$, where the median stellar mass of the sample is $\sim 10^{10}\,M_\odot$. Above $M_{\rm cut} = 10^{10}\,M_\odot$, the star formation rate is fully decoupled from both stellar mass and halo mass. Our results suggest that the switch in principal components occurs at $\sim 10^{10}\,M_\odot$, which is consistent with the stellar mass at which a significant change in the quenched galaxy fraction occurs.

5.4 Discussion of quenching modes

5.4.1 Environment quenching of satellite galaxies

Our results clearly indicate that halo environment plays an important role in the evolution of galaxies. For the whole sample of $z = 0$ EAGLE galaxies, the principal correlation within the population is between halo mass, stellar mass and star formation rate: more massive galaxies tend to live in higher mass haloes and have higher star formation rates. However, we identify the second principal component as a negative correlation between halo mass and star formation rate, with no stellar mass term. This points towards a predominant quenching mechanism that is driven by the halo environment, independent of stellar mass.

We find that this second component contains the most variance for satellites in high mass haloes. The first principal component of satellites in high mass haloes is dominated by the stellar mass - star formation rate correlation; for these galaxies, halo mass is less strongly coupled than for the population as a whole. This reflects the accretion histories of satellites, which have tightly correlated star formation

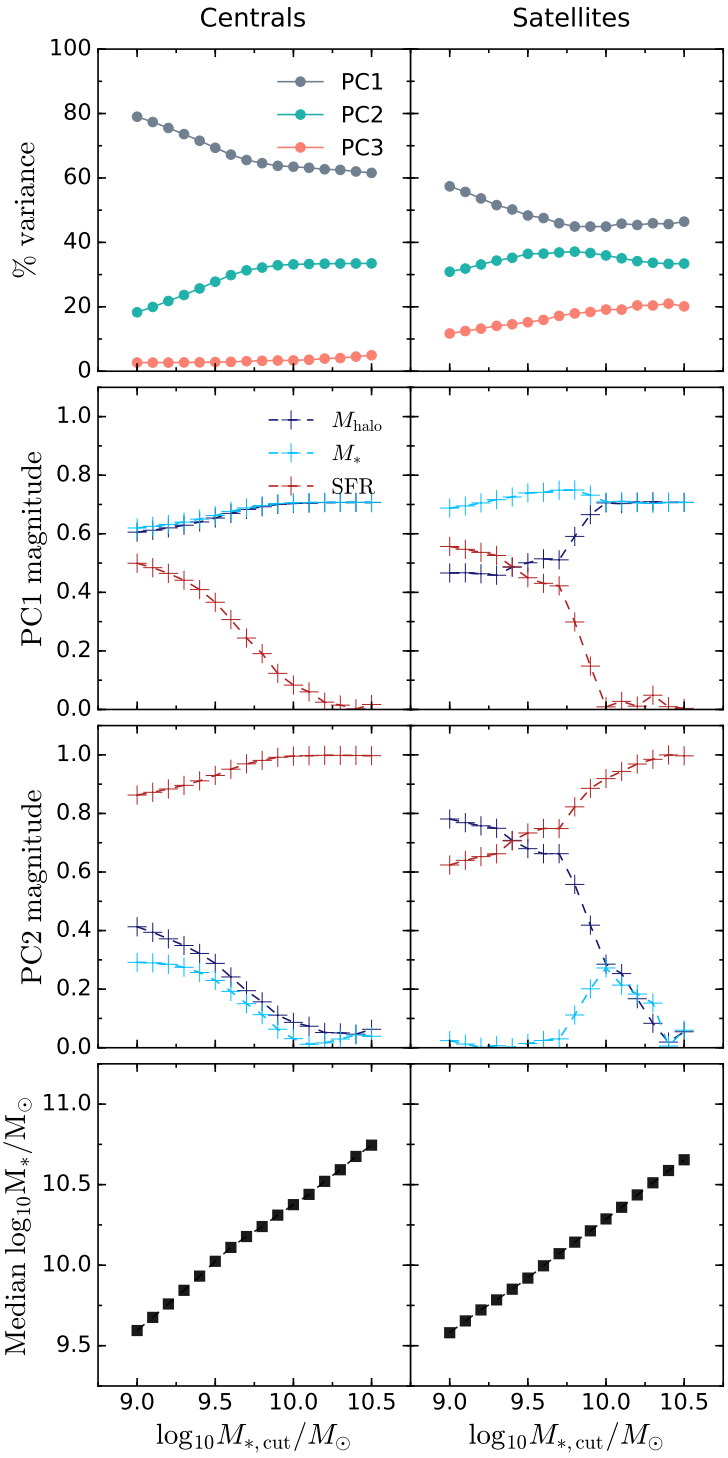

Figure 5.6 *Principal components of subsamples of central (left) and satellite (right) galaxies as a function of minimum stellar mass. EAGLE galaxies at* $z = 0$ *with* $M_* > M_{*,\rm cut}$ *in the halo mass range* $10^{10}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$ *are included in each bin. The median stellar mass of each subsample is plotted in the bottom panel. star formation rate decouples from stellar mass and halo mass above* $M_* \sim 10^{10}\,M_\odot$.

rate and stellar mass but have not grown stellar mass along with the group dark matter halo, but rather in their sub-halo. The halo mass dominates PC2, acting in opposition to the star formation rate, indicating that that environment is the dominant driver of quenching for these galaxies. This is in line with Wetzel et al. (2013), who argue that the majority of $z = 0$, $M_* < 10^{10}\,M_\odot$ passive galaxies were quenched as satellites, either within their current host halo or via pre-processing in another halo. Satellites in low mass haloes have similar principal components, but with a slightly larger contribution from stellar mass to PC2. The principal components and variance for satellites in both low and high mass haloes are fairly constant with redshift (see Figure 5.2), indicating that this halo-driven quenching acts from early times.

5.4.2 Quenching mechanisms for high mass galaxies

Comparing the results from Figure 5.2 and 5.5, it is clear that the principal components of the whole sample of satellite galaxies in high mass haloes (which is dominated by lower stellar mass satellites) and of the high stellar mass only satellite galaxy sample are very different. Satellite galaxies with high stellar masses have star formation rates decoupled from their stellar masses and halo masses in PC1. The similarity of the principal components of these high stellar mass satellite galaxies in high mass haloes to those of high mass centrals in the same haloes, and to high mass centrals in lower mass haloes, suggests that it is the stellar mass rather than the halo environment which is important in this decoupling. However, in all cases star formation rate is equally decoupled from the halo mass, so this remains difficult to prove.

Here, we consider whether these results could be consistent with work proposing that halo mass is also the underlying mechanism of stellar mass quenching. For this, it is important to consider the assembly histories of galaxies, since high stellar mass satellite galaxies are likely to have spent time forming stars building stellar mass as centrals within high mass haloes. As argued by Gabor & Davé (2015), quenching could have preceded satellite accretion and been driven by the halo mass of the previous halo. Given that the satellite's stellar mass will be tightly correlated with the mass of the previous halo, rather than that of the new halo, past halo mass quenching then looks like stellar mass quenching.

For the high stellar mass satellite galaxies, we investigate this by examining their halo histories. At each EAGLE timestep, we identify the progenitors of the $z = 0$ galaxies, via the EAGLE 'main branch' (see McAlpine et al. 2016,

but note that our results are the same if we instead manually select the most massive progenitor at each redshift). We find that 97% of $z = 0$ high stellar mass satellites in high mass haloes have primary progenitors that were central galaxies at $z < 1.5$. We collect the most recent central primary progenitors and perform the same principal component analysis on these (note that they span a range of redshifts, $0.1 < z < 1.5$, as different galaxies first become satellites at different times). The principal components we find (PC1, PC2, PC3 = [0.71, 0.71, -0.04], [0.02, 0.04, 1.0], [0.71, -0.70, 0.02], Var1, Var2, Var3 = 58.4%, 33.3%, 8.3%) are very similar to those of central galaxies in high mass haloes; thus, at the time that these galaxies became satellites, their star formation rate was already decoupled from both stellar and halo mass. Therefore, from this population we are unable to determine whether it is stellar mass or halo mass that drives the quenching of star formation.

More insight may be gained by looking at high stellar mass central galaxies in low mass haloes. Compared to low mass galaxies in equally massive haloes, star formation rate is less strongly coupled to halo mass and stellar mass in PC1 for these galaxies. As for the the other high mass galaxy subsamples, PC2 is dominated by SFR. Since the halo mass is low, these objects appear to present the most direct evidence for stellar mass-driven quenching.

However, it is important to consider how galaxies with such unusually high stellar-to-halo mass ratios formed. Gabor & Davé (2015) find a population of red central galaxies living in low mass haloes within their simulations, which comprise former satellite galaxies that were ejected from a more massive halo following halo-driven quenching. If this is the case for the bulk of these high stellar mass centrals in low mass haloes, then this would remove evidence for stellar mass being the driving factor. We therefore search the progenitors of EAGLE high mass galaxies in low mass haloes to determine whether our galaxies assembled this way. We find that only $\sim 17\%$ of $z = 0$ EAGLE galaxies have a primary progenitor at $z < 1.5$ that was a satellite. Excluding these galaxies does not lead to a significant change in the principal component analysis. This suggests that the decoupling of star formation rate in these galaxies is driven more directly by the high stellar masses of the galaxies than by their halo mass (although we cannot fully exclude that some other halo property, which also gives rise to the unusually high stellar mass-to-halo mass ratio, is responsible). The similarity of the principal components for all of these three high stellar mass samples then suggests that stellar mass driven quenching is important in all high stellar mass galaxies. Our analysis confirms that this becomes significant above $\sim 10^{10}\,M_{\odot}$.

5.5 Conclusions

In this Chapter we study the halo environments of galaxies in the EAGLE simulations, focusing on how dark matter halo mass relates to two key baryonic galaxy properties: stellar mass and star formation rate. We apply the statistical technique Principal Component Analysis to EAGLE galaxies, with comparison to observational data from SDSS. Our main results are presented here.

- We find a clear primary correlation between host halo mass, galaxy stellar mass and star formation rate. This correlation is particularly dominant for central galaxies in low mass haloes. It demonstrates the important role that dark matter haloes play in fuelling star formation in galaxies.

- We find strong evidence for environment-driven quenching in satellite galaxies via an anticorrelation between halo mass and star formation rate in the second principal component.

- We present evidence for an alternative mass-driven quenching mechanism at high stellar mass. This appears to be independent of environment and to set in at $\sim 10^{10}\,M_{\odot}$.

- Crucially, we find excellent agreement between the principal components derived for EAGLE simulated galaxies and observed galaxies drawn from SDSS at $z \sim 0$, for all sub-populations studied. This gives confidence in the validity of the results derived from EAGLE.

- Probing EAGLE galaxies back to $z = 4$, we find that the principal components of galaxies within each class do not evolve significantly with redshift, despite changes in the numbers of galaxies in each class and an overall shift towards lower star formation rates at low redshifts. The only exception is centrals in high mass haloes. For these galaxies, SFR becomes somewhat more decoupled towards low redshift. The overall redshift-independence of the principal components suggests that the physical mechanisms driving the evolution of galaxies do not evolve strongly over cosmic time.

Overall, it is clear that host dark matter haloes play a key role in fuelling and quenching star formation in galaxies at all redshifts. We show that this role differs for central and satellite galaxies in low/high mass dark matter haloes. However, within these sub-classes, the principal relations between halo mass, stellar mass and star formation rate, hold across cosmic time.

5.6 Tables of principal components at each redshift

Redshift	**PC1**	**PC1 Var**	**PC2**	**PC2 Var**	**PC3**	**PC3 Var**
Centrals, $10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$						
0.00	[0.59, 0.60, 0.54]	78.6%	[0.46, 0.31, -0.83]	14.4%	[0.67, -0.74, 0.09]	7.0%
0.10	[0.59, 0.59, 0.55]	79.6%	[0.44, 0.33, -0.83]	13.6%	[0.68, -0.73, 0.07]	6.8%
0.18	[0.59, 0.59, 0.55]	80.1%	[0.42, 0.36, -0.83]	13.2%	[0.69, -0.72, 0.04]	6.7%
0.27	[0.59, 0.59, 0.55]	81.5%	[0.43, 0.35, -0.83]	12.1%	[0.68, -0.73, 0.05]	6.4%
0.37	[0.58, 0.59, 0.56]	82.4%	[0.45, 0.33, -0.83]	11.0%	[0.67, -0.74, 0.07]	6.6%
0.50	[0.58, 0.59, 0.56]	83.5%	[0.44, 0.36, -0.82]	10.0%	[0.69, -0.73, 0.05]	6.5%
0.62	[0.58, 0.58, 0.57]	84.6%	[0.51, 0.29, -0.81]	9.0%	[0.64, -0.76, 0.13]	6.5%
0.74	[0.58, 0.59, 0.57]	84.7%	[0.59, 0.18, -0.79]	8.9%	[0.56, -0.79, 0.24]	6.4%
0.87	[0.58, 0.58, 0.57]	85.1%	[0.60, 0.17, -0.78]	8.4%	[0.55, -0.79, 0.25]	6.6%
1.00	[0.57, 0.58, 0.57]	85.3%	[0.68, 0.04, -0.73]	8.3%	[0.45, -0.81, 0.38]	6.5%
1.26	[0.57, 0.58, 0.58]	85.4%	[0.76, -0.12, -0.63]	8.3%	[0.30, -0.80, 0.52]	6.3%
1.49	[0.57, 0.58, 0.58]	84.8%	[0.78, -0.17, -0.60]	8.7%	[0.25, -0.79, 0.56]	6.5%
1.74	[0.57, 0.59, 0.58]	84.0%	[0.78, -0.16, -0.61]	9.2%	[0.26, -0.79, 0.55]	6.8%
2.01	[0.57, 0.58, 0.58]	84.9%	[0.81, -0.29, -0.50]	8.9%	[0.12, -0.76, 0.64]	6.2%
2.24	[0.57, 0.58, 0.58]	84.4%	[0.82, -0.31, -0.49]	9.2%	[0.11, -0.75, 0.65]	6.4%
2.48	[0.57, 0.59, 0.57]	82.6%	[0.73, -0.05, -0.68]	10.0%	[0.37, -0.81, 0.46]	7.4%
3.02	[0.58, 0.59, 0.57]	79.6%	[0.57, 0.20, -0.79]	11.8%	[0.58, -0.78, 0.22]	8.6%
3.53	[0.58, 0.59, 0.57]	78.6%	[0.57, 0.20, -0.79]	12.1%	[0.58, -0.78, 0.22]	9.3%
3.98	[0.56, 0.59, 0.58]	78.4%	[0.80, -0.24, -0.55]	12.4%	[0.18, -0.77, 0.61]	9.2%

Redshift	PC1	PC1 Var	PC2	PC2 Var	PC3	PC3 Var
Satellites, $10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$						
0.00	[0.53, 0.61, 0.59]	58.1%	[0.84, -0.28, -0.46]	23.7%	[0.11, -0.74, 0.66]	18.2%
0.10	[0.50, 0.62, 0.60]	58.2%	[0.86, -0.27, -0.44]	25.1%	[0.11, -0.74, 0.67]	16.7%
0.18	[0.48, 0.62, 0.61]	56.8%	[0.87, -0.30, -0.38]	26.3%	[0.05, -0.72, 0.69]	16.9%
0.27	[0.49, 0.63, 0.60]	58.5%	[0.86, -0.24, -0.45]	25.7%	[0.14, -0.74, 0.66]	15.8%
0.37	[0.50, 0.63, 0.60]	60.4%	[0.86, -0.24, -0.45]	24.8%	[0.14, -0.74, 0.66]	14.8%
0.50	[0.46, 0.64, 0.61]	59.9%	[0.88, -0.23, -0.42]	26.3%	[0.12, -0.73, 0.67]	13.9%
0.62	[0.49, 0.61, 0.62]	63.6%	[0.87, -0.37, -0.32]	23.9%	[-0.04, -0.70, 0.72]	12.4%
0.74	[0.47, 0.63, 0.62]	62.8%	[0.88, -0.27, -0.39]	25.3%	[0.08, -0.73, 0.68]	11.9%
0.87	[0.46, 0.63, 0.63]	63.5%	[0.89, -0.31, -0.33]	25.4%	[0.02, -0.71, 0.70]	11.1%
1.00	[0.48, 0.62, 0.62]	66.1%	[0.88, -0.34, -0.33]	23.7%	[-0.01, -0.70, 0.71]	10.2%
1.26	[0.48, 0.62, 0.63]	64.7%	[0.88, -0.37, -0.30]	24.0%	[-0.05, -0.69, 0.72]	11.3%
1.49	[0.47, 0.62, 0.63]	67.4%	[0.88, -0.36, -0.31]	23.7%	[-0.04, -0.70, 0.72]	8.9%
1.74	[0.43, 0.63, 0.64]	64.7%	[0.90, -0.33, -0.28]	26.2%	[-0.03, -0.70, 0.71]	9.1%
2.01	[0.45, 0.63, 0.63]	65.2%	[0.89, -0.30, -0.33]	25.4%	[-0.02, -0.71, 0.70]	9.4%
2.24	[0.41, 0.64, 0.65]	63.5%	[0.91, -0.30, -0.27]	27.4%	[-0.02, -0.70, 0.71]	9.1%
2.48	[0.48, 0.63, 0.62]	60.0%	[0.88, -0.30, -0.38]	25.4%	[0.05, -0.72, 0.69]	14.6%
3.02	[0.50, 0.61, 0.61]	68.5%	[0.86, -0.37, -0.34]	21.4%	[-0.02, -0.70, 0.71]	10.1%
3.53	[0.49, 0.64, 0.59]	61.9%	[0.84, -0.18, -0.51]	24.9%	[0.22, -0.75, 0.62]	13.1%
3.98	[0.57, 0.58, 0.58]	57.9%	[0.77, -0.15, -0.61]	21.4%	[0.27, -0.80, 0.54]	20.7%

Table 5.2 *Principal components of* $M_* > 10^9\,M_\odot$, $\mathrm{SFR} > 0\,\mathrm{M_\odot yr^{-1}}$, *central and satellite EAGLE galaxies in low mass haloes* ($10^{10}\,M_\odot < M_{\rm halo} < 10^{12}\,M_\odot$) *at each redshift. Vectors have ordering* $[\log_{10} M_{\rm halo}/M_\odot,\ \log_{10} M_*/M_\odot,\ \log_{10} \mathrm{SFR/M_\odot yr^{-1}}]$.

Redshift	**PC1**	**PC1 Var**	**PC2**	**PC2 Var**	**PC3**	**PC3 Var**
Centrals, $10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$						
0.00	[0.71, 0.71, 0.04]	61.1%	[-0.07, 0.01, 1.00]	33.4%	[0.71, -0.71, 0.06]	5.5%
0.10	[0.70, 0.71, 0.10]	60.8%	[-0.10, -0.03, 0.99]	33.1%	[0.70, -0.71, 0.05]	6.1%
0.18	[0.70, 0.70, 0.13]	61.1%	[-0.13, -0.06, 0.99]	32.9%	[0.70, -0.71, 0.05]	6.0%
0.27	[0.70, 0.70, 0.15]	60.8%	[-0.15, -0.07, 0.99]	32.8%	[0.70, -0.71, 0.06]	6.4%
0.37	[0.69, 0.70, 0.20]	61.2%	[-0.19, -0.09, 0.98]	32.3%	[0.70, -0.71, 0.07]	6.5%
0.50	[0.69, 0.70, 0.17]	60.1%	[-0.18, -0.07, 0.98]	32.7%	[0.70, -0.71, 0.08]	7.2%
0.62	[0.68, 0.69, 0.24]	60.7%	[-0.20, -0.14, 0.97]	31.7%	[0.70, -0.71, 0.04]	7.6%
0.74	[0.68, 0.69, 0.26]	60.4%	[-0.21, -0.16, 0.96]	31.4%	[0.70, -0.71, 0.03]	8.2%
0.87	[0.68, 0.69, 0.23]	59.2%	[-0.20, -0.13, 0.97]	31.9%	[0.70, -0.71, 0.05]	8.8%
1.00	[0.67, 0.69, 0.28]	59.7%	[-0.25, -0.15, 0.96]	31.2%	[0.70, -0.71, 0.07]	9.0%
1.26	[0.66, 0.67, 0.34]	60.3%	[-0.29, -0.19, 0.94]	29.9%	[0.69, -0.72, 0.07]	9.8%
1.49	[0.65, 0.67, 0.34]	60.3%	[-0.31, -0.17, 0.93]	30.0%	[0.69, -0.72, 0.10]	9.7%
1.74	[0.65, 0.67, 0.36]	60.5%	[-0.30, -0.21, 0.93]	29.5%	[0.70, -0.72, 0.06]	10.0%
2.01	[0.65, 0.66, 0.38]	60.8%	[-0.30, -0.23, 0.93]	28.9%	[0.70, -0.71, 0.04]	10.4%
2.24	[0.65, 0.66, 0.39]	60.3%	[-0.31, -0.24, 0.92]	28.7%	[0.70, -0.72, 0.05]	11.0%
2.48	[0.64, 0.65, 0.42]	61.3%	[-0.31, -0.28, 0.91]	27.4%	[0.70, -0.71, 0.02]	11.3%
3.02	[0.60, 0.65, 0.46]	59.6%	[-0.47, -0.18, 0.86]	26.7%	[0.64, -0.74, 0.19]	13.6%
3.53	[0.59, 0.65, 0.48]	56.6%	[-0.53, -0.13, 0.84]	27.6%	[0.61, -0.75, 0.27]	15.8%
3.98	[0.59, 0.62, 0.52]	59.1%	[-0.48, -0.24, 0.84]	24.2%	[0.65, -0.75, 0.15]	16.7%

Redshift	PC1	PC1 Var	PC2	PC2 Var	PC3	PC3 Var
Satellites, $10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$						
0.00	[0.39, 0.70, 0.60]	54.9%	[0.84, -0.02, -0.54]	32.6%	[0.37, -0.71, 0.60]	12.5%
0.10	[0.33, 0.71, 0.62]	54.2%	[0.87, 0.02, -0.49]	34.0%	[0.36, -0.70, 0.61]	11.8%
0.18	[0.34, 0.71, 0.61]	54.4%	[0.87, 0.01, -0.50]	33.8%	[0.36, -0.70, 0.61]	11.7%
0.27	[0.34, 0.71, 0.62]	53.6%	[0.87, 0.01, -0.50]	33.8%	[0.36, -0.70, 0.61]	12.6%
0.37	[0.30, 0.72, 0.63]	53.8%	[0.89, 0.03, -0.46]	34.3%	[0.35, -0.70, 0.62]	11.9%
0.50	[0.29, 0.72, 0.63]	53.9%	[0.89, 0.05, -0.46]	34.7%	[0.36, -0.69, 0.62]	11.4%
0.62	[0.29, 0.71, 0.64]	54.3%	[0.89, 0.04, -0.44]	34.3%	[0.34, -0.70, 0.63]	11.4%
0.74	[0.25, 0.72, 0.65]	54.3%	[0.91, 0.06, -0.41]	34.7%	[0.33, -0.69, 0.64]	11.0%
0.87	[0.22, 0.72, 0.66]	53.8%	[0.92, 0.08, -0.39]	35.2%	[0.34, -0.69, 0.64]	11.0%
1.00	[0.21, 0.72, 0.66]	55.0%	[0.92, 0.07, -0.38]	34.9%	[0.32, -0.69, 0.65]	10.1%
1.26	[0.26, 0.71, 0.66]	56.2%	[0.93, 0.01, -0.38]	33.6%	[0.28, -0.70, 0.65]	10.2%
1.49	[0.27, 0.70, 0.66]	56.6%	[0.93, -0.02, -0.36]	32.8%	[0.24, -0.71, 0.66]	10.6%
1.74	[0.21, 0.71, 0.67]	55.3%	[0.95, 0.02, -0.33]	33.7%	[0.25, -0.70, 0.67]	10.9%
2.01	[0.27, 0.71, 0.66]	54.6%	[0.93, -0.01, -0.37]	33.2%	[0.25, -0.71, 0.66]	12.2%
2.24	[0.21, 0.71, 0.67]	52.2%	[0.93, 0.05, -0.35]	34.2%	[0.28, -0.70, 0.66]	13.6%
2.48	[0.18, 0.72, 0.68]	52.1%	[0.94, 0.08, -0.33]	34.6%	[0.29, -0.69, 0.66]	13.3%
3.02	[0.41, 0.67, 0.61]	55.9%	[0.88, -0.11, -0.47]	29.9%	[0.25, -0.73, 0.63]	14.3%
3.53	[0.43, 0.70, 0.57]	51.5%	[0.81, -0.02, -0.59]	32.5%	[0.40, -0.72, 0.57]	16.0%
3.98	[0.46, 0.66, 0.59]	52.0%	[0.84, -0.11, -0.54]	29.3%	[0.29, -0.74, 0.60]	18.7%

Table 5.3 *Principal components of $M_* > 10^9\,M_\odot$, $\mathrm{SFR} > 0\,\mathrm{M_\odot yr^{-1}}$, central and satellite EAGLE galaxies in high mass haloes ($10^{12}\,M_\odot < M_{\rm halo} < 10^{14}\,M_\odot$) at each redshift. Vectors have ordering* $[\log_{10} M_{\rm halo}/M_\odot,\ \log_{10} M_*/M_\odot,\ \log_{10} \mathrm{SFR/M_\odot yr^{-1}}]$.

Chapter 6

Predictions for the spatial distribution of the dust continuum emission in $1 < z < 5$ star-forming galaxies

This chapter contains work that has been published as Cochrane R. K., Hayward, C. C., Anglés-Alcázar, D., Lotz, J., Parsotan, T., Ma, X., Kereš, D., Feldmann, R., Faucher-Giguère, C. A., Hopkins, P. F., 2019, Monthly Notices of the Royal Astronomical Society, Volume 488, Issue 2, p.1779-1789

In this Chapter, we present the first detailed study of the spatially-resolved dust continuum emission of simulated galaxies at $1 < z < 5$. We run the radiative transfer code SKIRT on a sample of submillimeter-bright galaxies drawn from the Feedback in Realistic Environments (FIRE) project. These simulated galaxies reach Milky Way masses by $z = 2$. Our modelling provides predictions for the full rest-frame far-ultraviolet-to-far-infrared spectral energy distributions of these simulated galaxies, as well as 25-pc-resolution maps of their emission across the wavelength spectrum. The derived morphologies are notably different in different wavebands, with the same galaxy often appearing clumpy and extended in the far-ultraviolet yet an ordered spiral at far-infrared wavelengths. The observed-frame 870-μm half-light radii of our FIRE-2 galaxies are $\sim 0.5-4$ kpc, consistent with existing ALMA observations of galaxies with similarly high redshifts and stellar masses. In both simulated and observed galaxies, the dust continuum emission is generally more compact than the cold gas and the dust mass, but

more extended than the stellar component. The most extreme cases of compact dust emission seem to be driven by particularly compact recent star formation, which generates steep dust temperature gradients. Our results confirm that the spatial extent of the dust continuum emission is sensitive to both the dust mass and SFR distributions.

6.1 Introduction

Observations suggest that the physical properties of star-forming (SF) galaxies at the peak of cosmic star formation ($z \sim 2$) differ greatly from those of the ordered disks and ellipticals in the local Universe. Galaxies at these redshifts display high star formation rates, believed to be driven by large molecular gas reservoirs (Tacconi et al., 2010, 2013, 2017; Papovich et al., 2016; Falgarone et al., 2017; Jiménez-Andrade et al., 2018) that arise due to steady accretion of cold gas along filaments of the cosmic web (Kereš et al., 2005; Dekel et al., 2009; Faucher-Giguère et al., 2011; Martin et al., 2016; Kleiner et al., 2017). Structurally, high-redshift galaxies are less ordered than their low-redshift counterparts, with star formation taking place within turbulent disks (Genzel et al., 2008; Kassin et al., 2012; Guo et al., 2015; Tadaki et al., 2018) that often harbour massive ultraviolet (UV)-bright clumps (Elmegreen et al., 2013; Guo et al., 2017; Soto et al., 2017). However, our measurements of the high-redshift Universe are largely reliant on data at rest-frame optical and UV wavelengths, which can be biased towards the least dust-obscured galaxies.

