What drives black hole and galaxy growth in the EAGLE simulation?

Stuart McAlpine

A Thesis presented for the degree of
Doctor of Philosophy

Department of Physics
Durham University
United Kingdom

November 2018

What drives black hole and galaxy growth in the EAGLE simulation?

Stuart McAlpine

Submitted for the degree of Doctor of Philosophy

November 2018

Abstract:

In this thesis we investigate what drives the evolution of actively accreting central supermassive black holes and unusually active strongly star-forming galaxies using the *Evolution and Assembly of GaLaxies and their Environments* (EAGLE) suite of cosmological hydrodynamical simulations. We find that many of our results are intimately tied to the complex evolutionary pathway taken by the central black holes within the simulation. This evolution can be separated into three distinct phases, each related to the mass of the host dark matter halo. In low mass haloes, stellar feedback dominates by driving an effective outflow and substantially hinders the growth of the central black hole. As haloes become more massive, the stellar feedback loses its efficiency, and the outflow stalls. This gives the first opportunity for the central black hole to grow, which is does so initially at a rapid rate. After this phase of rapid growth, the central black hole then becomes sufficiently massive to regulate the gas inflow onto the halo, resulting in both the star formation of the galaxy and any continued rapid growth of the central black hole to be substantially restricted via the outputted energy of an actively accreting supermassive black hole (referred to as an active galactic nuclei, or AGN). In Chapter 4 we discover that

this complex evolutionary behaviour is integral to understanding how the growth rates of galaxies and their black holes are related to each other throughout cosmic time. We use this behaviour to explain why the current observational studies report different relationships between galaxy and black hole growth rates depending on the initial selection method used. Finally, in Chapter 5 we find that the evolutionary state of the black hole is also closely connected with high star formation rates in lower mass galaxies ($M_* \lesssim 10^{11}$ $\mathrm{M}_\odot$, where M_* is the stellar mass of the galaxy). Such 'starbursting' galaxies are rare, and we argue that they are produced through a culmination of two coinciding events; (1) the galaxy must host an underdeveloped black hole (one that has not yet entered its rapid growth phase), thus ensuring that the galaxy has maintained a gas rich reservoir and contains a low mass black hole; and (2) the galaxy must undergo an interaction to kick-start the starburst process. This tells us that strongly star-forming galaxies are a predominately merger driven population that host undermassive black holes, making them fundamentally distinct from the 'typical' star-forming population.

Declaration

The work described in this thesis was undertaken between September 2014 and July 2018 while the author was a research student under the supervision of Prof. Richard G. Bower in the Department of Physics at Durham University. This work has not been submitted for any other degree at Durham University or any other University.

Some aspects of Chapter 2 are lifted directly from the EAGLE database release paper published in Astronomy & Computing and the EAGLE particle release paper of which I was a leading author:

- *The* EAGLE *simulations of galaxy formation: public release of halo and galaxy catalogues.*

  Stuart McAlpine, John C. Helly, Matthieu Schaller, James W. Trayford, Yan Qu, Michelle Furlong, Richard G. Bower, Robert A. Crain, Joop Schaye, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, Ian G. McCarthy, Adrian Jenkins, Yetli Rosas-Guevara, Simon D. M. White, Maarten Baes, Peter Camps, Gerard Lemson, *Astronomy and Computing* 2016, Volume 15, pp. 72-89.
  arXiv:1510.01320

- *The EAGLE simulations of galaxy formation: Public release of particle data*

  The EAGLE **team**

  arXiv:1706.09899

Chapters 3 to 5 have all been submitted for publication in the form of papers to Monthly Notices of the Royal Astronomical Society (MNRAS):

- *The rapid growth phase of supermassive black holes.*

  Stuart McAlpine, Richard G. Bower, David J. Rosario, Robert A. Crain, Joop Schaye, Tom Theuns.

  arXiv:1805.08293.

- *The link between galaxy and black hole growth in the* EAGLE *simulation.*

  Stuart McAlpine, Richard G. Bower, Chris M. Harrison, Robert A. Crain, Matthieu Schaller, Joop Schaye, Tom Theuns, *MNRAS* 2016, Volume 468, pp. 3395-3407.

  arXiv:1701.01122

- *Strongly star-forming galaxies in the* EAGLE *simulation: triggering, descendants and submillimetre properties.*

  Stuart McAlpine, Ian Smail, Mark A. Swinbank, James W. Trayford, Richard G. Bower, Robert A. Crain, Joop Schaye, Tom Theuns.

Contents

Abstract **ii**

1 Introduction **2**

1.1 Cosmology and galaxy evolution 2

1.1.1 The ΛCDM paradigm 3

1.1.2 The Friedmann equations 8

1.1.3 Galaxy formation and evolution 11

1.2 Simulating the Universe 13

1.2.1 Dark matter only simulations 14

1.2.2 Semi-Analytic modelling 14

1.2.3 Hydrodynamical simulations 15

1.2.4 Idealised and zoom simulations 18

1.2.5 Future work and upcoming simulations 19

1.3 Thesis outline 20

2 The EAGLE simulation suite **22**

2.1 Overview 22

2.1.1 The simulation runs 24

2.2 Gravity 25

2.3 Hydrodynamics 27

2.4 Subgrid models 29

2.4.1 Radiative processes and reionization 29

2.4.2 The interstellar medium and star formation 30

2.4.3 Stellar mass loss and evolution 31

2.4.4 Stellar feedback 32

2.4.5 Supermassive black holes 33

2.5 The simulation output 35

2.5.1 Halo, subhalo and galaxy identification 37

2.5.2 Merger trees 37

3 The rapid growth phase of supermassive black holes 39

3.1 Introduction 39

3.2 The EAGLE simulation 41

3.2.1 The phases of black hole growth 41

3.2.2 Black hole sample selection 42

3.2.3 Defining $n_{\rm dyn}$: the most proximate merger 45

3.3 Results 48

3.3.1 Properties of the black holes 48

3.3.2 Properties of the hosts at the start of the rapid growth phase 52

3.3.3 The proximity of mergers to the rapid growth phase . . . 54

3.4 Discussion 58

3.4.1 Stalling stellar feedback and the transition to the rapid growth phase of black holes 58

3.4.2 The role of galaxy mergers in triggering the rapid growth phase of black holes . . . 61
3.4.3 Observing the rapid growth phase of black holes . . . 62
3.4.4 The dependence on the model . . . 65
3.5 Conclusions . . . 66

4 The link between galaxy and black hole growth in the EAGLE simulation 68

4.1 Introduction . . . 68
4.2 The EAGLE simulation suite . . . 72
4.2.1 Constructing histories of individual galaxies . . . 72
4.2.2 The $M_{\rm BH}$–M_{200} relation . . . 74
4.2.3 Absolute calibration of SFRs . . . 76
4.3 Predictions from the integrated quantities . . . 77
4.4 Results . . . 78
4.4.1 Comparison to observations . . . 78
4.4.2 Understanding the BHAR-SFR relationship . . . 83
4.4.3 The connection to the host dark matter halo . . . 89
4.5 Discussion . . . 94
4.6 Conclusions . . . 99

5 Strongly star-forming galaxies in the EAGLE simulation: triggering, descendants and submillimetre properties 101

5.1 Introduction . . . 101
5.2 The EAGLE simulation . . . 104
5.2.1 Selecting strongly star-forming model galaxies . . . 105

5.3 Results 110

5.3.1 The nature of strongly star-forming galaxies 110

5.3.2 The submillimeter properties of strongly star-forming galaxies 121

5.4 Discussion 128

5.4.1 What triggers strongly star-forming galaxies? 128

5.5 Conclusions 133

6 Conclusions **136**

6.1 Thesis results & Future work 139

6.1.1 Future work 140

squa

Chapter 1

Introduction

1.1 Cosmology and galaxy evolution

The *Lambda cold dark matter* (ΛCDM) cosmogony is the currently accepted concordance, or 'standard', model of cosmology. Its framework consists of a Universe that has evolved from a 'Hot Big Bang' model, with a matter component dominated by a non-relativistic weakly interacting *dark matter* and contains a cosmological constant, Λ, associated with *dark energy*. As with all modern cosmological models, it is founded on the *cosmological principle*; which states that the distribution of material within the Universe is uniform (homogeneous) and that there are no preferred locations within the Universe (isotropic). It also assumes that Einstein's theory of general relativity the correct descriptor of gravity on large scales. The ΛCDM paradigm is the simplest model for which many aspects of our cosmos can be explained; such as the origin of the cosmic microwave background, the distribution of galaxies within the large scale structure, the initial abundances of the elements, and the accelerated expansion of the Universe.

1.1.1 The ΛCDM paradigm

The Hot Big Bang model

Much of the unveiling as to the true nature of our Universe came from the pioneering work of Edwin Hubble in the 1930's. He found that all distant galaxies are moving away from us, and that the velocity at which they are receding is approximately proportional to their distance. This result is encapsulated in what is known as *Hubble's law*,

$$\mathbf{v} = H_0\mathbf{R}, \tag{1.1.1}$$

where $\mathbf{v}$ is the recessional velocity, $\mathbf{R}$ is the proper (i.e., not comoving) distance from the galaxy to the observer and H_0 is the Hubble constant[1], which is observed to be ≈ 70 km s^{-1} Mpc^{-1} (where Mpc stands for megaparsecs). This provided the first observational based evidence that our Universe is *expanding*, seemingly with our galaxy at the epicentre of that expansion. However, it can be shown that all observers would derive an identical law regardless of their location within the Universe (due to its linear nature), thus conforming to the isotropy of the cosmological principle. The consequence of this expansion, is that the Universe must have been hotter and more dense in the distant past, eventually collapsing to a singularity known as the *Big Bang*. This is the first piece of evidence that strongly supports a 'Hot Big Bang' scenario. Two further fundamental aspects of our cosmos can be neatly described within the context of this model; the initial abundances of the elements, and the origin of the cosmic microwave background.

Primordial nucleosynthesis, or Big Bang nucleosynthesis, is responsible for producing the so-called 'light elements', ^{4}He, D, ^{3}He and ^{7}Li, in the first ≈ 20 minutes of the Universe. After this time, the Universe had expanded and cooled to the point where it was no longer hot enough for any significant fusion to occur, and the initial

[1]Or more strictly, it is the value of the time-varying Hubble parameter, $H(t)$, at the present day.

abundances of the elements was set. Alpher et al. (1948) used this theory to predict what the abundances of the elements in our Universe should be if such a process had taken place. To measure this, however, we are required to look in the Universe's most pristine environments, those that have been the least polluted by the *stellar* nucleosynthesis process, and remain as close as possible to their primordial state[2]. Observations of such 'metal-poor' systems have yielded a good agreement with the predicted abundances, however, there are potential tensions between the predicted abundances of ^{7}Li to those that are observed (see Steigman, 2007 for a review).

Approximately 380,000 years later, the Universe cooled to a point where is became energetically favourable for electrons and protons to combine ($\approx$ 3000 K or an average photon energy of $\approx$ 0.7 eV[3]), allowing the first neutral atoms to stably form. In the absence of the free electrons, the photons could suddenly travel almost completely unhindered as the Universe turns 'transparent', known to cosmologists as the 'period of recombination', or the 'surface of last scattering'. This 'relic radiation' was predicted to stretch and cool as the Universe expands, eventually shifting into the microwave region of the electromagnetic spectrum (Gamow, 1948; Alpher & Herman, 1948a,b). Its discovery did not come until almost 20 years later, when in 1964 Arno Penzias and Robert Wilson stumbled upon a background microwave signal that they could not explain, seemingly coming from every direction on the sky. It turned out to be the predicted radiation from recombination, and was labelled the *cosmic microwave background* (CMB).

Many precision satellites have since analysed the CMB in great detail; such as the COBE (Smoot et al., 1992), WMAP (Bennett et al., 2013) and Planck (Planck Collaboration et al., 2014) missions, finding it to be almost completely uniform, with an approximately constant temperature of 2.7 K in every direction of the sky. However it is not completely uniform, and yields fluctuations one part in 100,000,

[2]Note that deuterium is also 'destroyed' in stars and must be accounted for.

[3]Note that this is well below the ionization energy of hydrogen (13.6 eV), however at temperatures greater than $\approx$ 3000 K, for a primordial plasma in thermodynamic equilibrium with a 'photon-to-baryon' ratio of $\sim 10^{9}$, there remain a sufficient number of high energy photons in the blackbody tail to keep the hydrogen atoms ionized.

created from the tiny variations in the density of the Universe immediately after the Big Bang. These anisotropies reveal specific facts about the nature of our Universe; such as information about its curvature and matter content. Due to the remarkable agreement between the predicted origin of the CMB, and the wealth of precise observations, it remains the most compelling piece of evidence supporting the Hot Big Bang paradigm today.

Dark matter and dark energy

The two fundamental components of the ΛCDM paradigm are dark matter and dark energy. These two components alone dominate the energy budget of the Universe, yet, astonishingly, neither one has been directly detected. Instead, the evidence for their existence comes purely from theory and indirect inferences.

There have been various proposals suggesting the existence of an invisible matter since the beginning of the 20$^{\text{th}}$ century. In 1933 Fritz Zwicky studied the motions of galaxies around the nearby Coma Cluster, deducing that there was simply not enough mass within the cluster (estimated from the visible light) to account for the high orbital velocities of the galaxies, which, if correct, would have torn the cluster apart, inferring the need for some form of additional mass/gravity within the system to keep it together (Zwicky, 1933, 1937). In an analogous fashion, during the 1960s and 70s, Vera Rubin provided further strong evidence for a missing matter component by analysing the rotation curves of spiral galaxies, inferring that they could not produce their high rotational velocities given the amount of visible matter within the system (Rubin et al., 1980). Perhaps the strongest observational evidence for dark matter comes from gravitational lensing. Using Einstein's general theory of relativity we know that matter curves space, and that even light will trace that curvature. Therefore massive compact systems, such as galaxy clusters, can substantially deflect the path of background light in the same manor as magnification by a lens. By measuring this effect we can infer a *total* mass for the cluster, which is always found to be considerably more than is measured from the visible light

alone (e.g., Massey et al., 2010). Additionally, from a theoretical viewpoint, when attempting to reproduce the large scale structure of the Universe numerically via simulations, it is only possible by assuming that the matter content of the Universe is dominated by a collisionless dark matter (e.g., Springel et al., 2006).

The leading theory as to the nature of dark matter is that it is a massive[4] fundamental particle that only interacts through the weak force and by gravity. As it does not interact electromagnetically (i.e., it does not absorb or emit light), detecting these 'weakly interacting massive particles' (or WIMPs) is an extremely challenging task. One option is to attempt to produce the particle artificially within a particle accelerator, such as the Large Hadron Collider. These experiments energetically collide two protons to produce heavier particles, which, on rare occasions, are theorised to produce dark matter candidates ('detected' via the absence of energy within the collision, as the particle itself would not interact directly with the detector). To date no such particle has been created; however, these experiments can place strong constraints on the allowed properties of the dark matter candidates (e.g., Khachatryan et al., 2015). Direct detection is another method, whereby sensitive detectors are placed in isolation deep underground[5] in the hope that a collision between an incoming dark matter particle and the baryonic particles comprising the detector will be measured. Such interactions are predicted to be extremely rare, with an expected signal of $\approx 10^{-5}$–10^{1} events per kilogram of detector per day. The current generation of detectors have detection thresholds of around 1 event per kilogram of detector per day; however, even with the numerous examples of these detectors around the world, to date no signal has been discovered. And finally, as dark matter can self-annihilate, large concentrations of dark matter, such as at the centre of our own galaxy, are predicted to produce an excess signal at this annihilation energy. However, again, there has been no (confirmed) detection of this signal. The direct detection of dark

[4]These massive dark matter particles are thought to be produced via the thermal creation process in the early Universe. They are predicted to lie in the mass range $\approx$ 10-1000 GeV, estimated from the measured abundance of dark matter today, which constrains the self-annihilation cross section.

[5]In an attempt to remove as much 'noise' as possible from non-dark matter sources, such as cosmic rays.

matter would certainly be one of the crowning achievements in all of physics.

Even more mysterious than dark matter, is dark energy, the existence of which is inferred from the observed accelerated expansion of our Universe. By invoking general relativity, the expansion history of the Universe can be derived (see next section). This derivation classically produced three broad scenarios as to the fate of our Universe; (1) the energy that drives the outward expansion of the Universe is large enough to overcome the attraction of gravity, and it expands indefinitely (an 'open' Universe); or vice-versa, (2) where gravity wins and the Universe eventually collapses under its own mass beyond a critical time (a 'closed' Universe); or (3) there is a perfect amount of matter in the universe to balance the expansion and it asymptotes towards a static Universe (a 'flat' Universe). However, evidence emerged that our Universe is actually *accelerating* in its expansion; for example, from the inability of a 'simple' CDM theory to explain the abundance of structure on large scales found by the APM survey (Efstathiou et al., 1990), and from the observed apparent brightness of type-Ia supernovae[6] in distant galaxies (discovered independently by the Supernova Cosmology Project (Perlmutter et al., 1999) and the High-Z Supernova Search Team (Riess et al., 1998)). Thus the theories had to be redefined to incorporate an additional form of unknown energy which drives the accelerated expansion, which we coin dark energy, however its origin and nature are barely understood. A leading theory interprets dark energy as a *vacuum energy*, whereby virtual particle-antiparticle pairs come into existence and then self-annihilate shortly thereafter. However, the observed density of the cosmological vacuum energy density and the value predicted from quantum field theory are *significantly* discrepant, which is referred to as the 'vacuum catastrophe' (e.g., Adler et al., 1995; Rugh & Zinkernagel, 2000), and is one of the great open questions in modern physics today. Dark energy appears in the ΛCDM paradigm as a cosmological constant, Λ (see also the next section).