Only in the last few years have new facilities such as ALMA had the resolving power to resolve and probe the morphology of longer-wavelength emission from highly star-forming galaxies. The angular resolution of previous instruments such as SCUBA (Holland et al., 1999) was low, so it was not possible to determine the structural properties of high-redshift galaxies. Source confusion has also been a hindrance in the identification of fainter sources (e.g. with Herschel; Oliver et al., 2012; Scudder et al., 2016). ALMA has the potential to be particularly fruitful in identifying high-redshift galaxies, due to the so-called 'negative k-correction' at rest-frame far-infrared (FIR) fluxes (observed $\lambda \sim 850$-μm flux from galaxies with similar intrinsic spectra remains approximately constant across the redshift range $z \sim 1-6$, as we trace further up the Rayleigh-Jeans tail at higher redshift; see Figure 1.6). The most interesting physical insights will likely come from the combination of these new millimeter/sub-millimeter (mm/sub-mm) data with shorter-wavelength imaging. To this end, Dunlop et al. (2017) present 1.3mm

ALMA imaging of the Hubble UltraDeep Field (HUDF), previously mapped with HST's Wide Field Camera 3 to an unprecedented 5σ depth of 30AB mag (Bouwens et al., 2010; Oesch et al., 2010; Illingworth et al., 2013; Dunlop et al., 2013; Ellis et al., 2013), and also with the HST Advanced Camera for Surveys (ACS; Beckwith et al., 2006), over an area of 4.5 arcmin^2. Combining these new ALMA data with *Herschel* and *Spitzer* 24-μm photometry and fitting to a template spectral energy distribution (SED), Dunlop et al. (2017) find that $\sim$ 85% of the total star formation at $z \sim 2$ is enshrouded in dust (see also Bourne et al. 2017; Whitaker et al. 2017). They show that for high-mass galaxies ($M_* > 2 \times 10^{10}\,M_\odot$), which host $\sim$ 65% of the total star formation at this epoch, the star formation rate derived from long-wavelength emission is an extraordinary 200 times that derived from unobscured light. Bowler et al. (2018) demonstrate that dust-obscured star formation could be substantial even as early as $z \sim 7$.

In addition to finding substantially different measurements of star formation rates compared to dust-uncorrected short-wavelength light, studies of galaxies at longer wavelengths also present a different view of the morphologies of high-redshift galaxies. Unobscured emission (probed at short wavelengths) tends to be significantly more extended and clumpier than the rest-frame FIR emission. Barro et al. (2016), for example, find that the 345-GHz (ALMA Band 7) dust continuum emission of a $M_* = 10^{10.9}\,M_\odot$, SFR $= 500\,\mathrm{M_\odot yr^{-1}}$, z = 2.45 galaxy has a half-light radius which is half that of the rest-frame optical emission probed by HST. Hodge et al. (2016) imaged 16 $z \sim 2.5$ similarly massive, highly star-forming, luminous sub-mm galaxies at 0.16″ in the same ALMA band. Many of these galaxies display clumpy structures in HST's H_{160} and I_{814} bands, but their dust emission appears substantially smoother and more compact. The 870-μm radii obtained are small (median 1.8 ± 0.2 kpc), with no convincing evidence for clumpy dust emission at the ALMA resolution probed.

Molecular gas reservoirs have also been mapped for a handful of high-redshift, FIR-luminous galaxies. High-spatial-resolution studies show that molecular gas is compact, though it tends to be slightly more extended than the dust continuum emission. Tadaki et al. (2017) imaged two $z = 2.5$ galaxies at 345 GHz, obtaining 870-μm dust-emission radii of 1.2 ± 0.1 kpc and 1.3 ± 0.1 kpc, around half the size of the CO(J=3-2) emission. Consistent results have been found by other studies (e.g. Strandet, 2017; Calistro Rivera et al., 2018; Tadaki et al., 2018), though sample sizes remain small.

A physical understanding of these differences in spatial extent of emission in different wavebands is critical in order to make the best use of the unmatched

sensitivity and high spatial resolution of ALMA. This is currently difficult observationally, as only small samples of galaxies have been resolved at high resolution with multiple instruments, and such samples are often biased towards either the least dust-obscured systems (if selected in the UV) or the most compact, FIR-bright systems (if selected by FIR surface brightness). Interpretations are further complicated by uncertainty in what FIR/sub-mm fluxes actually probe. Frequently used relations between observed FIR/sub-mm fluxes and SFR (e.g. Kennicutt & Evans, 2012) do not fold in the shape of the full dust SED, which should reflect not only SFR but also dust mass, dust temperature, and the geometry of the source (e.g. Misselt et al., 2001; Hayward et al., 2011, 2012; Lanz et al., 2014; Safarzadeh et al., 2016; Kirkpatrick et al., 2017).

The primary aim of this Chapter is to understand which physical properties of high-redshift galaxies are probed by their dust continuum emission, in a spatially-resolved manner. We address this using simulations to study individual galaxies at high spatial resolution across the rest-frame UV-FIR spectrum. We select a sample of the most FIR-bright galaxies from the Feedback In Realistic Environments 2 (FIRE-2) simulations (Hopkins et al., 2018b)[1] presented in Anglés-Alcázar et al. (2017). We perform three-dimensional continuum radiative transfer on selected galaxy snapshots from these simulations, modelling the effects of dust attenuation and re-emission to predict the spatially-resolved multi-wavelength emission of these high-redshift FIR/sub-mm-bright simulated galaxies. We then compare the spatial extent of the dust continuum emission to various intrinsic physical properties of our simulated galaxies.

The structure of this Chapter is as follows. In Section 6.2, we discuss our method for selecting sub-mm-bright snapshots from the FIRE-2 simulations. We describe the radiative transfer modelling used to post-process these galaxy snapshots and present their simulated rest-frame UV-FIR spectral energy distributions in Section 6.3. We present predictions for the spatial extent of dust continuum emission in Section 6.4 and also compare our predictions to observational results. We also study which physical properties of high-redshift, dusty star-forming galaxies are best probed by the spatial extent of the dust continuum emission. We summarise our conclusions in Section 6.5.

[1] http://fire.northwestern.edu

6.2 A sample of simulated high-redshift galaxies

6.2.1 Galaxies in the FIRE-2 simulations

The Feedback In Realistic Environments (FIRE) project (Hopkins et al., 2014, 2018b) is a set of state-of-the-art hydrodynamical cosmological zoom-in simulations that explore the role of stellar feedback in galaxy formation and evolution. Stellar feedback must play an important role in galaxy formation. Without it, the galactic ISM would collapse on dynamical time-scales, leading to gravitational collapse, fragmentation and accelerated star formation. While galaxies simulated without stellar feedback thus rapidly convert all available gas into stars (e.g. Hopkins et al. 2011; see the review by Somerville & Davé 2015), the tight locus of observed galaxies on the Kennicutt-Schmidt relation (Kennicutt, 1998, see Section 1.2.2) implies that gas consumption timescales in real galaxies are long. Furthermore, both galaxy stellar mass functions (e.g. Ilbert et al., 2013; Muzzin et al., 2013; Davidzon et al., 2017) and the stellar mass-halo mass relation (e.g. Moster et al. 2010; Behroozi et al. 2013, Chapter 4) imply that galaxies convert only a small fraction of the universal baryon fraction into stars. Galactic outflows are therefore needed to regulate the mass of galaxies over time (Kereš et al., 2009; Anglés-Alcázar et al., 2014; Faucher-Giguère et al., 2015a; Anglés-Alcázar et al., 2017). These outflows are likely responsible for the observed enrichment of the circumgalactic medium (CGM) and intergalactic medium (IGM), and are also observed directly (e.g. Weiner et al. 2009; Steidel et al. 2010; see the review by Rupke 2018). These pieces of observational evidence demand that stellar feedback must be at play (Schaye et al., 2003; Oppenheimer & Davé, 2006; Faucher-Giguère et al., 2015b; Hafen et al., 2017). Stellar feedback is believed to be particularly important in low-stellar-mass galaxies, below the peak of the stellar mass-halo mass relation.

Various stellar feedback processes are thought to contribute. These processes include supernovae, protostellar jets, photo-heating, stellar mass loss from O- and AGB-stars and radiation pressure (see Dale, 2015, for a review). Importantly, these processes are believed to act non-linearly, and so modelling the stellar processes of even a single galaxy becomes a complex computational challenge. Only recently have cosmological zoom simulations achieved sufficient resolution to model these feedback processes directly.

The FIRE project employs two main techniques to explicitly model multi-channel stellar feedback. Firstly, the FIRE simulations resolve the formation

of giant molecular clouds (GMCs), and star formation takes place only in self-gravitating (according to the Hopkins et al. 2013 criterion), self-shielding molecular gas (following Krumholz & Gnedin 2011) at high densities ($n_H > 1000\,\mathrm{cm}^{-3}$ in the simulations we use here). Secondly, FIRE includes models for both energy and momentum return from the main stellar feedback processes, using directly the predictions of stellar population synthesis models without the extensive parameter tuning employed in other simulations. Specifically, once a star particle forms, the simulations explicitly follow several different stellar feedback mechanisms as detailed in Hopkins et al. 2018a, including (1) local and long-range momentum flux from radiation pressure (in both the initial UV/optical single-scattering regime and re-radiated light in the IR); (2) energy, momentum, mass and metal injection from SNe (Types Ia and II) and stellar mass loss (both OB and AGB); and (3) photo-ionization and photo-electric heating. Every star particle is treated as a single stellar population with known mass, age, and metallicity, and then all feedback event rates, luminosities and energies, mass-loss rates, and all other quantities are tabulated directly from stellar evolution models (STARBURST99; Leitherer et al. 1999), assuming a Kroupa (2002) initial mass function. The FIRE simulations succeed in generating galactic winds self-consistently (Faucher-Giguère et al., 2015a; Anglés-Alcázar et al., 2017), without relying on sub-grid hydrodynamic decoupling or delayed cooling techniques, and broadly reproducing many observed galaxy properties, including stellar masses, star formation histories and the 'main sequence' of star-forming galaxies (see Hopkins et al. 2014; Sparre et al. 2017), metallicities and abundance ratios (Ma et al., 2016; van de Voort et al., 2015), as well as morphologies and kinematics of both thin and thick disks (Ma et al., 2018a).

For this Chapter, we study the central galaxies of four massive haloes originally selected and simulated by Feldmann et al. (2016, 2017) with the original FIRE model (Hopkins et al., 2014) as part of the MASSIVEFIRE suite. The haloes were selected to have dark matter halo masses of $M_{\rm halo} \sim 10^{12.5}\,M_\odot$ at $z = 2$. The central galaxies of these haloes have stellar masses of $7 \times 10^{10} - 3 \times 10^{11}\,M_\odot$ at $z = 2$ (detailed in Table 6.1), with a variety of formation histories; see Feldmann et al. (2017) for details. The simulations in the present study are drawn from Anglés-Alcázar et al. (2017), who re-simulated these four massive haloes down to $z = 1$ with the upgraded FIRE-2 physics model (Hopkins et al., 2018b) and with a novel on-the-fly treatment for the seeding and growth of supermassive black holes (SMBHs). Note that these simulations do not include feedback from the accreting SMBHs. Compared to FIRE, FIRE-2 simulations are run with a new,

Name	$\log_{10} M_{\rm halo}/M_\odot$	$\log_{10} M_*/M_\odot$	$\log_{10} M_{\rm gas}/M_\odot$	SFR/M$_\odot$yr^{-1}	$r_{1/2}$/kpc
A1	12.45	11.24	10.22	65	0.73
A2	12.56	11.46	10.51	168	0.98
A4	12.49	11.10	10.33	63	0.81
A8	12.41	10.85	10.45	88	0.91

Table 6.1 *Properties of the four simulated FIRE-2 haloes at $z = 2$. $M_{\rm halo}$ denotes the total mass of the central halo at $z = 2$. M_*, $M_{\rm gas}$, and* SFR *denote the stellar mass, gas mass, and star formation rate of the halo's central galaxy at $z = 2$, all calculated within $0.1R_{\rm vir}$. $r_{1/2}$ is the half-mass radius, calculated using the stellar mass within $0.1R_{\rm vir}$.*

more accurate hydrodynamics solver (a mesh-free Godunov solver implemented in the GIZMO[2] code; Gaburov & Nitadori 2011; Hopkins 2015). They also feature improved treatments of cooling and recombination rates, gravitational softening and numerical feedback coupling, and they adopt a higher density threshold for star formation (Hopkins et al., 2018a). The mass resolution of our simulations is $3.3 \times 10^4\,M_\odot$ for gas and star particles and $1.7 \times 10^5\,M_\odot$ for dark matter particles.

6.2.2 Selection of sub-mm-bright galaxy snapshots at $z > 2$

We wish to simulate galaxies that are representative of those typically observed with ALMA at high spatial resolution at high redshifts, which implies that we should select those that are likely to have high 850-μm flux densities ($f_{850} > 1\,$mJy). Performing radiative transfer on each of the ~ 300 redshift snapshots available to predict sub-mm fluxes and then selecting the brightest would be unnecessarily computationally intensive. The first step in our analysis is therefore to select redshift snapshots for each of the four galaxies for which we expect the sub-mm flux to be particularly bright, using simply the SFR and dust mass at each redshift. We adopt the following equation, derived from fits to the observed-frame 850-μm flux densities of simulated galaxies computed via dust radiative transfer and presented in Hayward et al. (2013):

$$f_{850} = 0.81\,{\rm mJy} \times \left(\frac{\rm SFR}{100\,{\rm M_\odot yr^{-1}}}\right)^{0.43} \left(\frac{\rm M_{dust}}{10^8\,{\rm M_\odot}}\right)^{0.54}, \tag{6.1}$$

where we estimate the dust mass for the present purposes using $M_{\rm dust} = 0.01 M_{\rm gas}$, where $M_{\rm gas}$ is the total gas mass within $0.1R_{\rm vir}$. The agreement between these predictions and the 850-μm flux densities that result from the radiative transfer

[2] http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html

modelling is correct to within a factor of ~ 2 for fluxes above $\sim 1\,\mathrm{mJy}$. Future work will involve running radiative transfer over a larger number of snapshots and re-fitting this formula. We select the ~ 20 snapshots with the highest predicted f_{850} for each simulated halo (named A1, A2, A4 & A8)[3], within the redshift range $1 < z < 5$. A couple of snapshots were excluded from the analysis after performing radiative transfer due to poorly defined centres and extremely diffuse dust emission. We do not attempt to produce a complete sample of sub-mm bright galaxies from the FIRE-2 simulations; instead, our selection is sufficient to yield a sample of sub-mm bright snapshots for which we can perform radiative transfer and study multi-wavelength properties.

6.3 Radiative transfer methods and results

6.3.1 Performing radiative transfer with SKIRT

Modelling dust and its emission in galaxies is a difficult computational problem (see Steinacker et al., 2013, for a comprehensive review). The process of radiative transfer is non-local in space (photons can propagate long distances before interacting with dust), and it is coupled in terms of both direction and wavelength. The distribution of dust in (both real and simulated) galaxies is far from a simple screen; instead, it is necessary to model the three-dimensional distribution of sources of radiation (stars and AGN) and dust.

In this work, we make use of the Stellar Kinematics Including Radiative Transfer (SKIRT)[4] Monte Carlo radiative transfer code (Baes et al., 2011; Camps & Baes, 2015), which has also been used to model dust attenuation and emission in the EAGLE simulations (see Trayford et al., 2017; Camps et al., 2018; McAlpine et al., 2019). Monte Carlo radiative transfer codes like SKIRT treat the radiation field from stars (and sometimes AGN) as a flow of photons through the dusty medium of a galaxy to compute the effects of dust absorption, scattering, and re-emission of the absorbed light, including dust self-absorption. We are able to perform the radiative transfer in post-processing because the light-crossing time is short compared to dynamical times on resolved scales (such that the dust and stellar geometry does not change appreciably over a light-crossing time).

We extract gas and star particles from the FIRE-2 simulations at each of our chosen snapshots. The coordinate system is rotated to align with the angular

[3] Our analysis focuses on the central galaxies of each halo. Throughout, we will refer to the central galaxy of e.g. halo A1 simply as 'galaxy A1'.

[4] http://www.skirt.ugent.be

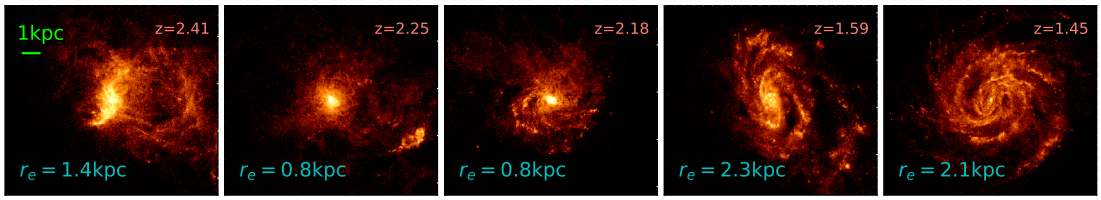

Figure 6.1 *870-μm observed-frame flux maps for a subsample of snapshots of the central galaxy of FIRE-2 halo A4 predicted using the* SKIRT *radiative transfer code and plotted on the same logarithmic flux scale. These cutouts span* $10\,\mathrm{kpc} \times 10\,\mathrm{kpc}$ *(proper distances). The predicted dust continuum emission displays a range of morphologies as the ordered disk develops. This emission is compact, spanning half-light radii of* $\sim 1 - 2\,\mathrm{kpc}$ *(these are shown in blue).*

momentum vector of the gas particles within $0.1R_{\rm vir}$ prior to input to SKIRT, so that for disk galaxies, a viewing angle of 0 degrees corresponds to face-on. For gas particles with temperature $< 10^6\,\mathrm{K}$, we compute dust masses using the metallicity of the gas particles and a dust-to-metals mass ratio of 0.4 (Dwek, 1998; James et al., 2002). We assume that dust is destroyed in gas particles with temperature $> 10^6\,\mathrm{K}$ (Draine & Salpeter, 1979; Tielens et al., 1994). We use a Weingartner & Draine (2001) Milky Way dust prescription to model a mixture of graphite, silicate and PAH grains. Star particles are assigned Bruzual & Charlot (2003) SEDs according to their ages and metallicities (note that our results are unchanged if we instead use the STARBURST99 SED templates presented by Leitherer et al. 1999). We use an octree dust grid, in which cell sizes are adjusted according to the dust density distribution. We impose the condition that no dust cell may contain more than 0.0001% of the total dust mass of the galaxy, which yields excellent convergence of the integrated SED.

We also specify a number of properties relating to the SED output. We define a wavelength grid with ~ 100 discrete wavelengths, spaced uniformly in log(wavelength) between rest-frame UV and FIR wavelengths and including each of the ALMA bands. To model the flux that would be received by an observer on Earth, we place seven detectors at $z = 0$, uniformly spaced at different inclinations with respect to the disk plane of the galaxy. These detectors have pixel sizes corresponding to a proper length of 25 pc, and the field of view is set to 10% of the virial radius for each galaxy snapshot studied. Note that the pixel scale of the images should not be confused with the resolution of the simulation; this varies across the galaxy, and is higher than 25 pc in denser regions but lower in general ISM gas.

The focus on resolved dust emission in the present work is complementary to

other radiative transfer studies of FIRE galaxies with SKIRT. Liang et al. (2018) analyzed how well galaxy-integrated, single-band dust continuum emission traces molecular gas. The implications of defining galaxy-integrated dust temperatures in different ways, for instance with regard to various scaling relations, are studied in Liang et al. (2019). Finally, Ma et al. (2019) make predictions for the dust extinction and emission properties of $z > 5$ galaxies using a sample of 34 FIRE-2 haloes, including those first presented in Ma et al. (2018b).

6.3.2 Morphology across the wavelength spectrum

Our radiative transfer modelling enables us to track the emission from each of the galaxies at multiple epochs. We compute images of each of the galaxy snapshots at every wavelength on our grid, spanning the rest-frame FUV to FIR. We show an example of the high quality of our maps of the sub-mm flux in Figure 6.1, to illustrate the wide range of morphologies displayed by a single galaxy evolving in the redshift range $z = 2.41 - 1.45$. We also find that the same galaxy can look vastly different in the different wavebands. We illustrate this qualitatively with five representative wavelengths for the central galaxy of halo A1 (hereafter 'galaxy A1') at $z = 4.38$ in Figure 6.2 and galaxy A2 at $z = 2.95$ in Figure 6.3. We also show for comparison the spatial distributions of four key galaxy properties: total stellar mass, cool, dense gas mass ($T < 300\,\mathrm{K}, n_H > 10\,\mathrm{particles/cm^{-3}}$), dust mass, and recently formed stars (age $< 100\,\mathrm{Myr}$).

Galaxy A1 at $z = 4.38$ has a SKIRT-predicted observed-frame 850-μm flux density of 0.79 mJy. At this snapshot, the galaxy exhibits very clumpy FUV emission (rest-frame 1500Å), spanning a few kpc. This emission is aligned with the cool gas and exhibits peaks where this gas is densest. The optical emission (rest-frame 4000Å) is also clumpy. Longwards of 24-μm the emission becomes more regular and centrally concentrated, resembling the total and recently formed stellar mass distributions more closely.

Galaxy A2 at $z = 2.95$ is brighter, with a SKIRT-predicted observed-frame 850-μm flux density of 1.45 mJy. It is also substantially more extended in all wavebands. The rest-frame 1500Å and 7000Å emission is again clumpy, bearing little resemblance to the ordered bulge plus spiral structure that is clear from the SFR and stellar, gas and dust mass maps. Interestingly, the peak of the short-wavelength emission occurs in a region to the right of centre that is largely free of dust (see Figure 6.4). It appears that the clumpiness of the FUV emission is driven by the structure of the dust, with FUV emission tracing holes in the dust. This is just one example, and such offsets are common in the sample, as

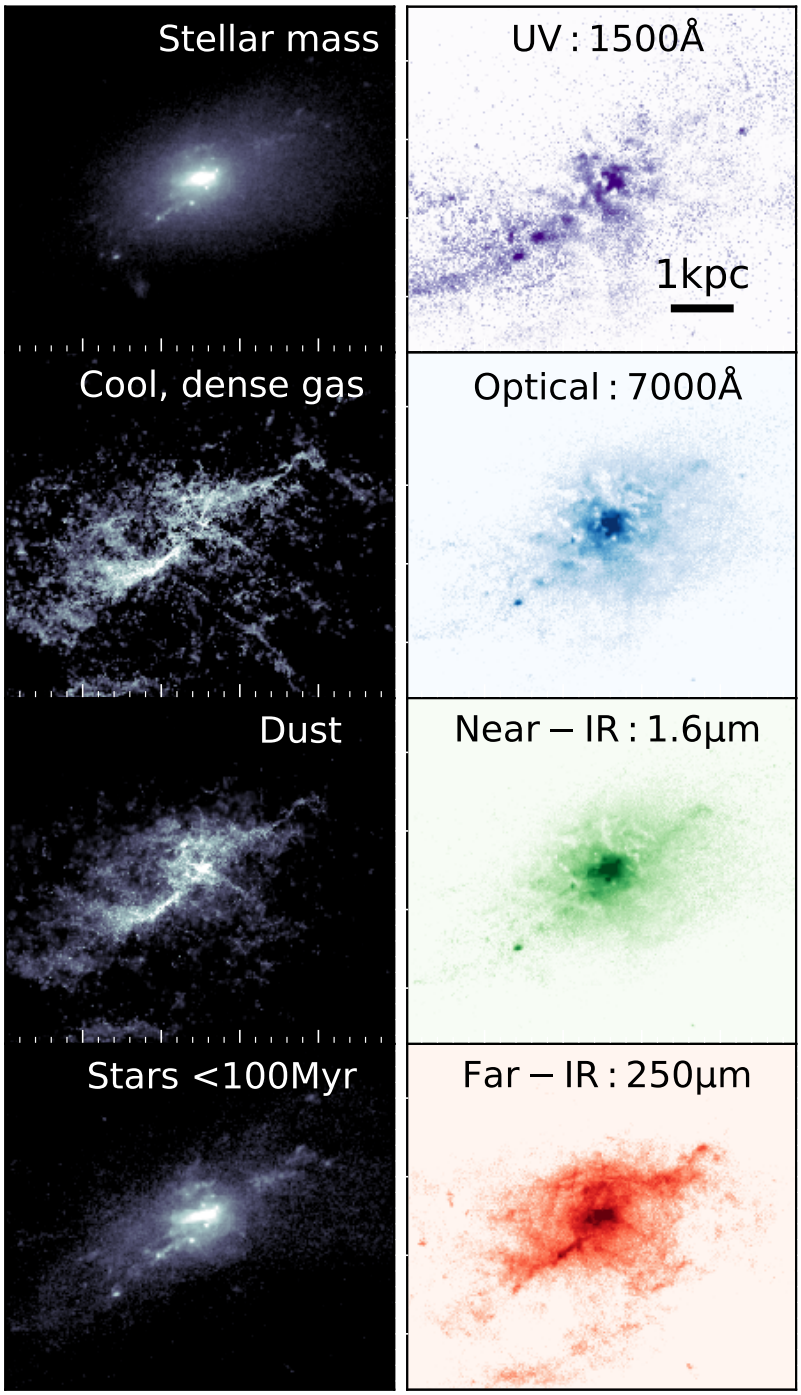

Figure 6.2 *The wavelength-dependent morphology of galaxy A1 at* $z = 4.38$*. The left panels show the projected distributions of stellar mass, cold, dense gas mass, dust mass, and stars formed within* 100 Myr *(intrinsic properties of the galaxy). The right panels show the* SKIRT*-predicted images at different rest-frame wavelengths. The morphology is strongly dependent on the wavelength. The galaxy appears clumpy and extended in the rest-frame UV but more ordered at longer wavelengths. The colour scales are logarithmic and span the* $70^{\text{th}} - 99^{\text{th}}$ *percentiles of the flux distribution of each panel, to highlight the qualitative differences in morphology.*

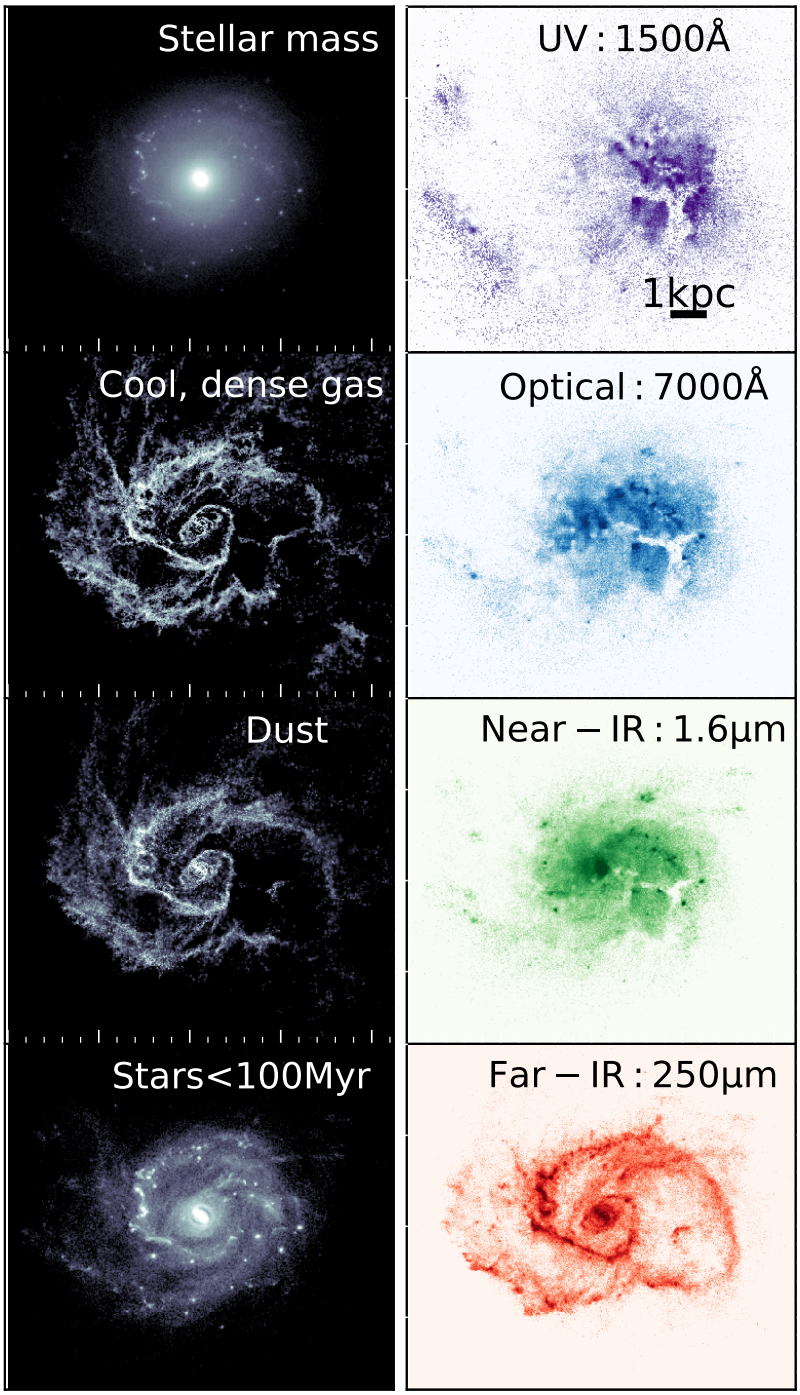

Figure 6.3 *The wavelength-dependent morphology of galaxy A2 at $z = 2.95$, with panels as described in Figure 6.2. The UV and optical light is significantly offset from the peak of the stellar mass and SFR, appearing to trace holes in the dust.*

in observations (e.g. Chen et al. 2017; Calistro Rivera et al. 2018, Chapter 7). Future work will explore these offsets, as well as the resolved multi-wavelength emission on a pixel-by-pixel basis.