[6]If a white dwarf star is able to accrete mass from a binary partner it can exceed the Chandrasekhar mass and ignite as a type-Ia supernova. These make for ideal 'standard candle' distance indicators, as the supernovae explosion has a well defined intrinsic brightness.

1.1.2 The Friedmann equations

The behaviour of a homogeneous and isotropic Universe can be described by perhaps the most important equation in cosmology, the *Friedmann equation*. It is obtained by inserting the Friedmann-Robertson-Walker metric into the Einstein Field Equation, yielding

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho(t) + \frac{kc^2}{R_0 a^2} + \frac{\Lambda}{3}, \tag{1.1.2}$$

where R_0 is the radius of curvature of the Universe as the present day, $k = +1, -1$ or 0 is a curvature constant that describes the geometry of the Universe, Λ is the cosmological constant, $\rho(t)$ is the *total* density (from all sources of energy, not just matter) and a is the scale factor; a dimensionless factor between 0 and 1 that parametrizes the relative expansion of the Universe. From a Newtonian viewpoint, where our Universe is described simply by an expanding sphere, the scale factor is the ratio of the radius of the Universe at a given time to the radius of the Universe today, $a = R(t)/R_0$. Note that the scale factor is a time dependent quantity, $a \equiv a(t)$, however the (t) is traditionally dropped for clarity. The Friedmann equation relates the rate of growth of the Universe to the matter, energy and curvature contained within it.

In order to solve the Friedmann equation, we require an additional relation between the two unknowns; the scale factor and the density. By invoking the first law of thermodynamics, $dE = -PdV + dQ$ (where E is the internal energy, P is the pressure, V is the volume and Q is the total heat) and by understanding that the expansion of a homogeneous Universe is an *adiabatic* process ($dQ = 0$), we arrive at the *fluid equation*,

$$\dot{\rho}(t) + 3\frac{\dot{a}}{a}\left(\rho(t) + \frac{P}{c^2}\right) = 0, \tag{1.1.3}$$

which has no dependence on the cosmological constant, Λ, however it has introduced

an additional unknown, the pressure, P. Note that P is also a time dependent quantity, $P \equiv P(t)$, however the (t) is again traditionally dropped for clarity. Finally, by combining eqs. (1.1.2) and (1.1.3) we obtain the *acceleration equation*,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3P}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{1.1.4}$$

that describes the rate of change of the expansion of the Universe.

Solving the Friedmann equations

In order to solve the Friedmann equations directly, we require one final relation, between the pressure P and the density $\rho(t)$, invoked via the equation of state of a perfect fluid,

$$P = \omega\rho(t)c^2, \tag{1.1.5}$$

where ω is a dimensionless constant. The combination of these four equations allows us to understand how a particular universe will evolve with time.

As an example, if we substitute eq. (1.1.5) into eq. (1.1.3) we find that in a universe dominated by matter ($\omega = 0$) the density falls as $\rho(t) \propto a^{-1/3}$ (as we would obtain from the classical Newtonian viewpoint), in a radiation dominated universe ($\omega = 1/3$) this increases to $\rho(t) \propto a^{-1/4}$ (the extra dimensionality is from the redshifting of the radiation) and in a universe dominated by a cosmological constant ($\omega = -1$) the density remains constant ($\rho(t) = \text{const}$). In reality of course, there is a mixing between the components, with each component dominating at a different time.

The Friedmann equations can also be solved to see how the scale factor evolves as a function of time, whether or not the expansion (or contraction) of the Universe is speeding up or slowing down and what role the geometry plays in the eventual fate of the Universe.

The critical density and the density parameter

By realising that $H(t) = \dot{R}/R = \dot{a}/a$ (from eq. (1.1.1)) we can rewrite the Friedmann equation in terms of the Hubble parameter, from which the *critical density* is defined,

$$\rho_{\rm c}(t) = \frac{3H(t)^2}{8\pi G}, \tag{1.1.6}$$

which is the mean density required to create a flat ($k = 0$) Universe in the absence of a cosmological constant ($\Lambda = 0$). The critical density of the Universe today is observed to be $\rho_{\rm c,0}(t) \approx 1.4 \times 10^{11}$ M$_\odot$ Mpc^{-3} or ≈ 1 hydrogen atom per cubic meter. In models without a cosmological constant, the critical density represents the tipping point between a Universe that will eventually collapse under its own gravity ($\rho(t) > \rho_{\rm c}(t)$) and those that will expand indefinitely ($\rho(t) < \rho_{\rm c}(t)$).

Cosmologists generally prefer to describe the density of the Universe in terms of the *density parameter*; the ratio of the total density to the critical density, i.e.,

$$\Omega(t) = \frac{\rho(t)}{\rho_{\rm c}(t)} = \rho(t) \times \frac{8\pi G}{3H(t)^2}. \tag{1.1.7}$$

Once substituted, this yields an alternate form of the Friedmann equation,

$$1 - \Omega(t) = -\frac{kc^2}{R_0^2 a(t)^2 H(t)^2}, \tag{1.1.8}$$

thus showing that if $\Omega(t) > 1, \Omega(t) = 1$ or $\Omega(t) < 1$ at any time (i.e., if the Universe is more, equal or less dense than the critical density), it remains so at all times, as the sign of the right hand side of the equation cannot change during expansion. By solving this equation for the present day, i.e.,

$$R_0 = \frac{c}{H_0}|1 - \Omega_0|^{-1/2}, \tag{1.1.9}$$

we see that if Ω_0 is very close to 1 (which we believe it to be), then the curvature radius of the Universe is very large, and therefore would be immeasurable on small

scales.

The cosmological constant

At the time, Einstein was convinced that the Universe was static. Yet frustratingly, no solution to the Friedmann equations could allow for such a Universe to exist; eventually it must either expand, or contract. He therefore added a term to his field equation called the cosmological constant, Λ, the purpose of which was to *counteract* the inevitable expansion or contraction of the Universe. He later realised in the wake of Hubble's discovery of an expanding Universe that the term was no longer required, and subsequently dropped it, claiming it to be his "greatest blunder".

However, when the expansion of the Universe was discovered to be *accelerating*, the Friedmann equations could once again not yield a solution, and the cosmological constant was reintroduced. By looking at the Friedmann equation (eq. (1.1.2)), the addition of the cosmological constant is equivalent to adding a new component of the total density,

$$\rho_\Lambda = \frac{\Lambda}{8\pi G}, \tag{1.1.10}$$

which, if constant, implies from the fluid equation that

$$P_\Lambda = -\frac{\Lambda}{8\pi G}. \tag{1.1.11}$$

The cosmological constant therefore provides a constant energy density resulting in a *negative* pressure, driving the accelerated expansion.

1.1.3 Galaxy formation and evolution

It is now widely accepted that galaxies form within the centres of gravitationally bound dark matter structures, known as 'haloes' (see Benson, 2010 for a review of

galaxy formation theory). Dark matter is initially distributed near uniformly in the primordial Universe (as seen from the near uniformity of the CMB), however, the regions that are slightly overdense will concentrate together via gravity, eventually becoming non-linear and undergoing gravitational collapse, thus decoupling them from the Hubble flow. The smaller scale overdensities collapse first[7], forming the primordial haloes, which continue to grow either through the smooth accretion of material or by mergers with neighbouring haloes (White & Rees, 1978). This process results in the hierarchical build-up of structure within the Universe.

Pristine baryonic material will then be gravitationally attracted towards the deep potential wells of these dark matter haloes. If the virial temperature of the halo exceeds the temperature of the incoming gas it will shock, and the kinetic energy of the infalling gas will be thermalized. If the cooling times of the halo are long compared to the dynamical times, this shock generally occurs at the virial radius, however, in the other limit of short cooling times, the shock must instead form at much smaller radii (e.g., White & Frenk, 1991). This prevents the gas clouds from immediately fragmenting into stars (e.g., Rees & Ostriker, 1977). How efficiently this gas then cools is predicted to depend on the mass of the halo; the gas in haloes of mass 10^{10}–10^{12} $M_{\odot}$ cools so efficiently it collapses at the free-fall rate and quickly fragments into stars, with the gas in larger haloes taking much longer to cool. White & Frenk (1991) introduced the transition mass between the between rapid and slow cooling regimes at the point where the cooling and dynamical times at the halo virial radius become equal. This result has been confirmed by a number of hydrodynamical simulations (e.g., Benson et al., 2001; Helly et al., 2003). As angular momentum is conserved, the cooled gas then goes on to form a galactic disk where the majority of star formation occurs (e.g., Fall & Efstathiou, 1980; Mo et al., 1998).

Many complex processes over a multitude of scales then occur. Individual stars

[7]This is assuming that the dark matter is 'cold' (low streaming velocities); if the dark matter was 'warm' or 'hot', haloes would initially form more massive and fragment into smaller structures forming a top-down scenario, as the larger streaming velocities would wipe out fluctuations on smaller scales.

and star clusters are observed to form from cool molecular gas at a rate that is proportional to the gas surface density (Kennicutt, 1998). Elements heavier than helium (referred to as 'metals') are synthesised within these stars via nuclear fusion and go on to enrich the surrounding interstellar medium (ISM) from stellar winds. In addition to this, the stars that explode as type II and type Ia supernova enrich the ISM further and also deposit energy (in the form of heat) into the surrounding gas. The collective energy output from multiple star clusters drives an outflow and disrupts the galaxy's gas via turbulence, which eventually establishes a regulatory balance with the large scale inflow of gas onto the galaxy (e.g., White & Frenk, 1991). Finally, supermassive black holes must also form (Rees, 1966), and are observed to be interlinked with their galaxy hosts (e.g., Magorrian et al., 1998). These compact objects release energy into the surrounding ISM through jets, winds and radiation from the accretion process (see Fabian, 2012 for a review). This energy output is significant, and is thought to curb any continued star formation within the galaxy, creating the 'red and dead' massive galaxies that we see today (e.g., Bower et al., 2006).

The culmination of these processes is collectively refereed to as *galaxy evolution*.

1.2 Simulating the Universe

There have been many attempts to collectively model the processes that govern galaxy evolution; yet, from the complexities and non-linearities involved, it is not feasible to achieve purely analytically. Instead, numerical simulations have been adopted in an attempt to approach the problem. These simulations can broadly be divided into two types; dark matter only simulations, where the influence of baryonic material is ignored; and hydrodynamical simulations, where the evolution of both baryonic and non-bayonic materials are traced self-consistently through the entirety of cosmic time, each with the goal of recreating a realistic 'virtual universe' by the present day. As the contents of this thesis are based around the results from the

hydrodynamical cosmological EAGLE simulation, the discussion here will concentrate on numerical simulations of cosmological scales.

1.2.1 Dark matter only simulations

The simplest treatment comes in the form of dark matter only (DMO) simulations, whereby the Universe is modelled purely as a collisionless dark matter fluid that evolves under the influence of Newton's laws in an expanding background and ignores the behaviour of the gas, the formation of stars, black holes, etc . Many different groups have worked on this method over the past 30 years, yielding many discoveries. The earliest examples were extremely limited by their computing power, only capable of tracing the evolution of 32^3 particles within a relatively small volume (Davis et al., 1985; Frenk et al., 1988). Yet even from their limited size, these simulations were crucial in establishing the legitimacy of 'cold' dark matter as a working model. As the computing power increased, so to did the volume of the simulations; producing the publicly successful Millennium Simulation (Springel et al., 2005b), and then the Millennium-XXL Simulation (Angulo et al., 2012), which now traces the evolution of 6720^3 particles within a vast cosmologically representative volume. These simulations have made important contributions towards our understanding of the large scale structure of the Universe and the formation of the 'cosmic web'.

1.2.2 Semi-Analytic modelling

In an attempt to take these simulations a step further, and incorporate galaxies, semi-analytic models (SAM) were introduced. This technique uses the underlying dark matter distribution from the DMO simulation as a foundation, in an attempt to describe an evolving galaxy population via a set of coupled differential equations. These models are calibrated to reproduce a chosen subset of observational data, such as the $z = 0$ stellar mass function. As the entirety of this technique is performed in post processing, it is extremely computationally inexpensive, allowing for a rapid

exploration of the parameter space in the models (Henriques et al., 2009; Bower et al., 2010) and has been successful in producing many self-consistent SAMs of galaxy evolution that can rigorously match many observational datasets (e.g., Benson et al., 2003; Bower et al., 2006). Additionally, as they can be run on such large underlying DMO volumes, the galaxy catalogues produced from SAMs are essential to create mock catalogues for the calibration of upcoming large scale surveys, such as EUCLID (Laureijs et al., 2011).

However, there are a number of shortcomings to using such a technique; the input parameters to the models can only be made dependent on the properties of the dark matter haloes that are assumed to be hosting the galaxies; as the modelling is run 'after-the-fact', the galaxy evolution process can never influence the underlying dark matter distribution; there is usually no information about the internal structures of galaxies; they can be 'over calibrated' to too many parameters, yielding little predictive power; and perhaps most crucially, the results are strongly dependent on the complexity of the models chosen and the simplifying assumptions that feed into them.

1.2.3 Hydrodynamical simulations

The ultimate aim, then, is to produce simulations with a full hydrodynamic treatment, thus self-consistently tracing the evolution of the baryons in conjunction with the dark matter through cosmic time. Two hydrodynamics schemes are generally favoured; Smoothed Particle Hydrodynamics (SPH) and Adaptive Mesh Refinement (AMR), each with their pros and cons (see Price, 2012 for a review). Hydrodynamical simulations are now able to create galaxy populations to the levels of accuracy of SAMs, however with many distinctive advantages; the galaxies now contain spatial information, the evolution of the baryons can now influence the underlying dark matter distribution, and the bayonic effects; such as the ram-pressure stripping of gas and the influence of stellar and black hole feedback upon the gas can now be

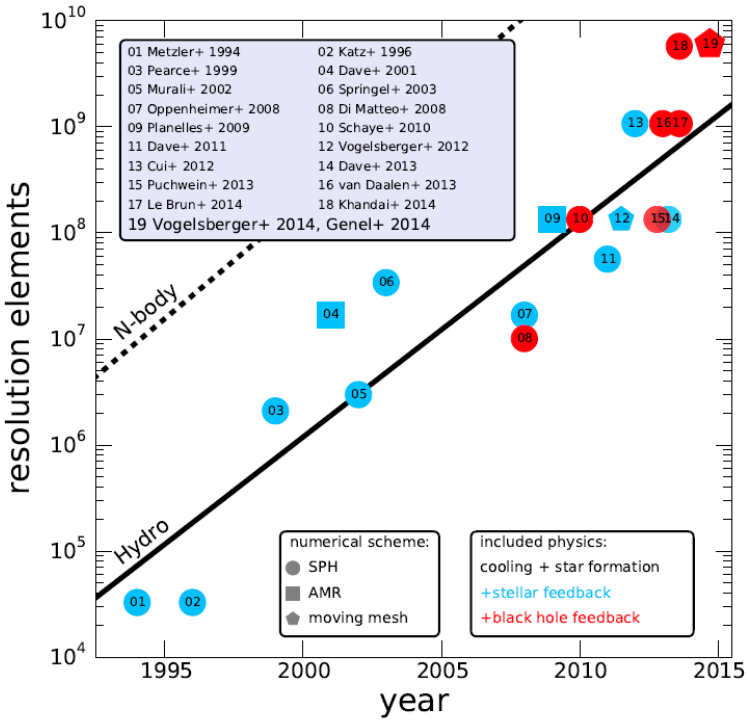

Figure 1.1: This figure is lifted directly from Genel et al. (2014) and shows the number of hydrodynamical resolution elements of simulations, as a function of their publication date. Note this only shows simulations that have completed down to $z = 0$. The different symbols and colours indicate different hydrodynamical schemes and included physics as shown by the legends. Astonishingly, hydrodynamical simulations have increased in the number of their resolution elements by $\approx$ *five* orders of magnitude over the last 20 years.

directly resolved and do not need to be added as a separate equation.

However, there are still many astrophysical processes that cannot be directly resolved even by hydrodynamical simulations, particularly those simulating cosmological volumes. For example; gas cooling, star formation and the accretion onto supermassive black holes. Instead, these essential processes are implemented as a series of 'subgrid' models, the free parameters of which are calibrated to a chosen subset of observational data.

The development of gravity codes, hydrodynamic solvers, subgrid models and computing resources have progressed at a rapid rate over the last 20 years (see Figure 1.1), to the point where we can now model representative cosmological volumes[8] with full hydrodynamical treatments; such as the ILLUSTRIS and ILLUSTRIS-TNG simulations (Vogelsberger et al., 2014; Pillepich et al., 2018), the EAGLE simulation (Schaye et al., 2015; Crain et al., 2016) and the HORIZON-AGN simulation (Dubois et al., 2016). Each of these works adopt their own chosen subgrid modelling techniques and hydrodynamic schemes, yet still match remarkably well to both calibrated and non-calibrated observational datasets, making them the state-of-the-art of simulations today. This generation has pioneered a golden age of hydrodynamical cosmological simulations, proving their capability to rapidly expand our knowledge of galaxy evolution.

In this thesis we utilise the hydrodynamical cosmological EAGLE simulation to investigate the evolution of rare phenomena within the Universe; such as actively accreting supermassive black holes, and strongly star-forming submillimetre galaxies. A more in depth overview of the EAGLE simulation is presented in Chapter 2.

[8]These simulations now achieve volumes of $(100\text{cMpc})^3$ and contain a few tens of thousands of resolved galaxies, allowing us to study a range of galaxy and halo masses, formation histories and environments simultaneously in one self-consistent simulation.

1.2.4 Idealised and zoom simulations

It is not always necessary to simulate the entirety of a galaxy population, much can also be learned from the study of a small collection of galaxies; such as binary merging systems, isolated galaxy groups or even entire clusters. One method used to achieve this is through a manually constructed set of idealised initial conditions, which has traditionally been used in the case of binary merging systems (e.g., White & Rees, 1978). Simulations such as these are relatively computationally inexpensive, and can therefore explore a wide parameter space of merger scenarios (such as the impact parameter, galaxy mass ratio, initial velocities, etc) and model dependencies (the different between those run with and with out supermassive black holes, for example). However, results from these studies must be treated with caution, as they do not evolve self consistently within a cosmological environment. For this reason, the 'zoom-in' method is generally preferred. This takes a selected region of interest from a low resolution parent cosmological volume (usually DMO) and resimulates the chosen region at a much higher resolution; these can be run as DMO simulations (e.g., Springel et al., 2008), but they are generally resimulated with the full hydrodynamics (e.g, Hopkins et al., 2014). This method is distinctly advantaged over the idealised method described above, as the initial conditions are naturally set, and the resimulated region evolves under the influence of the full cosmological environment (ensuring that the resimulated galaxies experience large scale inflows and hierarchical build-up). They are, however, more computationally expensive to perform than the idealised simulations.

A substantial amount of information has been gained from the simulations using these techniques. For example, they provide ideal testbeds for the development of new subgrid models, and have yielded the crucial importance of both stellar (e.g., Agertz et al., 2013; Schaye & Dalla Vecchia, 2008) and black hole feedback (e.g., Springel et al., 2005b; Hopkins et al., 2008). The computational expense saved via this method ensures that the subgrid models are robustly tested before being applied

to the full cosmological simulations. The primary advantage of such simulations, however, is their capability to accurately probe the small scale structure within galaxies, with resolutions that can be many orders of magnitude greater than those achieved by cosmological simulations. For example, the AURIGA (Grand et al., 2017) and APOSTLE (Sawala et al., 2016) projects take advantage of such high resolutions to investigate the satellite populations of Milky Way analogues within the local group, providing valuable incite to the 'missing satellites' (e.g., Moore et al., 1999) and 'too big to fail' (e.g., Boylan-Kolchin et al., 2011) problems.

1.2.5 Future work and upcoming simulations

Looking towards the next generation of hydrodynamical cosmological simulations, there will be two key areas of focus.

- The first is to improve upon the subgrid modelling techniques, as the prescriptions for these models yield a dramatic influence over the simulated galaxy population. This can be achieved by revisiting the physics surrounding these processes from first principles, or by examining the shotcomings of the current generation of simulations and attempting an improved recalibrated strategy. For example, the baryon fractions of massive galaxies are commonly too high, creating an excess in the amount of star formation over what is observed. This could be improved upon by revisiting the AGN feedback model within zoom simulations of individual clusters where AGN feedback is dominant. Additional improvements will also come from updated simulation codes, by exploring newer and more efficient schemes to implement the gravity and hydrodynamics. For example, the parallel task-based code SWIFT (Schaller et al., 2018) promises speed-ups of an order of magnitude over previous codes, which will yield many new opportunities to running larger and more accurate simulations in the future.

- The second improvement is to increase the volume of the Universe that is represented. The volumes of the current generation ($\approx (100$ cMpc$)^3$)[9] are still not large enough to capture the rarest (and most easily observed) phenomena within our Universe; such as high redshift quasars, strongly star-forming submillimeter galaxies and massive clusters; nor are they large enough to adequately compare to observational surveys in order to produce reliable mock catalogues. For these reasons, larger simulations are required (by a factor of $\approx$ 10–100). Efficiently producing and analysing the vast amounts of data from these simulations will be the greatest challenge for the next generation.

1.3 Thesis outline

The remainder of this thesis is organised as follows. Chapter 2 outlines the workings of the cosmological hydrodynamical EAGLE simulation, including an overview of the subgrid models. The analysis of this thesis is formed entirely from the results of this simulation. Chapter 3 investigates the rapid growth phase of supermassive black holes, a period of accelerated black hole growth that occurs within $\sim L_*$ (those with halo masses $M_{200} \sim 10^{12}$ M$_\odot$) galaxies that is embedded between two regulatory states of the galaxy host: in sub L_* galaxies efficient stellar feedback is found to regulate the gas inflow onto the galaxy and significantly reduces the growth of the central black hole, while in galaxies more massive than L_* efficient AGN feedback is found to regulate the gas inflow onto the galaxy and curbs further non-linear black hole growth. We find that this creates a complex evolutionary pathway for the black holes within the simulation, and has many knock-on implications for the way galaxies evolve. Chapter 4 goes on to investigate the coevolution of black hole and galaxy growth throughout cosmic time. We find that the simulation can reproduce the alternate trends found empirically between the galaxy and black hole

[9]Throughout this thesis we use the following notation; co-moving distances are preceded by the letter 'c', e.g., cMpc, and proper distances are preceded by the letter 'p', e.g., pkpc.

growth rates when selecting the initial sample via different methods. We argue that this discrepancy is not created due to the short variability timescales of black hole accretion (which has been commonly assumed), and it is instead linked to the complex interplay between galaxies and their black holes throughout cosmic time, due to the multiple phases of black hole growth described in Chapter 3. Finally, in Chapter 5 we investigate the origin and evolution of the most strongly star-forming galaxies (those with star formation rates greater than 80 $M_{\odot}$ yr^{-1}). We find, that at lower stellar masses ($M_{*} \lesssim 10^{11}$ $M_{\odot}$) these systems are rare, and are undergoing a different star formation process to the 'typical' star-forming population. We argue that the majority of strongly star-forming galaxies are produced through a culmination of two coinciding events; (1) the galaxy must host an underdeveloped black hole, thus ensuring that the galaxy has maintained a gas rich reservoir and contains a low mass black hole; and (2) the galaxy must undergo an interaction to kick-start the starburst process. This tells us that strongly star-forming galaxies are predominantly a merger driven population hosting undermassive black holes, making them fundamentally different from the typical star-forming population.

Chapter 2

The EAGLE simulation suite

The "Evolution and Assembly of GaLaxies and their Environment" (EAGLE) project is a suite of hydrodynamical cosmological simulations that follow the formation and evolution of galaxies within a series of cosmologically representative volumes. The largest simulation, measuring $(100\ \mathrm{cMpc})^3$ and containing over 7 billion resolution elements, was completed in January 2014 and took ≈ 4.5 Million CPU hours to reach $z = 0$. A comprehensive detailing of the simulation suite is given by Schaye et al. (2015) and Crain et al. (2015), however, as the entirety of the work presented in this thesis is formed from the analysis of these simulations, this chapter is included as a reference to the reader as an overview of the simulation suite.

2.1 Overview

The EAGLE simulation suite is a set of cosmological hydrodynamical simulations in cubic, periodic volumes ranging from 25 to 100 cMpc per side that track the evolution of both baryonic (gas, stars and supermassive black holes) and non-baryonic (dark matter) elements from a starting redshift of $z = 127$ to the present day. All simulations adopt a flat ΛCDM cosmology with parameters taken from the *Planck* mission (Planck Collaboration et al., 2014) results: $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.04825$, $\sigma_8 = 0.8288$, $n_s = 0.9611$, $Y = 0.248$ and $H_0 = 67.77$ km s^{-1}

Mpc^{-1} (i.e., $h = 0.6777$). The initial conditions were generated using second-order Lagrangian perturbation theory (Jenkins, 2010) and the phase information is taken from the public PANPHAISA Gaussian white noise field (Jenkins, 2013). Full details of how the initial conditions were made are given in Appendix B of Schaye et al. (2015). The simulation suite was run with a modified version of the GADGET-3 Smoothed Particle Hydrodynamics (SPH) code (last described by Springel, 2005), and includes a full treatment of gravity and hydrodynamics. The modifications to the SPH method are collectively referred to as ANARCHY (Dalla Vecchia, (*in prep.*), see also Appendix A of Schaye et al. (2015) and Schaller et al. (2015a)), and include the $\mathcal{C}_2$ kernel of Wendland (1995), the Hopkins (2013) pressure-entropy formulation of SPH, the time-step limiters introduced by Durier & Dalla Vecchia (2012), the artificial viscosity switch of Cullen & Dehnen (2010) and a weak thermal conduction term of the form proposed by Price (2008). The effects of this state-of-the-art formulation of SPH on the galaxy properties is explored in detail by Schaller et al. (2015a).

The astrophysical processes that operate below the simulation resolution are treated by a series of 'subgrid' prescriptions (a description of the subgrid models is given in Section 2.4). Because of our limited understanding of these processes and because of the limited resolution of the simulations, the subgrid source and sink terms involve poorly constrained parameters whose values must be determined by comparison of the simulation results to a subset of the observational data. In the case of EAGLE, the subgrid parameters were calibrated for feedback from star formation and AGN by using three properties of galaxies at redshift $z = 0$, specifically the galaxy stellar mass function, the galaxy size – stellar mass relation, and the black hole mass – stellar mass relation. The calibration strategy is described in detail by Crain et al. (2015) who also presented additional simulations to demonstrate the effect of parameter variations.

Once the simulations have been calibrated using a subset of the observational data, they can be validated by comparison to additional datasets. Studies have so far

shown that the simulations broadly reproduce a variety of other observables such as the $z = 0$ Tully-Fisher relation, specific star formation rates and the column density distribution of intergalactic C IV and O VI (Schaye et al., 2015), the H I and H_2 properties of galaxies (Bahé et al., 2016; Lagos et al., 2015), the column density distribution of intergalactic metals (Schaye et al., 2015), galaxy rotation curves (Schaller et al., 2015b), the $z = 0$ luminosity function and colour-magnitude diagram (Trayford et al., 2015), the evolution of the galaxy stellar mass function (Furlong et al., 2015b), the high-redshift H I column density distribution (Rahmati et al., 2015) and the AGN luminosity function (Rosas-Guevara et al., 2016). To date, over 100 refereed scientific papers have been produced that directly utilize data from the EAGLE simulation suite.

2.1.1 The simulation runs

Table 2.1 summarises the the fiducial simulations of the EAGLE suite, including the comoving cubic box length, baryonic and non-baryonic particle masses and gravitational softening lengths. Together these parameters determine the dynamic range and resolution that can be achieved by the simulation. The simulation name includes a suffix to indicate the simulation box length in comoving megaparsecs (e.g., L0100) and the cube root of the initial number of particles per species (e.g., N1504). Simulations with the same subgrid model as the primary run (L0100N1504) are denoted with the prefix "Ref-". As discussed in Schaye et al. (2015), the "Recal-" higher-resolution simulations, such as Recal-L0025N0752, use values of the subgrid parameters that have been recalibrated following the same procedure used for the reference simulation to improve the fit to the $z = 0$ galaxy stellar mass function. These recalibrated higher resolution simulations allow for the *weak* convergence of the code to be tested[1]. To a similar end, higher resolution non-recalibrated simulations,

[1]Two simulations run at alternate resolutions are said to 'weakly' converge if they produce the same result after a recalibration of the subgrid model parameters. By contrast, these two simulations would 'strongly' converge if they produce the same result without the need for any recalibration. See Schaye et al. (2015) for a discussion on weak and strong convergence.

such as Ref-L0025N0752, allow for the *strong* convergence of the code to be tested. These simulations use all the subgrid parameters of the reference model, but at a higher mass resolution.

Finally, many simulations with alternate subgrid parameters to the reference model were also run to investigate their influence on the properties of galaxies. For example, AGNdT9-L0050N0752 uses a higher AGN heating temperature and increased black hole accretion viscosity parameter, $C_{\rm visc}$. As discussed by Schaye et al. (2015), this results in a better match to the properties of diffuse gas in galaxy group haloes, but has only a small effect on the properties of galaxies.

2.2 Gravity

Dark matter interacts only through gravity, and is treated as a collisionless fluid. Its motion is therefore governed by the collisionless Boltzmann equation in an expanding Universe (e.g., Binney & Tremaine, 1987) coupled with the Poisson equation,

$$\nabla^2\Phi(\mathbf{r}) = 4\pi G a^2(\rho(\mathbf{r}) - \bar{\rho}), \tag{2.2.1}$$

where G is the gravitational constant, $\Phi(r)$ is the *peculiar* gravitational potential, ρ is the density, $\bar{\rho}$ is the mean density and a is the expansion factor. These coupled equations are complex, and are instead solved in the simulation by sampling the phase-space density field with a number of N tracer particles.

The simplest approach is to compute the gravity force on a given particle i of mass $m_{\rm i}$ using the conventional Newtonian approach,

$$\mathbf{F}(\mathbf{r}_{\rm i}) = \sum_{j=1,j\neq i}^{N-1} \frac{G m_{\rm i} m_{\rm j}}{|\mathbf{r}_{\rm j} - \mathbf{r}_{\rm i}|^3}(\mathbf{r}_{\rm j} - \mathbf{r}_{\rm i}). \tag{2.2.2}$$

However, as this is essentially a N^2 problem ($N(N-1)$), it is computationally extremely challenging to achieve for large values of N. In reality, this sum is approximated in order to substantially reduce the computational expense.

Identifier	L [cMpc]	N	$m_{\rm g}$ [$\rm M_\odot$]	$m_{\rm dm}$ [$\rm M_\odot$]	$\epsilon_{\rm com}$ [ckpc]	$\epsilon_{\rm phys}$ [pkpc]	$n_{\rm H,0}$ [$\rm cm^{-3}$]	$n_{\rm n}$	$C_{\rm visc}$	$\Delta \rm T_{AGN}$ [K]
Ref-L0025N0376	25	2×376^3	1.81×10^6	9.70×10^6	2.66	0.70	0.67	2/ln10	2π	$10^{8.5}$
Ref-L0025N0752	25	2×752^3	2.26×10^5	1.21×10^6	1.33	0.35	0.67	2/ln10	2π	$10^{8.5}$
Recal-L0025N0752	25	2×752^3	2.26×10^5	1.21×10^6	1.33	0.35	0.25	1/ln10	$2\pi\times10^3$	$10^{9.0}$
Ref-L0050N0752	50	2×752^3	1.81×10^6	9.70×10^6	2.66	0.70	0.67	2/ln10	2π	$10^{8.5}$
AGNdT9-L0050N0752	50	2×752^3	1.81×10^6	9.70×10^6	2.66	0.70	0.67	2/ln10	$2\pi\times10^2$	$10^{9.0}$
Ref-L0100N1504	100	2×1504^3	1.81×10^6	9.70×10^6	2.66	0.70	0.67	2/ln10	2π	$10^{8.5}$

Table 2.1: Parameters describing the fiducial simulations within the EAGLE suite. From left-to-right the columns show: simulation name suffix; comoving box size; total number of particles; initial baryonic particle mass; dark matter particle mass; comoving Plummer-equivalent gravitational softening length; maximum physical softening length and the subgrid model parameters that vary: $n_{\rm H,0}$, $n_{\rm n}$, $C_{\rm visc}$ and $\Delta \rm T_{AGN}$ (see section 4 of Schaye et al. (2015) for an explanation of their meaning).

This is achieved through a combination of the tree and particle mesh (PM) techniques (tree-PM). Long range and periodic forces are computed by mapping the density field onto a regular mesh and solving the Poisson equation (eq. (2.2.2)) directly in Fourier space. Short range forces are obtained using the octree approach of Barnes & Hut (1986). A cubical root node that encompasses the full mass distribution is recursively divided into eight daughter nodes until one ends up with 'leaf' nodes containing single particles. Particles belonging to each node then have their gravity represented collectively via a single multipole force. The total gravity force on a given particle is then obtained by 'walking' the tree, starting with the root node. A decision is made whether or not the multipole expansion of the node provides an accurate enough representation of the force acting upon the particle. If 'yes', the multipole force is used and the walk along this branch of the tree is terminated, if 'no', the node is 'opened' and each of the daughter nodes are considered in turn. The force from the node upon the considered particle is *not* considered accurate enough, and 'opened', when

$$\frac{GM}{r^2}\left(\frac{l}{r}\right)^2 \leq \alpha|\mathbf{a}|, \tag{2.2.3}$$

where r is the distance to the node, l is the node size, M is the total node mass, α is a tolerance parameter (≈ 0.7) and $\mathbf{a}$ is the total acceleration of the particle at the previous simulation timestep.

The gravity forces acting upon the non-dark matter tracer particles (i.e., the gas, star and black hole particles, see next section) are computed in the same manner.