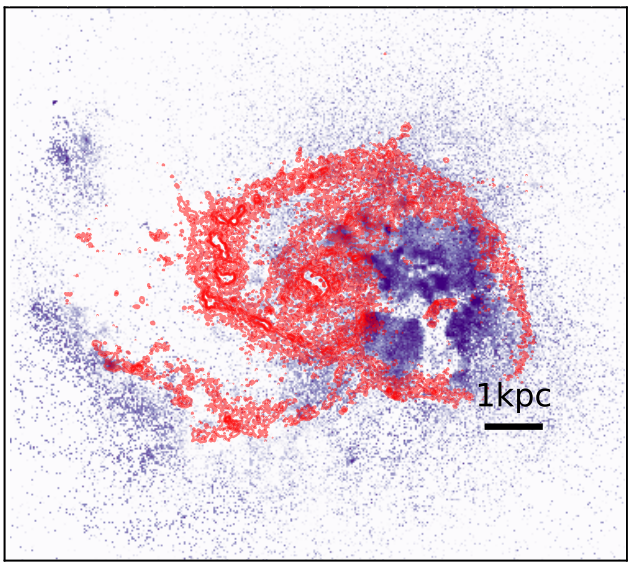

Figure 6.4 *The rest-frame* 1500 Å *image of galaxy A2 at* $z = 2.95$*, with 250-*μm *contours overlaid in red. The long- and short-wavelength emission occupy strikingly different spatial regions. While there is recent star formation across the extent of the galaxy disk, light at short wavelengths does not escape from regions of high dust density. This leads to a spatial offset between the FUV and dust continuum emission.*

6.4 Quantifying the spatial extent of dust continuum emission

6.4.1 The dependence of size on FIR wavelength

In this section, we quantify the sizes of sub-mm-bright FIRE-2 galaxies as a function of wavelength. An example of an observed-frame SED predicted by SKIRT, overplotted with the wavelengths corresponding to eight ALMA bands, is shown in Figure 6.5 (top panel). We extract the SKIRT image at each of these wavelengths and derive an 'effective radius' for the predicted emission. This is the radius that contains half of the total flux (calculated within a ~ 12 kpc radius), derived using a circular aperture centered on the flux-weighted center of the emission in each band. In each case, an error bar is derived from the standard deviation of the effective radius measurements at seven different galaxy/detector inclinations.

In the bottom panel of Figure 6.5, we show an example of our results. The effective radius of the emission varies with wavelength, with longer wavelength FIR emission spanning a greater spatial extent. We also overplot the effective

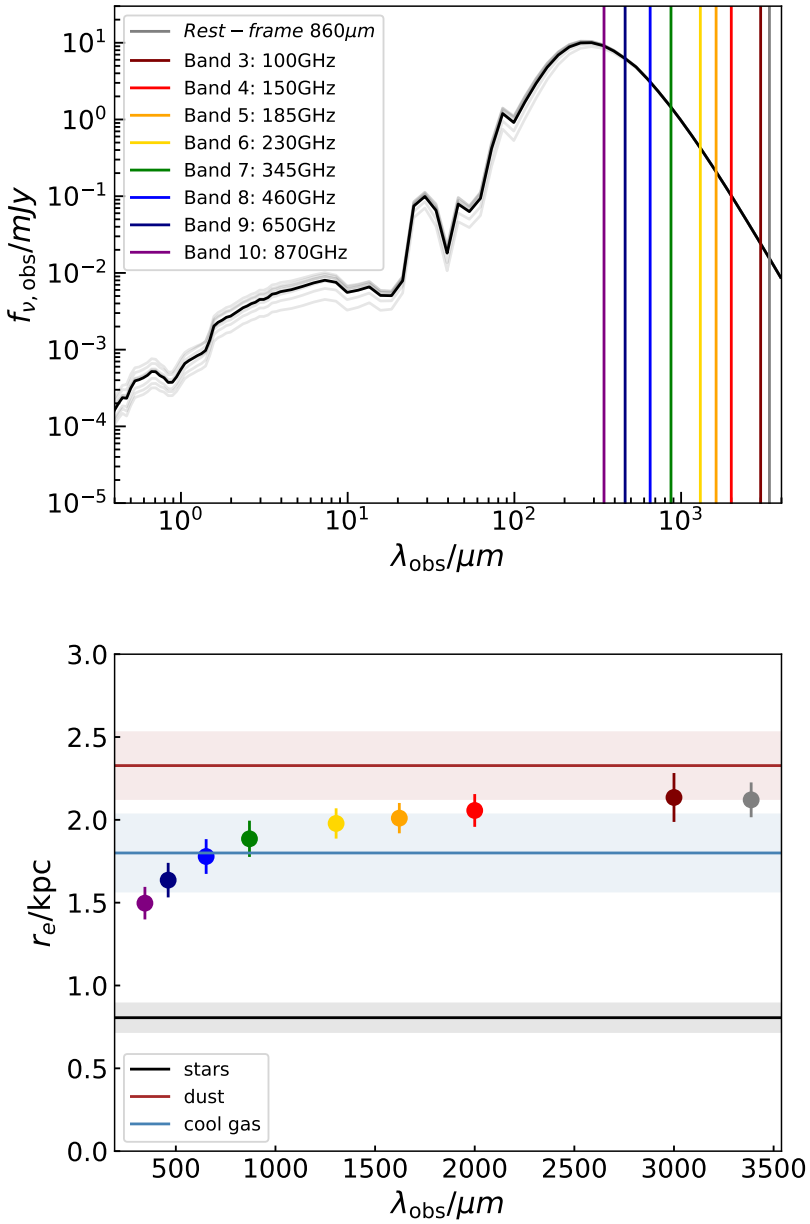

Figure 6.5 *Top: observed-frame SED of galaxy A2 at $z = 2.95$, with the centres of ALMA bands overlaid. The grey lines show the SED at different inclinations, with the mean shown in black. Bottom: the half-light radii of the emission at the wavelengths of centres of each of these bands compared to those of the galaxy stellar mass, cold gas mass and dust mass. Error bars and shaded regions are derived considering seven different viewing angles. Shorter-wavelength FIR emission is closer in size to the stellar component, whereas longer-wavelength emission traces the cold gas and dust.*

radii of several key physical quantities of the galaxy: stellar mass, dust mass and cool, dense gas mass. We define cool, dense gas using the criteria $T < 300\,\mathrm{K}, n_H > 10\,\mathrm{cm}^{-3}$. This has been shown to be a reasonable proxy for molecular gas by Orr et al. (2018). In this example, and in general, the cool gas mass and the dust mass are more spatially extended than the stellar mass. The sizes of the dust continuum emission tend to be between those of the dust mass and the stellar mass. Shorter-wavelength FIR emission, corresponding to hotter dust, tends to be more compact, whereas longer-wavelength FIR emission better traces the extended cold gas and dust.

The 870-μm (345-GHz) observed-frame emission probed by ALMA Band 7 is frequently used to study the dust continuum emission of high-redshift galaxies (e.g. Barro et al., 2016; Chen et al., 2017; Simpson et al., 2017b). For simplicity, from here on, we consider only the 870-μm flux in our discussion of the spatial extent of the dust continuum emission. However, we note that we consistently find that emission at longer rest-frame FIR wavelengths is more extended, as shown in Figure 6.5.

6.4.2 Agreement with compact dust continuum observations

In Figure 6.6, we compare the sizes of both our predicted 870-μm dust emission and the cool, dense gas to the following ALMA observations of galaxies at similar redshifts: 860-μm emission and 12CO(J $= 4-3$) line flux for a SFR $> 1000\,\mathrm{M}_\odot$/yr, $M_* \sim 10^{11}\,M_\odot$ galaxy at $z = 4.3$, from Tadaki et al. (2018) (grey star); 870-μm emission and 12CO(J $= 3-2$) line flux for two massive ($M_* \sim 10^{11}\,M_\odot$) galaxies at $z = 2.5$, from Tadaki et al. (2017) (black stars); the effective radii for stacked ALMA maps of 870-μm and 12CO(J $= 3-2$) emission for 16 ALESS galaxies at $z = 2.5 \pm 0.2$, from Calistro Rivera et al. (2018) (black circle); 870-μm emission and 12CO(J $= 3-2$) line flux for a SFR $\sim 500\,\mathrm{M}_\odot$/yr, $M_* \sim 2\times10^{11}\,M_\odot$ galaxy at $z = 2.12$, from Chen et al. (2017) (black square); FIR emission (from APEX/LABOCA and *Herschel*) and 12CO(J $= 3-2$) at $z = 6.9$, from Strandet (2017) (black triangle); stacked 870-μm and [CII] emission for $z \sim 4.5$ galaxies, from Cooke et al. (2018) (purple star); 1.1mm (from NOEMA) and Hα emission, mapped for a $M_* \sim 7\times10^{10}\,M_\odot$ at $z = 1.25$, from Nelson et al. (2019) (purple circle). Our derived 870-μm effective radii are $\sim 0.5-4\,\mathrm{kpc}$. This is in excellent agreement with observations at a range of redshifts (e.g. Ikarashi et al., 2015; Iono et al., 2016; Simpson et al., 2015; Hodge et al., 2016). In line with the observations, the snapshots of each of the four FIRE-2 galaxies tend to lie below the 1-1 line (dashed black), i.e. the dust continuum emission is

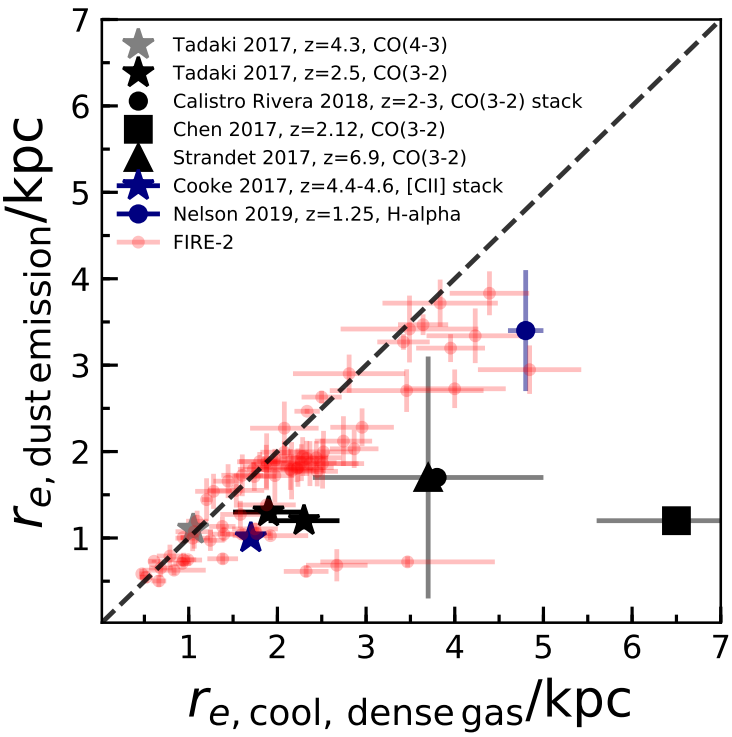

Figure 6.6 *The effective radius of the 870-μm dust emission versus that of the cool, dense gas ($T < 300\,\mathrm{K}, n_H > 10\,\mathrm{particles/cm^{-3}}$), for FIRE-2 galaxy snapshots (red) and selected observational measurements. The dashed black line shows the 1-1 relation. For all FIRE-2 galaxies, the dust emission is more compact than the cool, dense gas at almost all redshift snapshots studied, consistent with the observational sample (see text). The effective sizes of dust emission and cool gas of our simulated galaxies agree well with those of the observed galaxies, although the slope of the relation between the two sizes seems to be somewhat steeper for the simulations than for the (small) sample of observed galaxies. However, the slope of the relation for simulated galaxies is sensitive to the temperature and density cuts made to select cool, dense gas.*

more compact than the 'molecular' gas. Our predictions agree fairly well with the observational results for the small number of high-redshift galaxies that have been observed in both dust continuum and CO at high spatial resolution. However, the few observations that do exist tend to have slightly more compact dust emission at fixed molecular gas effective radius than our predictions. This could be due to the definition of 'molecular' gas within the simulation (different temperature and density cuts yield slightly different 'molecular' gas sizes, and hence slopes; 'very dense gas' is much more compact, as shown in Figure 6.7), or due to the selection of observational samples.

6.4.3 How does the dust continuum emission trace the physical properties of galaxies?

The dust continuum emission is frequently used as an indicator of both SFR and dust mass. However, until now, the spatial extent of these physical and observable quantities has not been studied consistently with simulated galaxies. Motivated by the clear differences in effective radii of the dust continuum emission, stellar mass, dust mass, and gas mass found for individual FIRE-2 galaxies (see the right-hand panel of Figure 6.5), we set out to identify which physical properties are best reflected by the spatial extent of the dust continuum emission.

In Figure 6.7, we plot the effective radius of the simulated observed-frame 870-μm dust continuum emission versus those of eight different physical quantities derived directly from the FIRE-2 simulations. The 870-μm emission is more compact than the total gas component, but only slightly more compact than the cool, dense gas. The dust emission is also more compact than the dust mass distribution for all FIRE-2 snapshots. We also consider how the spatial extent of dust emission correlates with that of the total stellar mass, recently formed stellar mass (within 10 Myr and 100 Myr) and instantaneous SFR. We find that the emission from dust is more extended than all of these stellar quantities and than the densest gas ($n_H > 100\,\mathrm{particles/cm^3}$). The extent of the dust emission appears to correlate more tightly with the extent of the most recently formed stars than the total stellar mass. This reflects the role of young stars as the primary heating source for the dust grains that reprocess their short-wavelength light, as will be discussed in detail in the next two subsections. Note also that stellar mass appears to assemble inside-out, with the more recently formed stars spanning a larger half-mass radius than the total stellar component.

6.4.4 The role of star formation in determining the extent of the dust continuum emission

The observed-frame submm emission depends on the effective dust temperature, which is sensitive to the luminosity absorbed by dust, the dust mass, and possibly the geometry of the system (e.g. Misselt et al., 2001; Hayward et al., 2011; Safarzadeh et al., 2016). Consequently, we expect that the spatial extent of the dust continuum will depend on those of the young stars and dust. To quantify this dependence, we use a Principal Component Analysis (PCA). As described in Chapter 5, PCA is a statistical technique used to describe the variance within a dataset. Variables are converted into a set of uncorrelated,

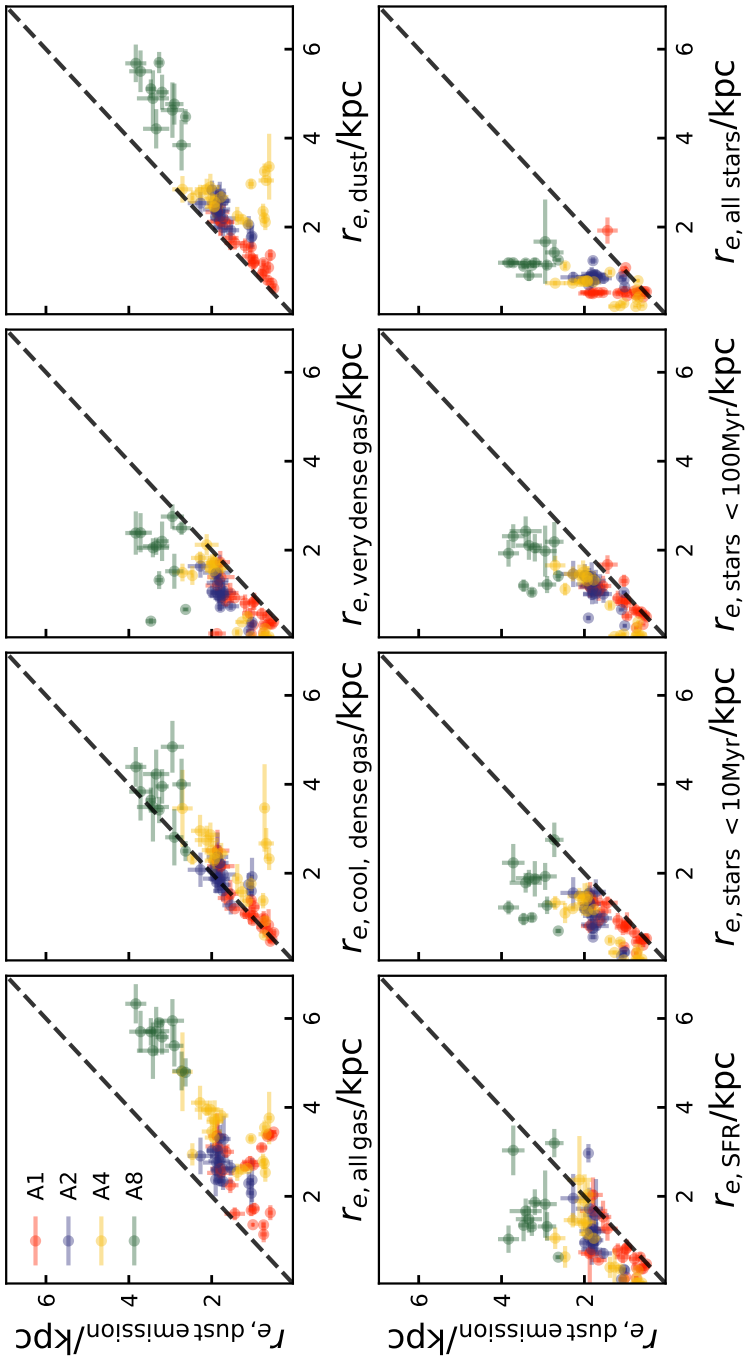

Figure 6.7 *The effective radius of the simulated observed-frame 870-μm dust continuum emission versus those of all gas, cool gas ($T < 300\,\mathrm{K}, n_H > 10\,\mathrm{particles/cm^3}$), dust, very dense gas ($n_H > 100\,\mathrm{particles/cm^3}$), instantaneous star formation rate, recently formed stars (age $< 10\,\mathrm{Myr}, < 100\,\mathrm{Myr}$) and all stars, derived directly from the FIRE-2 simulation snapshots within $0.1R_{\mathrm{vir}}$. Data points are colour-coded according to their haloes. The dashed black lines show a 1-1 relation. The 870-μm emission is more compact than the cool gas and dust but more extended than the very dense gas and stellar components.*

orthogonal principal components. The first component reveals the direction of maximum variance, which describes the key correlation between variables within the dataset. Successive components account for less of the variance of the population. Some latter components may be dominated by noise, leaving the data decomposed into fewer dimensions. PCA has been used in a number of recent galaxy evolution studies, including in studies of the mass-metallicity relation (Lagos et al., 2016; Bothwell et al., 2016; Hashimoto et al., 2018) and in Chapter 5 to study the quenching of galaxies within the EAGLE simulations.

Here, we use PCA to study the relationship between the effective radii of the dust component, SFR and dust continuum emission. We use the PCA python tool *scikit.learn* to perform the PCA analysis. Each variable is normalized to its mean and scaled to unit variance before performing the PCA. For each snapshot studied, we construct a vector of the form $[r_{\rm eff,\ dust\ emission}, r_{\rm eff,\ dust}, r_{\rm eff,\ stars\ <100Myr}]$ and supply all of these vectors to the PCA.

We find that the first principal component, [0.62, 0.56, 0.55], almost entirely describes the dataset, accounting for 83% of the variance of the sample. This primary correlation indicates that all three variables – the effective radii of dust emission, dust mass and recently formed stars – are positively correlated. The effective radius of the dust emission exhibits a strong correlation with the effective radius of the dust mass itself and with the effective radius of the recently formed stars. This result confirms our physical intuition that the sub-mm flux density (and thus the effective radius of the dust continuum emission) is sensitive to both the SFR and dust mass (Hayward et al., 2011; Safarzadeh et al., 2016). The second principal component, [0.03, 0.69,-0.73], accounts for 15% of the scatter, essentially all of the rest. This component represents the scatter in the relation between the effective radii of the dust mass and the recently formed stars. Note that there is, therefore, little scatter in the strong correlations between the effective radius of the dust emission and the dust mass and between the effective radius of the dust emission and the recently formed stars.

6.4.5 The physical drivers of compact dust emission: dust temperature gradients

In Section 6.4.4, we found that the effective radius of the dust continuum emission is correlated with both the effective radius of the dust mass itself, and the effective radius of the recently formed stars. In this final section, we investigate the physical drivers of these correlations, in particular the role of dust heating. In Figure 6.8, we plot the effective radius of the 870-μm dust continuum

emission against that of the recently formed stars, normalising both quantities by the effective radius of the dust mass. If the dust continuum emission traced the dust mass directly, all points would lie at $r_{e,\mathrm{dust\,emission}}/r_{e,\mathrm{dust}} = 1$. While $r_{e,\mathrm{dust\,emission}}/r_{e,\mathrm{dust}}$ is close to 1 for most snapshots, we find a clear slope in this relation: galaxies with low $r_{e,\mathrm{dust\,emission}}/r_{e,\mathrm{dust}}$ (which have particularly compact dust emission with respect to the total dust mass distribution) also have recent star formation which is much more compact than the dust. This reflects the dual role of the dust mass and star formation in determining the spatial extent of the dust continuum emission.

We have identified that the spatial scales of dust and recently formed stars play roles in determining the extent of the dust continuum emission, with very compact dust emission appearing to correlate with compact recent star formation. Now, we use the dust mass-weighted temperatures of the grid cells output by SKIRT to explore the physical drivers of particularly compact dust emission. The temperature of the dust in our simulated galaxies varies significantly as a function of distance from the galaxy's centre. Typical dust temperatures are $\sim 20-40\,\mathrm{K}$ outside the central kpc, and $\sim 40-70\,\mathrm{K}$ inside the central kpc. Here, we study the relationship between dust temperature gradients and the spatial extent of dust emission. For each snapshot, we construct thin concentric shells about the halo centre, out to the effective radius calculated for the dust mass. For each shell, we calculate the mean dust temperature, T_{dust}, and mean dust particle position, r_{dust}. We then fit the following relation: $\log_{10} T_{\mathrm{dust}} = \nabla T_{\mathrm{dust}} \log_{10} r_{\mathrm{dust}} + C$, where the units of T_{dust} are K and the units of r_{dust} are kpc. This yields a power-law index, ∇T_{dust}, which we call the dust temperature gradient index, for each galaxy snapshot. The snapshots in Figure 6.8 (top left panel) are colour-coded by this gradient. It is clear that those galaxies with compact dust emission and recent star formation also have the steepest dust temperature gradients. The other panels of Figure 6.8 show ratios of the effective radii of the dust mass, dust continuum emission and recently formed stars. For gradients flatter than $\nabla T_{\mathrm{dust}} \sim -0.2$, the dust continuum emission broadly traces the extent of both recently formed stars and dust mass. For steeper gradients, below $\nabla T_{\mathrm{dust}} \sim -0.2$, the recently formed stars are substantially more compact than the dust mass, causing the ratio of absorbed luminosity to dust mass to be higher in the galaxy center than at larger radii, leading to the steep dust temperature gradients. At these low values of ∇T_{dust}, the spatial extent of the dust continuum emission is not a good approximation of the distribution of dust mass or recently formed stars.

These results confirm that the spatial extent of the dust continuum emission

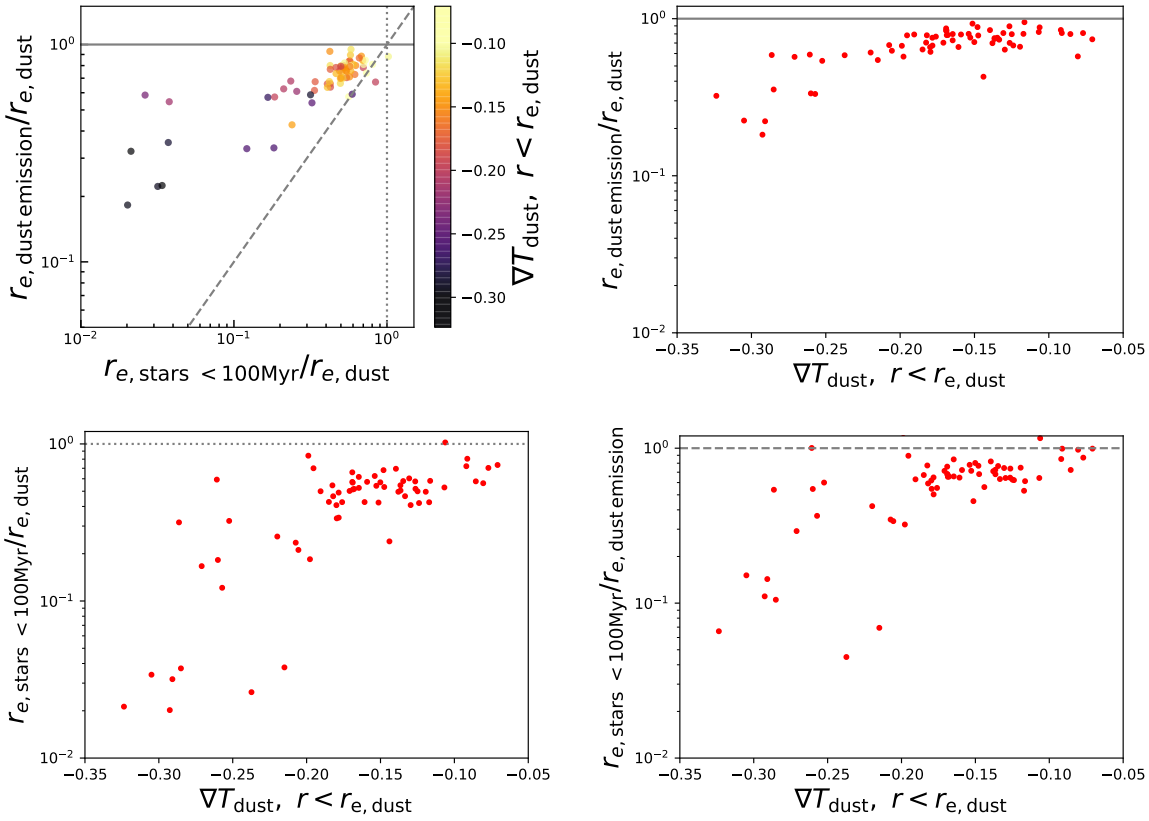

Figure 6.8 *The role of dust temperature gradient in determining the size of the observed-frame 870-μm dust continuum emission. Top left: the x-axis shows the effective radius of recently formed stars* (age $< 100\,\mathrm{Myr}$)*, divided by effective radius of the dust mass itself, for each snapshot studied in this work. The y-axis shows the effective radius of the observed-frame dust emission, divided by effective radius of the dust mass. The solid grey line shows* $r_{e,\,\mathrm{dust\ emission}} = r_{e,\,\mathrm{dust}}$*, the dashed grey line shows* $r_{e,\,\mathrm{dust\ emission}} = r_{e,\,\mathrm{stars}\,<100\,\mathrm{Myr}}$*, and the dotted grey line shows* $r_{e,\,\mathrm{dust}} = r_{e,\,\mathrm{stars}\,<100\,\mathrm{Myr}}$*. Each snapshot is colour-coded by the dust temperature gradient, calculated within concentric shells on the three-dimensional dust grid. Snapshots where dust continuum emission is particularly compact with respect to the dust mass distribution also display even more compact recent star formation. This is associated with steeper dust temperature gradients. The top right and bottom panels show ratios of the effective radii of dust emission, dust mass and recently formed stars, against dust temperature gradient. For gradients steeper than* $\nabla T_{\mathrm{dust}} \sim -0.2$*, dust continuum emission is a poor tracer of the spatial extent of the dust mass or recently formed stars.*

is sensitive to the scales of recent star formation because of the effects of dust heating. Thus, inferring spatial scales of star formation or dust mass from observations of dust continuum emission is non-trivial. This may be done more robustly using measurements of dust temperature gradients within galaxies. Studies of local galaxies show that dust temperature distributions can be derived from spatially-resolved imaging in multiple bands (e.g. Galametz et al., 2012).