2.3 Hydrodynamics

Gas within the simulation is assumed to be an inviscid fluid, treated by a Smoothed Particle Hydrodynamics (SPH) scheme (Lucy, 1977; Gingold & Monaghan, 1977; Monaghan, 1992). Analogous to the dark matter, tracer particles are used to dis-

cretize the fluid, and their SPH quantities are obtained by interpolating between their SPH neighbours using a smoothing kernel function, $W(|\mathbf{r} - \mathbf{r}_\mathrm{i}|, h)$, where h is the smoothing length.

The smoothing lengths of gas particles are predicted as per the method of Hopkins (2013). First, the SPH density for each particle is defined as

$$\rho_i = \sum_{j=1}^{N} m_j W(|\mathbf{r}_\mathrm{j} - \mathbf{r}_\mathrm{i}|, h_i), \tag{2.3.1}$$

where m_j is the mass of each other particle and $W(|\mathbf{r}_\mathrm{j} - \mathbf{r}_\mathrm{i}|, h_i)$ is the value of the kernel at that location. The smoothing length is predicted by knowing that it must yield a proportionality to this density, $h_i \propto \rho_i^{-1/3}$, such that the relationship $(4\pi/3)h_i^3\rho_i = m_i N_\mathrm{ngb}$ holds true for a given choice of N_ngb, referred to as the 'effective neighbour number' (see Hopkins (2013) and Appendix A1 of Schaye et al. (2015) for details). N_ngb is chosen to be 58 for gas particles. The smoothing lengths for star and black hole particles are also predicted from the neighbouring gas particles. However as they are not gas particles themselves, the smoothing length is now computed ensuring that the relation $(4\pi/3)h_i^3 \sum_{j=1}^{N} W_{ij}(h_i) = N_\mathrm{ngb}$ holds true for a given choice of N_ngb. For stars N_ngb is chosen to be 48 (to speed up the calculation), and 58 for black holes.

One can then deduce the equations of motion from these particle densities (see Price, 2012 for a review). We start with the Lagrangian (L) of the system of particles, whereby it is assumed that the potential (thermal) energy can be expressed through an equation of state that connects the internal energy (u) of a particle to its density (ρ) and entropy (s). Then, coupled with the 1st law of thermodynamics (to link to the pressure, P), the Euler-Lagrange equation can be employed to reveal the standard SPH expression for the equations of motion (Monaghan, 1992),

$$\frac{d\mathbf{v}_i}{dt} = -\sum_j m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2} \right) \nabla_i W_{ij}(h). \tag{2.3.2}$$

By coupling the density and entropy of the gas particles through an equation of state, the pressure P can be derived, which allows for the computation of the forces that will act upon the particles.

2.4 Subgrid models

Processes not resolved by the numerical scheme are implemented as subgrid source and sink terms in the differential equations. For each process, schemes were adopted that are as simple as possible and that only depend on the local hydrodynamic properties. This last requirement differentiates EAGLE from most other cosmological, hydrodynamical simulation projects (e.g. Oppenheimer et al., 2010; Puchwein & Springel, 2013; Vogelsberger et al., 2014; Khandai et al., 2015) and ensures that galactic winds develop without pre-determined mass loading factors and directions, without any direct dependence on halo or dark matter properties.

The subgrid models adopted by the EAGLE simulation are an improved version of those used in the GIMIC and OWLS simulations (Crain et al., 2009; Schaye et al., 2010). Here we only give a brief review, for a comprehensive detailing of the models and the calibration strategy of the free parameters see Schaye et al. (2015) and Crain et al. (2015), respectively.

2.4.1 Radiative processes and reionization

The photoheating and radiative cooling rates of the gas particles are computed on an element-by-element (for the 11 elements H, He, C, N, O, Ne, Mg, Si, S, Ca and Fe) basis by interpolating the tabulated rates from Wiersma et al. (2009a). The rates, which are a function of density, temperature and redshift, are produced using the radiative transfer code CLOUDY (Ferland et al., 1998) by assuming that the gas is optically thin and in ionisation equilibrium, and is exposed to the CMB and the evolving UV/X-ray background from Haardt & Madau (2001).

To mimic the effects of hydrogen and helium II reionization, the time-dependent, spatially-uniform ionizing background from Haardt & Madau (2001) is turned on at $z = 11.5$. In addition, to account for the boost in the photoheating rates during the epochs of reionization (relative to the optically thin rates computed for the uniform background), 2 eV of thermal energy per proton mass is also injected. For hydrogen reionization, this energy is injected instantaneously at $z = 11.5$ (to be consistent with the Planck constraints) and quickly raises the gas particle temperature to $\sim 10^4$ K. For helium II reionization, this energy is injected over a range of redshifts drawn from a Gaussian centered on $z = 3.5$ with a width of $\sigma(z) = 0.5$. These values were chosen to broadly agree with the thermal history of the intergalactic medium as measured by Schaye (2004).

2.4.2 The interstellar medium and star formation

The physics and resolution to model the interstellar medium (ISM), particularly the cold gas phase ($T \ll 10^4$ K, $n_{\rm H} > 0.1$ cm^{-3}), is not captured by the simulation. A temperature floor, $T_{\rm eos}(\rho)$ is therefore imposed, corresponding to a polytropic equation of state $P_{\rm eos} \propto \rho^{4/3}$. This is normalised to $T_{\rm eos} = 8\times10^3$ K at $n_{\rm H} = 0.1$ cm^{-3}, a temperature typical for the warm ISM. The fiducial choice of 4/3 is adopted to ensure that the Jeans mass, and the ratio of the Jeans length to the SPH kernel, is independent of density, and thus, helps to prevent the spurious fragmentation of high density gas (Schaye & Dalla Vecchia, 2008).

A minimum density threshold for star formation is implemented, above which a cold phase (and with it stars) would be expected to form. The metallicity-dependent threshold proposed by Schaye (2004) is used,

$$n_{\rm H}^* = 10^{-1}{\rm cm}^{-3}\left(\frac{Z}{0.002}\right)^{-0.64}, \tag{2.4.1}$$

where Z is the metallicity of the gas particle (i.e., the fraction of the gas particle's mass from elements more massive than helium). As the Schaye (2004) relation

diverges at low metallicities, an upper limit of $n_{\rm H}^* = 10\,{\rm cm}^{-3}$ is imposed. Additionally, to prevent star formation in low overdensities at high redshift, this threshold must also exceed 57.7 times the cosmic mean.

For the gas particles that meet the density threshold criteria, their star formation rate is computed following the pressure-dependent law from Schaye & Dalla Vecchia (2008),

$$\dot{m}_* = m_{\rm g} A (1\ {\rm M}_\odot {\rm pc}^{-2})^{-n} \left(\frac{\gamma}{G} f_{\rm g} P\right)^{(n-1)/2}, \tag{2.4.2}$$

where G is the gravitational constant, $m_{\rm g}$ is the gas particle mass, γ is the ratio of specific heats (set to 5/3), $f_{\rm g}$ is the mass fraction in gas (set to unity) and P is the total pressure. This is derived from the observed Kennicutt-Schmidt star formation law, $\dot{\Sigma}_* = A(\Sigma_{\rm g}/1\ {\rm M}_\odot\ {\rm pc}^{-2})^n$ (Kennicutt, 1998), by assuming that the gas is self-gravitating. This implementation, compared to those with a pure density dependence, has the advantage that the free parameters A and n are then set directly by the observations, and forgoes their requirement to be tuned.

Star particles are formed stochastically. The probability that a gas particle is converted into a star particle is $\min(\dot{m}_* \Delta t / m_{\rm g}, 1)$, where Δt is the simulation timestep.

2.4.3 Stellar mass loss and evolution

Once formed, star particles are treated as simple stellar populations with a Chabrier (2003) initial mass function in the range 0.1–100 $\rm M_\odot$. Following the implementation from Wiersma et al. (2009b), the metallicity-dependent lifetimes from Portinari et al. (1998) are used to compute which stellar masses within the simple stellar population will reach the end of their main-sequence phase after each simulation timestep. This then dictates what fraction of the star particle's mass is lost to its neighbouring gas particles through the winds of AGB and massive stars, and type Ia and type II supernovae. The resulting enrichment of the interstellar medium by these processes

is tracked individually for the eleven elements, to be used in the computation of the radiative heating and cooling rates of the gas as mentioned above.

2.4.4 Stellar feedback

The energy released by core collapse supernovae, referred to here as 'stellar feedback', is injected thermally and *stochastically* following the methods from Dalla Vecchia & Schaye (2012). By implementing the feedback stochastically, the mean energy injection per unit of stellar mass formed over the course of the simulation can remain fixed, whilst the energy *per feedback event* can be strictly controlled. This method is invoked as it has traditionally been difficult to overcome the catastrophic radiative losses associated with thermal feedback within hydrodynamical simulations (Katz et al., 1996). This 'overcooling' has typically been attributed to a lack in numerical resolution, as the simulations cannot produce cold dense clouds, star formation is not sufficiently clumpy, and the resulting feedback energy is then distributed too smoothly. Furthermore, the feedback energy is distributed over multiple gas particle neighbours of an equivalent mass to the SSP, which dilutes the potential maximum temperature that could be achieved by the surrounding ISM in reality. Therefore the gas particles heated by stellar feedback events are raised by an adequately large fixed temperature value, denoted ΔT.

Once a star particle has reached the age 3×10^7 yr it is eligible for its one and only feedback event (corresponding to the maximum lifetime of stars that explode as core collapse supernovae). At this time, $\approx 10^{51}$ erg of thermal energy per massive star ($M_* > 6$ M$_\odot$) is released, of which, a fraction $f_{\rm th}$ will couple to the gas. Each gas particle within the star particle's kernel has the same probability of receiving energy from the star particle during this event. As it is required that the mean injected energy to the neighbouring gas particles balances the available energy produced by the star particle ($f_{\rm th}\epsilon_{\rm SNII}m_*$, where $\epsilon_{\rm SNII}$ is the total energy per unit stellar mass provided by type II supernovae and m_* is the star particle mass), the probability

that a neighbouring gas particle is heated during the feedback process is

$$p = f_{\rm th} \frac{\epsilon_{\rm SNII}}{\Delta\epsilon} \frac{m_*}{\Sigma_{\rm i=0}^{N_{\rm ngb}} m_{\rm i}}, \tag{2.4.3}$$

where m_i is the gas particle mass, $N_{\rm ngb}$ is the number of gas particles within the star particles kernel and $\Delta\epsilon$ is the amount of thermal energy per unit mass given to each heated gas particle (which is proportional to the fixed temperature increase ΔT). Therefore the greater the value of ΔT the more intermittent the stellar feedback becomes (as it becomes probabilistically less likely), which can lead to a potential under-sampling of the feedback cycle, however if the value of ΔT is too low, radiative losses will dominate. To provide the best balance, the selected value of ΔT for the EAGLE reference model is $\Delta T = 10^{7.5}$ K. The expectation value for the number of heated gas particles per star particle is

$$\langle N_{\rm heat} \rangle \approx 1.3 f_{\rm th} \left(\frac{\Delta T}{10^{7.5} {\rm K}} \right)^{-1}. \tag{2.4.4}$$

The value of $f_{\rm th}$ is not fixed (however it is ≈ 1), and is a function of the metallicity and density of the gas particle, and is one of the primary tunable parameters of the simulation (see Crain et al., 2015 for full details).

2.4.5 Supermassive black holes

Following the original prescription from Springel et al. (2005a), the modeling of BHs can be divided into three primary processes: seeding, growth and feedback.

- **BH Seeding**. Firstly, seed BHs of mass 1.475×10^5 $M_\odot$ are injected into haloes more massive than 1.475×10^{10} $M_\odot$ that do not already contain a BH. The haloes are identified 'on-the-fly' using the 'Friends Of Friends' algorithm (described in the next section) at intervals of $\Delta a = 0.005a$, where a is the expansion factor. If these conditions are met, the most bound gas particle of

the halo is converted into a seed BH. BHs enter the simulation with two masses: a particle mass ($\hat{m}_{\rm BH}$), which is used in the computation of the gravitational forces, and a subgrid mass ($m_{\rm BH}$), which is used in the computation of the subgrid models. The initial particle mass of the BH is inherited from the converted gas particle, and the initial subgrid mass, referred to as the 'true' mass of the BH, is the BH seed mass. As the simulation cannot model the dynamical friction acting upon BHs with masses much lower than the gas particle mass, the positions of BHs are manually migrated towards the local potential minimum of the neighbours within their kernel until they reach 100 times the initial gas particle mass.

- **BH Growth**. BHs are allowed to grow via the accretion of the neighbouring gas particles, and by BH–BH mergers. BHs accrete mass at a rate that is the minimum between a modified Bondi & Hoyle (1944) rate:

$$\dot{m}_{\rm BH} = \frac{4\pi G^2 m_{\rm BH}^2 \rho}{(c_{\rm s}^2 + v^2)^{3/2}} \times \min[(c_{\rm s}/v_\phi)^3/C_{\rm visc}, 1], \tag{2.4.5}$$

and the Eddington rate

$$\dot{m}_{\rm edd} = \frac{4\pi G m_{\rm BH} m_{\rm p}}{\epsilon_{\rm r}\sigma_{\rm T} c}. \tag{2.4.6}$$

For these equations, G is the gravitational constant, $m_{\rm p}$ is the mass of the proton, $\sigma_{\rm T}$ is the Thomson cross section, c is the speed of light, $\epsilon_{\rm r}$ is the radiative efficiency of the accretion disk set to 0.1, $m_{\rm BH}$ is the BH mass and ρ, $c_{\rm s}$ and v are the density, sound speed and relative velocity of the gas surrounding the BH, respectively. The accretion onto BHs is not considered to be spherically symmetric, and accounts for the circular velocity of the surrounding gas, v_ϕ (Rosas-Guevara et al., 2015). $C_{\rm visc}$ is a free parameter of the simulation, thought of as the effective viscosity of the accretion disk, and is set to 2π for the EAGLE reference model. Unlike previous simulations (e.g., Booth & Schaye, 2009),

EAGLE does not use an α 'boost factor' in the accretion rate calculation. BHs are merged if they are separated by a distance less than both the smoothing kernel of the BH and three gravitational softening lengths, and if their relative velocity is smaller than the circular velocity at the distance between them.

- **BH Feedback**. AGN feedback is implemented via a single mode, whereby energy is injected thermally and stochastically into the neighbouring gas particles in a manner analogous to the method for stellar feedback. Each BH has an energy reservoir, $E_{\rm BH}$, that increases by $\epsilon_{\rm f}\epsilon_{\rm r}\dot{m}_{\rm BH}\Delta t$ after each simulation timestep. Here, $\epsilon_{\rm f}$ is the feedback efficiency (i.e., the fraction of energy that couples to the gas), and is set to 0.15. This occurs until the reservoir has stored enough energy to heat at least one gas particle by $\Delta T_{\rm AGN} = 10^{8.5}$ K. Each gas particle within the BHs kernel is then eligible to receive energy from the BH. For each gas particle neighbour, the heating probability is

$$P = \frac{E_{\rm BH}}{\Delta\epsilon_{\rm AGN} N_{\rm ngb}\langle m_{\rm g}\rangle}, \tag{2.4.7}$$

where $\Delta\epsilon_{\rm BH}$ is the change in internal energy per unit mass corresponding to the temperature increase $\Delta T_{\rm AGN}$, $N_{\rm ngb}$ is the number of gas particles within the BHs kernel and $\langle m_{\rm g}\rangle$ is their mean mass. The timestep of the simulation is limited (Durier & Dalla Vecchia, 2012) for BHs such that the expectation value is $P < 0.3$. Because of the uncertainty of using accretion rates calculated at the previous timestep, P is also capped to 0.3. Once the heating process is complete, the energy reservoir of the BH is reduced by the expectation value for the injected energy.

2.5 The simulation output

The memory footprint of cosmological simulations is simply too large to continuously store to disk (the Ref-100 EAGLE volume occupied over 20 Tb of memory). Therefore,

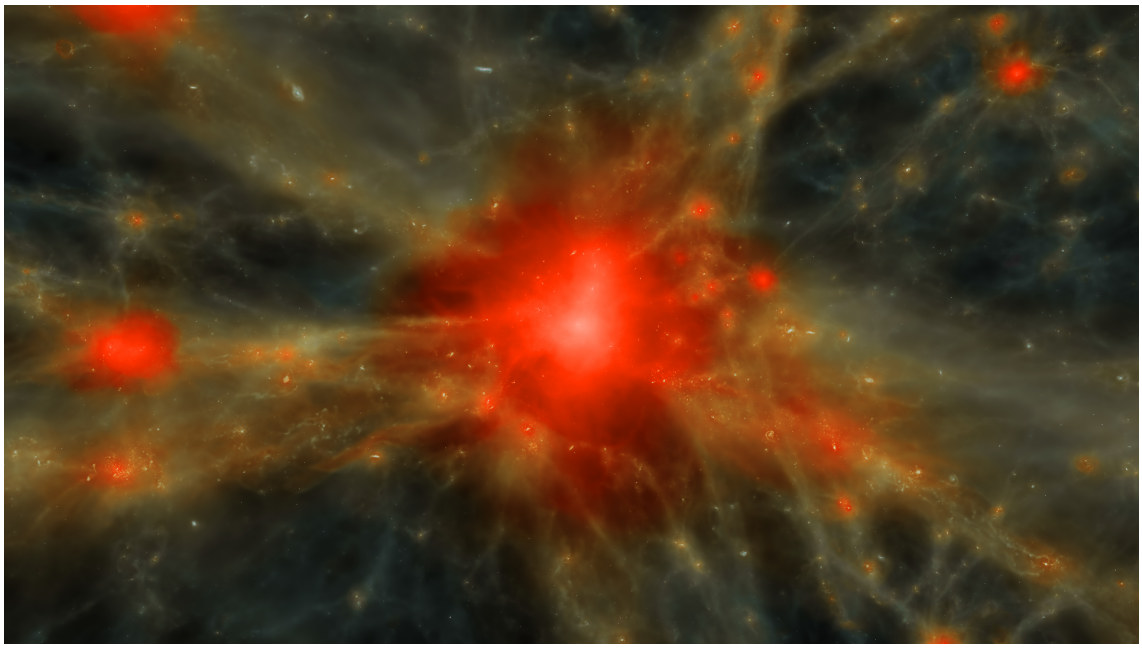

Figure 2.1: The projected gas density of one of the largest haloes within the Ref-L0100N1504 simulation. The image is colour coded by the gas temperature (white/red $\sim 10^6$ K to blue $\sim 10^4$ K).

for each simulation, 29 'snapshot' outputs are produced between redshift 20 and 0. These contain the full particle data at the discrete redshift, and have typical temporal spacings of $\approx$ 400–800 Myr between each output. For the Ref-L0100N1504 simulation an additional 400 data-lite 'snipshots' were also produced for the same redshift interval. These contain less information per particle compared to the snapshot output, but have a higher time resolution, with typical temporal spacings of $\approx$ 20–80 Myr between each output.

A novel method of displaying the simulation output is in the form of a virtual image, such as the one shown in Figure 2.1. This shows the projected gas density of one of the largest haloes within the Ref-L0100N1504 simulation, coloured by the temperature of the gas. These images demonstrate the vast dynamic range of spatial scales that are reproduced by cosmological simulations, whilst also providing a useful analysis tool for inspecting phenomena all the way from large scale structure down to individual astrophysical objects.

2.5.1 Halo, subhalo and galaxy identification

The raw particle data themselves are not required for many comparisons with observations. Instead, the simulation outputs are processed individually to locate bound structures which are identified with galaxies and their associated dark matter haloes. The processing steps are described in detail by Schaye et al. (2015). In brief, overdensities of dark matter are identified using the "Friends-of-Friends" (FoF) method (Davis et al., 1985) adopting a linking length of 0.2 times the average inter-particle spacing. Baryonic particles are then assigned to the same FoF-halo as their closest dark matter neighbour. Self-bound "subhaloes", which can contain both baryonic and dark matter, are later identified using the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009) using all particle species.

It is important to note that particles are not shared between subhaloes, so that the correspondence between particles and subhaloes is unique. The baryonic component of each subhalo is identified as a galaxy. A FoF halo may contain several subhaloes; the subhalo that contains the particle with the lowest value of the gravitational potential is defined to be the *central galaxy* (which is almost always the most massive) while any remaining subhaloes are classified as *satellite galaxies*.

For this work, the standard EAGLE mass conventions are adopted. Halo mass, M_{200}, is defined as the total mass enclosed within r_{200}, the radius at which the mean enclosed density is 200 times the critical density of the Universe. Galaxy mass, M_*, is defined as the total stellar content bound to a subhalo within a spherical aperture with radius 30 proper kiloparsecs (pkpc), as per Schaye et al. (2015).

2.5.2 Merger trees

As galaxies rarely evolve in isolation, they are subject to mergers with neighbouring galaxies. This adds serious complexity to tracing the history of an individual galaxy from the present-day to its formation and as such a *merger tree* must be constructed to connect galaxies across simulation outputs. Descendant subhaloes, and hence

galaxies, are identified using the D-Trees algorithm (Jiang et al., 2014), with a complete description of its adaptation to the eagle simulations provided in Qu et al. (2017). In essence, the algorithm traces subhaloes using the $N_{\rm link}$ most bound particles of any species and identifies the subhalo that contains the majority of these particles at the next output time as the subhalo's descendant. It is defined that $N_{\rm link} = \min(100, \max(0.1N_{\rm galaxy}, 10))$, where $N_{\rm galaxy}$ is the total number of particles in the parent subhalo. This process allows for the identification of descendants even in the case where most of the particles have been stripped and it minimises the misprediction of mergers during fly-bys (Fakhouri & Ma, 2008; Genel et al., 2009).

The galaxy with the most $N_{\rm link}$ particles at the next output is identified as the single *descendant* of a galaxy, while a descendant galaxy can have multiple progenitors. However, the *main progenitor*, corresponding to the main branch of the merger tree, is defined as the progenitor with the largest 'branch mass', i.e., the mass summed across all earlier outputs, as proposed by De Lucia & Blaizot (2007). This definition of the main progenitor, as opposed to the simple definition of the progenitor with the largest mass, is used to avoid main branch swapping in the case of similar-mass mergers, as explained by Qu et al. (2017). Note that because the progenitor with the largest branch mass determines the main branch of the tree, main branch galaxies do not necessarily always correspond to the central galaxy of a given halo.

There are two further aspects of the merger trees that must be kept in mind when analysing the simulation:

- A galaxy can disappear from a snapshot but reappear at a later time (e.g. if one galaxy passes through another one). To account for this, descendants are identified using up to 5 snapshots at later times.
- Care must be taken when determining mass ratios, for example in the case of mergers, as galaxies can lose or gain mass due to the definition of the subhaloes.

Both of these relatively rare cases are considered further by Qu et al. (2017), who discuss their impact on the assembly of galaxy mass.

Chapter 3

The rapid growth phase of supermassive black holes

3.1 Introduction

Feedback from star formation, including stellar winds, radiation pressure and supernovae, plays a key role in galaxy evolution. Collectively described as 'stellar feedback', the energy injection into the surrounding interstellar medium can eject material from the galaxy via an outflow (see Veilleux et al., 2005 for a review). In the absence of this process, many observed phenomena within the galaxy population simply cannot be reproduced by current models: such as the relatively low percentage of baryons that eventually convert into stars ($\approx 10\%$, e.g., Fukugita et al., 1998), the flattening of the faint-end slope of the luminosity function (e.g., White & Rees, 1978; Dekel & Silk, 1986; Benson et al., 2003), the formation of exponential disks (e.g., Binney et al., 2001; Scannapieco et al., 2008), the formation of dark matter cores (e.g., Navarro et al., 1996), the cosmic star formation history (e.g., White & Frenk, 1991) and the chemical enrichment of the intergalactic medium (e.g., Aguirre et al., 2001).

At masses below $\sim L_*$ ($M_{200} \sim 10^{12}$ M$_\odot$), galaxies maintain a quasi-equilibrium, with the star formation rate and the associated supernovae-driven outflow balancing the

rate of the cosmic inflow (e.g., White & Frenk, 1991; Finlator & Davé, 2008; Bouché et al., 2010; Schaye et al., 2010). However, as galaxies evolve past $\sim L_*$, stellar feedback becomes unable to effectively remove material from the galaxy, and the equilibrium breaks (e.g., Benson et al., 2003; Hopkins et al., 2014; Keller et al., 2016). A further source of energy is therefore required to balance against the cosmic inflow and restore the quasi-equilibrium, which is commonly attributed to the feedback from the central supermassive back hole (BH, e.g., Croton et al., 2006; Bower et al., 2006; Booth & Schaye, 2010).

Beyond affecting the continued production of stars within the galaxy, it is plausible that stellar feedback can also significantly hinder the growth of the central supermassive BH in sub $\sim L_*$ galaxies, where stellar feedback remains able to drive an effective outflow, and starve the inner regions of fuel for BH accretion. This result is indeed found by many current hydrodynamical simulations (e.g., Dubois et al., 2015; Anglés-Alcázar et al., 2017; Bower et al., 2017; Habouzit et al., 2017). The critical point at which the stellar feedback driven outflows begin to stall will naturally be linked to the first meaningful period of BH growth. However, the critical mass scale at which this transition occurs, the triggering mechanism, and the growth of the BH during this time, remain uncertain.

In this study we utilise the EAGLE cosmological hydrodynamical simulation (Schaye et al., 2015; Crain et al., 2015) to investigate the evolution of 1,888 massive BHs ($M_{\rm BH} \geq 10^7$ M$_\odot$) and the host galaxies during the rapid growth phase. This large sample of BHs allows us for the first time to link the stalling of stellar feedback driven outflows to the initiation of rapid BH growth in statistical detail, and measure the importance of external events, such as galaxy–galaxy mergers, to this period of BH evolution.

The paper is organised as follows. In Section 3.2 we briefly describe the EAGLE simulations, our BH sample selection, how we define the time of the rapid growth phase and how we define the 'most proximate' merger. Section 5.3 contains our main results, Section 5.4 outlines our discussion and in Section 4.6 we present our

conclusions.

3.2 The EAGLE simulation

A full overview of the EAGLE simulation suite can be found in Chapter 2. For this chapter we are interested in the evolution of massive BHs ($M_{\rm BH} \geq 10^7 {\rm M}_\odot$), and therefore restrict our study to the largest simulation, Ref-L0100N1504, which contains the greatest number of these objects.

We define the completion time of a galaxy–galaxy merger as the cosmic time of the first simulation output where two galaxies that were previously identified as separate individually bound objects are now identified as a single bound object by the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009). There are 200 simulation outputs between redshifts $z = 20$ and $z = 0$ at intervals of 40 to 80 Myr. Mergers are classified by the stellar mass ratio, $\mu = M_{*,1}/M_{*,2}$, where $M_{*,2}$ is the stellar mass of the most massive member of the binary. They are considered major if $\mu \geq \frac{1}{4}$, minor if $\frac{1}{10} \leq \mu < \frac{1}{4}$ and either major or minor if $\mu \geq \frac{1}{10}$. To account for the effect of stellar stripping during the later stages of the interaction, the stellar masses are computed when the in-falling galaxy had its maximum mass (e.g., Rodriguez-Gomez et al., 2015; Qu et al., 2017). To account for the resolution of the simulation, mergers are only considered 'resolved' when $M_{*,2} \geq 10^8$ M$_\odot$ (≈ 100 stellar particles).

3.2.1 The phases of black hole growth

BHs in the EAGLE simulation transition through three distinct phases of growth, governed by the mass (or more strictly the virial temperature) of the host dark matter halo. As we will repeatedly use the terminology adopted by previous studies, we briefly revisit their meaning here. For a more comprehensive description of these phases and how they affect the observable properties of galaxies and their central

BHs see Chapter 4, for a physical interpretation of these phases see Bower et al. (2017) (see also Dubois et al., 2015; Anglés-Alcázar et al., 2017 for related, but different, interpretations).

1. *The stellar feedback regulated phase*: the buoyant outflows created by stellar feedback efficiently regulate the gas content of galaxies residing in low-mass haloes ($M_{200} \ll 10^{12}$ M$_\odot$). As a consequence, the central density of gas in these systems remains low, resulting in only limited growth of the central BH. In this phase BHs tend to remain close to the seed mass[1].

2. *The non-linear/rapid black hole growth phase*: as haloes evolve towards M_{200} $\sim 10^{12}$ M$_\odot$ the virial temperature of the halo surpasses that of the stellar outflow, causing them to stall (as they can no longer buoyantly rise). This gives the first opportunity for a high gas density to build up in the galaxy centre. Now the central BH is able to grow nearly unhindered, doing so initially at a highly non-linear rate, arising since Bondi-like accretion is proportional to the mass of the BH squared (Bondi & Hoyle, 1944). We will interchangeably refer to this phase of evolution as either the 'non-linear' or 'rapid growth' phase.

3. *The AGN feedback regulated phase*: after the rapid growth phase, the central BH has become massive ($\gtrsim 10^7$ M$_\odot$). It can now effectively regulate the gas inflow onto the halo via efficient AGN feedback. Therefore in massive haloes ($M_{200} \gtrsim 10^{12}$ M$_\odot$) regulatory equilibrium is once again restored, and the specific growth of the BH retires to a lower rate.

3.2.2 Black hole sample selection

Our sample comprises all BHs more massive than 10^7 M$_\odot$ at $z = 0$. We only consider BHs more massive than this as they have likely completed the non-linear phase and

[1] $M_{\rm BH[seed]} = 1.48 \times 10^5$ M$_\odot$ for the reference model.

will have entered the AGN feedback regulated phase. This ensures that the three phases of growth outlined in Section 3.2.1 can be robustly identified. A lower mass cut would contaminate the sample with a large number of BHs still undergoing the non-linear phase. We estimate this mass cut via an inspection of the BH mass–halo mass relation (see Figure 4.2), selecting the pivot point that marks the transition from a supra-linear to $\approx$ linear relation between the two properties. This yields a total sample of 1,888 BHs.

Identifying the non-linear phase of black hole growth

To segregate the BHs within our sample into the three evolutionary phases outlined in Section 3.2.1, we require a robust identification of the beginning and end of the non-linear phase. BHs enter the non-linear growth phase at $\approx$ the seed mass, as growth is curtailed in the preceding stellar feedback regulated phase (see Chapter 4). The specific black hole accretion rate (sBHAR[2], the accretion rate of the BH normalised by the BH mass, i.e., $\dot{M}_{\rm BH}/M_{\rm BH}$) during the non-linear phase is naturally large, due to the high $\dot{M}_{\rm BH}$ and the relatively low $M_{\rm BH}$ over this period. Therefore, to first order, the peak of the sBHAR history provides a good estimate for when the non-linear growth phase is occurring. We then estimate the extent of the non-linear phase by tracing the $\log_{10}M_{\rm BH}$ history in each direction, starting from the sBHAR peak. When the gradient, $\mathrm{d}(\log_{10}M_{\rm BH})/\mathrm{d}t$, shallows below a critical value, we take these thresholds to be the start and end points of non-linear growth, $t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$ respectively. We find the value $\mathrm{d}(\log_{10}\,M_{\rm BH})/\mathrm{d}t = 0.25$ dex Gyr^{-1} provides a robust separation of the three phases for our BH sample; however the results are insensitive to the choice of this value.

In Figure 3.1 we illustrate these steps for two randomly selected BHs (one represented by a solid line in each panel and the other by a dashed line in each panel). The top panel shows the 50 Myr time-averaged sBHAR history, highlighting our starting point,

[2]As instantaneous BH activity is highly variable (see Figure 4.1), the value of $\dot{M}_{\rm BH}$ used in all our sBHAR calculations is the 50 Myr time-averaged rate.

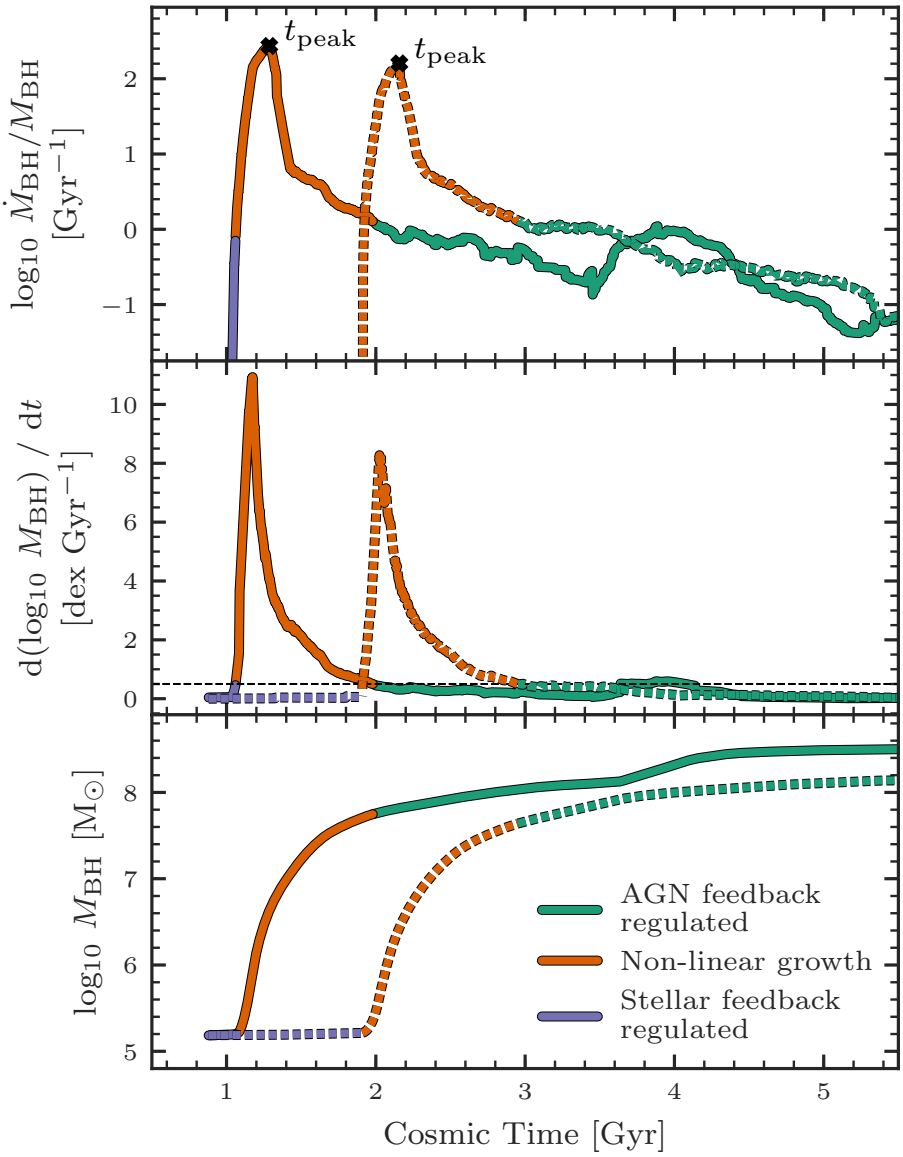

Figure 3.1: Two illustrative examples of how the start and end points ($t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$ respectively) of the non-linear phase of BH growth are computed. Each panel is plotted as a function of cosmic time. *Top panel*: the 50 Myr time-averaged sBHAR ($\dot{M}_{\rm BH}/M_{\rm BH}$), annotated with the maximum value, $t_{\rm peak}$. *Middle panel*: the gradient of $\log_{10}M_{\rm BH}$, d($\log_{10}M_{\rm BH}$)/dt. Where the gradient crosses the threshold value of d($\log_{10}M_{\rm BH}$)/dt = 0.25 dex Gyr^{-1} (shown as a horizontal dashed line) in each direction, starting from $t_{\rm peak}$, defines the times $t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$. *Bottom panel*: The BH mass. Each line is colour coded via the identified phase of BH evolution, as indicated by the legend.

the maximum value, $t_{\rm peak}$. The middle panel shows the gradient of the $\log_{10}M_{\rm BH}$ history, highlighting our threshold value of $\mathrm{d}(\log_{10}M_{\rm BH})/\mathrm{d}t = 0.25$ dex Gyr^{-1} as a horizontal dashed line. Where the histories first intersect with this threshold both backwards and forwards from the value $t_{\rm peak}$, defines $t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$ respectively. Finally, the bottom panel shows the BH mass history. Each line is colour coded via the identified phase of evolution: purple lines represent the stellar feedback regulated phase ($t < t_{\rm NLG[start]}$), orange lines the non-linear growth phase ($t_{\rm NLG[start]} \leq t \leq t_{\rm NLG[end]}$) and green lines the AGN feedback regulated phase ($t > t_{\rm NLG[end]}$).