6.5 Conclusions

We have performed radiative transfer modelling on a subset of rest-frame FIR-bright redshift snapshots of four massive galaxies drawn from the FIRE-2 simulations presented in Anglés-Alcázar et al. (2017). The simulated galaxies have stellar masses $7 \times 10^{10} - 3 \times 10^{11}\,M_{\odot}$, and reside in dark matter haloes of mass $M_{\rm halo} \sim 10^{12.5}\,M_{\odot}$ at $z = 2$. Our modelling yields full rest-frame FUV-FIR SEDs as well as maps of the emission in ~ 100 wavebands, resolved to 25 pc scales. We find clear differences between the morphologies of the same galaxies in the different wavebands, with shorter-wavelength emission (UV/optical) frequently appearing clumpy and extended. We find that this is due to the emitted short-wavelength light tracing 'holes' in the dust distribution. At rest-frame FIR wavelengths, the galaxies tend to have more regular morphologies.

The primary focus of our study is the spatial extent of the dust continuum emission. For the snapshots we study, the FIRE-2 galaxies have observed-frame 870-μm fluxes of up to ~ 2 mJy at $z \sim 1 - 5$. These simulated FIR-bright galaxies exhibit very compact dust continuum emission, with effective radii of $\sim 0.5 - 4$ kpc, in line with existing observations of star-forming galaxies at these redshifts. At rest-frame FIR wavelengths (which we can probe with ALMA for high-redshift galaxies), longer-wavelength emission tends to be more extended because the dust tends to be hotter in the central regions of galaxies, and the shorter-wavelength emission is dominated by hotter grains.

We also compare the spatial extent of the 870-μm dust continuum emission to that of several key intrinsic physical quantities, including the dust, gas and stellar components. In both simulated and observed galaxies, the dust continuum emission is more compact than the cool, dense gas and the dust, but more extended than the stellar component. Extracting only recently formed (age < 100 Myr) stars from the simulations enables us to study the role of recent star formation in determining the spatial extent of the dust emission. We find that in some snapshots, the simulated galaxies exhibit extremely compact

dust emission ($\sim 0.5\,\mathrm{kpc}$). This emission appears to be driven by particularly compact recent star formation. Overall, the spatial extent of the dust continuum emission correlates with those of both the dust mass and the recently formed stars. Physically, this makes sense: dust that emits strongly in the FIR is predominantly heated by young massive stars (as these simulations do not contain AGN), and when the recent star formation is particularly compact, the central regions of the galaxy have steeper dust temperature gradients and consequently more compact emission. In such systems, constraints on the dust temperature gradient are necessary to infer the spatial extents of the young stars and dust reliably. This work thus motivates multiband ALMA observations to constrain the dust temperature gradients of observed galaxies.

Chapter 7

Resolving a dusty, star-forming galaxy at $z = 2.2$ with HST, ALMA and SINFONI

In this chapter we present high spatial resolution imaging of SHiZELS-14, a massive ($M_* \sim 10^{11}\,M_\odot$), dusty, star-forming galaxy at $z = 2.24$. Our $\sim 0.15''$, matched-resolution data comprise three different widely used tracers of star formation: the Hα emission line (from SINFONI/VLT), rest-frame UV continuum (from *HST* F606W imaging), and the rest-frame far-infrared (from ALMA), as well as the rest-frame optical (from *HST* F140W imaging). SHiZELS-14 displays a compact, dusty centre, as well as extended emission in both Hα and the rest-frame FIR. The UV emission is spatially offset from the extended dust emission, and appears to trace holes in the dust distribution. We find that the dust attenuation varies across the spatial extent of the galaxy, reaching a peak of $A_{\mathrm{H}\alpha} \sim 5$ in the most dusty regions. Global star formation rates inferred using standard calibrations to the different tracers vary from $\sim 10 - 1000\,M_\odot \mathrm{yr}^{-1}$, and are particularly discrepant in the galaxy's dusty centre. This galaxy highlights the biased view of galaxy evolution provided by UV data in the absence of long-wavelength data.

7.1 Introduction

As discussed in Chapters 1 and 6, star formation within individual galaxies is known to become more extreme towards high redshift. While highly luminous dusty galaxies are rare at $z = 0$ and known as 'ultra luminous infrared galaxies'

(ULIRGs, with total infrared luminosities $L_{\rm TIR} > 10^{12-13}\,L_{\odot}$; Smail et al. 1997; Barger et al. 1998), galaxies with typical ULIRG luminosities are more common around the peak of cosmic star formation. Submillimeter galaxies (SMGs) are ULIRGs at high redshift with bright submillimeter fluxes corresponding to star formation rates of $\sim 1000\,M_{\odot}$/yr. Chapman et al. (2005) found that the volume density of SMGs increases by a factor of ~ 1000 between $z = 0$ and $z = 2.5$, with the redshift distribution peaking at $z \sim 2.2$ (broadly in agreement with Smith et al. 2017, who find that 68 per cent of SMGs reside in the redshift range $1.07 < z < 3.06$, though note that brighter SMGs seem to lie at higher redshifts; Hayward et al. 2013; Koprowski et al. 2014). SMGs at $1 < z < 5$ appear to account for $\sim 20-30$ per cent of the total comoving star formation rate density at these redshifts (Swinbank et al., 2014; Smith et al., 2017). Even in less FIR-luminous galaxies, a significant amount of star formation at high redshift is obscured by dust. Dunlop et al. (2017) combined long- and short-wavelength data from two of the astronomy community's premier observatories: ALMA (probing the dust continuum emission at 1.3 mm) and Hubble Space Telescope (Wide Field Camera 3, probing rest-frame UV), in the well-studied Hubble Ultra Deep Field (Bouwens et al., 2010; Oesch et al., 2010; Illingworth et al., 2013; Dunlop et al., 2013; Ellis et al., 2013). As noted in Chapter 6, these complementary data enabled them to show that $\sim 85\%$ of the total star formation at $z \sim 1-3$ is enshrouded in dust, with star formation in massive galaxies being most highly attenuated.

While studies of wide fields are important in tracking the evolving properties of star-forming galaxies and the build-up of stellar mass in the Universe, understanding the physical processes of star formation within individual galaxies requires higher angular resolution. Resolved studies of distant star-forming (SF) galaxies tend to be based on observations from near-infrared integral field unit spectrographs, which probe rest-frame optical emission lines such as Hα and [OIII] at $z \sim 2$ (e.g. Genzel et al., 2008; Swinbank et al., 2012a; Reddy et al., 2015; Stott et al., 2016; Simons et al., 2017), or from *HST* at rest-frame UV wavelengths (e.g. Wuyts et al., 2012; Fisher et al., 2017). These largely converge on a physical picture in which star formation takes place within massive clumps embedded in turbulent disk structures (Genzel et al., 2008; Elmegreen et al., 2013; Guo et al., 2015, 2017; Soto et al., 2017). Emission at these short wavelengths is, however, strongly attenuated by dust, and star formation in the dustiest regions is hidden from view. The significant global obscuration of star formation quantified by Dunlop et al. (2017) at $z < 4$ therefore suggests that our understanding of galaxy evolution from short-wavelength studies is likely to be highly biased by

dust, even when the measurements are at the high spatial resolution required to resolve sub-galactic scales.

Recent work made possible by new submillimeter interferometers, in particular ALMA and NOEMA which offer both high sensitivity and spatial resolution, has focused on characterising the spatially-resolved properties of SMGs at long wavelengths. As discussed in Chapter 6, dust continuum emission and CO emission appear very compact for these galaxies, with typical half-light radii $\sim 1-2\,\mathrm{kpc}$ (Tadaki et al., 2016, 2017, 2018; Hodge et al., 2016; Oteo et al., 2017; Strandet, 2017; Calistro Rivera et al., 2018). A number of studies have shown that these sizes are comparable to the optical sizes of $z \sim 2$ compact quiescent ellipticals, galaxies that must have formed a huge amount of stellar mass and then quenched early (Krogager et al., 2014; Onodera et al., 2015; Belli et al., 2016). This, together with the large stellar masses of SMGs ($M_* \sim 10^{11} M_\odot$) has fuelled speculation that the SMGs detected at $z \sim 3-6$ are the progenitors of $z = 2$ massive ellipticals (e.g. Toft et al., 2014; Oteo et al., 2017; Gómez-Guijarro et al., 2018). Other work has focused on SMGs as tracing a rapid phase of bulge-building, with the compact dust-obscured central regions hiding extreme star formation (e.g. Tadaki et al., 2016; Nelson et al., 2019).

However, observations of compact dust continuum sizes appear to be at odds with the extended, clumpy structures traced by *HST* imaging (Barro et al., 2016; Hodge et al., 2016). In some cases, kpc-scale offsets have been found between FIR and UV emission (Tadaki et al., 2016; Chen et al., 2017; Calistro Rivera et al., 2018, see also the simulation work in Chapter 6). These offsets could bias interpretations of global measurements. Indeed, Simpson et al. (2017a) argue that attenuation in the dusty regions of SMGs is so great that essentially all the co-located stellar emission is obscured at optical-to-near-infrared wavelengths; for $\sim$ 30 per cent of their sample, the data available at these wavelengths is insufficient to put constraints on photometric redshifts and stellar masses. Chen et al. (2017) show that this can be catastrophic for inferences of the infrared emission from the IRX $-$ β relation, the empirical law that relates the spectral slope at 1600 Å (β) to the ratio of IR-to-UV luminosity (Meurer et al., 1999).

Overall, it has become clear that drawing conclusions from single-wavelength surveys, especially in the rest-frame UV, is subject to substantial bias and uncertainty, even at high angular resolution. Spatially-resolved observations of the dust continuum emission of SMGs are most powerful when performed in combination with short-wavelength observations. In this Chapter, we present multi-wavelength, $0.15''$ resolution imaging of SHiZELS-14, a highly star-forming,

Hα-selected galaxy at $z = 2.24$. This galaxy is one of the most massive of the SHiZELS parent sample ($M_* \sim 10^{11.1\pm0.1} M_\odot$), with the largest of all SHiZELS Hα-derived half-light radii (4.6 ± 0.4 kpc) (Swinbank et al., 2012a,b). Our observations comprise matched-resolution imaging of the Hα emission line (from SINFONI/VLT), rest-frame UV and optical continuum (from *HST*), and the rest-frame far-infrared (from ALMA). We find bright, extended structures in the multiwavelength imaging, with clear clumps in Hα and extended dust continuum emission. Given this extended structure and the high signal-to-noise that results from its high star formation rate (SFR), we have been able to resolve star formation on sub-kpc scales.

The structure of this Chapter is as follows. In Section 7.2, we provide an overview of the data available for our study of SHiZELS-14. We review the high quality, but less well-resolved multi-wavelength data available from imaging of the COSMOS field, and present new 0.15″ resolution imaging from SINFONI/VLT, *HST* and ALMA, with details of the astrometric calibration. In Section 7.3, we present the global properties of SHiZELS-14 that may be inferred from spectral energy distribution fitting. We show that the SFRs derived using widely-used flux-SFR calibrations vary wildly between wavebands. In Section 7.4, we present maps of the spatially-resolved SFRs inferred from different SFR indicators, and derive a spatially-resolved dust attenuation map. We show that the structures and SFRs vary because of the spatial distribution of dust within the galaxy. In Section 7.5 we summarise the main results of this Chapter.

7.2 Overview of SHiZELS-14 data

Our galaxy is drawn from the HiZELS sample used in Chapters 3 and 4. As well as providing the sample sizes for population studies such as these, HiZELS has also provided parent samples for more detailed follow-up observations (Sobral et al., 2013b; Magdis et al., 2016; Stott et al., 2016). A subset of bright Hα emitters ($f_{\mathrm{H}\alpha} > 0.7 \times 10^{-16}\,\mathrm{erg\,s^{-1}cm^{-2}}$) residing within 30″ of bright natural guide stars ($R < 15$) were targeted for IFU spectroscopy of the Hα line using adaptive optics with the SINFONI Integral Field Unit on the VLT (see Swinbank et al., 2012a,b; Molina et al., 2017). This campaign, known as SINFONI-HiZELS (SHiZELS), yielded high-resolution spectral maps for 20 galaxies at $z = 0.84$, $z = 1.47$ and $z = 2.23$ at $\sim 0.15''$ (~ 1 kpc) resolution.

In order to understand the nature of the clumpy Hα emission seen in many of the SHiZELS galaxies and the role of dust attenuation, we complement these

data with imaging at similar angular resolution but different wavelengths. Nine HiZELS galaxies were targeted at $\sim 0.1''$ resolution with ALMA Bands 6 & 7, to map the dust continuum emission. UVIS Imaging in the F606W and F140W filters obtained during *HST* Cycle 24 completes this dataset. We now have FIR-UV-Hα matched-resolution observations of a handful of HiZELS galaxies. Since these galaxies are Hα-selected, they are likely to be more representative of the high-redshift SF population than UV-selected samples, which target the largest, bluest and least dusty galaxies (see Oteo et al., 2015).

Here, we present data for SHiZELS-14, which is the brightest, most extended and more extreme object in our sample. These properties enable a detailed investigation of the multi-wavelength extended structures of this galaxy. In the following subsections, we provide details of the new high-resolution imaging we have recently obtained. We also describe the existing data available for our multi-wavelength characterisation of this galaxy. SHiZELS-14 is a $z = 2.24$ source, with high stellar mass ($M_* \sim 10^{11} M_\odot$; Swinbank et al. 2012a; Laigle et al. 2016), and a star formation rate of many hundreds of solar masses per year.

7.2.1 Resolving Hα emission with SINFONI

Full details of the SHiZELS SINFONI observations are presented in Swinbank et al. (2012a,b) and Molina et al. (2017). We present only the most important aspects here. Fourteen galaxies with Hα fluxes in the range $0.7 - 1.6 \times 10^{-16}\,\mathrm{erg\,s^{-1}cm^{-2}}$ residing within $40''$ of bright stars with R-band magnitude $12 < R < 15$ were selected for natural guide star AO observations with the SINFONI IFU as part of the initial campaign. These observations took place between September 2009 and April 2011, in good seeing and photometric conditions ($\sim 0.6''$), with exposure times in the range $3.6 - 13.4$ ks. This yielded sub-kpc resolution Hα maps of nine galaxies. Further observations were carried out between October 2013 and September 2014, in similar conditions, with exposure times in the range $4.8 - 9.6$ ks (Molina et al., 2017). This increased the size of the successfully-resolved SHiZELS sample to 20 galaxies (6 at $z = 0.84$, 8 at $z = 1.47$ and 6 at $z = 2.23$). SHiZELS-14, the focus of this Chapter, was the only $z = 2.23$ source resolved in the initial Swinbank et al. (2012a,b) campaign. The exposure time for this source was 12.0 ks.

Data reduction and analysis procedures are outlined in full in Swinbank et al. (2012a,b) (see also Molina et al. 2017; Gillman et al. 2019). In summary, the SINFONI ESOREX data reduction pipeline was used to perform extraction, flat fielding and wavelength calibration, and to create the data cube for each exposure.

These data cubes were then stacked and combined using an average with a 3σ clip, to reject cosmic rays. Flux calibration was performed using observations of standard stars taken immediately before/after science exposures, which were reduced in the same way. Hα and [NII] $\lambda\lambda$ 6548, 6583 emission lines were fitted on a pixel-by-pixel basis, using a χ^2 minimisation procedure. This yielded intensity, velocity, and velocity dispersion maps. An angular resolution of $\sim 0.15''$ was achieved. The spectral resolution of the instrument is $\lambda/\Delta\lambda \sim 4500$.

The Hα flux derived from the SINFONI observations of SHiZELS-14 is $1.6 \pm 0.1 \times 10^{-16}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. The half-light radius is $4.6 \pm 0.4\,\mathrm{kpc}$ (Swinbank et al., 2012a). The Hα image is shown in the bottom left-hand panel of Figure 7.1.

7.2.2 Resolving UV and optical light with *HST*

We targeted nine SHiZELS galaxies with resolved Hα maps with 18 orbits of *HST* during Cycle 24 (Program 14719). Five of these targets lie at $z = 1.47$, and four at $z = 2.23$. Each source was targeted in one orbit (2700 s exposure) with each of the WFC3/UVIS Imaging F606W filter and the WFC3/IR Imaging F140W filter. Orbits were split into a 3-point dither pattern in the UVIS channel, as a compromise between maximising sensitivity and sub-sampling the PSF. Since angular resolution was preferred over sensitivity in the IR channel, a 4-point dither pattern was used for these orbits. At $z = 2.24$, the filters correspond to the rest-frame near-UV at 1900 Å, and the rest-frame optical at 4350 Å for SHiZELS-14. Our observations were designed to span the 4000 Å break, and therefore sample both young and old stellar populations, in line-free regions of the galaxy spectrum. The *HST* images are shown in the top panels of Figure 7.1.

7.2.3 Resolving the rest-frame far-infrared emission with ALMA

Eight sources with resolved Hα maps, four at $z = 1.47$ and four at $z = 2.23$, were targeted with ALMA with Bands 7 and 6 respectively, to map dust continuum emission at rest-frame frequency 840 GHz. Observations were taken during ALMA Cycles 1, 2 and 3 and combined to obtain deep images. Images were manually cleaned down to $3\,\sigma$ (rms $\sim 25\,\mu$Jy/beam) at the source positions. The Band 6 image of SHiZELS-14 has the highest signal-to-noise of all the sources, showing a distinct extended structure with a $16\,\sigma$ peak.

We used Briggs (robust=0) visibility weighting, which assigns higher weights to

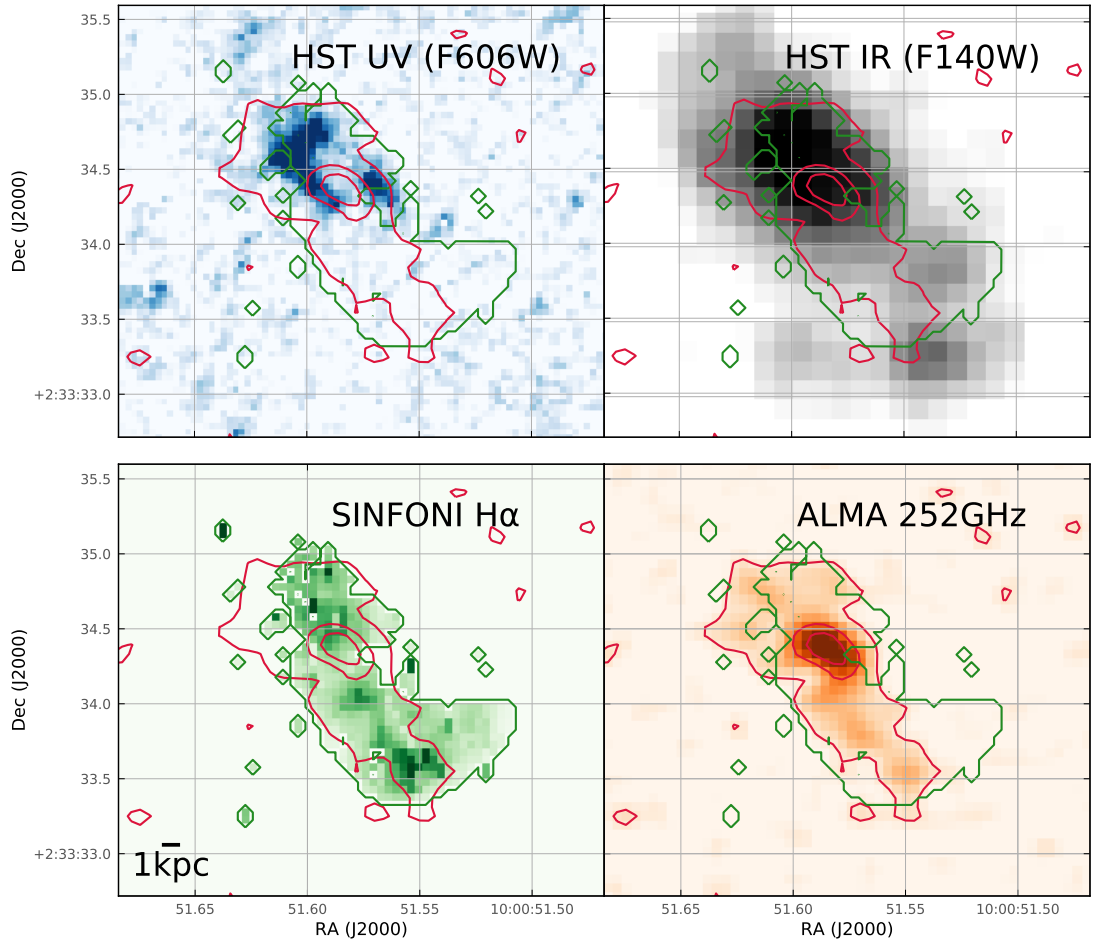

Figure 7.1 *Astrometry-calibrated, high-resolution observations of SHiZELS-14 in the rest-frame UV (HST F606W filter; top left), rest-frame optical (HST F140W filter; top right),* Hα *(SINFONI/VLT; bottom left) and dust continuum (ALMA; bottom right). The red contours on all panels outline the ALMA dust continuum emission at 50, 200, and 300 μJy/beam. The green SINFONI contours outline the 3σ emission as described in Section 7.2.1. The emission imaged by SINFONI, ALMA and the HST F140W filter span the same extended region, but display very different morphologies. The peak of the dust emission probed by ALMA is clearly offset from the peaks of the short-wavelength emission. This is particularly striking for the F606W rest-frame UV emission, which is concentrated in regions with little dust emission and does not extend down to the southern regions that are clearly probed by the other bands.*

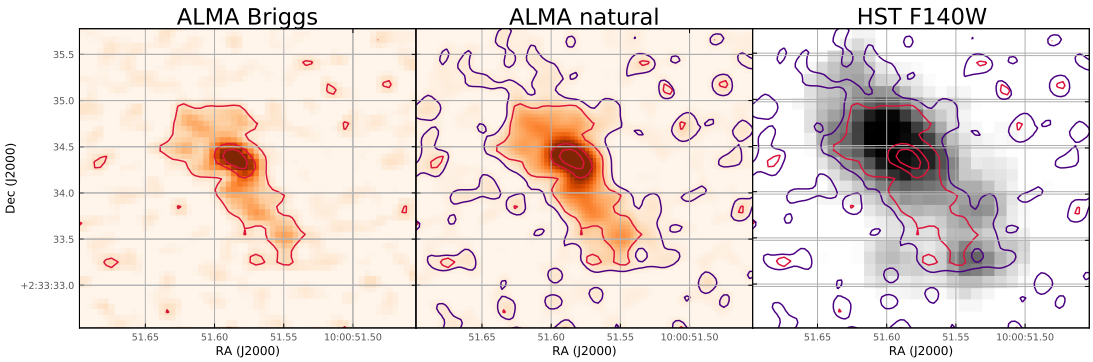

Figure 7.2 252 GHz *ALMA imaging of SHiZELS-14, reduced with different weightings. Left: image produced directly from original ALMA data products, with red contours showing 50, 200, and 300 μJy/beam, as in Figure 7.1. Centre: image produced from data re-reduced using natural weighting, with purple contours showing 25 μJy/beam and red contours outlining the contours of the Briggs-reduced image. Both ALMA images are plotted on the same flux scale. Right: HST F140W image with contours of the two ALMA weightings overlaid. The slightly lower angular resolution natural-weighted image shows flux towards the top left and the bottom right, in the regions with extended F140W flux.*

longer baselines, producing an image with higher angular resolution (see the image in the bottom right-hand panel of Figure 7.1). We have checked how changing the visibility weighting changes the reduced ALMA image. We re-reduced the ALMA data using a natural weighting, which weights visibilities only by the rms noise (see middle panel of 7.2). This method minimises the noise level but provides poorer angular resolution, given that the density of visibilities falls towards the outskirts of the uv-plane. Using the re-reduced, lower angular resolution natural-weighted image, we probe to slightly lower flux density per beam. This will be used to assess the quality of our astrometric calibration in Section 7.2.5.

SHiZELS-14 has an observed-frame 252 GHz flux density of 2.7 ± 0.15 mJy. It displays a compact, ~ 3 kpc core of dust emission, with extended emission contributing substantially to the flux. Its half-light radius is notably larger, 3.9 ± 0.1 kpc, due to this extended faint emission. This is consistent with the large half-light radius derived from Hα.

7.2.4 Radio observations from COSMOS-VLA

We make use of the deep existing radio observations in the COSMOS field made available by the VLA-COSMOS surveys. The VLA-COSMOS Large Project (Schinnerer et al., 2007) surveyed 2 square degrees in VLA A-array and C-array

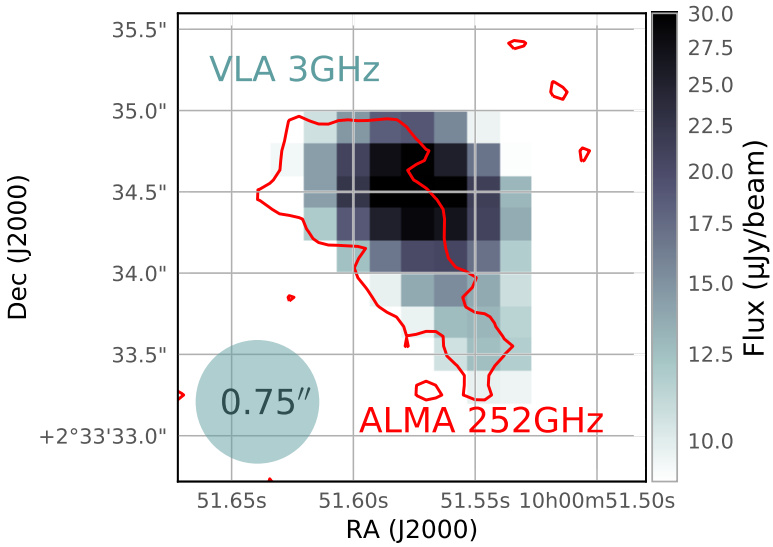

Figure 7.3 *Radio image from the VLA-COSMOS* 3 GHz *Large Project (dark blue; no astrometric correction applied), with higher angular resolution ALMA Band 6 contours overlaid. The astrometric alignment is good, though comparison of the convolved ALMA image with the lower resolution radio image indicates that there is an offset of* ~ 1 *VLA pixel (*$0.2''$*).*

configurations at 1.4 GHz (20 cm). The project yielded images with rms noise $\sim 10-15\,\mu$Jy/beam at angular resolution $1.5'' \times 1.4''$. The VLA-COSMOS Deep project (Schinnerer et al., 2010) added further A-array observations at 1.4 GHz in the central region of the COSMOS field. The VLA-COSMOS 3 GHz GHz Large Project (Smolčić et al., 2017) subsequently surveyed 2.6 square degrees at a wavelength of 10 cm with the upgraded A-array, reaching a mean rms depth of $\sim 2.3\,\mu$Jy/beam at $0.75''$ angular resolution.

SHiZELS-14 is one of the sources detected by these VLA surveys (see Figure 7.3). The measured flux densities are $S_{1.4\mathrm{GHz}} = 119 \pm 27\,\mu$Jy and $S_{3\mathrm{GHz}} = 68 \pm 4\,\mu$Jy. The lower angular resolution of the radio images limits our ability to probe resolved structure, but the source is still extended in the radio. We also use the total flux density to derive a star formation rate. We will obtain higher angular resolution observations at $4-8$ GHz with the VLA in A array configuration during Cycle 2019A (see Chapter 8).

7.2.5 Astrometric calibration

Accurate astrometric alignment is critical when comparing multi-wavelength emission on these small angular scales. However, due to the small fields of view of both the SINFONI and ALMA data, aligning the images is non-trivial. Here,

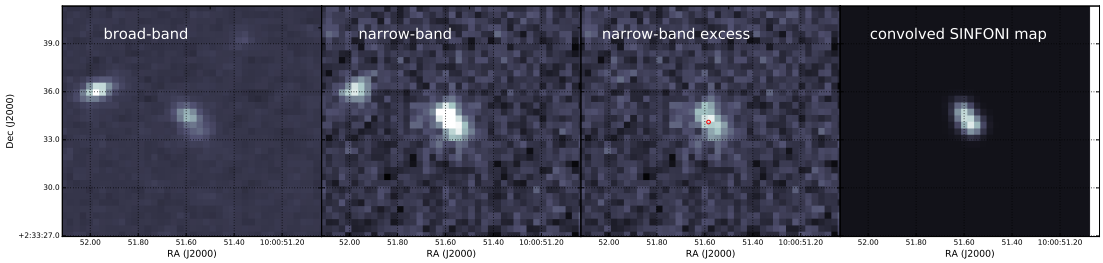

Figure 7.4 *The alignment of the SINFONI image. The SINFONI image was iteratively shifted and convolved to the resolution of the HiZELS broad-band-subtracted narrow-band image, which had been aligned to the Two Micron All-Sky Survey. Subtracting the SINFONI image from the narrow-band excess image enabled a χ^2 fit to define the optimal alignment.*

we describe alignment of the images.