3.2.3 Defining $n_{\rm dyn}$: the most proximate merger

To aid in establishing galaxy–galaxy mergers as potential triggering mechanisms for the non-linear phase in Section 3.3.3, we introduce $n_{\rm dyn}$, defined as the number of dynamical times between the start of the non-linear growth phase and the completion time of the most proximate (i.e., closest in time) merger, i.e.,

$$n_{\rm dyn} = \frac{t_{\rm NLG[start]} - t_{\rm merger}}{t_{\rm dyn}}, \tag{3.2.1}$$

where $t_{\rm NLG[start]}$ is the onset time of non-linear growth defined in Section 3.2.2, $t_{\rm merger}$ is the completion time of the most proximate host galaxy merger and $t_{\rm dyn}$ is the dynamical time. We define the the dynamical time as the free-fall time of the dark matter halo, i.e.,

$$t_{\rm dyn} \equiv \left(\frac{3\pi}{32G(200\rho_{\rm crit})}\right)^{1/2}, \tag{3.2.2}$$

where $\rho_{\rm crit}$ is the critical density of the Universe at $t_{\rm NLG[start]}$. For reference, $t_{\rm dyn} \approx 1.6$ Gyr at $z = 0$, ≈ 0.5 Gyr at $z = 2$ and ≈ 0.2 Gyr at $z = 5$. Thus negative (positive) values of $n_{\rm dyn}$ indicate that the most proximate merger completed after (before) the rapid growth phase began. We compute $n_{\rm dyn}$ separately for the most proximate

major merger ($t_{\rm merger}(\mu \geq \frac{1}{4})$, denoted $n_{\rm dyn[maj]}$), minor merger ($t_{\rm merger}(\frac{1}{10} \leq \mu < \frac{1}{4})$, denoted $n_{\rm dyn[min]}$) and either a major or minor merger ($t_{\rm merger}(\mu \geq \frac{1}{10})$, denoted $n_{\rm dyn[all]}$).

High values of $n_{\rm dyn}$ are capped to ± 10 dynamical times as mergers with $|n_{\rm dyn}| > 10$ are unlikely to have had an influence on the non-linear period. The BHs hosted in galaxies that did not experience any merger of a particular classification throughout their lifetime (and therefore have no valid value of $t_{\rm merger}$) are assigned the value $n_{\rm dyn} = 10$ to still contribute to the normalisation of the merger rate.

Creating a control sample of $n_{\rm dyn}$

To ascertain the significance of mergers in proximity to the non-linear phase, we require a control sample. Therefore for each BH's value of $n_{\rm dyn[maj]}$, $n_{\rm dyn[min]}$ and $n_{\rm dyn[all]}$ we construct ten associated control values. These are obtained by recomputing $n_{\rm dyn[maj]}$, $n_{\rm dyn[min]}$ and $n_{\rm dyn[all]}$ in ten random control galaxies using the $t_{\rm NLG[start]}$ value of the source galaxy (overriding the native value of $t_{\rm NLG[start]}$ in the control galaxies). The control galaxies are selected only on stellar mass (required to be within ± 0.5 dex of the source galaxy) and redshift, and therefore yield the expectation values of $n_{\rm dyn[maj]}$, $n_{\rm dyn[min]}$ and $n_{\rm dyn[all]}$ that would be obtained for a galaxy of that mass, at that epoch, solely from the background merger rate, with no regard to the activity of the BH. For any collection of $n_{\rm dyn}$ values, such as the distributions in Figure 3.6, we combine their associated control values to create ten control samples. Any deviations from the $n_{\rm dyn}$ distributions of the controls indicates the relative prevalence of mergers around the rapid growth phase over the background rate.

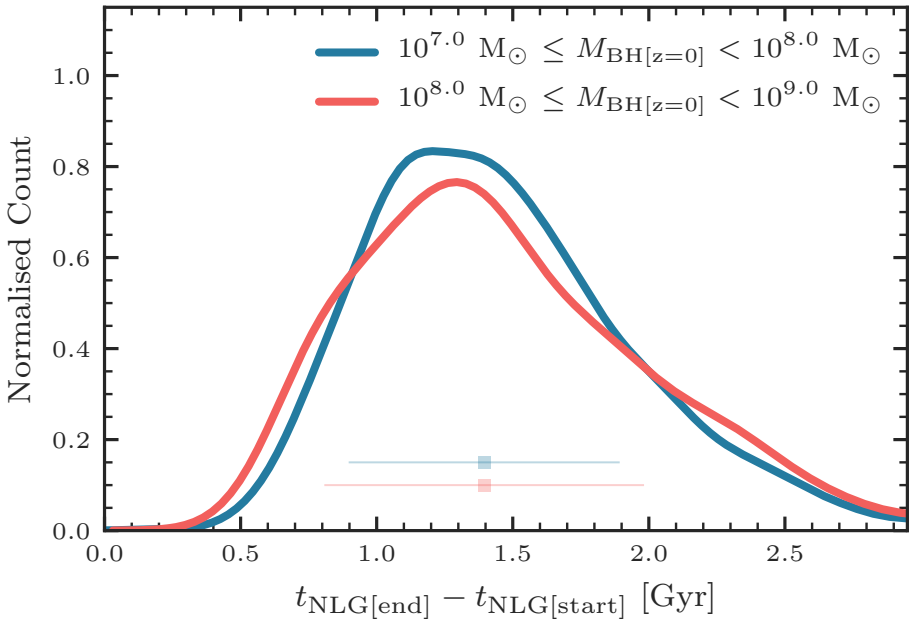

Figure 3.2: The distribution of non-linear growth durations (i.e., $t_{\mathrm{NLG[end]}} - t_{\mathrm{NLG[start]}}$) for the BHs within our sample, separated into two present day BH mass ranges: 10^7 M$_\odot$ $\leq M_{\mathrm{BH[z=0]}} < 10^8$ M$_\odot$ (red line) and 10^8 M$_\odot$ $\leq M_{\mathrm{BH[z=0]}} < 10^9$ M$_\odot$ (blue line). The median values and 10^{th}–90^{th} percentile ranges for each distribution are indicated by error bars ($1.4^{+0.6}_{-0.9}$ Gyr and $1.4^{+0.5}_{-0.7}$ for the upper and lower BH mass ranges, respectively). The median period of time BHs spend within the non-linear phase is insensitive to the eventual mass of the BH over this range.

3.3 Results

3.3.1 Properties of the black holes

We begin with investigating the properties of the BHs within our sample in relation to their rapid growth phase. Figure 3.2 shows the distribution of the non-linear phase durations (i.e., $t_{\rm NLG[end]} - t_{\rm NLG[start]}$), separated into two present day BH mass ranges: 10^7 M$_\odot$ $\leq M_{\rm BH[z=0]} < 10^8$ M$_\odot$ (red line) and 10^8 M$_\odot$ $\leq M_{\rm BH[z=0]} < 10^9$ M$_\odot$ (blue line). Both distributions are relatively narrow and broadly symmetric in their shape. The median duration of the rapid growth phase for the upper and lower present day BH mass ranges are almost identical ($1.4^{+0.6}_{-0.9}$ Gyr and $1.4^{+0.5}_{-0.7}$ respectively, the error values outline the $10^{\rm th}$–$90^{\rm th}$ percentile ranges). Therefore the median period of time spent within the non-linear phase is insensitive to the eventual BH mass over this range.

Further properties of the rapid growth phase are investigated in Figure 3.3. Here we show, from top to bottom, the onset redshift of the non-linear phase, the fraction of the BHs lifetime that was spent in the three evolutionary phases and the fraction of the total final BH mass that was accumulated, via both mergers and accretion, in the three evolutionary phases, each as a function of the final BH mass.

Starting with the top panel, we find today's most massive BHs began their non-linear phase, on average, the earliest ($z \approx 2$ for $M_{\rm BH[z=0]} = 10^7$ M$_\odot$ increasing to $z \approx 6$ for $M_{\rm BH[z=0]} = 10^9$ M$_\odot$). This result is expected, as these BHs, which are hosted by some of the most massive haloes today (see Chapter 4), will tend to have reached the critical halo virial temperature for non-linear growth at earlier epochs than their lower mass counterparts. The fraction of a BHs lifetime spent in the rapid growth phase is low, and relatively constant for all the BHs within our sample ($\approx 15\%$, see middle panel). Most of the duration of massive BH life is spent in the AGN feedback regulated phase (between ≈ 60 and 90% of their lifetimes). The fraction of the total BH mass that is accumulated in the non-linear phase is not constant; it

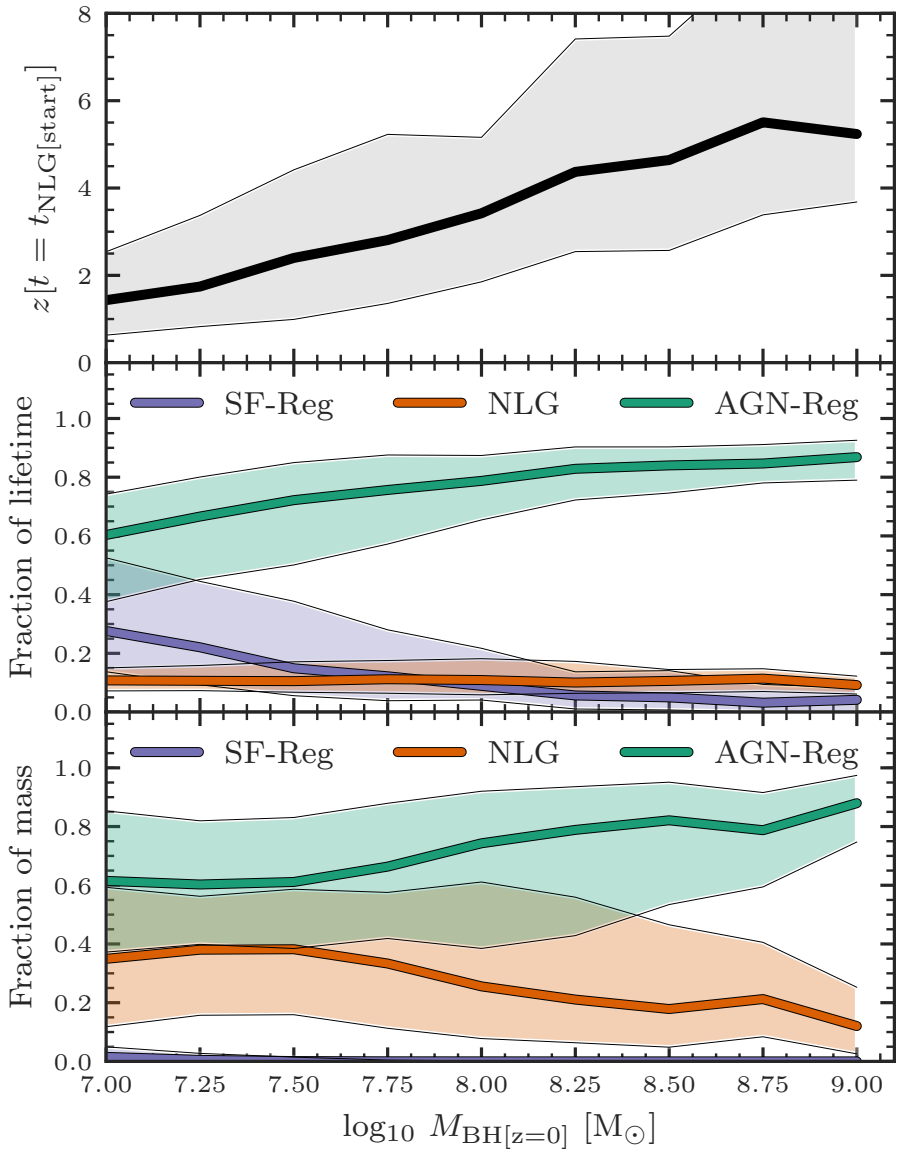

Figure 3.3: Properties of the BHs within our sample in relation to their rapid growth phase. The solid lines are the median values and the shaded regions outline the 10^{th}–90^{th} percentile ranges. Each property is plotted as a function of the final BH mass. *Top panel*: the onset redshift of the rapid growth phase. *Middle panel*: the fraction of the BHs lifetime that was spent in the three evolutionary phases. *Bottom panel*: the fraction the total BH mass that was accumulated, via both mergers and accretion, in the three evolutionary phases.

accounts for $\approx 30\%$ of the final mass for $M_{\rm BH[z=0]} = 10^7$ M$_\odot$ and decreases to $\approx 5\%$ for $M_{\rm BH[z=0]} = 10^9$ M$_\odot$ (see bottom panel). Regardless of the time BHs spent in the stellar feedback regulated phase, which is only non-negligible for the lowest-mass BHs we study, almost no mass is accumulated, due to the quenching of BH growth via efficient stellar feedback.

Therefore, the earlier BHs undergo their non-linear growth phase, the less contribution this phase has to the present day mass. Regardless of when this phase begins, it is generally short lived relative to the lifetime of the BH.

Black hole activity during the rapid growth phase

The accretion activity of the BHs within our sample during their rapid growth phase is investigated in Figure 3.4. For each BH, we divide the non-linear phase into four equal time segments[3] between $t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$ and measure the mean bolometric AGN luminosity ($L_{\rm AGN}$[4], top panel) and the mean Eddington rate ($\lambda_{\rm edd}$[5], bottom panel) for each quarter. This allows us to measure the comparative trends of BH activity throughout each segment of the rapid growth phase. The BHs are separated by the redshift at which they began their non-linear phase (i.e., $z[t = t_{\rm NLG[start]}]$).

The general evolutionary trend for both the AGN luminosity and the Eddington rate through the non-linear phase is very similar for each redshift range. The AGN luminosity in the 1$^{\rm st}$ quarter initiates at a relatively low rate ($\sim 10^{42}$ erg s^{-1}), steadily increases towards the 3$^{\rm rd}$ quarter ($\sim 10^{44}$ erg s^{-1}) and remains approximately at this level through to the 4$^{\rm th}$ quarter. This behaviour is consistent with the scenario of a growing BH embedded within a relatively constant source of fuel. The Eddington

[3]Note that the absolute time intervals of the quarters will be different for each BH due to the varying range of non-linear growth durations (see Figure 3.2).

[4]Defined as $L_{\rm AGN} = \epsilon_r \dot{M}_{\rm BH} c^2$, where ϵ_r is the radiative efficiency of the accretion disk, which is assumed to be 0.1 (Shakura & Sunyaev, 1973).

[5]Defined as $\lambda_{\rm edd} = \dot{M}_{\rm BH}/\dot{M}_{\rm edd}$ where $\dot{M}_{\rm BH}$ is the accretion rate of the BH and $\dot{M}_{\rm edd}$ is the Eddington limit. The BH accretion rate in the EAGLE reference model is capped to the Eddington limit over h (i.e., the maximum allowed value of $\lambda_{\rm edd} = 1/h = 1.48$).

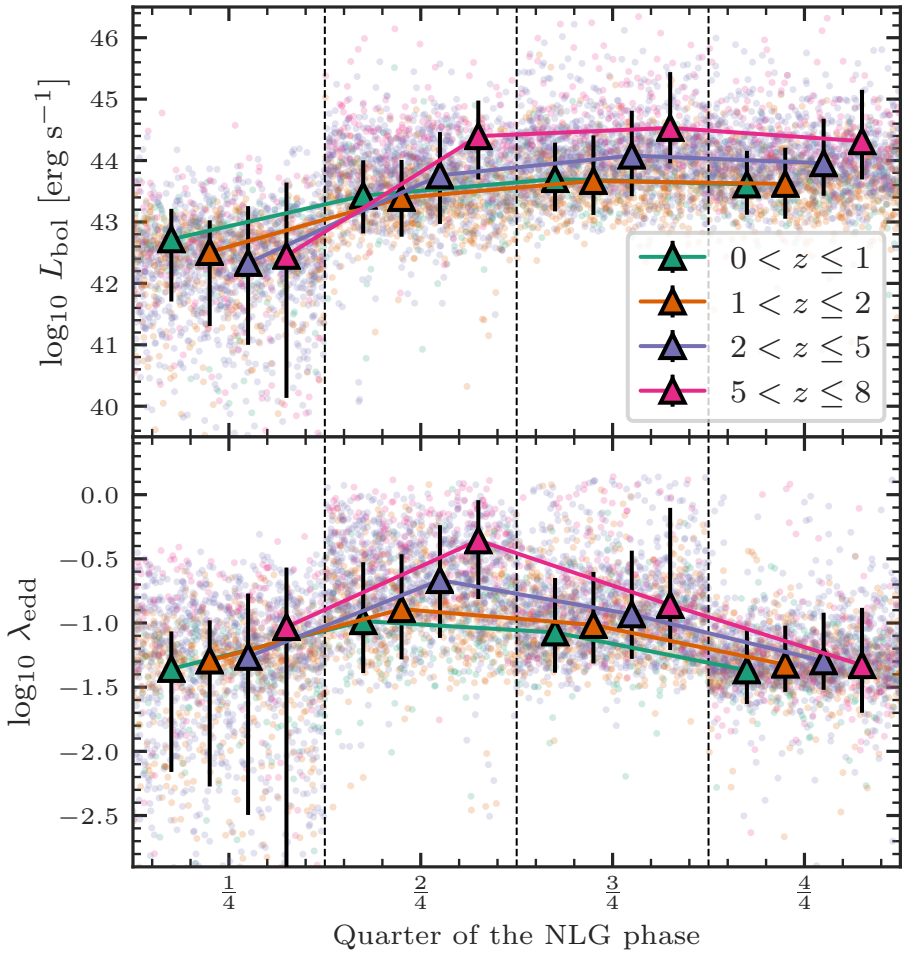

Figure 3.4: The accretion activity of the BHs within our sample during their rapid growth phase. For each BH, the non-linear phase is divided into four equal time segments between $t_{\rm NLG[start]}$ and $t_{\rm NLG[end]}$, and the mean AGN luminosity (top panel) and the mean Eddington rate (bottom panel) is computed for each quarter. The solid circles are the mean values for each individual BH at a given quarter, coloured by the redshift at which they started their non-linear phase (i.e., $z[t = t_{\rm NLG[start]}]$), as indicated by the legend. We assign each BH a random scatter along the x-axis of each quarter bin, for clarity. The solid triangles indicate the median values of the four bins, with the error bars outlining the $10^{\rm th}$–$90^{\rm th}$ percentile range. The median values are offset from each other along the x-axis, for clarity. The bolometric luminosity increases from the beginning to the end of the non-linear phase. The Eddington rate peaks at approximately 50% of the way through the rapid growth phase. These trends are epoch independent, however at higher redshift the mean values increase in each property.

rate similarly begins at a relatively low level in the 1st quarter ($\lambda_{\rm edd} \sim 10^{-1.5}$), evolves towards a peak in the 2nd and 3rd quarters ($\lambda_{\rm edd} \sim 10^{-0.5}$), and finally reduces back to values similar to that of the 1st quarter. This remains consistent with the picture seen in the panel above: the AGN luminosity of a growing BH tends to a constant rate in the later states of non-linear growth. For each of the two properties, the mean values increase with increasing redshift, indicating that the BHs that underwent their rapid growth phase at higher redshift are on average more luminous and closer to the Eddington limit than their counterparts at lower redshift. If we examine the individual mean Eddington rate values (background coloured circles), we find that it is extremely rare to sustain continued growth at the Eddington limit for any period during the non-linear phase.

3.3.2 Properties of the hosts at the start of the rapid growth phase

We now turn to the galaxies and dark matter haloes that host the BHs within our sample at the onset of their rapid growth phase. Figure 3.5 shows, from left to right, the galaxy stellar mass, halo mass and halo virial temperature[6], each at the time $t = t_{\rm NLG[start]}$, as a function of the redshift at which the rapid growth phase began. There is a distinct negative trend visible in the first two panels, with both the host galaxy and halo mass decreasing as the redshift increases ($M_* \approx 10^{10.5}$ M$_\odot$ at $z \approx 0$ decreasing to $M_* \approx 10^{9}$ M$_\odot$ at $z \approx 6$ and $M_{200} \approx 10^{12.4}$ M$_\odot$ at $z \approx 0$ decreasing to $M_{200} \approx 10^{11.2}$ M$_\odot$ at $z \approx 6$). There appears, therefore, to be no fixed galaxy or halo mass at which non-linear BH growth initiates, instead, the rapid growth phase of BHs appears to ubiquitously initiate when the host halo reaches a critical virial temperature ($T_{\rm vir} \approx 10^{5.6}$ K, see right panel). This is consistent with the physical scenario outlined in Section 3.2.1, whereby the buoyancy of the stellar feedback

[6]The virial temperature of the halo is defined as $T_{\rm vir} = \mu m_{\rm p} V_{\rm c}^2 / 5 k_{\rm b}$, where μ is the mean molecular weight of the gas in the halo (assumed to be 0.59 for a primordial gas), $m_{\rm p}$ is the mass of the proton, $k_{\rm b}$ is the Boltzman constant and $V_{\rm c} = GM_{200}/r_{200}$ is the virial velocity (**?**).

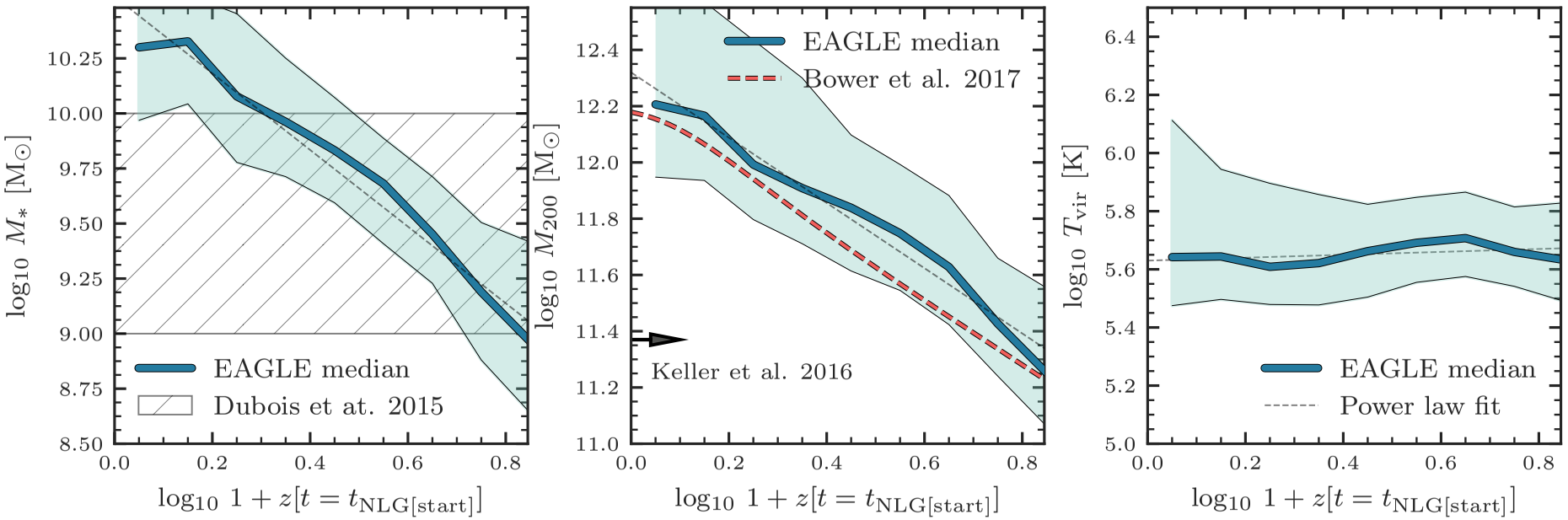

Figure 3.5: The galaxy stellar mass (left panel), halo mass (middle panel) and halo virial temperature (right panel) of the hosts of the BHs within our sample at the beginning of their rapid growth phase ($t = t_{\mathrm{NLG[start]}}$) as a function of the redshift at which their rapid growth began. The solid lines indicate the median values, with the shaded regions outlining the 10^{th}–90^{th} percentile ranges. Single power law fits to the median trends are indicated by dashed black lines. The BHs starting their rapid growth phase at low redshift do so in haloes and galaxies approximately an order of magnitude more massive than their high-redshift counterparts, indicating that there is no fixed halo or galaxy mass at which the rapid growth phase initiates, instead, BHs enter their rapid growth phase at a $\approx$ constant critical halo virial temperature ($T_{\mathrm{vir}} \approx 10^{5.6}$ K). Included in the left and middle panels are three theoretical predictions for the stellar/halo mass(es) at which stellar feedback can no longer efficiently regulate the gas content of the galaxy, and stalls, marking the theoretical transition point to the non-linear phase of BH growth, see Section 3.4.1 for a discussion.

driven outflows stall at a critical halo virial temperature, allowing the gas density within the centre of the galaxy to rise, triggering the rapid growth phase.

We fit each of the median trends with a single power law relation using the python module LMFIT[7], indicated on the figure as dashed black lines. The 1σ errors on the median values inserted into the fitting routine are computed from bootstrap resampling. The fits are: $\log_{10}(M_*/M_\odot) = (-1.74 \pm 0.11)\log_{10}(1+z) + 10.53 \pm 0.06$, $\log_{10}(M_{200}/M_\odot) = (-1.16 \pm 0.07)\log_{10}(1+z) + 12.32 \pm 0.04$ and $\log_{10}(T_{\rm vir}/K) = (0.05 \pm 0.04)\log_{10}(1+z) + 5.63 \pm 0.02$, from the left to right panels, respectively.

3.3.3 The proximity of mergers to the rapid growth phase

We conclude this section by investigating the physical connection between the start of the non-linear phase of BH growth and galaxy mergers. Figure 3.6 shows the distributions of $n_{\rm dyn[maj]}$ (green lines), $n_{\rm dyn[min]}$ (orange lines) and $n_{\rm dyn[all]}$ (purple lines) for each BH contained within our sample (see Section 3.2.3 for their definitions). The BHs are separated into those that began their rapid growth phase in the redshift ranges $0.0 \leq z < 1.0$, $1.0 \leq z < 2.0$ and $2.0 \leq z < 4.0$[8], shown in the top, middle and bottom panels, respectively. These distributions reveal the characteristic proximity in time between galaxy–galaxy mergers of the host and the onset of the rapid growth phase of the central BH.

Starting with the top panel, we find that the distribution of $n_{\rm dyn[all]}$ (purple line) strongly peaks just before the value $n_{\rm dyn[all]} = 0$ (indicated by a vertical black line). The abundance of quantitatively low values of $|n_{\rm dyn[all]}|$ indicates that for these BHs there is often either a major or minor merger during this phase of their evolution. Additionally, the preference for negative values tells us that the most proximate merger generally completes *after* the non-linear phase has begun. If we were to

[7]https://lmfit.github.io/lmfit-py/

[8]We note that whilst there are galaxies that begin their non-linear phase at $z > 4$ (see Figure 3.3), we limit our merger analysis to $z < 4$ to ensure we adequately resolve minor mergers ($M_{*,2} \geq 10^8$ $M_\odot$, see Section 3.2) for all galaxies, as the host galaxies of the BHs beginning their rapid growth at $z < 4$ have masses $M_* \geq 10^9$ $M_\odot$ (see Figure 3.5).

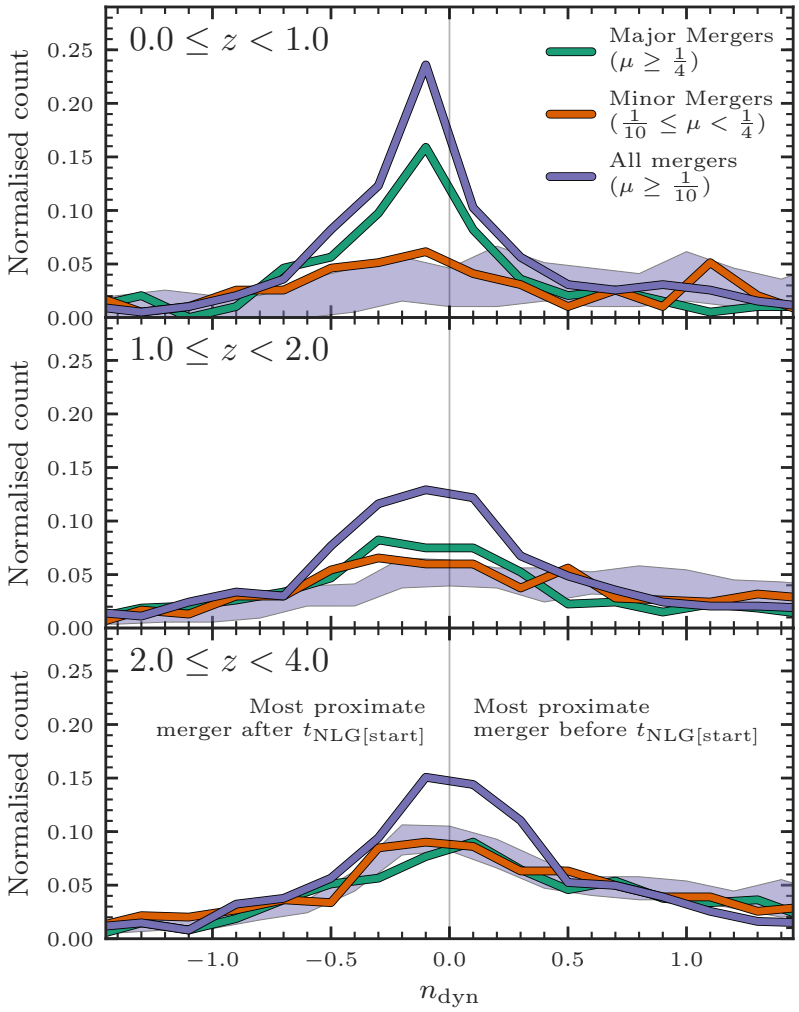

Figure 3.6: The distributions of $n_{\mathrm{dyn[maj]}}$ (green lines), $n_{\mathrm{dyn[min]}}$ (orange lines) and $n_{\mathrm{dyn[all]}}$ (purple lines) for each BH contained within our sample (the number of dynamical times since the most proximate in time merger, see Section 3.2.3 for definitions). The BHs are separated into those that began their rapid growth phase in the redshift ranges $0.0 \leq z < 1.0$ (top panel), $1.0 \leq z < 2.0$ (middle panel) and $2.0 \leq z < 4.0$ (bottom panel). The shaded regions outline the 10^{th}–90^{th} percentile range of the control distributions for $n_{\mathrm{dyn[all]}}$ (see Section 3.2.3). These reveal the predicted distribution of $n_{\mathrm{dyn[all]}}$ that would be produced solely from the background merger rate. The distributions are normalised by the total number of BHs in that redshift range, including those with host galaxies that experienced no mergers of the particular classification in their lifetimes (see Section 3.2.3). The significant peak just before the value $n_{\mathrm{dyn[all]}} = 0$ in the upper panel, relative to the control, shows that mergers commonly trigger this phase of BH evolution at low redshift (almost exclusively from major mergers). At higher redshifts the peak lowers and the distribution broadens, with the distributions falling closer to that of the control sample.