We assume that the ALMA image is on the International Celestial Reference System (ICRS), and then check the quality of the alignment to the other images. Although calibration errors and self-calibration processes can lead to astrometric offsets at the pixel level, there are no other detected sources within the ALMA field of view for us to align with. As described in Section 7.2.3, we re-reduced the image at slightly lower angular resolution, to recover fainter flux densities at the outskirts of the galaxy. As shown in the middle panel of Figure 7.2, there is faint flux towards the bottom right, tracing the regions that show extended F140W flux. This, combined with the excellent alignment of the SINFONI and ALMA images, gives us confidence in the alignment of the ALMA image.

We then align all other images to the ICRS. The SINFONI image was aligned to the same reference frame as the main HiZELS survey. We used a broad-band-subtracted narrow-band image, which had been aligned to the Two Micron All-Sky Survey (2MASS), which itself uses the ICRS. We shifted the SINFONI image by sub-pixel quantities, and convolved down to the resolution of the broad-band image (see Figure 7.4). Subtracting the images enabled a χ^2 fit to define the optimal alignment. Based on these comparisons, we are able to achieve an accuracy on the SINFONI image alignment of $\sim 0.2''$.

We calibrated the astrometry of the *HST* images by aligning to 2MASS directly. For both the F606W and F140W images, we extracted sources within the field of view using the SEXTRACTOR software (Bertin & Arnouts, 1996). We matched those within $1''$ of catalogued 2MASS sources to derive histograms of offsets in RA and Dec. The peaks of these histograms were selected as the offset to be applied to each of the *HST* images. The offsets applied in RA and Dec were $-0.041''$ and $-0.25''$ for *HST* F606W, and $-0.21''$ and $-0.12''$ for F140W.

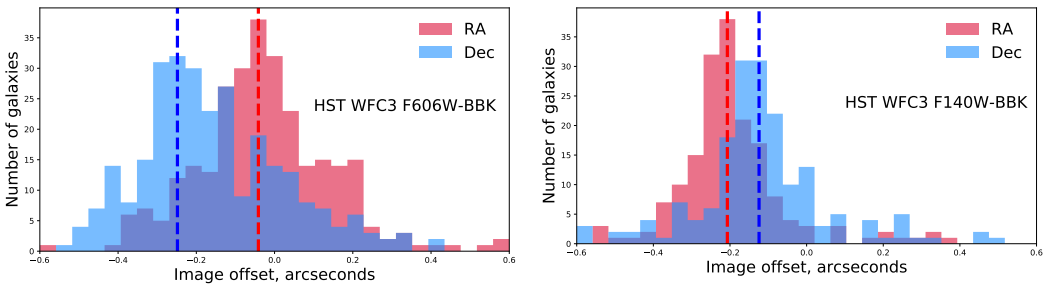

Figure 7.5 *The offsets between sources in the two HST images (prior to astrometric calibration) and broad-band K images from HiZELS. The RA and Dec offsets applied are shown by the dashed lines (respectively, $-0.041''$ and $-0.25''$ for HST F606W, and $-0.21''$ and $-0.12''$ for F140W).*

Based on the widths of the histograms, we estimate that the alignment is correct to well within $0.2''$.

7.2.6 Comparison of morphologies

Figure 7.1 shows our four resolved maps after these small astrometric corrections were applied. The emission in all bands is aligned along the same axis. However, the peak of the dust emission probed by ALMA is clearly offset from the peaks of the FUV and Hα emission. These offsets are far larger than the residual astrometric uncertainties. The dust emission is centrally concentrated, whereas there are a number of Hα peaks along the extended region where dust emission is faint. There is a peak in the emission from both *HST* bands towards the top left of the image, yet no detectable dust emission. This is in line with the excess in the CFHT u-band emission (compared to the Subaru $r+$ and $g+$ bands) shown in Figures 7.6 and 7.7. Such offsets are also seen in simulated dusty galaxies, as shown in Chapter 6.

7.2.7 Optical/IR data from the COSMOS field

A wealth of lower resolution data exists for this galaxy due to its privileged location within the well-imaged COSMOS field (Scoville et al., 2007). At NUV-optical wavelengths, COSMOS was imaged in the u*-band from the Canada-France-Hawaii Telescope (CFHT/MegaCam), and in 6 broad bands (B, V, g, r, i, z+), 12 medium bands (IA427, IA464, IA484, IA505, IA527, IA574, IA624, IA679, IA709, IA738, IA767, and IA827), and two narrow bands (NB711 and NB816), all from the COSMOS-20 survey (Subaru Suprime-Cam; Taniguchi et al. 2007, 2015). Y-band imaging was obtained with HSC (Miyazaki et al., 2012). In the NIR, Y,

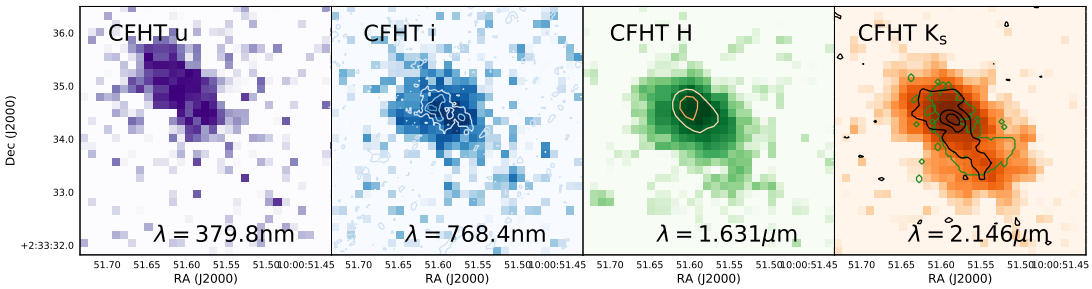

Figure 7.6 *NUV-NIR imaging of SHiZELS-14 from CFHT. These observations are seeing-limited, with angular resolution* $\sim 0.8 - 1.0''$. *We overlay contours from our resolved imaging campaigns on relevant panels. Overplotted on the CFHT i-band image are contours from HST F606W imaging (blue). The contours on the CFHT H-band image are from HST F140W imaging (orange). Both SINFONI* Hα *(green) and ALMA dust continuum emission (black) contours are overplotted on the CFHT* K_s*-band image.*

J, H, & K_s data are provided by the UltraVISTA-DR2 release (McCracken et al., 2015), which uses the VIRCAM instrument on the VISTA telescope. These are supplemented by H and K WIRCAM data (McCracken et al., 2010). Mid-IR data are drawn from IRAC channels 1, 2, 3 and 4 (3.6 μm, 4.5 μm, 5.8 μm and 8.0 μm), collected by the *Spitzer* Large Area Survey with Hyper-Suprime-Cam Subaru (HSC) project (SPLASH survey; Lin et al. 2017; Capak et al. in prep).

Laigle et al. (2016) improve upon the earlier work of Ilbert et al. (2013), collating these observations and providing an NIR-selected photometric redshift catalogue. We use the 3″ diameter aperture fluxes extracted for SHiZELS-14 by Laigle et al. (2016). We provide a summary here and tabulate these measurements in Table 7.3.

7.2.8 Data at mid-IR and far-IR wavelengths

At mid-IR and far-IR wavelengths, we draw data from *Spitzer* and *Herschel* imaging. We adopt the 24 μm flux density from the *Spitzer* Multiband Imaging Photometer (MIPS; Rieke et al. 2004). The *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al. 2012) targeted COSMOS at wavelengths 100 − 500 μm. The survey used *Herschel*-Spectral and Photometric Imaging Receiver (SPIRE) at 250 μm, 350 μm and 500 μm and the *Herschel*-Photodetector Array Camera and Spectrometer (PACS) at 100 μm and 160 μm. One of the main aims of the *Herschel* Extragalactic Legacy Project (HELP)[1] was to develop the advanced statistical tools needed to de-blend the low-resolution data from

[1] http://herschel.sussex.ac.uk

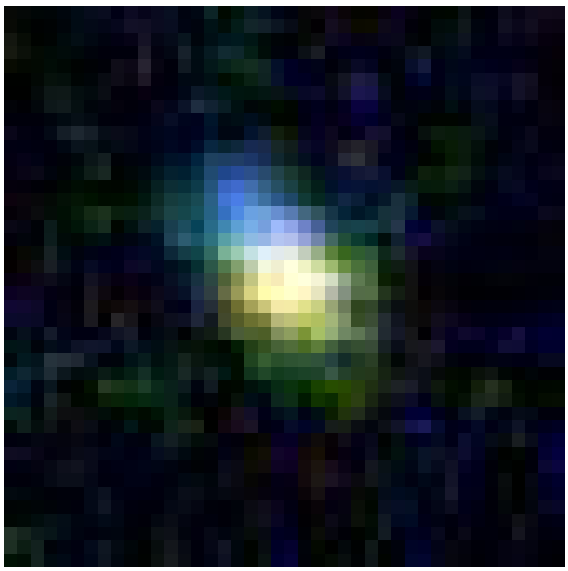

Figure 7.7 *Pseudo-colour image constructed using Subaru r_+ & g_+ bands and CFHT u band. There is a clear excess of u-band light towards the top left of the galaxy.*

Herschel, in order to assign fluxes to components (Hurley et al., 2017; Pearson et al., 2017). We use these publicly available, catalogued flux densities for SHiZELS-14.

We adopt the ALMA Band 7 flux density measured by Scoville et al. (2014). The total 350 GHz flux density is 4.67 ± 0.75 mJy, and the peak flux density is 1.93 ± 0.32 mJy.

7.3 Global properties of SHiZELS-14 from Spectral Energy Distribution fitting

Before examining the resolved structures of SHiZELS-14, we place these into context by deriving the global properties of the galaxy. As discussed in Section 1.1.6, SED fitting provides a powerful basis for estimating galaxy properties from photometry. Previous photometric SED fitting of SHiZELS-14 has provided estimates for its stellar mass and star formation rate. Laigle et al. (2016) used LEPHARE to derive the stellar masses of galaxies in the COSMOS field, using the fluxes in their catalogues. They used a library of synthetic spectra generated using the Stellar Population Synthesis model (Bruzual & Charlot 2003; BC03), and a Chabrier (2003) initial mass function. The star formation history supplied to the fit is a combined exponentially-declining SFH and delayed SFH. They input two metallicities (solar and half-solar), emission lines following Ilbert et al. (2009),

and two attenuation curves: the starburst curve of Calzetti et al. (2000) and a curve with a slope of $\lambda^{0.9}$ (Arnouts et al., 2013). $E(B-V)$ was allowed to vary up to 0.7. The stellar mass derived for SHiZELS-14 from the best-fitting BC03 template was $\log_{10} M_*/M_\odot = 11.23 \pm 0.05$. The star formation rate derived from the template fit was SFR $= 200 \pm 30\,\mathrm{M}_\odot/\mathrm{yr}$. Scoville et al. (2014) also derived a star formation rate, using a combination of the rest-frame UV continuum and the IR, from *Herschel* PACS and SPIRE data, following Scoville et al. (2013), obtaining SFR $= 525\,\mathrm{M}_\odot\mathrm{yr}^{-1}$.

7.3.1 Fitting the NUV-MIR SED with BAGPIPES

The SED fits of Laigle et al. (2016) were optimised to provide fits to over half a million galaxies. Given that we study only one galaxy here, we can afford to perform more detailed fitting. In particular, we extend the range of parameters input to the SED fits. We use the SED fitting code Bayesian Analysis of Galaxies for Physical Inference and Parameter EStimation, BAGPIPES (Carnall et al., 2018, 2019), to re-fit the short-wavelength photometric data collated by Laigle et al. (2016) and the two ALMA datapoints. Our fitting makes use of the 2016 version of the BC03 simple stellar population (SSP) templates, with a Kroupa (2002) IMF (note that the difference between a Kroupa and Chabrier IMF is negligible). Nebular emission is computed using the CLOUDY photoionization code (Ferland et al., 2017), following Byler et al. (2017). CLOUDY is run using each SSP template as the input spectrum, with the ionisation parameter, U, also allowed to vary. Dust grains are included using CLOUDY's 'ISM' prescription, which implements a grain-size distribution and abundance pattern that reproduces the observed extinction properties for the ISM of the Milky Way. We select a Calzetti et al. (2000) dust attenuation curve. Dust emission includes both a hot dust component from HII regions and a grey body component from the cold, diffuse dust.

For our SED fit of SHiZELS-14, we include the Megacam/CFHT u^* band, all the Suprime-Cam/Subaru bands, VIRCAM/VISTA bands, the four *Spitzer*/IRAC bands, one *Spitzer*/MIPS band, 5 *Herschel*-HerMES bands, and two ALMA bands (see Table 7.3 for the fluxes). We impose a wide dust attenuation prior, $A_v = [0, 6]$, which gives the code the option to fit a high degree of attenuation. We fit three parameters that affect the shape of the dust SED, following Draine & Li (2007): U_{min}, the lower limit of the starlight intensity; γ, the fraction of stars at U_{min}; and q_{PAH}, the mass fraction of polycyclic aromatic hydrocarbons. Our priors on these parameters are broad, to allow the model to fit hot dusty galaxies:

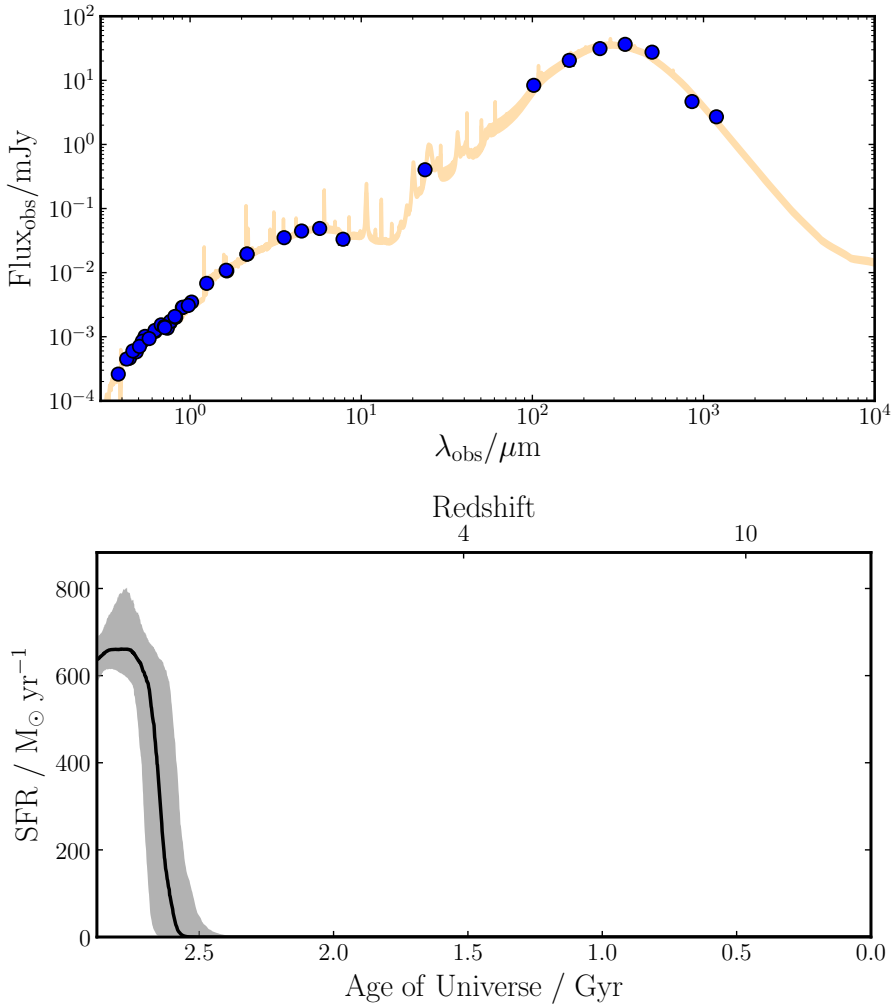

Figure 7.8 *Top: data presented in Section 7.3, fitted with the BAGPIPES code (Carnall et al., 2018), using a wide A_v prior and a 'double power law' star formation history. A minimum flux uncertainty of* 10% *is imposed. Bottom: posterior estimate of the star formation history of SHiZELS-14. The fitting favours a recent burst of star formation, with* $\mathrm{SFR} = 690 \pm 110\,\mathrm{M_\odot/yr}$ *over the last* $\sim 2 \times 10^8$ *years. The derived stellar mass is* $\log_{10} M_*/M_\odot = 11.1 \pm 0.1$.

Basic property	Measurement	Source
RA (J2000)	10 00 51.6	Swinbank+12
Dec (J2000)	+02 33 34.5	Swinbank+12
$z_{\mathrm{H}\alpha}$	2.2418	Swinbank+12
Derived property	**Measurement**	**Source**
$\log_{10} M_{*,SED}/M_{\odot}$	11.1 ± 0.1	This thesis
$\log_{10} M_{\mathrm{gas}}/M_{\odot}$	10.1 ± 0.4	Swinbank+12
$\log_{10} M_{\mathrm{dust}}/M_{\odot}$	9.1 ± 0.1	This thesis
$\log_{10} L_{\mathrm{TIR}}/L_{\odot}$	12.81 ± 0.02	This thesis
$\mathrm{SFR}_{\mathrm{TIR}}/\mathrm{M}_{\odot}\mathrm{yr}^{-1}$	950 ± 50	This thesis
$r_{e,\,\mathrm{H}\alpha}/\mathrm{kpc}$	4.6 ± 0.4	Swinbank+12
$r_{e,\,\mathrm{FIR}}/\mathrm{kpc}$	3.9 ± 0.1	This thesis

Table 7.1 *Summary of properties of SHiZELS-14. Full details of SFRs derived using different methods are presented separately in Table 7.2.*

$U_{\mathrm{min}} = [0, 25]$, $\gamma = [0, 1]$, and $q_{\mathrm{PAH}} = [0, 10]$. We also fit η, the multiplicative factor on A_V for stars in birth clouds, using the range $\eta = [1, 5]$. We allow metallicity to vary in the range $Z = [0, 2.5]\, Z_{\odot,\mathrm{old}}$, where $Z_{\odot,\mathrm{old}}$ denotes solar models prior to Asplund et al. (2009). We fix the redshift at $z = 2.2418$, as this is known from the SINFONI spectrum.

We experiment with various SFH parametrisations, which yield very similar fits to the spectrum and consistent values for stellar mass $\log_{10} M_*/M_{\odot} = 11.1 \pm 0.1$ (in good agreement with the estimate of Laigle et al. 2016). All parametrisations, even those allowing multiple bursts, favour a recent (at $z = 2.24$), rapid burst of star formation in which the vast majority of the stellar mass is formed. In Figure 7.8, we plot a representative fit to the photometry with corresponding SFH. This particular model uses a double power law SFH parametrisation. The posterior estimate for the star formation rate is $\mathrm{SFR} = 690 \pm 110\,\mathrm{M}_{\odot}/\mathrm{yr}$ (substantially higher than the estimate of Laigle et al. 2016), and the estimated specific star formation rate (sSFR) is $\log_{10} \mathrm{sSFR}/\mathrm{yr}^{-1} = -8.25^{+0.12}_{-0.10}$. Note that the SFR is more sensitive than the stellar mass to the parametrisation of the SFH and the data included in the fit. For example, if we exclude data at wavelengths longer than $24\,\mu$m, the fit favours a shorter burst of recent star formation, which increases the 100 Myr SFR average. The posterior estimate for the dust attenuation in the V-band is $A_v = 1.8 \pm 0.1$. We will return to this high attenuation in Section 7.4.

7.3.2 Fitting the dust SED

We fit the MIR-FIR SED of SHiZELS-14 using data from ALMA, *Herschel* and *Spitzer* separately. We parametrise the emission from cold and warm dust using a simple two-body model:

$$f_\nu(\mathrm{mJy}) = \mathrm{A_{warm}}\lambda^{-\beta_{\mathrm{warm}}}\mathrm{B}_\nu(\mathrm{T_{warm}}) + \mathrm{A_{cold}}\lambda^{-\beta_{\mathrm{cold}}}\mathrm{B}_\nu(\mathrm{T_{cold}}), \tag{7.1}$$

where A_{warm} and A_{cold} are normalisations and $B_\nu(T)$ is the Planck function, from dust grains radiating at rest frequency ν_{rest}, at temperature T. All wavelengths were input at their rest-frame. In line with the literature, we have fixed $\beta = 2$ for both the cold and warm dust components, to minimise the number of fitting parameters. We use the EMCEE MCMC python package (Foreman-Mackey, 2016), with 300 walkers and 5000 steps. This yields posterior estimates: $\log_{10} A_{\mathrm{warm}} = 5.4 \pm 0.3$, $T_{\mathrm{warm}} = 64 \pm 6\,\mathrm{K}$, $\log_{10} A_{\mathrm{cold}} = 7.6 \pm 0.1$, and $T_{\mathrm{cold}} = 28 \pm 2\,\mathrm{K}$. The best-fitting model is shown in Figure 7.9. Note that there is a known strong correlation between β_{cold} and T_{cold}, and a 5-parameter fit that allows β_{cold} to vary favours a higher β_{cold} and a lower T_{cold}.

7.3.3 The inappropriateness of the IRX $-$ β relation

The IRX $-$ β relation (Calzetti et al., 1994; Meurer et al., 1999) between the ratio of the FIR and UV luminosity ($\mathrm{IRX} = \mathrm{L_{FIR}}/\mathrm{L_{1600}}$) and the spectral slope (β, where $f_\lambda \propto \lambda^\beta$) evaluated at 1600 Å is a popular method used to infer SFRs where only rest-frame UV luminosities are available. This generally works well for samples of galaxies with low dust content (especially at very high redshift). However, individual galaxies show a large amount of scatter around this relation, and it has been shown that this method is not appropriate highly star-forming galaxies, particularly those with geometrical offsets (e.g. Casey et al., 2014; Chen et al., 2017; Narayanan et al., 2018).

Using the data available, we can derive both IRX and β for SHiZELS-14. We use the publicly available *HST* I-band image ($\lambda_{\mathrm{mean}} = 13970$ Å, rest-frame $\lambda_{\mathrm{mean}} = 4309$ Å), along with our own F606W images ($\lambda_{\mathrm{mean}} = 6001$ Å, rest-frame ($\lambda_{\mathrm{mean}} = 1851$ Å), to calculate β. Adopting our derived $\beta = -0.5 \pm 0.1$, and applying the relation $A_{1600} = 4.43 + 1.99\beta$, we derive $A_{1600} = 3.4 \pm 0.2$. Correcting the global SFR inferred from the FUV flux accordingly yields $\mathrm{SFR} = 300^{+70}_{-50}\,\mathrm{M_\odot yr^{-1}}$. This is a substantially lower SFR estimate than inferred from the SED fitting. We calculate IRX using the TIR luminosity derived in Section

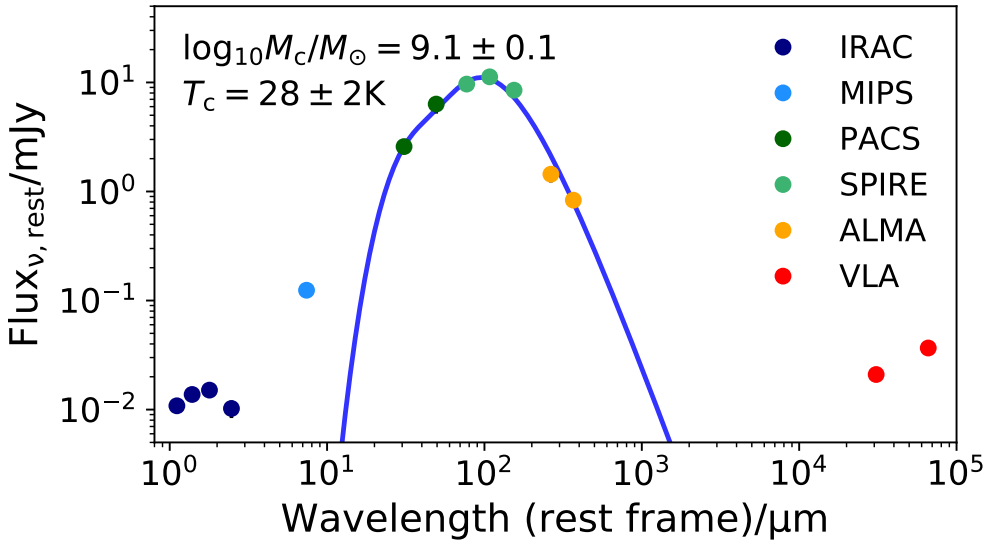

Figure 7.9 *SHiZELS-14 dust SED, constructed using collated archival data and the new ALMA 252 GHz data. A two grey body model parametrisation provides a good fit to both the cold and warm dust components, yielding temperature ($T_{\mathrm{cold}} = 28 \pm 2\,K$) and dust mass ($\log_{10} M_{\mathrm{cold}} = 9.1 \pm 0.1\,M_\odot$) estimates. Integrating the $8 - 1000\,\mu m$ emission provides a TIR-based SFR ($SFR_{8-1000\mu\mathrm{m}} = 950 \pm 50\,M_\odot/\mathrm{yr}$).*

7.3.4, and the rest-frame 1851 Å luminosity. Globally, the galaxy has $\log_{10}\mathrm{IRX} = 2.09 \pm 0.06$. In combination with the derived β, this places it ~ 0.7 dex above the Meurer et al. (1999) relation. This highlights that the galaxy has a higher TIR luminosity than expected from the derived UV slope. This is likely to be because the UV and FIR fluxes are not co-located, as shown in Figure 7.1. SHiZELS-14 highlights that the IRX relation may not provide reliable estimates of the FIR emission for the most dusty galaxies.

7.3.4 Calculation of cold dust mass and TIR luminosity

The dust mass is:

$$M_{\mathrm{dust}} = \frac{1}{1+z}\frac{S_{\mathrm{obs}} D_L^2}{\kappa_{\mathrm{rest}} B_\nu(T_{\mathrm{cold}})} \tag{7.2}$$

where $S_{\mathrm{obs}} = 3.2$ mJy is the flux density within a 50 pixel diameter, D_L is the luminosity distance, $\nu_{\mathrm{rest}} = 816.48$ GHz and κ_{rest} is the mass absorption coefficient at the rest frequency.

Calculating the dust mass within the MCMC fit enables us to fold in the correlations between fitted parameters. We used $\kappa_{850} = 0.07 \pm 0.02\,\mathrm{m}^2/\mathrm{kg}$ (James

et al., 2002), which gives $\log_{10} M_{\rm dust}/M_\odot = 9.1 \pm 0.1$. This provides a high dust-to-stellar mass ratio, $\log_{10} M_{\rm dust}/M_* = -1.9 \pm 0.2$, which is comparable to the ratios derived by Calura et al. (2017) for SMGs of stellar mass $\sim 10^{11}\,M_\odot$ at $z \sim 1-3$. We also integrate the two-body fits at wavelengths $8-1000\,\mu$m within the MCMC fit, obtaining an estimate for the total IR luminosity, $\log_{10}(L_{\rm TIR}/{\rm erg\,s^{-1}}) = 46.39 \pm 0.02$, and $\log_{10}(L_{\rm TIR}/L_\odot) = 12.81 \pm 0.02$. The TIR-based SFR is $950 \pm 50\,M_\odot {\rm yr}^{-1}$.

7.3.5 Global SFR estimation

In Table 7.2 we present global SFR estimates from global measurements in different wavebands, using the calibrations of Kennicutt & Evans (2012) and assuming a Kroupa (2002) IMF. It is clear that applying standard SFR calibrations to flux measurements at different wavelengths provides little consensus on the total star formation rate of this galaxy. SFRs derived from indicators that probe dust-obscured emission are orders of magnitude higher than from those from dust-unobscured emission. This suggests that the discrepancy in global SFR derived from different SFR indicators is due to the highly dusty nature of this galaxy. In the following section, we explore the differences in the spatially-resolved SFRs, derived at different wavelengths.

7.3.6 Position on the IR-radio relation

As discussed in Section 1.1.5, the ratio of IR to radio luminosity (e.g. Appleton et al., 2004) is frequently employed to separate radio-loud AGN from star-forming galaxies. Following Sargent et al. (2010), we use the following equation with the TIR luminosity calculated in Section 7.3.4:

$$q_{\rm TIR} = \log\left(\frac{L_{\rm TIR}}{3.75 \times 10^{12}{\rm W}}\right) - \log\left(\frac{L_{1.4{\rm GHz\,rest}}}{{\rm WHz^{-1}}}\right). \tag{7.3}$$

The rest-frame 1.4 GHz luminosity is:

$$L_{1.4{\rm GHz,\,rest}} = \frac{4\pi D_L^2}{(1+z)^{1+\alpha}}\left(\frac{\nu_{1.4{\rm GHz}}}{\nu_{\rm obs}}\right)^{\alpha} S_{1.4{\rm GHz,\,obs}} = 10^{24.54\pm0.08}\,{\rm WHz^{-1}}. \tag{7.4}$$

We assume a spectral index $\alpha = -0.7$. This gives $q_{\rm TIR} = 2.28 \pm 0.10$. This is in line with the mean $q_{\rm TIR}$ for the radio-selected sample of Sargent et al. (2010)

SFRs from individual tracers		
Waveband (Instrument)	Formula for $\log_{10}(\mathrm{SFR}/\mathrm{M}_\odot\mathrm{yr}^{-1})$	$\mathrm{SFR}/\mathrm{M}_\odot\mathrm{yr}^{-1}$
FUV (*HST* F606W)	$\log_{10}(\nu L_\nu/\mathrm{erg\ s}^{-1}) - 43.17$	13 ± 1
"	corrected using A_{1600} derived from β, with $\beta = -0.5 \pm 0.1$	300^{+70}_{-50}
"	corrected using $A_{UV} = 4.5 \pm 0.2$, derived from scaled $A_{\mathrm{H}\alpha}$	820^{+170}_{-140}
Hα (SINFONI/VLT)	$\log_{10}(L_{\mathrm{H}\alpha}/\mathrm{erg\ s}^{-1}) - 41.27$	33 ± 2
"	$L_{\mathrm{H}\alpha}$ corrected using 1 mag dust extinction	83 ± 5
"	$L_{\mathrm{H}\alpha}$ corrected using M_*-dependent dust extinction of Garn et al. (2010), $\log_{10} M_*/M_\odot = 11.1$	168 ± 11
"	$L_{\mathrm{H}\alpha}$ corrected using $A_{\mathrm{H}\alpha} = 2.1 \pm 0.1$, from SED fitting	227 ± 20
$\mathrm{TIR}_{8-1000\mu\mathrm{m}}$ (dust SED fit)	$\log_{10}(L_{\mathrm{TIR}}/\mathrm{erg\ s}^{-1}) - 43.1$	950 ± 50
Radio (1.4GHZ, VLA, Kennicutt & Evans 2012 conversion)	$\log_{10}(L_{1.4\mathrm{GHz,\,rest}}/\mathrm{erg\ s}^{-1}\,\mathrm{Hz}^{-1}) - 28.2$	2010 ± 170
Radio (1.4GHZ, VLA, Bell & Kennicutt 2001 conversion)	$\log_{10}(L_{1.4\mathrm{GHz,\,rest}}/\mathrm{erg\ s}^{-1}\,\mathrm{Hz}^{-1}) - 28.43$	1180 ± 100
SFRs from combinations of tracers		
FUV + TIR	$L_{\nu,\mathrm{corr}} = L_{\nu,\mathrm{obs}} + 0.27 L_{\mathrm{TIR}}$, L_ν − SFR conversion above	440 ± 20
Hα + TIR	$L_{\mathrm{H}\alpha,\mathrm{corr}} = L_{\mathrm{H}\alpha,\mathrm{obs}} + 0.0024 L_{\mathrm{TIR}}$, $L_{\mathrm{H}\alpha}$ − SFR conversion above	330 ± 20
FUV + radio	$L_{\mathrm{FUV,corr}} = L_{\mathrm{FUV,obs}} + 4.2 \times 10^{14} L_{1.4\mathrm{GHz}}$	920 ± 80

Table 7.2 *The global SFR of SHiZELS-14, derived from different combinations of SFR indicators, using a Chabrier (2003) IMF. Inferring an SFR from dust-uncorrected fluxes at short-wavelengths yields SFRs of* $< 50\,M_\odot$/yr. *Applying either standard dust correction corresponding to* $A_{\mathrm{H}\alpha}$*=1 or a stellar mass dependent dust correction provides a modest increase in SFR. However, no dust correction brings the UV or* Hα*-inferred SFRs into line with the high values derived from the TIR and radio (*$\mathrm{SFR} = 1000 - 2000\,\mathrm{M}_\odot$/yr*).*

$(2.17^{+0.1}_{-0.28})$. We confirm this using q derived from the 24μm flux:

$$q_{24} = \log\left(\frac{S_\nu(24\,\mu\mathrm{m})}{S_\nu(1.4\,\mathrm{GHz})}\right) \tag{7.5}$$

Using observed-frame fluxes (i.e. without applying K-corrections), we obtain $q_{24} = 0.52 \pm 0.07$. Appleton et al. (2004) obtain $q_{24,\mathrm{mean}} = 0.84 \pm 0.28$ for star-forming galaxies, which they show to remain broadly constant to $z = 1$, though note that Sargent et al. (2010) find a slight decreasing trend of q_{24} with increasing redshift with larger samples.

Both q values for SHiZELS-14 are $\sim 1\sigma$ from the mean relations derived for star-forming galaxies. We find no evidence that SHiZELS-14 is host to a radio-loud AGN. We will explore the radio-IR relation in more detail and in a spatially resolved manner in future work with planned VLA $4-8$ GHz observations.

7.3.7 The lack of evidence for AGN activity

As discussed in Section 7.3.5, the star formation rates derived from the UV, Hα and FIR differ greatly. In this section, we investigate whether the presence of an active galactic nucleus (AGN) could be a factor in this. In this scenario, the extreme dust continuum emission towards the centre of the galaxy could be powered by heating from a central AGN, rather than a compact region of star formation. Since different types of AGN emit in different wavebands (see Heckman & Best 2014 for a review), identification of AGN requires a multi-wavelength approach. Here, we use some of the key methods for AGN identification to hunt for signs of AGN activity.

No sign of X-ray emission

X-ray emission probes the accretion disk corona very close to a supermassive black hole. The Chandra COSMOS-Legacy Survey (Civano et al., 2016) imaged 2.2 deg^2 of the COSMOS field in the wavelength range $0.5-10$ keV. SHiZELS-14 lies in the outer region of the COSMOS-Legacy field, where effective exposure times are ~ 80 ks (compared to ~ 160 ks for the deeper inner region). At this depth, SHiZELS-14 is undetected.

No mid-infrared excess

Obscured AGN are characterised by a strong mid-infrared (rest-frame $\sim 3-30\,\mu$m) excess, produced by a dusty obscuring torus. Our SED shows no sign of

such an excess, with the single data point in this wavelength range, SPIRE 24 μm (rest-frame 7.4 μm) well-fitted by an SED constructed without AGN templates. Fitting the SED with CIGALE, which does include emission from AGN, provides no evidence of an AGN ($f_{\mathrm{AGN,\,best}} = 0.001$). In addition to this, the temperature derived from fits to the dust SED is 28 ± 2 K, well within the normal range for star-forming galaxies.

7.4 Resolved star formation rates and dust attenuation

7.4.1 Resolved star formation rates

In Figure 7.10, we present maps of SFR per square arcsecond, derived for each of the four SFR indicators using the flux-SFR calibrations of Kennicutt & Evans (2012) and Bell & Kennicutt (2001). It is clear that the SFRs derived using these standard calibrations differ across the entire spatial extent of the galaxy. To investigate this more quantitatively, we derive star formation rate radial profiles by applying Kennicutt & Evans (2012) calibrations to the rest-frame FUV F606W, Hα, and TIR flux maps (see Figure 7.11, thick dashed lines). The three profiles are wildly discrepant, with the TIR-based SFR profile increasing sharply towards the centre where the other two flatten. Without any corrections for dust attenuation, the FUV and Hα-derived SFRs are lower than the FIR-derived SFR across the radial extent of the galaxy. The FUV profile broadly follows the Hα profile in shape, but with a different normalisation. The FUV is more strongly attenuated by dust, and yields the lowest dust-uncorrected SFRs across the galaxy. Thus, the discrepancy between the SFRs derived globally cannot be attributed solely to the compact dusty centre of the galaxy, though this is where the measurements are most discrepant. Instead, it appears that short-wavelength light is suffering substantial attenuation across the galaxy.

We also show the affects of applying a dust correction of $A_{\mathrm{H}\alpha} = 2$, which is a conservative estimate of the dust attenuation in the outskirts of the galaxy (as shown in Figure 7.12), in order to bring the SFR profiles into better agreement. A_{UV} is calculated from $A_{\mathrm{H}\alpha}$ according to the Calzetti et al. (2000) law. These dust corrections bring the outer regions of the FUV and Hα profiles further towards agreement (see transparent solid lines). However, it is clear that the TIR-derived SFR is still much higher in the centre, at radii less than 2 kpc. This reflects strong central star formation and a steep gradient in dust attenuation across the galaxy.

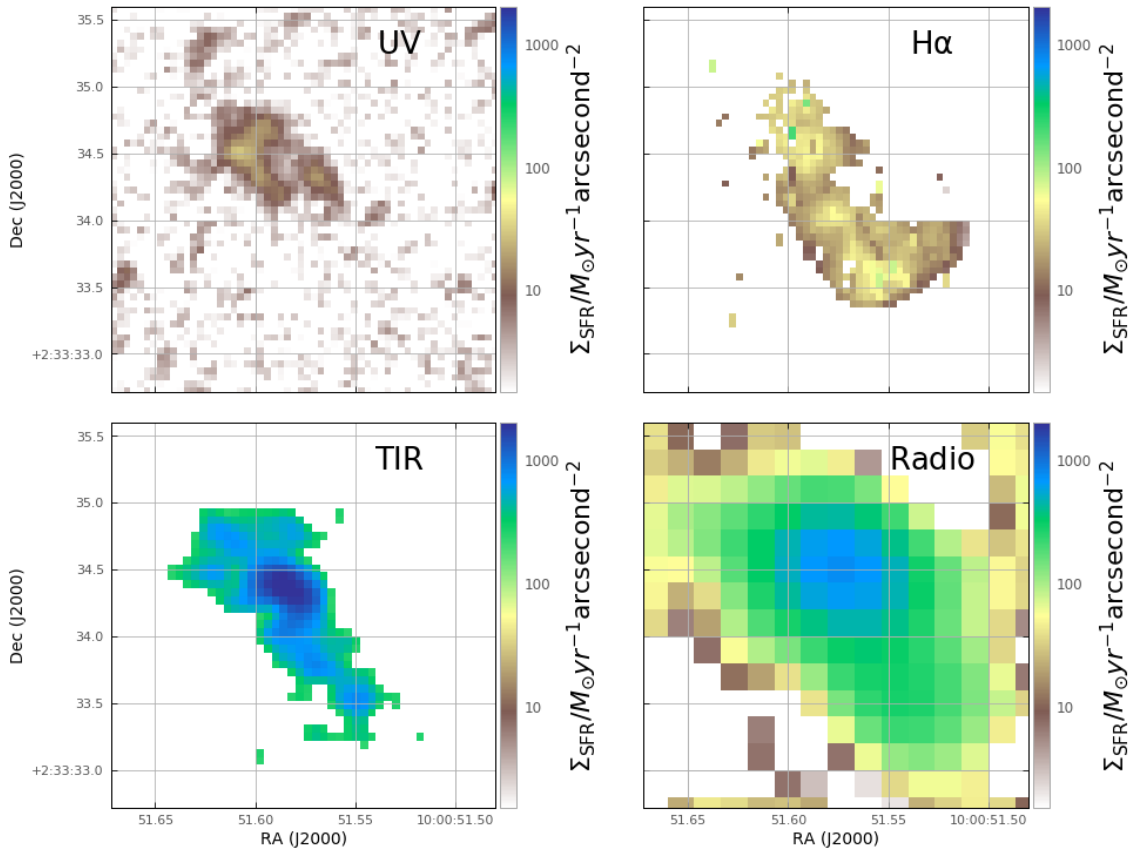

Figure 7.10 *Maps of SFR surface density, derived for each of the four SFR indicators using the flux-SFR calibrations of Kennicutt & Evans (2012) (top two and bottom left panels) and Bell & Kennicutt (2001) (bottom right panel). We have made a flux cut on the TIR map to avoid an overly noisy image. We plot the maps on the same log scale, to compare the SFRs directly. It is clear that the derived SFRs differ across the spatial extent of the galaxy, not simply in its dusty centre.*

7.4.2 Inferring dust attenuation using Hα and FIR maps

In Figure 7.10, we showed that the SFR surface densities derived in different wavebands from dust-uncorrected fluxes are wildly different. We can use this to estimate the spatially-resolved dust attenuation. In the left-hand panel of Figure 7.12, we present the ratio of the TIR-dervived SFR to the Hα-derived SFR (with no dust correction applied). We can also use this ratio of the fluxes to estimate $A_{H\alpha}$ in a spatially-resolved way, as follows. Folding in a dust-correction to the Hα flux, and then equating the two SFRs:

$$\mathrm{SFR/M_{\odot}yr^{-1}} = \mathrm{L_{TIR}} \times 10^{-43.41} = \mathrm{L_{H\alpha}} \times 10^{-41.27} \times 10^{0.4\mathrm{A_{H\alpha}}} \quad (7.6)$$

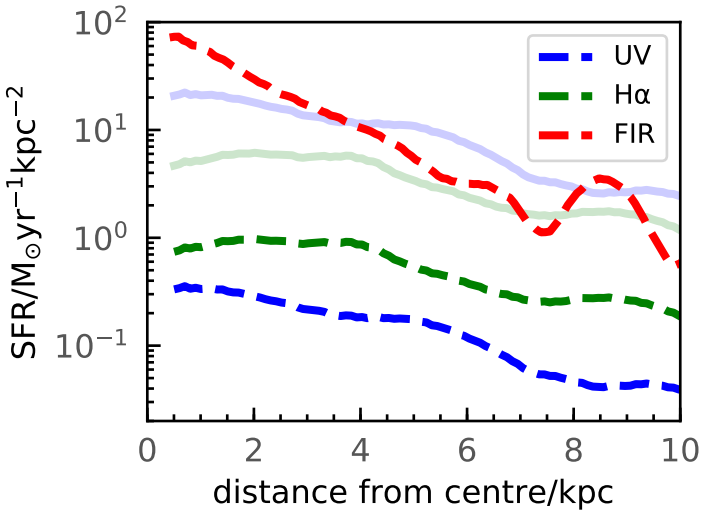

Figure 7.11 *Star formation rate surface density profiles derived using rest-frame FUV F606W,* Hα*, and rest-frame FIR flux map (scaled to the SFR derived from fits to the dust SED). The profiles are centred on the flux-weighted centre of the* Hα *emission. The thick dashed lines show the surface densities derived using Kennicutt & Evans (2012) calibrations, with no dust corrections applied. The solid transparent lines show the profiles derived using an* $A_{\mathrm{H}\alpha} = 2$ *(*$A_{\mathrm{UV}} = 4.5$*, derived using a Calzetti et al. 2000 law) correction.*

yields an expression for $A_{\mathrm{H}\alpha}$,

$$A_{\mathrm{H}\alpha} = 2.5 \log_{10}\left(\frac{\mathrm{L_{TIR}}}{\mathrm{L_{H}\alpha}}\right) - 5.35. \tag{7.7}$$

Note that this method assumes that Hα and FIR flux are tracing only recently formed stars, and sensitive to star formation on the same timescales. Although this breaks down if there is significant contribution to the FIR flux from older stellar populations (e.g. Hayward et al., 2014), our SED fits present a picture of very recent (at $z = 2.24$) star formation, and so this assumption is not unreasonable.

We plot the spatially resolved $A_{\mathrm{H}\alpha}$ in the right-hand panel of Figure 7.12. $A_{\mathrm{H}\alpha}$ substantially exceeds $A_{\mathrm{H}\alpha} = 1$, the canonical value applied to global studies, across the spatial extent of the galaxy. In the most dusty central region, it reaches a peak of $A_{\mathrm{H}\alpha} \sim 5$.

We can then predict the UV flux from the Hα flux, using $A_{\mathrm{H}\alpha}$ and:

$$I_{\mathrm{int,H}\alpha} = I_{\mathrm{obs,H}\alpha} \times 10^{0.4 A_{\mathrm{H}\alpha}} = \frac{10^{41.27}}{10^{43.17}} \times I_{\mathrm{obs,UV}} \times 10^{0.4 A_{UV}}. \tag{7.8}$$

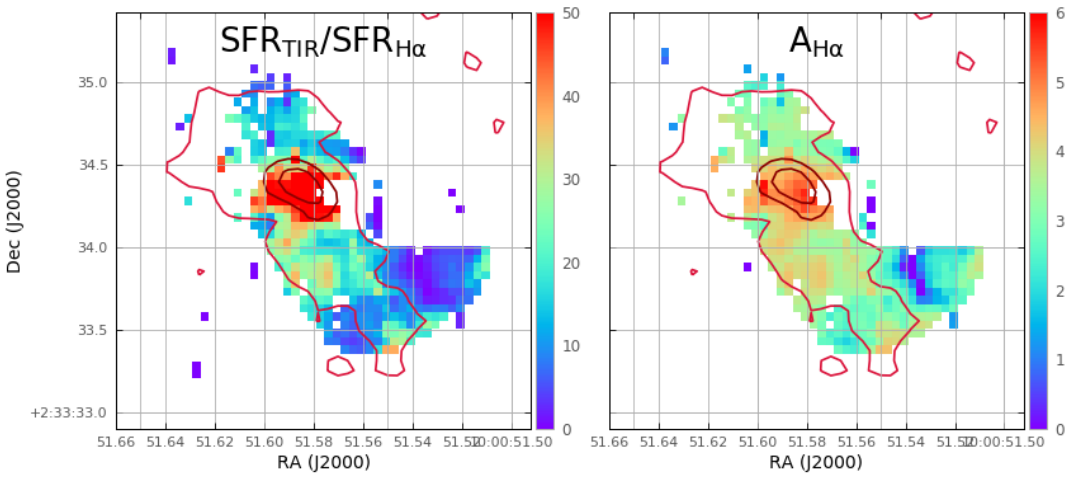

Figure 7.12 *Left: the ratio of TIR-derived SFR to* Hα*-derived SFR without any correction for dust attenuation. The TIR-derived SFR is larger than that derived from* Hα *across the full extent of the galaxy, but is discrepant by a factor of* ~ 50 *in the dusty central region. Right: the dust attenuation* $A_{H\alpha}$ *derived from this ratio.* $A_{H\alpha}$ *varies across the galaxy, within a broad range* $A_{H\alpha} \sim 2-6$*. Surveys such as HiZELS often assume a modest global dust correction of* $A_{H\alpha} = 1$*, but the dust attenuation of SHiZELS-14 derived here is well in excess of this value. ALMA contours are overlaid on both panels.*

The predicted UV flux, along with the observed one, is shown in Figure 7.13. Given the high levels of dust attenuation, the predicted flux is below the 1σ noise level of the UV image ($\sigma = 2.3 \times 10^{18}\,\mathrm{erg\,s^{-1}cm^{-2}}$), so we wouldn't expect to see UV from the recent star formation that the Hα is tracing. This implies that the UV flux that we do observe is tracing star formation on longer timescales. This is consistent with the peak of the stellar mass lying towards the top left of the Hα flux (see the F140W image).

7.5 Conclusions

In this Chapter, we have presented a study of SHiZELS-14, a $z = 2.24$ galaxy originally identified by HiZELS via its Hα emission. SHiZELS-14 was one of the galaxies selected for high spatial resolution follow-up, due to its proximity to a guide star (for adaptive optics observations), rather than any special properties. However, this galaxy has some intriguing features when resolved at high spatial resolution.

The global properties of SHiZELS-14 suggest that it is highly star-forming. SED fits to photometric data favour a strong burst of star formation within $\sim 200\,\mathrm{Myr}$ of $z = 2.24$ and a stellar mass of $10^{11.1\pm0.1}\,M_{\odot}$. Fitting the dust SED

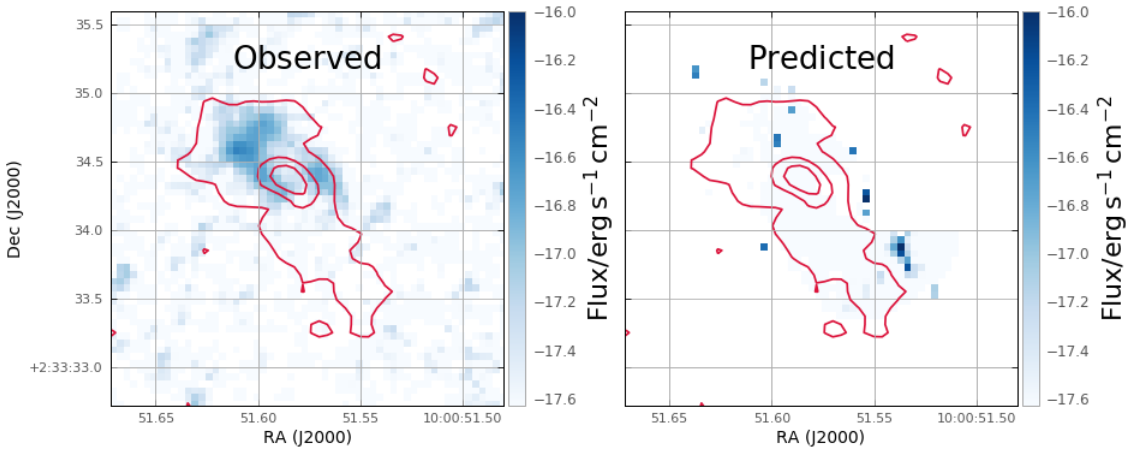

Figure 7.13 *Left: observed UV flux. Right: UV flux predicted from the combination of the* Hα *map and the spatially-resolved* $A_{\mathrm{H}\alpha}$ *shown in Figure 7.12. The flux scale is the same for both panels, with lower limit set at the* 1σ *noise level of the HST F606W observations. We do not expect to detect UV flux from the star formation traced by* Hα*, given the dust attenuation we have estimated. ALMA contours are overlaid on both panels.*

with a two grey-body model yields a dust mass of $M_{\mathrm{dust}} = 10^{9.1\pm0.1}\,M_{\odot}$ and a TIR luminosity of $\log_{10}(L_{\mathrm{TIR}}/\mathrm{erg\,s^{-1}}) = 46.39 \pm 0.02$. This bright IR emission places it in the category of a ULIRG. Unlike many similarly IR-bright galaxies, SHiZELS-14 lies on the radio-IR relation expected for a star-forming galaxy and shows no strong evidence of AGN activity.

FUV, Hα and FIR emission are all used to infer SFR, individually and in combination. We investigate the agreement of widely-used SFR calibrations, globally and in a spatially-resolved manner. Without any dust corrections, the SFRs inferred from FUV and Hα are $13\pm1\,M_{\odot}$/yr and $33\pm2\,M_{\odot}$/yr, respectively. The SFR inferred from the TIR emission is $950 \pm 50\,M_{\odot}$/yr, and the radio-derived SFR is also in the region $1000 - 2000\,M_{\odot}$/yr. Thus, SFR inferred from short wavelength light is orders of magnitude lower than that inferred at longer wavelengths. Even after applying a $A_{\mathrm{H}\alpha} = 1$ dust correction to the Hα emission, the corrected SFR is still only $83 \pm 5\,M_{\odot}$/yr. This suggests that SHiZELS-14 is affected by a large degree of dust attenuation, in line with its substantial dust mass and FIR flux.

We present kpc-scale imaging in the rest-frame FUV and optical (from *HST*), at FIR-wavelengths (from ALMA), and of the Hα emission line (from SINFONI, on the VLT). The range of wavelengths probed enables us to detect both unattenuated and dust-reprocessed emission. SHiZELS-14 shows striking, extended emission in both Hα and the FIR, with Hα-derived half-light radius

4.6 ± 0.4 kpc. This distinguishes it from many SMGs studied at similar redshifts, which tend to be more compact (and often show signatures of AGN activity).

The high spatial resolution of our data enables us to study emission on kpc scales, and compare SFRs in a spatially-resolved manner. We show that the SFR surface density maps derived from UV, Hα, TIR and radio emission are discrepant across the the extent of the galaxy. Comparison of the Hα and TIR maps enables us to map the dust attenuation. We find high levels of dust attenuation across the galaxy, with $A_{\mathrm{H}\alpha} \sim 2-3$ in the outskirts, rising to $A_{\mathrm{H}\alpha} \sim 5$ in the central region. This is in line with studies of SMGs that derive very small effective radii of $1-2$ kpc from the dust continuum emission. The UV flux that is observed appears to arise from older stellar populations, rather than trace the recent star formation, which is strongly affected by dust extinction. We also place our galaxy on the IRX-β plane. Global measurements place the galaxy at high $\beta \sim -0.5$, in line with other dusty galaxies. It lies ~ 0.7 dex above the widely-used IRX-β relation derived by Meurer et al. (1999); for its β slope, it displays substantially higher FIR flux than the canonical law implies. This work highlights the importance of studying galaxies at multiple wavelengths and demonstrates the biases that can be introduced by assuming that calibrations derived using samples of relatively dust-poor galaxies will be appropriate for extremely dusty systems.

Instrument/Telescope (Survey)	Filter	Measurement (μJy)
MegaCam/CFHT	u^*	0.261 ± 0.038
Suprime-Cam/Subaru	B	0.464 ± 0.028
	V	1.012 ± 0.053
	r	1.208 ± 0.047
	i_+	1.656 ± 0.050
	z_+	2.867 ± 0.160
	z_{++}	2.874 ± 0.073
	IA427	0.449 ± 0.078
	IA464	0.599 ± 0.086
	IA484	0.580 ± 0.077
	IA505	0.707 ± 0.092
	IA527	0.850 ± 0.064
	IA574	0.936 ± 0.104
	IA624	1.252 ± 0.075
	IA679	1.532 ± 0.125
	IA709	1.437 ± 0.092
	IA738	1.360 ± 0.096
	IA767	1.737 ± 0.124
	IA827	2.011 ± 0.144
	NB711	1.392 ± 0.164
	NB816	2.071 ± 0.148
HSC/Subaru	$Y_{\rm HSC}$	3.076 ± 0.234
VIRCAM/VISTA (UltraVISTA-DR2)	Y	3.471 ± 0.072
	J	6.804 ± 0.103
	H	10.636 ± 0.159
	K_s	19.396 ± 0.140
WIRCam/CFHT	$K_{\rm sw}$	19.547 ± 0.777
	$H_{\rm w}$	10.892 ± 0.705
Spitzer/IRAC (SPLASH)	IRAC$_{3.6\mu{\rm m}}$	35.121 ± 0.340
	IRAC$_{4.5\mu{\rm m}}$	44.614 ± 0.327
	IRAC$_{5.8\mu{\rm m}}$	48.898 ± 3.574
	IRAC$_{8\mu{\rm m}}$	33.191 ± 6.082
Spitzer/MIPS	24μm	403.0 ± 17.0
Herschel-HerMES/ Oliver+12 HELP catalogue values	100μm	8.368 ± 0.924 (mJy)
	160μm	20.528 ± 3.744 (mJy)
	250μm	31.309 ± 2.175 (mJy)
	350μm	36.486 ± 2.523 (mJy)
	500μm	27.543 ± 2.710 (mJy)
ALMA Band 6,	252GHz	2700 ± 150
ALMA Band 7, Scoville+14	350GHz	4670 ± 750
VLA, Smolčić+17	3GHz	68.0 ± 4.1
VLA, Schinnerer+10	1.4GHz	122 ± 13

Table 7.3 *Compilation of existing measurements of SHiZELS-14, with source. Unless otherwise stated, the data are taken these from the tables of Laigle et al. (2016), adopting their values calculated within a 3″ diameter aperture.*

Chapter 8

Conclusions and future work

8.1 Conclusions

Here I review the main findings of this thesis. In Chapter 2, I outline the clustering and HOD fitting techniques used in Chapters 3 & 4 to quantify galaxy environment. I then use the pioneering cosmological hydrodynamical simulation EAGLE to study the parametrisation of the HOD. I present a flexible HOD parametrisation and show that this is suitable for a variety of stellar mass and SFR-selected galaxy samples.

In Chapters 3 & 4, I present a study of the luminosity and stellar mass-dependent clustering of $0.8 < z < 2.2$ star-forming galaxies selected from the HiZELS survey. I performed clustering analyses of identically-selected Hα-emitting galaxies in 3 narrow redshift slices (at $z = 0.8$, $z = 1.47$ and $z = 2.23$), from HiZELS, a deep, near-infrared narrow-band survey. The HiZELS samples span the peak in the cosmic star formation rate density, identifying typical star-forming galaxies at each epoch. Narrow band samples have well-defined redshift distributions and are therefore ideal for clustering analyses.

I measured the clustering of the three complete samples initially, before splitting each of these by Hα luminosity. I first used simple power law fits to the two-point correlation functions, but then sought to relate these to dark matter halo environments more robustly. I therefore implemented a Halo Occupation Distribution (HOD) model fitting procedure, using the code HALOMOD. The HOD fits provide a superior fit to the two-point correlation function. This also enabled me to consider both the two-halo term of the angular correlation function, which describes the large-scale clustering of host haloes, and the one-halo term, which occurs due to multiple galaxies in the same dark matter halo.

I found that the clustering strength, r_0, and the bias of galaxy populations relative to the clustering of dark matter increase linearly with Hα luminosity (and, by implication, star formation rate) at all three redshifts, as do the host dark matter halo masses of the HiZELS galaxies. The typical galaxies in my samples are star-forming centrals, residing in haloes of mass $M_{\rm halo} \sim$ a few times $10^{12}\,M_{\odot}$. I found a remarkably tight redshift-independent relation between the Hα luminosity scaled by the characteristic luminosity, $L_{\rm H\alpha}/L^{*}_{\rm H\alpha}(z)$, and the host dark matter halo mass of central galaxies. Simple analytic modelling suggests that this is consistent with a model in which the dark matter halo environment is a strong driver of galaxy star formation rate and therefore of the evolution of the star formation rate density in the Universe.

I then distinguished the stellar mass and star formation rate (SFR) dependence of the clustering of these galaxies. At high stellar masses ($M_*/M_{\odot} \gtrsim 2 \times 10^{10}$), where HiZELS selects galaxies close to the so-called star-forming main sequence, the clustering strength is observed to increase strongly with stellar mass (in line with the results of previous studies of mass-selected galaxy samples) and also with SFR. These two dependencies are shown to hold independently. At lower stellar masses, however, where HiZELS probes high specific SFR galaxies, there is little or no dependence of the clustering strength on stellar mass, but the dependence on SFR remains: high-SFR low-mass galaxies are found in more massive dark matter haloes than their lower SFR counterparts. I argued that this is due to environmentally driven star formation in these systems. I applied the same selection criteria to EAGLE galaxies, to see whether these trends were consistent with their predictions. I found that, in EAGLE, the high-SFR low-mass galaxies are central galaxies in more massive dark matter haloes, in which the high SFRs are driven by a (halo-driven) increased gas content.

In Chapter 5, I present work on the quenching of star formation in EAGLE galaxies, in particular the dual roles of stellar mass and halo environment. I characterised the connection between $M_{\rm halo}$, M_* and SFR across redshift using Principal Component Analysis. This enabled me to identify the key axes of correlation between these physical quantities, for the full galaxy sample and split by satellite/central and low/high halo mass. The first principal component of the $z = 0$ EAGLE galaxy population is a positive correlation between $M_{\rm halo}$, M_* and SFR. This component is particularly dominant for central galaxies in low mass haloes. The second principal component, most significant in high mass haloes, is a negative correlation between $M_{\rm halo}$ and SFR, indicative of environmental quenching. For galaxies above $M_* \sim 10^{10}\,M_{\odot}$, however, the SFR is seen to

decouple from the $M_{\rm halo}$–M_* correlation; this result is found to be independent of environment, suggesting that mass quenching effects are also in operation. I found extremely good agreement between the EAGLE principal components and those of SDSS galaxies, lending confidence to EAGLE's predictions.

Extending this study to EAGLE galaxies in the range $z = 0 - 4$, I found that, although the relative numbers of galaxies in the different subsamples change, their principal components do not change significantly with redshift. This indicates that the physical processes that govern the evolution of galaxies within their dark matter haloes act similarly throughout cosmic time.

In Chapter 6, I present the first detailed study of the spatially-resolved dust continuum emission of simulated high-redshift galaxies. I ran the SKIRT radiative transfer code on a subsample of far-infrared/submillimeter-bright snapshots of Milky Way mass, high redshift galaxies drawn from the FIRE-2 simulations. These simulations are the state-of-the-art in zoom-in cosmological hydrodynamical simulations. This modelling provides predictions for the full rest-frame far-ultraviolet-to-far-infrared Spectral Energy Distributions of these simulated galaxies, as well as maps of their emission across the wavelength spectrum, resolved to 25 pc spatial scales.

The derived morphologies are notably different in different wavebands, with the same galaxy appearing clumpy in the far-ultraviolet yet regular and spiral-like at far-infrared wavelengths. The observed-frame 870 μm half-light radii of our FIRE-2 galaxies are $\sim 0.5 - 4$ kpc, consistent with existing ALMA observations of high redshift galaxies. In both simulated and observed galaxies, the dust continuum emission is more compact than the cold gas, but more extended than the stellar component. I show that the most extreme cases of compact dust emission are driven by particularly compact recent star formation, which can drive steep dust temperature gradients.

In Chapter 7, I present a study of the small-scale properties of a $z = 2.24$ galaxy within the COSMOS field, originally identified by its Hα emission. The global properties of this galaxy, SHiZELS-14, suggest that it is highly star-forming, with no indication of AGN activity. SED fits to photometric data favour a strong burst of star formation within ~ 200 Myr of $z = 2.24$ and a stellar mass of $10^{11.1\pm0.1}\,M_\odot$. Fitting the dust SED with a two grey-body model yielded an estimate for the galaxy's dust mass, $M_{\rm dust} = 10^{9.1\pm0.1}\,M_\odot$. The estimated TIR luminosity, $\log_{10}(L_{\rm TIR}/L_\odot) = 12.81 \pm 0.02$, places it in the category of a ULIRG.

Resolving this galaxy at kpc-scales at rest-frame FUV, optical, and FIR-wavelengths, as well as with the Hα emission line, I have been able to probe both

dust-attenuated and dust-unattenuated emission. This galaxy shows striking, extended emission in both Hα and the FIR, with Hα-derived half-light radius 4.6 ± 0.4 kpc. This distinguishes it from many SMGs studied at similar redshifts, which tend to be more compact. I show that the SFR is highly sensitive to the wavelength used to infer it. While standard UV-SFR calibrations yield SFR $\sim 10\,M_\odot$/yr, similar FIR-SFR calibrations suggest SFR $\sim 1000\,M_\odot$/yr. I show that these discrepancies are likely to be due to the galaxy's high dust content and high levels of dust attenuation.

8.2 Future Work I: Studies of the radio continuum emission of SHiZELS-14

Radio wavelengths are perhaps the most powerful tool for studying cosmic star formation, since they probe both obscured and unobscured star formation. Existing surveys such as VLA-COSMOS (Schinnerer et al., 2016; Smolčić et al., 2017) and SKA pathfinders such as LOFAR cover large areas of sky, yielding galaxy samples of unparalleled size. Synchotron radio emission (which dominates the radio spectrum below ~ 30 GHz) has long been established as a reliable tracer of young ($< 3 \times 10^7$ yr), massive ($M > 5\,M_\odot$) stars (Condon, 1992). As a result, widely-used relations exist to calibrate global radio flux to SFR (e.g. Bell 2003, Murphy et al. 2011), although questions remain about the linearity of this relation (e.g Brown et al., 2018) and its cosmic evolution (e.g. Calistro Rivera et al., 2017). The remarkably tight FIR-radio correlation (see Section 1.1.7) has been shown to hold (with debated evolution in normalisation) out to at least $z = 2$ (Ibar et al., 2008; Ivison et al., 2010; Magnelli et al., 2015; Delhaize et al., 2017) for large samples of SF galaxies and radio-quiet AGN. However, global measurements of galaxies mix regions of high and low dust obscuration, making physical interpretations of these relations difficult.

Since only long-wavelength radio emission traces star formation independently of dust, I have obtained VLA time to image SHiZELS-14 at angular resolution comparable to the UV, Hα and FIR data analysed in Chapter 7. Imaging at radio wavelengths, without the uncertainties of dust attenuation, is of particular interest for this source, as the current radio-derived SFR is ~ 10 times higher than is estimated from the Hα emission without accounting for dust. The observations will use the A-array configuration to map SHiZELS-14 at $0.33''$ resolution (robust weighting) and $\sim 1\,\mu$Jy/beam rms depth at $4 - 8$ GHz. The data will be $\sim 2\times$ higher angular resolution than the VLA-COSMOS 3 GHz Large

Project, and $\sim 3\times$ deeper in integrated S/N. This excellent resolution and S/N will be sufficient to characterise the structure of the galaxy's radio emission on comparable scales to our other imaging. I outline the science that the data will enable below.

8.2.1 Mapping star formation in a dust-free manner and calibrating SFR indicators at the peak of cosmic star formation.

Our existing data provides both dust-unobscured (UV & Hα) and dust-obscured (FIR) emission measurements for $\sim$ 100 pixels across the large spatial extent of SHiZELS-14. Unlike other galaxies resolved at this redshift, SHiZELS-14 has a large dust mass ($M_{\rm dust} = 10^{9.1\pm0.1}\,M_{\odot}$), and, using current calibrations, the FIR-derived SFR exceeds the UV and Hα-derived SFR by at least an order of magnitude. With the proposed VLA data, we will be able to obtain a reliable map of the total SFR in a completely dust-independent manner, at matched kpc-scale resolution. We will compare this radio map against SFR maps derived from combinations of the other datasets, to calibrate the four different indicators (which probe different SFR timescales) to SFR accurately, as has been done in the local Universe (e.g. Brown et al., 2018). In particular, spatially resolved SFRs will be derived from the other indicators by scaling and summing the dust-obscured (FIR) and dust-unobscured (UV or Hα) emission for each pixel, using the radio emission to constrain the scale-factors and compare these to previously-derived values. Exploiting the different dust-extinction of the UV and Hα, we will also dust-correct the Hα emission spatially, for comparison with the radio. The comparisons will allow an accurate determination of the radio-SFR relation, which critically underpins the goal of using radio emission as a SFR calibrator at high redshifts with existing and upcoming deep radio surveys.

8.2.2 Studying the spatially-resolved radio-FIR relation.

The short-lived massive stars whose dust clouds emit reprocessed light in the FIR later accelerate cosmic rays when they explode as supernova, resulting in synchotron emission at radio wavelengths. The observed FIR-radio correlation thus reflects the fact that both trace the process of star formation in a galaxy. However, the particular tightness of the relation across orders of magnitude is puzzling: dust cannot be a perfect calorimeter, as many galaxies show substantial UV emission, and energetic cosmic ray electrons may escape from small galaxies

(Bell et al., 2003). Furthermore, the evolution of inverse Compton losses from CMB photons and changing magnetic fields make the suggested lack of observed redshift evolution quite extraordinary. There is therefore an urgent need to improve our understanding of the physical processes that conspire to maintain this tight FIR-radio relation. Spatially resolved studies at high redshifts are key to this but rarely performed. Using the SHiZELS-14 FIR data from ALMA and the radio imaging proposed here, we will be able to study the high-redshift, resolved FIR-radio correlation. This has only previously been done for very local galaxies, where, interestingly, the smallest scale at which the FIR-radio correlation holds appears to vary galaxy-to-galaxy (Tabatabaei et al., 2013).

8.3 Future Work II: Mapping star formation and stellar mass assembly at $z = 1.5 - 2.2$

Having studied SHiZELS-14 in detail, I will exploit the full dataset of ~ 10 star-forming galaxies at at the peak of cosmic star formation, imaged at matched sub-kpc resolution in multiple wavebands as part of the SHiZELS survey.

8.3.1 Star formation across the spatial extent of SHiZELS galaxies

For dusty galaxies like SHiZELS-14, the FIR-derived SFRs vastly exceed the UV and Hα-derived SFRs, as shown in Chapter 7. Long-wavelength data is critical in obtaining robust measurements of the total SFR. The combination of dust-unobscured and dust-obscured emission will allow us to test flux-SFR calibrations that were derived for galaxies in the nearby Universe and answer a key question: do these hold for high redshift galaxies in a spatially resolved way? Using the ratio of $\frac{\mathrm{SFR_{H\alpha}}}{\mathrm{SFR_{UV}}}$ together with maps of dust emission, I will also test whether clumps observed in short-wavelength data are genuine or tracing holes in a clumpy dust distribution. This is crucial for understanding the nature of star formation and hence the physical drivers of galaxy evolution in the distant Universe.

8.3.2 How does stellar mass assemble within galaxies?

The build-up of stellar mass at the peak of cosmic star formation is an important topic in the field. In particular, the morphological change from disky to spheroidal that appears to accompany the quenching of SFR in individual galaxies is not well

understood. One popular theory is that of 'inside-out growth', whereby a massive stellar bulge forms from rapid central star formation (e.g. Tacchella et al., 2015). Resolved studies of the relation between galaxy stellar mass and SFR at early epochs can test such theories, and are thus critical for understanding quenching and morphological transformations.

I will exploit data gathered using the F140W filter, which better traces light from old stellar populations (e.g. Wuyts et al., 2012). In combination with existing 12-band UV-optical-IR photometry for each galaxy, the two HST images will enable the decomposition of the images into high resolution maps of the young and older stellar populations separately. I will also be able to refine my UV SFR estimates, and distinguish between clumpiness due to star formation in large-scale instabilities, or due to merging activity.

8.3.3 Preparations for the upcoming James Webb Space Telescope

The James Webb Space Telescope (JWST; launch planned for 2021) will be a transformative infrared telescope. One of its primary science goals is to study the assembly of galaxies, from the first galaxies in the Epoch of Reionisation, through to the present day. I am part of a collaboration (PI Best) planning to use NIRCam to survey the Hα emission of $z > 6$ galaxies, in a similar vein to the lower redshift HiZELS survey. One of the parallel goals of this survey is to map the distributions of ionised gas in approximately 1000 $z \sim 2$ star-forming galaxies at sub-kpc resolution. The quality of the data will be similar to the SINFONI Hα observations of the HiZELS galaxies, but the sample size will permit a far more detailed statistical analysis. Combining the data with continuum observations of old stellar populations from JWST's F200W filter and matched-resolution rest-frame UV data from the existing HST ACS/WFC3 data, I will investigate how the relationship between UV and ionised gas structures depends upon host galaxy properties, such as mass, morphology and star formation rate.

I will lead a subset of the preparations for this survey and analysis of the early data. The details of this will depend on the key science questions driving the survey by the time the telescope launches. I anticipate using updated versions of the FIRE-2 simulations to simulate the kinematics of ionised gas in $z \sim 2$ star-forming galaxies, to make predictions for JWST and inform the survey strategy.

8.4 Future Work III: Simulating distant galaxies

8.4.1 Combining observations and simulations to understand discrepant galaxy size measurements.

I plan to extend my simulation work with FIRE-2 to compare the spatial extent of the intrinsic properties of simulated galaxies (gas, dust and stellar components, current and past SFR) to the sizes of their multi-wavelength emission. I will identify combinations of observations that provide most insight into key physical processes and the spatial scales on which they act.

8.4.2 Simulated and observed spatially resolved scaling relations.

The mismatch in spatial extent of various observed and inferred quantities such as dust continuum emission, molecular gas and stellar mass has important implications for galaxy scaling relations. One such relation is the Kennicutt-Schmidt (K-S) relation, which describes the long-observed relation between the global gas surface density and SFR surface density of spatially unresolved galaxies (Schmidt, 1959; Kennicutt, 1998, see Section 1.2.2). While this has been shown to hold over at least three orders of magnitude in SFR density for local, star-forming galaxies (Kennicutt, 1998), it remains unclear whether individual regions of galaxies lie on the same locus of points, particularly at high redshift. Some data suggest that this may not be the case, particularly for extremely star-forming regions (see, for example, the compact star-forming galaxy presented by Chen et al. 2017).

Using my radiative transfer models of FIRE-2 galaxies, I will correlate the emission at different wavelengths to the physical properties of the galaxy, all on pixel-scales. This will enable me to study spatially-resolved scaling relations such as $\mathrm{flux_{FIR}} - \mathrm{SFR}$, $\mathrm{flux_{FIR}} - \mathrm{M_{dust}}$ and $\mathrm{flux_{FIR}} - \mathrm{M_{gas}}$, and also to determine the minimum scales to which scaling relations derived from spatially-unresolved measurements hold.

8.5 Future Work IV: Galaxy surveys with LOFAR and WEAVE-LOFAR

As part of the LOFAR Surveys group, I have been involved in the cross-matching of radio-identified sources with optical and infrared counterparts led here in Edinburgh. I plan to exploit the deep radio data in Elais-N1, Boötes and Lockman Hole, to study the properties of radio-identified star-forming galaxies from the local Universe to high redshift. These deep fields benefit from high quality ancillary data in the near- and mid-infrared, far-infrared, radio, ultraviolet and optical. Early work will be based on photometric redshifts (Duncan et al., 2019). My later projects will make use of spectroscopic redshifts from WEAVE.

I intend to conduct a study of extragalactic star formation based on the low-frequency data. At full 10 μJy depth, we expect to detect Milky-Way luminosity galaxies back to $z \sim 1$, and more highly star-forming galaxies out to $z \sim 5$. I plan to use these deep data from LOFAR to construct radio luminosity functions of star-forming galaxies binned by redshift and then derive a dust-independent estimate of the star formation rate density as a function of redshift. Following this, I will perform a more detailed analysis of the roles of stellar mass and environment in driving star formation, as has been done at low redshift.

"The earth makes a sound as of sighs and the last drops fall from the emptied cloudless sky. A small boy, stretching out his hands and looking up at the blue sky, asked his mother how such a thing was possible. Fuck off, she said."

- Samuel Beckett, *The End*

Appendix A

Basic cosmology

In this Appendix, I present a brief overview of some of the cosmological bases for studies of galaxy evolution. More thorough derviations and explanations are provided in Longair (2008) and Hogg (1999).

A.1 The FRW metric

For a homogeneous and isotropic universe, the Friedmann-Robertson-Walker (FRW) metric is used:

$$ds^2 = c^2 d\tau^2 = c^2 dt^2 - R(t)^2 \left[\frac{d\chi^2}{1 - kr^2} + [S(\chi)]^2 (d\theta^2 + sin^2\theta d\phi^2) \right] \quad \text{(A.1)}$$

where:

- dτ is an interval of proper time
- dt is an interval of cosmic time, as measured by a comoving fundamental observer (with zero velocity with respect to the Hubble flow).
- $d\chi$ is an interval of χ, a comoving radial coordinate
- R(t) is the scale factor.
- k reflects the curvature of the Universe. For a closed universe, $k = +1$, for an open Universe, $k = -1$. For a universe with zero curvature, $k = 0$, $\chi = S(\chi) = \frac{r}{R(t)}$.

A.2 Friedmann's equations

Solving the Einstein Field Equations for an FRW metric yields Friedmann's equations. These relate the expansion rate of the universe, $\frac{\dot{R}}{R}$, to the curvature of the Universe, and the acceleration of this expansion, $\frac{\ddot{R}}{R}$, to the pressure and density:

$$\left(\frac{\dot{R}}{R}\right)^2 - \frac{8\pi G\rho}{3} - \frac{\Lambda}{3} = -\frac{kc^2}{R^2} \tag{A.2}$$

$$\frac{\ddot{R}}{R} + \frac{4\pi G\rho(1+\epsilon)}{3} - \frac{\Lambda}{3} = 0 \tag{A.3}$$

where:

- ϵ describes the ratio of pressure to density. For a mass-dominated universe, $\epsilon = 0$, and for a radiation-dominated universe, $\epsilon = 1$.
- Λ is the Cosmological constant, which corresponds to the component of curvature not caused by matter.
- ρ is the density of the Universe.

These equations are often written in terms of the Hubble parameter:

$$H = \frac{\dot{R}}{R}. \tag{A.4}$$

where 'Hubble's constant' is $H_0 = \frac{\dot{R}_0}{R_0}$, the present day value of the Hubble parameter. It is often written in terms of the dimensionless scaling h, where $H_0 = 100h\text{kms}^{-1}\text{Mpc}^{-1}$.

If we require the Universe to be flat for some critical mass-energy density, ρ_c, we obtain:

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{A.5}$$

Ω is the ratio of the actual density of the Universe to the critical density:

$$\Omega = \frac{\rho}{\rho_c}. \tag{A.6}$$

These densities and ratios evolve with epoch. Often the present-day ratios of the mass, dark energy and curvature densities to the critical density are expressed in terms of Hubble's constant:

$$\Omega_M = \frac{8\pi G\rho}{3H_0^2}; \quad \Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}; \quad \Omega_K = 1 - \Omega_M - \Omega_\Lambda = 0 \tag{A.7}$$

In a mass-dominated Einstein de Sitter model, $\Omega_M = 1$, $\Omega_\Lambda = \Omega_K = 0$. Recent evidence consistently points towards best-fitting parameters of approximately $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_K = 0$, and $H_0 = 70\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. These are the values we use in this thesis. The best constraints on the cosmological parameters come from fitting the Type 1a supernovae luminosity-redshift relation (Perlmutter et al., 1999; Schmidt et al., 1998), and the power spectrum of the cosmic microwave background (CMB, e.g. Spergel et al. 2003). Lahav & Liddle (2014) compile these results, estimating that $H_0 = 67.8 \pm 0.8\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $\Omega_M = 0.308 \pm 0.010$, $\Omega_\Lambda = 0.692 \pm 0.010$. These are consistent with the most recent measurements of the temperature and polarisation anisotropies from the Planck Collaboration (2015).

A.3 Distance measures

A.3.1 Redshift

We can use the FRW metric to derive some formulae for redshift. Consider a photon travelling in the radial direction ($d\theta = d\phi = 0$) in a flat universe ($k = 0, \chi = S(\chi) = \frac{r}{R(t)}$).

$$c^2 dt^2 = R(t)^2 dr^2. \tag{A.8}$$

We integrate between the time the photon is emitted, t_e, and the time it is observed, t_o, and use its periodic nature:

$$\frac{1}{c}\int_0^{r_e} dr = \int_{t_e}^{t_o} \frac{dt}{R(t)} = \int_{t_e+\Delta t_e}^{t_o+\Delta t_o} \frac{dt}{R(t)} \tag{A.9}$$

$$\int_{t_e}^{t_e+\Delta t_e} \frac{dt}{R(t)} = \int_{t_0}^{t_o+\Delta t_o} \frac{dt}{R(t)} \tag{A.10}$$

$$\frac{\Delta t_o}{\Delta t_e} = \frac{R(t_o)}{R(t_e)}. \tag{A.11}$$

Relating the period of the photon to its frequency enables us to relate the scale factor to the redshift:

$$\frac{R_o}{R_e} = \frac{f_e}{f_o} = \frac{\lambda_o}{\lambda_e} = 1 + z. \tag{A.12}$$

A.3.2 Line of sight comoving distance

The line of sight comoving distance, D_C, remains constant when two objects moving with the Hubble flow. It is the proper distance divided by the ratio of the scale factor then to the scale factor now.

$$D_C = \frac{c}{H_0} \int_0^z \frac{dz}{\sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 \Omega_m(1+z)^3 + \Omega_\Lambda}}, \tag{A.13}$$

where we call this denominator $E(z)$.

A.3.3 Angular diameter distance

The angular diameter distance, D_A, is the ratio of an object's physical transverse size to its angular size on the sky. For a flat universe:

$$D_A = \frac{D_C}{1+z}. \tag{A.14}$$

A.3.4 Luminosity distance

The luminosity distance, D_L, relates bolometric flux, S, to bolometric luminosity, L:

$$D_L = \sqrt{\frac{L}{4\pi S}} = (1+z)^2 D_A. \tag{A.15}$$

For flux that is not bolometric but emitted within a certain waveband, ν_1:

$$D_L = \sqrt{\frac{L_{\nu 1}(1+z)}{4\pi S}}. \tag{A.16}$$

An object's absolute magnitude, M, is related to its apparent magnitude, m, by:

$$M = -m + 5\log\left(\frac{D_L}{10\,\mathrm{pc}}\right) + K, \tag{A.17}$$

where K is the 'k-correction', applied when we consider flux in a certain waveband (rather than bolometric flux), since the frequency of emitted flux, ν_1, is observed at redshifted frequency ν_0.

$$K = -2.5\log\left[(1+z)\frac{L_{\nu 1}}{L_{\nu 0}}\right] \tag{A.18}$$

A.3.5 Comoving volume

The comoving volume, V_C, is a volume in which the comoving number densities of non-evolving objects remain constant with redshift. This is critical for studies of the evolution of the luminosity function.

$$dV_C = D_H \frac{(1+z)^2 D_A^2}{E(z)} d\Omega dz \tag{A.19}$$

Bibliography

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

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

Abraham R. G., Van den Bergh S., 2001, Science, 293, 1273

Alatalo K., et al., 2016, ApJS, 224, 38

Alves J. F., Lada C. J., Lada E. A., 2001, Nature, 409, 159

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

Anglés-Alcázar D., Faucher-Giguère C. A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017, MNRAS, 472, L109

Appleton P. N., et al., 2004, ApJS, 154, 147

Arnouts S., Schiminovich D., Ilbert O., Tresse L., Milliard B., Treyer M., Bardelli S., Budavari T., 2005, ApJ, 619, 43

Arnouts S., et al., 2013, A&A, 558, A67

Artale M. C., et al., 2017, MNRAS, 470, 1771

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

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARAA, 47, 481

Azeez J. H., Hwang C. Y., Abidin Z. Z., Ibrahim Z. A., 2016, Scientific Reports, 6, 1

Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, ApJS, 196

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

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

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

Bardeen J., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barger A. J., Cowie L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nature, 394, 248

Barro G., et al., 2016, ApJL, 827, L32

Beckwith S., Stiavelli M., Koekemoer A. M., Caldwell J. A. R., Ferguson H. C., Hook R., 2006, ApJ, 132, 1729

Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Bekki K., 2014, MNRAS, 438, 444

Bell E., Kennicutt R. C., 2001, ApJ, 548, 681

Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJSS, 149, 289

Bellhouse C., et al., 2017, ApJ, 844, 49

Belli S., Newman A. B., Ellis R. S., 2016, ApJ, 834, 1

Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bicker J., Fritzev. Alvensleben U., 2005, A&A, 443, L19

Bigiel F., Leroy A., Walter F., Brinks E., De Blok W. J., Madore B., Thornley M. D., 2008, ApJ, 136, 2846

Bigiel F., et al., 2011, ApJL, 730, L13

Birnboim Y., Dekel A., 2003, MNRAS, 345, 349

Birrer S., Lilly S., Amara A., Paranjape A., Refregier A., 2014, ApJ, 793, 12

Blumenthal G. R., Faber S. M., Rees M. J., Primack J. R., 1984, Nature, 311, 517

Boselli A., Gavazzi G., 2006, PASP, 118, 517

Bothwell M. S., Maiolino R., Cicone C., Peng Y., Wagg J., 2016, A&A, p. A48

Bouche N., et al., 2007, ApJ, 671, 303

Bourne N., et al., 2017, MNRAS, 467, 1360

Bouwens R. J., Illingworth G. D., Franx M., Ford H., 2007, ApJ, 670, 928

Bouwens R. J., et al., 2010, ApJL, 709, 133

Bouwens R. J., et al., 2015, ApJ, 803, 1

Bouwens R. J., Smit R., Labbe I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 12

Bower R. G., Benson a. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645

Bower R., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, MNRAS, 465, 32

Bowler R. A. A., Bourne N., Dunlop J. S., Mclure R. J., Mcleod D. J., 2018, MNRAS

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150

Brandt W. N., Hasinger G., 2004, ARAA, 43, 97

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Brown T., Cortese L., Catinella B., Kilborn V., 2018, MNRAS, 473, 1868

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bryant J. J., et al., 2015, MNRAS, 447, 2857

Buat V., et al., 2012, A&A, 545, A141

Bullock J. S., Wechsler R. H., Somerville R. S., 2002, MNRAS, 329, 246

Bundy K., et al., 2015, ApJ, 798, 7

Burgarella D., Buat V., Iglesias-Páramo J., 2005, MNRAS, 360, 1413

Butcher H., Oemler A., 1978, ApJ, 219, 18

Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, ApJ, 840, 44

Calistro Rivera G., Lusso E., Hennawi J. F., Hogg D. W., 2016, ApJ, 833, 98

Calistro Rivera G., et al., 2017, MNRAS, 3488, 3468

Calistro Rivera G., et al., 2018, ApJ, 863, 56

Calura F., et al., 2017, MNRAS, 465, 54

Calzetti D., 2001, PASP, 113, 1449

Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, ApJ, 429, 582

Calzetti D., Armus L., Bohlin R., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Calzetti D., Liu G., Koda J., 2012, ApJ, 752, 1

Camps P., Baes M., 2015, Astronomy and Computing, 9, 20

Camps P., et al., 2018, ApJSS, 234, 20

Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, Nature, 506, 63

Caputi K. I., Dole H., Lagache G., McLure R. J., Dunlop J. S., Puget J.-L., Le Floc'h E., Pérez-González P. G., 2006, A&A, 454, 143

Caputi K. I., et al., 2007, ApJ, 660, 97

Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379

Carnall A. C., et al., 2019, eprint arXiv:1903.11082

Casey C. M., Narayanan D., Cooray A., 2014, Physics Reports, 541, 45

Chabrier G., 2003, PASP, 115, 763

Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772

Chen C.-C., et al., 2017, ApJ, 846, 108

Cirasuolo M., McLure R. J., Dunlop J. S., Almaini O., Foucaud S., Simpson C., 2010, MNRAS, 401, 1166

Civano F., et al., 2016, The Astrophysical Journal, 819, 62

Cochrane R. K., Best P. N., 2018, MNRAS, 480, 864

Cochrane R. K., Best P. N., Sobral D., Smail I., Wake D. A., Stott J. P., Geach E., 2017, MNRAS, 469, 2913

Cochrane R. K., Best P. N., Sobral D., Smail I., Geach J. E., Stott J. P., Wake D. A., 2018, MNRAS, 475, 3730

Cochrane R. K., et al., 2019, MNRAS, 488, 1779

Cohen J., 1976, ApJ, 203, 587

Coil A. L., et al., 2008, ApJ, 672, 153

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000a, MNRAS, 319, 168

Cole S., et al., 2000b, MNRAS, 326, 255

Colin P., Klypin A., Kravtsov A., Khokhlov A., 1999, ApJ, 523, 32

Condon J. J., 1992, ARA&A, 30, 575

Connolly A. J., Szalay A. S., Dickinson M., Brunner R. J., 1997, ApJ, 486, L11

Conroy C., 2013, ARA&A, 51, 393

Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201

Conroy C., White M., Gunn J. E., 2010, ApJ, 708, 58

Contreras S., Baugh C. M., Norberg P., Padilla N., 2013, MNRAS, 432, 2717

Cooke E. A., et al., 2018, ApJ, 861, 100

Cooper M., Newman J., Madgwick D., Gerke B., Yan R., Davis M., 2005, ApJ, 634, 833

Cooray A., Sheth R., 2002, Physics Reports, 372, 1

Cordier D., Pietrinferni A., Cassisi S., Salaris M., 2007, ApJ, 133, 468

Coupon J., et al., 2015, MNRAS, 449, 1352

Cowie L., Songaila A., Hu E., Cohen J., 1996, ApJ, 112, 839

Crain R. A., et al., 2015, MNRAS, 450, 1937

Da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595

Da Cunha E., Charlot S., Dunne L., Smith D., Rowlands K., 2011, Proceedings of the International Astronomical Union, 7, 292

Daddi E., et al., 2007, ApJ, 670, 156

Daddi E., et al., 2010, ApJ, 713, 686

Dale J. E., 2015, New Astronomy Reviews, 68, 1

Darvish B., Mobasher B., Sobral D., Scoville N., Aragon-Calvo M., 2015, ApJ, 805, 121

Darvish B., Mobasher B., Sobral D., Rettura A., Scoville N., Faisst A., Capak P., 2016, ApJ, 825, 113

Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98

Davidzon I., et al., 2017, A&A, 605, A70

Davis M., Geller M. J., 1976, ApJ, 208, 13

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dekel A., Birnboim Y., 2006, MNRAS, 368, 2

Dekel A., et al., 2009, Nature, 457, 451

Delhaize J., et al., 2017, A&A, 602, A4

Dickey J. M., Salpeter E. E., 1984, ApJ, 284, 461

Diederik Kruijssen J. M., Longmore S. N., 2014, MNRAS, 439, 3239

Dolley T., et al., 2014, ApJ, 797, 125

Dotter A., 2016, ApJS, 222, 8

Draine B., 2003, ARAA, 41, 241

Draine B., Li A., 2001, ApJ, 551

Draine B. T., Li A., 2007, ApJ, 657, 810

Draine B. T., Salpeter E. E., 1979, ApJ, 231, 77

Dressler A., 1980, ApJ, 236, 351

Driver S. P., et al., 2018, MNRAS, 475, 2891

Duncan K. J., et al., 2019, A&A, 622, A3

Dunlop J. S., 2013, in , The First Galaxies, Astrophysics and Space Science Library. p. 223 (arXiv:1205.1543), doi:10.1007/978-3-642-32362-1˙5

Dunlop J. S., et al., 2013, MNRAS, 432, 3520

Dunlop J. S., et al., 2017, MNRAS, 466, 861

Durkalec A., et al., 2015, A&A, 576, L7

Dwek E., 1998, ApJ, 1, 643

Eales S., et al., 2017, MNRAS

Eardley E., et al., 2015, MNRAS, 448, 3665

Elbaz D., et al., 2007, A&A, 468, 33

Elbaz D., et al., 2011, A&A, 553, 119

Ellis R. S., et al., 2013, ApJL, 763, 8

Elmegreen B. G., Elmegreen D. M., Sánchez Almeida J., Muñoz-Tuñón C., Dewberry J., Putko J., Teich Y., Popinchalk M., 2013, ApJ, 774, 86

Erb D. K., Pettini M., Adelberger K. L., 2006, ApJ, 647, 128

Fakhouri O., Ma C. P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267

Falgarone E., et al., 2017, Nature, 548, 430

Faucher-Giguère C. A., Kereš D., Ma C. P., 2011, MNRAS, 417, 2982

Faucher-Giguère C. A., Hopkins P. F., Kerěs D., Muratov A. L., Quataert E., Murray N., 2015a, MNRAS, 449, 987

Faucher-Giguère C. A., Hopkins P. F., Kerěs D., Muratov A. L., Quataert E., Murray N., 2015b, MNRAS, 449, 987

Feldmann R., Gnedin N. Y., Kravtsov A. V., 2012, ApJ, 758

Feldmann R., Hopkins P. F., Quataert E., Faucher-Giguère C. A., Kerěs D., 2016, MNRAS, 458, L14

Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguére C. A., Kereš D., 2017, MNRAS, 470, 1050

Ferland G. J., et al., 2017, Revista Mexicana de Astronomia y Astrofisica, 53, 385

Fisher D. B., et al., 2017, MNRAS, 464, 491

Foreman-Mackey D., 2016, JOSS, 24

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Forster Schreiber N. M., Genzel R., Newman S. F., Kurk J. D., Lutz D., Tacconi L. J., Wuyts S., 2014, ApJ, 787, 38

French K. D., Yang Y., Zabludoff A., Narayanan D., Shirley Y., Walter F., Smith J. D., Tremonti C. A., 2015, ApJ, 801, 1

French K. D., Zabludoff A. I., Yoon I., Shirley Y., Yang Y., Smercina A., Smith J. D., Narayanan D., 2018, ApJ, 861, 123

Freundlich J., et al., 2013, A&A, 553, A130

Freundlich J., et al., 2019, A&A, 622, A105

Furlong M., et al., 2015, MNRAS, 450, 4486

Gabor J. M., Davé R., 2015, MNRAS, 447, 374

Gaburov E., Nitadori K., 2011, MNRAS, 414, 129

Galametz M., et al., 2012, MNRAS, 425, 763

Galliano F., Dwek E., Chanial P., 2008, ApJ, 672, 214

García-Benito R., et al., 2015, A&A, 576, A135

Garn T., Best P. N., 2010, MNRAS, 409, 421

Garn T., et al., 2010, MNRAS, 402, 2017

Geach J. E., 2019, eprint arXiv:1903.04779

Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, MNRAS, 388, 1473

Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, MNRAS, 426, 679

Geach J. E., et al., 2018, ApJL, 864, L1

Genel S., et al., 2014, MNRAS, 445, 175

Genzel R., et al., 2008, ApJ, 687, 59

Genzel R., et al., 2010, MNRAS, 407, 2091

Gillman S., et al., 2019, MNRAS, 486, 175

Giovanelli R., Haynes M. P., 1985, ApJ, 292, 404

Goddard D., et al., 2017, MNRAS, 466, 4731

Gómez-Guijarro C., et al., 2018, ApJ, 856, 121

Gomez P. L., et al., 2003, ApJ, 584, 210

González Delgado R. M., et al., 2017, A&A, 607, A128

Groth E. J., Peebles P., 1977, ApJ, 217, 385

Gruppioni C., et al., 2013, MNRAS, 432, 23

Gunn J. E., Gott J., 1972, ApJ, 176, 1

Guo Y., et al., 2015, ApJ, 800, 39

Guo Y., et al., 2017, ApJ, 853, 108

Hafen Z., et al., 2017, MNRAS, 469, 2292

Harikane Y., et al., 2016, ApJ, 821, 123

Harikane Y., Ouchi M., Ono Y., Saito S., Behroozi P. S., More S., Shimasaku K., 2017, PASP

Hartley W. G., et al., 2010, MNRAS, 407, 1212

Hashimoto T., Goto T., Momose R., 2018, MNRAS, 475, 4424

Hatfield P. W., Lindsay S. N., Jarvis M. J., Hauessler B., Vaccari M., Verma A., 2016, MNRAS, 459, 2618

Hatfield P. W., Bowler R. A., Jarvis M. J., Hale C. L., 2018, MNRAS, 477, 3760

Hayward C. C., Kere D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743

Hayward C. C., Jonsson P., Kereš D., Magnelli B., Hernquist L., Cox T. J., 2012, MNRAS, 424, 951

Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, MNRAS, 428, 2529

Hayward C. C., et al., 2014, MNRAS, 445, 1598

Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nature, 428, 625

Heckman T., Best P., 2014, ARAA, 52, 589

Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, ApJS, 129, 493

Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109

Hodge J. A., et al., 2016, ApJ, 833, 1

Hogg D. W., 1999, eprint arXiv:astro-ph/9905116

Holland W. S., et al., 1999, MNRAS, 303, 659

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142

Hopkins A. M., Onnolly A. J. C., Ram L. E. C., 2001, ApJ, 122, 288

Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950

Hopkins P. F., Narayanan D., Murray N., 2013, MNRAS, 432, 2647

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C. A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hopkins P. F., et al., 2018a, MNRAS, 477, 1578

Hopkins P. F., et al., 2018b, MNRAS, 480, 800

Hurley P. D., et al., 2017, MNRAS, 464, 885

Husemann B., et al., 2013, A&A, 549, A87

Ibar E., et al., 2008, MNRAS, 386, 953

Ibar E., et al., 2013, MNRAS, 434, 3218

Ikarashi S., et al., 2015, ApJ, 810, 133

Ilbert O., et al., 2009, ApJ, 690, 1236

Ilbert O., et al., 2013, A&A, 556, A55

Illingworth G. D., et al., 2013, ApJS, 209, 6

Iono D., et al., 2016, ApJL, 829, L10

Ivison R. J., et al., 2010, A&A, 518, 31

James A., Dunne L., Eales S., Edmunds M. G., 2002, MNRAS, 335, 753

Janowiecki S., Catinella B., Cortese L., Brown T., Wang J., 2017, MNRAS, 466, 4795

Jiménez-Andrade E. F., et al., 2018, A&A, 615, A25

Karim a., et al., 2011, ApJ, 730, 61

Kashino D., et al., 2017, ApJ, 843, 138

Kassin S. A., et al., 2012, ApJ, 758, 106

Kauffmann G., 1998, MNRAS, 297, 23

Kauffmann G., Heckman T., 2003, MNRAS, 341, 33

Kauffmann G., et al., 2003, MNRAS, 346, 1055

Kennicutt R. C., 1983, ApJ, 272, 54

Kennicutt R. C., 1998, ApJ, 498, 541

Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531

Kennicutt R. C., Tamblyn P., Congdon C. W., 1994, ApJ, 435, 22

Kennicutt R. C., et al., 2003, PASP, 115, 928

Kennicutt R. C., et al., 2007, ApJ, 671, 333

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2

Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160

Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121

Khostovan A. A., et al., 2018, MNRAS, 478, 2999

Kirkpatrick A., et al., 2017, ApJ, 843, 71

Kleiner D., Pimbblet K. A., Jones D. H., Koribalski B. S., Serra P., 2017, MNRAS, 466, 4692

Koprowski M. P., Dunlop J. S., Michałowski M. J., Cirasuolo M., Bowler R. A., 2014, MNRAS, 444, 117

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Koyama Y., et al., 2013, MNRAS, 434, 423

Kravtsov A. V., Klypin A. A., 1999, ApJ, 520, 437

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottloeber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35

Krogager J. K., Zirm A. W., Toft S., Man A., Brammer G., 2014, ApJ, 797

Kroupa P., 2002, Science, 295, 82

Krumholz M. R., Gnedin N. Y., 2011, ApJ, 729, 36

Lacy J. H., Baas F., Townes C. H., Geballe T. R., 1980, ApJ, 241, 132

Lagos C. D. P., et al., 2016, MNRAS, 459, 2632

Lahav O., Liddle A. R., 2014, in , eprint arXiv:1401.1389. (arXiv:1401.1389), http://arxiv.org/abs/1401.1389

Laigle C., et al., 2016, ApJS, 224, 1

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lanz L., Hayward C. C., Zezas A., Smith H. A., Ashby M. L., Brassington N., Fazio G. G., Hernquist L., 2014, ApJ, 785

Larson R., 1980, MNRAS, 190, 321

Lawrence A., et al., 2007, MNRAS, 379, 1599

Lee S.-K., 2015, ApJ, 810, 90

Lee N., et al., 2015, ApJ, 801, 80

Leitherer C., et al., 1999, ApJS, 123, 3

Lequeux J., Peimbert M., 1979, A&A, 80, 155

Lewis I., et al., 2002, MNRAS, 334, 673

Li C., White S. D. M., 2009, MNRAS, 398, 2177

Li C., et al., 2015, ApJ, 804, 125

Liang L., Feldmann R., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Hayward C. C., Quataert E., Scoville N. Z., 2018, MNRAS, 88, 83

Liang L., et al., 2019, eprint arXiv:1902.10727

Licquia T. C., Newman J. A., 2015, ApJ, 806, 96

Lilly S. J., Tresse L., Hammer F., Crampton D., Le Fèvre O., 1995, ApJ, 455, 108

Lilly S. J., Le Fèvre O., Hammer F., Crampton D., 1996, ApJ, 460, L1

Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119

Limber D. L., 1953, ApJ, 117, 134

Lin L., et al., 2017, ApJ, 851, 18

Longair M. S., 2008, Galaxy Formation. Springer, Cambridge

Longhetti M., Saracco P., 2009, MNRAS, 394, 774

Ma C.-P., Fry J. N., 2000, ApJ, 543, 503

Ma C.-J., et al., 2015, ApJ, 806, 257

Ma X., Hopkins P. F., Faucher-Giguère C. A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140

Ma X., et al., 2018a, MNRAS, 478, 1694

Ma X., et al., 2018b, MNRAS, 478, 1694

Ma X., et al., 2019, eprint arXiv:1902.10152

Maccacaro T., Caccianiga A., Della Ceca R., Wolter A., 1998, Astron. Nachr, 319, 15

Madau P., Dickinson M., 2014, ARAA, p. 76

Madau P., Ferguson H. C., Dickinson M., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1338

Magdis G. E., et al., 2016, MNRAS, 456, 4533

Magnelli B., et al., 2013, A&A, 553, A132

Magnelli B., et al., 2015, A&A, 573, A45

Magorrian J., et al., 1998, ApJ, 115, 2285

Mainieri V., et al., 2011, A&A, 535, A80

Maiolino R., et al., 2012, MNRAS, 425, L66

Man A., Belli S., 2018, Nature Astronomy, 2, 695

Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115

Maraston C., Daddi E., Renzini A., Cimatti A., Dickinson M., Papovich C., Pasquali A., Pirzkal N., 2006, ApJ, 652, 85

Marigo P., et al., 2017, ApJ, 835, 77

Marin F., 2018, MNRAS, 479, 3142

Martin D. C., Matuszewski M., Morrissey P., Neill J. D., Moore A., Cantalupo S., Prochaska J. X., Chang D., 2015, Nature, 524, 192

Martin D. C., Matuszewski M., Morrissey P., Neill J. D., Moore A., Steidel C. C., Trainor R., 2016, ApJ, 824, L5

Masters K. L., et al., 2010, MNRAS, 405, 783

Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, MNRAS, 465, 2381

McAlpine S., et al., 2016, Astronomy and Computing, 15, 72

McAlpine S., et al., 2019, eprint arXiv:1901.05467

McCracken H. J., et al., 2010, , 708, 202

McCracken H. J., et al., 2015, MNRAS, 449, 901

McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937

McLure R. J., et al., 2018, MNRAS, 479, 25

Mehlert D., Saglia R. P., Bender R., Wegner G., 2003, A&AS, 141, 449

Meurer G., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64

Michel-Dansac L., Lambas D. G., Alonso M. S., Tissera P., 2008, MNRAS, 386, L82

Misselt K. A., Gordon K. D., Clayton G. C., Wolff M. J., 2001, ApJ, 551, 277

Miyazaki S., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, pp 327–335, doi:10.1117/12.926844, https://doi.org/10.1117/12.926844

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Molina J., Ibar E., Swinbank A. M., Sobral D., Best P. N., Smail I., Escala A., Cirasuolo M., 2017, MNRAS, 466, 892

Moster B. P., Somerville R. S., Maulbetsch C., Bosch F. C. V. D., Macci A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Moutard T., et al., 2016, A&A, 590, A103

Murray S. G., 2015, PhD thesis

Murray S., Power C., Robotham A., 2013, Astronomy and Computing, 3, 23

Mutch S. J., Croton D. J., Poole G. B., 2013, MNRAS, 435, 2445

Muzzin A., et al., 2012, ApJ, 746, 188

Muzzin A., et al., 2013, ApJ, 777, 18

Narayanan D., Davé R., Conroy C., Thompson R., Geach J., Johnson B. D., 2018, MNRAS, 474, 1718

Nelson E. J., et al., 2019, ApJ, 870, 130

Newman J. A., et al., 2012, ApJS, 208, 5

Noeske K. G., et al., 2007, ApJ, 660, L43

Norberg P., et al., 2001, MNRAS, 328, 64

Norberg P., et al., 2002, MNRAS, 332, 827

Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, 396, 19

Oemler A., 1977, Highlights of Astronomy, Vol. 4

Oesch P. A., et al., 2010, ApJL, 709, L21

Oliver S. J., et al., 2012, MNRAS, 424, 1614

Onodera S., et al., 2010, ApJL, 722, L127

Onodera M., et al., 2015, ApJ, 808, 161

Oppenheimer B. D., Davé R., 2006, MNRAS, 373, 1265

Orr M. E., et al., 2018, MNRAS, 478, 3653

Oteo I., Sobral D., Ivison R. J., Chandra X.-r., 2015, MNRAS, 452, 2018

Oteo I., et al., 2017, ApJ, 850, 170

Pannella M., et al., 2015, ApJ, 807, 141

Panter B., Heavens A. F., Jimenez R., 2003, MNRAS, 343, 1145

Papovich C., Dickinson M., Ferguson H. C., 2002, ApJ, 559, 620

Papovich C., et al., 2016, Nature Astronomy, 1, 3

Pascarelle S. M., Lanzetta K. M., Fernandez-Soto A., 1998, ApJ, 508, L1

Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

Pearson W. J., Wang L., van der Tak F. F. S., Hurley P. D., Burgarella D., Oliver S. J., 2017, A&A, 603, A102

Pearson W. J., et al., 2018, A&A, 615, A146

Peebles P., 1982, ApJ, 263, L1

Peng Y.-j., et al., 2010, ApJ, 721, 193

Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4

Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Pereira-Santaella M., et al., 2016, A&A, 587, A44

Pérez-González P., et al., 2005, ApJ, 630, 82

Pérez-González P., et al., 2008, ApJ, 675, 234

Perlmutter S., Aldering G., Goldhaber G., Knop R. A., Nugent P., Castro P. G., 1999, ApJ, 517, 565

Planck Collaboration 2016, A&A, 594, A13

Poggianti B. M., et al., 2017, ApJ, 844, 48

Puglisi A., et al., 2016, A&A, 586, A83

Qin J., Zheng X. Z., Wuyts S., Pan Z., Ren J., 2019, MNRAS, 485, 5733

Reddy N. A., et al., 2015, ApJ, 806, 259

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Renzini A., Peng Y., 2015, ApJL, 801, L29

Riechers D. A., et al., 2018, ApJ, 872, 7

Rieke G. H., et al., 2004, ApJS, 154, 25

Rodighiero G., et al., 2011, ApJL, 739, 2

Rodighiero G., et al., 2014, MNRAS, 443, 19

Rodríguez-Puebla A., Primack J. R., Behroozi P., Faber S. M., 2016, MNRAS, 455, 2592

Rowlands K., et al., 2018, MNRAS, 480, 2544

Rupke D. S. N., 2018, Galaxies, 6, 138

Safarzadeh M., Hayward C. C., Ferguson H. C., Somerville R. S., 2016, ApJ, 818, 62

Safarzadeh M., Hayward C. C., Ferguson H. C., 2017, ApJ, 840, 15

Saintonge A., et al., 2013, ApJ, 778, 2

Salpeter E. E., 1955, ApJ, 121, 161

Sánchez S. F., et al., 2012, A&A, 538, A8

Santini P., Rosario D. J., Shao L., Lutz D., Maiolino R., Alexander D. M., Altieri B., Andreani P., 2012, A&A, 540, 109

Sargent M. T., et al., 2010, ApJS, 186, 341

Schaye J., Aguirre A., Kim T.-S., Theuns T., Rauch M., Sargent W. L. W., 2003, ApJ, 596, 768

Schaye J., et al., 2015, MNRAS, 446, 521

Schechter P., 1976, ApJ, 203, 297

Schinnerer E., et al., 2007, ApJS, 172, 46

Schinnerer E., et al., 2010, ApJS, 188, 384

Schinnerer E., et al., 2016, ApJ, 833, 112

Schmidt M., 1959, ApJ, 129, 243

Schmidt M., Suntzeff N. B., Phillips M. M., Schommer R. A., Chlocchitti A., 1998, ApJ, 507, 46

Scoville N., et al., 2007, ApJS, 172, 1

Scoville N., et al., 2013, ApJS, 206, 3

Scoville N., et al., 2014, ApJ, 783, 84

Scoville N., et al., 2017, ApJ, 837, 150

Scudder J. M., Oliver S., Hurley P. D., Griffin M., Sargent M. T., Scott D., Wang L., Wardlow J. L., 2016, MNRAS, 460, 1119

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Simon P., 2007, A&A, 473, 711

Simon P., Hetterscheidt M., Wolf C., Meisenheimer K., Hildebrandt H., Schneider P., Schirmer M., Erben T., 2009, MNRAS, 398, 807

Simons R. C., et al., 2017, ApJ, 843, 46

Simpson J. M., et al., 2015, ApJ, 799, 81

Simpson J. M., et al., 2017a, ApJ, 839, 58

Simpson J. M., et al., 2017b, ApJL, 844, 6

Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5

Smail I., et al., 2014, ApJ, 782, 19

Smith D. J., Hayward C. C., Jarvis M. J., Simpson C., 2017, MNRAS, 471, 2453

Smolčić V., et al., 2017, A&A, 602, A2

Snaith O. N., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2014, ApJL, 781, L31

Sobral D., et al., 2009, MNRAS, 398, 75

Sobral D., Best P. N., Geach J. E., Smail I., Cirasuolo M., Garn T., Dalton G. B., Kurk J., 2010, MNRAS, 1563, 1551

Sobral D., Best P. N., Smail I., Geach J. E., Cirasuolo M., Garn T., Dalton G. B., 2011, MNRAS, 411, 675

Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926

Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013a, MNRAS, 428, 1128

Sobral D., et al., 2013b, ApJ, 779, 139

Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 427, 3516

Sobral D., et al., 2015, MNRAS, 451, 2303

Sobral D., Kohn S. A., Best P. N., Smail I., Harrison C. M., Stott J., Calhau J., Matthee J., 2016a, MNRAS, 457, 1739

Sobral D., Stroe A., Koyama Y., Darvish B., Calhau J., Afonso A., Kodama T., Nakata F., 2016b, MNRAS, 458, 3443

Somerville R. S., Davé R., 2015, ARAA, 53, 51

Soto E., et al., 2017, ApJ, 837, 6

Sparre M., Hayward C. C., Feldmann R., Faucher-Giguère C. A., Muratov A. L., Kereš D., Hopkins P. F., 2017, MNRAS, 466, 88

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJSS, 214, 15

Spergel D., et al., 2003, ApJS, 148, 175

Springel V., et al., 2005, Nature, 435, 629

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Steinacker J., Baes M., Gordon K., 2013, ARAA, pp 63–105

Stern D., et al., 2005, ApJ, 631, 163

Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, ApJ, 683, 597

Stott J. P., et al., 2016, MNRAS, 457, 1888

Strandet M., 2017, ApJL, 842, L15

Sullivan M. A. R. K. S., Obasher B. A. M., Han B. E. N. C., Ram L. A. C., Llis R. I. E., Reyer M. A. T., Opkins A. N. H., 2001, ApJ, 558, 72

Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., Mccarthy I. G., Crain R. A., Theuns T., 2012a, MNRAS, 426, 935

Swinbank A. M., Smail I., Sobral D., Theuns T., Best P. N., Geach J. E., 2012b, ApJ, 760, 130

Swinbank A. M., et al., 2014, MNRAS, 438, 1267

Symeonidis M., Giblin B. M., Page M. J., Pearson C., Bendo G., Seymour N., Oliver S. J., 2016, MNRAS, 459, 257

Tabatabaei F. S., Berkhuijsen E. M., Frick P., Beck R., Schinnerer E., 2013, A&A, 557, 129

Tacchella S., Carollo C. M., Renzini A., Schreiber N. M. F., Lang P., 2015, Science, 348, 314

Tacconi L. J., Genzel R., Neri R., Cox P., Cooper M. C., Shapiro K., Bolatto A., 2010, Nature, 463, 781

Tacconi L. J., et al., 2013, ApJ, 768, 74

Tacconi L. J., et al., 2017, ApJ, 853, 179

Tadaki K.-i., et al., 2016, ApJ, 834, 1

Tadaki K.-i., et al., 2017, ApJL, 841, L25

Tadaki K., et al., 2018, Nature, 560, 613

Tal T., et al., 2014, ApJ, 789, 164

Taniguchi Y., et al., 2007, ApJS, 172, 9

Taniguchi Y., et al., 2015, PASJ, 67, 1

Tielens A. G. G. M., McKee C. F., Seab C. G., Hollenbach D. J., 1994, ApJ, 431, 321

Tinker J. L., Wechsler R. H., Zheng Z., 2010a, ApJ, 709, 67

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010b, ApJ, 724, 878

Toft S., et al., 2014, ApJ, 782

Trayford J. W., et al., 2017, MNRAS, 470, 771

Tremonti C. A., Heckman T. M., Kauffmann G., Brinchmann J., White S. D. M., Seibert M., Peng E. W., Schlegel D. J., 2004, ApJ, 613, 898

Trussler J., Maiolino R., Maraston C., Peng Y., Thomas D., Goddard D., Lian J., 2018, eprint arXiv:1811.09283

Vito F., et al., 2014, MNRAS, 441, 1059

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Wake D. A., et al., 2011, ApJ, 728, 46

Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435

Wechsler R. H., Grossi M. A. K., Primack J. R., Blumenthal G. R., Dekel A., 1998, ApJ, 506, 19

Weinberg D. H., Dave R., Katz N., Hernquist L., 2004, ApJ, 601, 1

Weiner B. J., et al., 2009, ApJ, 692, 187

Weingartner J. C., Draine B. T., 2001, ApJ, 548, 296

Wetzel A. R., Tinker J. L., Conroy C., Bosch F. C. V. D., 2013, MNRAS, 432, 336

Whitaker K. E., Pope A., Cybulski R., Casey C. M., Popping G., Yun M. S., 2017, ApJ, 850, 208

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wild V., Walcher C. J., Johansson P. H., Tresse L., Charlot S., Pollo A., Le Fèvre O., De Ravel L., 2009, MNRAS, 395, 144

Wilkinson A., et al., 2017, MNRAS, 464, 1380

Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879

Woo J., et al., 2013, MNRAS, 428, 3306

Wuyts S., et al., 2012, ApJ, 753

Wyder T. K., et al., 2009, ApJ, 696, 1834

Yang X., Mo H. J., Jing Y., Van Den Bosch F. C., 2005, MNRAS, 358, 217

Yang X., Mo H. J., Bosch F. C. V. D., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153

York D., Adelman J., 2000, ApJ, 120, 1579

Zehavi I., et al., 2005, ApJ, 630, 1

Zehavi I., et al., 2011, ApJ, 736, 59

Zheng Z., et al., 2005, ApJ, 633, 791

Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

Zu Y., Mandelbaum R., 2016, MNRAS, 457, 4360

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

van de Voort F., Quataert E., Hopkins P. F., Kereš D., Faucher-Giguére C. A., 2015, MNRAS, 447, 140

van der Kruit P. C., 1971, A&A, 15, 110

van der Kruit P. C., 1973, A&A, 29, 263