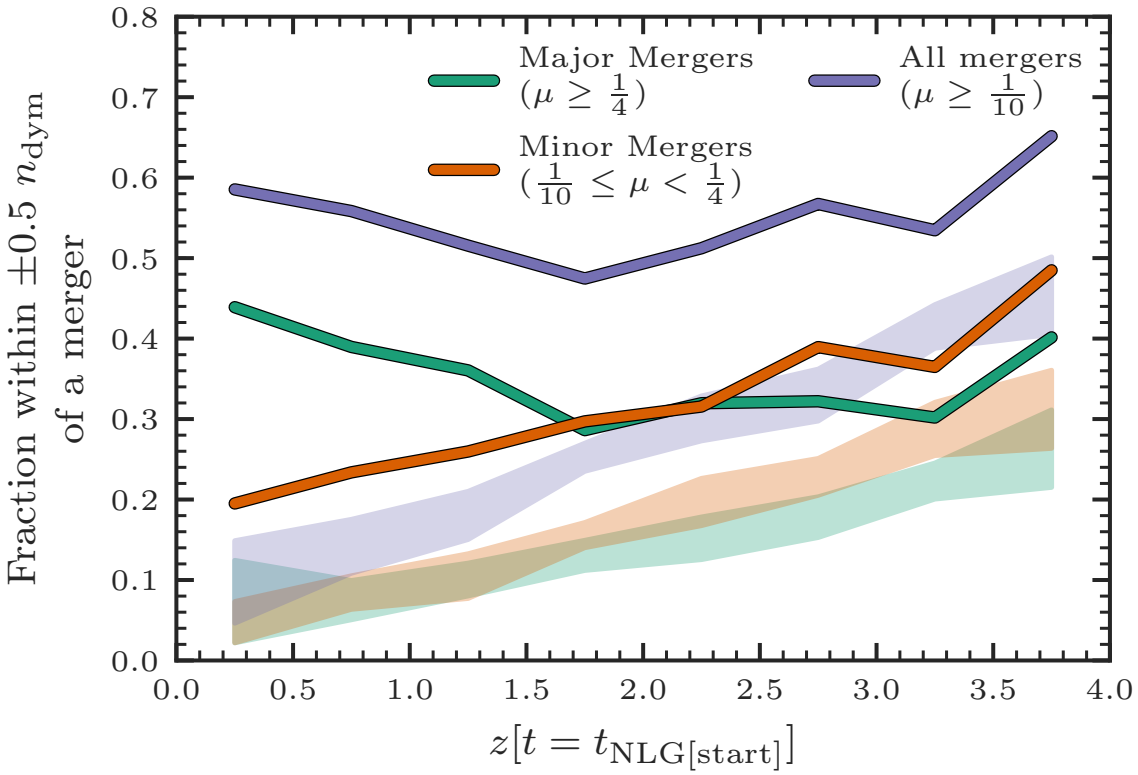

Figure 3.7: The fraction of BHs within our sample that began their rapid growth phase within ±0.5 dynamical times of a major merger (green line), minor merger (orange line) and either a minor or major merger (purple line) as a function of the redshift at which the rapid growth phase began. The fractions that would be expected from the background merger rate for similar mass galaxies (with no regard for BH activity) are shown as shaded regions. A substantial excess of BHs at low redshift are more proximate in time to a merger than the control prediction. Therefore mergers, almost exclusively major mergers, are triggering the rapid growth phase for a large fraction of the BHs at lower redshift.

attribute the most proximate merger as the triggering mechanism, it would indicate that the rapid growth phase initiates during the initial period of the interaction and well before the final coalescence of the two galaxies. If we consider minor and major mergers independently (orange and green lines), we find that most of the peak counts for all mergers is contributed by major mergers, rather than minor mergers. As we move to higher redshifts, in the middle and bottom panels, we find the distribution broadens and the peak lowers and shifts closer to the value $n_{\rm dyn} \approx 0$.

However, it is difficult to attribute any significance to these peaks without also knowing the expected distribution of $n_{\rm dyn[min]}$, $n_{\rm dyn[maj]}$ and $n_{\rm dyn[all]}$ that would arise just from the background merger rate, regardless of BH activity. For this we additionally show the $10^{\rm th}$–$90^{\rm th}$ percentile range of the ten control samples (see Section 3.2.3) for $n_{\rm dyn[all]}$ as a shaded region in each panel. For clarity, we exclude the control samples for the remaining two merger classifications from this figure, but note that they are indistinguishable from the control distribution that is plotted. Relative to the control, it is clear that the enhancement around the value $n_{\rm dyn[all]} \approx 0$ is a significant deviation from what is expected from the background rate, particularly at low redshift.

To measure this enhancement more clearly, we present Figure 3.7. This shows the fraction of BHs that began their non-linear phase within ± 0.5 dynamical times of a merger as a function of the redshift at which the non-linear phase began for the same three merger classifications. We additionally show the predicted fractions from our control sets as shaded regions. The behaviour first hinted towards in Figure 3.6 is now much more apparent. There is always an excess above the control, indicating that mergers are more common around the start of the rapid growth phase than one would predict from the background rate. At low redshift ($z \approx 0$) the excess is substantial; $\approx 60\%$ of the BHs starting their rapid growth phase at this time are found to be within ± 0.5 dynamical times of either a minor or major merger, when only $\approx 10\%$ would be expected to be so from the background rate. It therefore appears that mergers, primarily major mergers, are strong drivers of the rapid growth

phase for many BHs at low redshift. We discuss this result further in Section 3.4.2.

3.4 Discussion

3.4.1 Stalling stellar feedback and the transition to the rapid growth phase of black holes

Whilst a number of current hydrodynamical simulations have reported a link between efficient stellar feedback and the substantial reduction of BH growth in low-mass systems, it remains unclear exactly when, and how, the transition between stalling stellar feedback and the onset of rapid BH growth occurs.

Dubois et al. (2015) study the growth of an individual dark matter halo (10^{12} $M_\odot$ at $z = 2$) by means of a high-resolution cosmological zoom in, taken from the SETH simulation suite using the adaptive mesh refinement code RAMSES (Teyssier, 2002). They find that at redshifts $z > 3.5$ the galaxy's central reservoir of gas is sufficiently disrupted via efficient stellar feedback[9] to substantially restrict the accretion onto the central BH. After the galaxy has accumulated sufficient mass, they witness a decline in the ability of stellar feedback to disrupt the gas, and the central BH transitions into a rapid growth phase. They argue that this transition is directly linked to the balance between the momentum-driven stellar wind and the escape velocity of the central bulge. From this they predict the theoretical mass scale above which these winds can no longer escape the bulge, leading to a rise in the central gas density, which in turn feeds the central BH. They state the escape velocity for a bulge of mass 10^9 $M_\odot$ at a fixed bulge radius of 100 pc is $\approx$ 270 km s^{-1}, approximately equal to that achieved by a supernovae Sedov blast wave (see their Equation 1). This is indeed the bulge mass found by their simulation around which the rapid BH growth begins. The prediction that stellar feedback begins to stall ubiquitously at a

[9]This is only true when their delayed cooling prescription for stellar feedback is used (Teyssier et al., 2013). When performing similar tests with a kinetic stellar feedback model (Dubois & Teyssier, 2008), they only find a very limited effect on the growth of the central BH.

constant bulge mass and bulge radius (i.e., a constant density) implies the existence of a critical mass that is independent of epoch, contrary to our findings in Figure 3.5. We show this bulge mass (converted to a range of total stellar masses assuming a bulge to total stellar mass ratio of between 0.1 and 1.0) as a hatched region in the left panel of Figure 3.5. We note that, from the study of a single halo, one cannot capture the variation of the critical mass with time and halo properties. Indeed, the assumption of a fixed bulge density is potentially a key assumption that leads to a redshift-independent critical mass, though we know that bulges at high redshift are denser than those in the local Universe.

Keller et al. (2016) investigated the evolution of 18 isolated Milky Way-like disc galaxies from the MUGS2 simulation suite (Stinson et al., 2010), performed using the smoothed particle hydrodynamics code GASOLINE2 (Wadsley et al., 2017). They find that supernovae alone cannot regulate the incoming gas flow to systems with virial masses $> 10^{12}$ $M_{\odot}$, which can result in a runaway production of stars in the central bulge. This stalling is attributed to the shutdown of galactic winds from a deepening potential well (mass loading factors fall from a relatively constant level of $\eta \sim 10$ below the critical mass scale, to $\eta < 1$ just above). They report a redshift-independent central baryonic mass of $10^{10.0\pm0.1}$ $M_{\odot}$ and halo mass of $10^{11.37\pm0.08}$ $M_{\odot}$ at which the stellar feedback begins to stall. This halo mass is indicated by an arrow in the middle panel of Figure 3.5. Although a universal and non-evolving critical mass is again in conflict with our findings (see Figure 3.5), we note that only a moderate range of present day galaxy masses are explored in the simulation set of Keller et al. (2016) ($M_{*[z=0]}$ =0.5-20.8$\times10^{10}$ $M_{\odot}$). Furthermore, there is evidence of a varying critical halo mass even within this limited mass range (see their Figures 7 & 8). Perhaps most importantly, as no prescription for BHs is included for these simulations, they are unable to directly investigate the link between stalling stellar feedback and the rapid growth phase. The runaway production of stars seen in systems above this critical mass, however, strongly suggests that AGN feedback (and thus a massive BH) is required to curb continued galaxy growth.

Anglés-Alcázar et al. (2017) perform a set of high-resolution cosmological hydrodynamic simulations of quasar-mass halos ($M_{\rm halo}(z = 2) \approx 10^{12.5}$ M$_\odot$) using the FIRE-2 simulation code (**?**). These simulations model stellar feedback by supernovae, stellar winds, and radiation, and BH growth using a gravitational torque-based prescription (see also Section 3.4.4), however no AGN feedback is implemented. They discover that early BH growth in low mass galaxies is extremely limited by bursty stellar feedback continuously evacuating gas from the galactic nuclei ($\lesssim$ 100 pc). The BHs during this time remain significantly undermassive relative to their galaxy host. Only as the galaxy approaches $M_{\rm bulge} \sim 10^{10}$ M$_\odot$ does BH growth start to become more efficient, as the nuclear stellar potential begins to retain a significant gas reservoir, and the star formation becomes less bursty. In this more massive regime, the BHs are then seen to rapidly converge onto the $M_{\rm BH}$–$M_{\rm bulge}$ scaling relation. Analogous to Dubois et al. (2015), they attribute this transition to the increased escape velocity of the bulge now exceeding that of the stellar feedback-driven winds, and also suggest the possibility of a redshift independent critical mass.

Bower et al. (2017) provide a different explanation. They develop a simple analytical model that describes the interaction between buoyant, high entropy star formation driven outflows and the rate of the cosmic gas inflow. In low mass systems ($M_{200} \lesssim 10^{12}$ M$_\odot$) the adiabat of this outflow exceeds that of the haloes diffuse corona, and can buoyantly escape. This ensures that the central gas densities within the galaxy remains low, and the central BH is deprived fuel. In massive systems a hot corona forms, and the star formation-driven outflows are no longer buoyant relative to their surroundings. This triggers a high density build up of gas within the central regions of the galaxy, and a subsequent non-linear response from the central BH. The critical halo mass predicted for this transition is given by their Equation 5, which we show in the middle panel of Figure 3.5 as a red dashed line. There is a good agreement between the analytical prediction and that of our findings, reproducing the redshift dependence, with only a small offset in the normalisation between the two trends. We note, that whilst the model of Bower et al. (2017) was validated against the EAGLE

simulation, it was independently derived, and not calibrated using the simulation results.

To summarise, we find that the critical galaxy/halo mass at which stellar feedback stalls and rapid BH growth begins is not constant, and decreases with increasing redshift. Instead, we find that rapid BH growth phase initiates at an approximately constant halo virial temperature (see Figure 3.5). This is contrary to some previous predictions, where an epoch-independent single critical mass has been reported. But, we understand this as limitations of these works due to a limited range of simulated parameters, or because AGN feedback was not included in these simulations.

3.4.2 The role of galaxy mergers in triggering the rapid growth phase of black holes

In the paradigms set out by the studies in the previous section, the primary factor in transitioning from efficient to inefficient stellar feedback-driven outflows is the secular evolution of the bulge/galaxy/halo. That is, when the host system becomes sufficiently massive, the stellar winds/outflows become trapped via a deepening potential well or hot corona. However, the rapid growth phase of BHs may also, or exclusively, be triggered by galaxy–galaxy interactions.

Dubois et al. (2015) found for the evolution of a single halo (discussed in the previous section) that the rapid growth phase of the central BH was likely triggered by a major merger. In Section 3.3.3 we found a strong connection between the onset time of non-linear growth ($t = t_{\mathrm{NLG[start]}}$) and the most proximate merger, regardless of the redshift at which non-linear growth began. Approximately 60% of the BHs within our sample initiated their rapid growth phase within ± 0.5 dynamical times of either a minor or major merger ($> 40\%$ a major merger, see Figure 3.7). At lower redshifts ($z \approx 0$), the merger fractions were much greater than those expected from the background merger rate ($\approx 60\%$ versus $\approx 10\%$), whereas at higher redshifts ($z \approx 4$), the merger fractions fell much closer to the background rate ($\approx 60\%$ versus $\approx 45\%$,

as the background merger rate increases with increasing redshift Rodriguez-Gomez et al., 2015; Qu et al., 2017). We could interpret this in two ways: (1) a galaxy's central BH at lower redshift increasingly *requires* a major disturbance to initiate its rapid growth phase, derived from the increasing excess in the merger fractions above the control sample, or (2) mergers are always important for triggering the rapid growth phase, derived from the universally high merger fractions, and the fact that all galaxies ubiquitously experience mergers more frequently at higher redshifts is inconsequential.

In either case, galaxy interactions appear to be important triggering mechanisms for the rapid growth phase, at least in the low redshift Universe. From this one may conclude that mergers can act as catalysts to accelerate the transition from stalling stellar feedback to the rapid growth phase, however, the relatively low spread in halo virial temperatures at which the rapid growth phase initiates would suggest that this is not the case (see right panel of Figure 3.5). It appears, then, that whilst the non-linear phase may be initiated through a strong interaction, a characteristic halo virial temperature remains essential for rapid BH growth to occur.

3.4.3 Observing the rapid growth phase of black holes

We explore the considerations needed to validate the non-linear phase in observations of the BH population in Figure 3.8. This figure shows the median bolometric AGN luminosity (top panel) and the median Eddington rate (bottom panel) for all the BHs within the EAGLE volume as a function of the host galaxy stellar mass at six different redshifts. Here we see the familiar imprint of the three phases of BH evolution: before the critical halo virial temperature BHs are effectively inactive, the luminosities and Eddington rates then increase by many orders of magnitude over a narrow stellar mass window around the critical halo virial temperature, and finally the luminosities and Eddington rates come to settle to an approximately constant median rate after the critical halo virial temperature, though with very

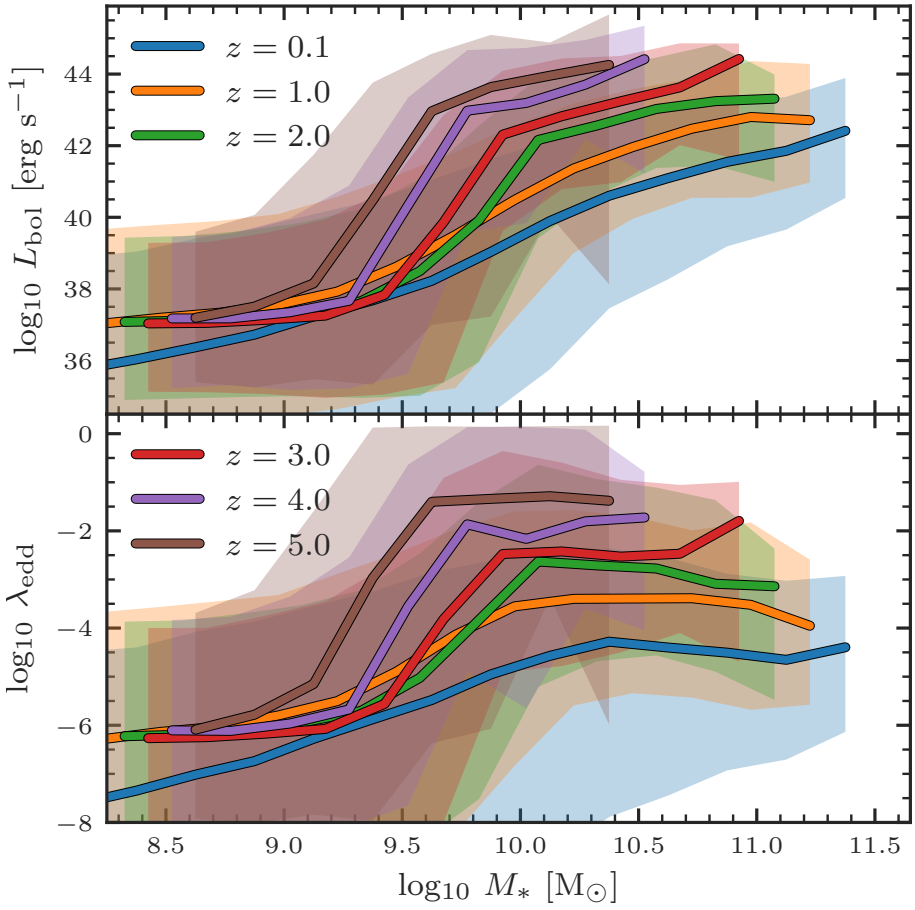

Figure 3.8: The median bolometric AGN luminosity (upper panel) and the median Eddington rate (lower panel) as a function of the host galaxy stellar mass for six redshifts, as indicated in the legend. These are computed from all galaxies at the stated epoch, and not only those hosting the BHs contained within the massive BH sample outlined in Section 3.2.2. The shaded regions outline the $10^{\rm th}$–$90^{\rm th}$ percentile range.

large scatters. As we saw in Figure 3.5, the critical *mass* marking this transition reduces with increasing redshift. Figure 3.8 also shows that the increase in AGN luminosity and Eddington rate during the non-linear phase is larger at high redshift.

One could then in principle observe evidence of the rapid growth phase in two ways: attempt to discover the transition between inactive BHs and moderately active BHs in low-mass galaxies, or find the transition between a steep and shallow relationship for the median $L_{\rm bol}$ and $\lambda_{\rm edd}$ around the critical halo virial temperature. The pivot *mass* in each case is predicted to decrease as the redshift increases. However, the spread of many orders of magnitude in the AGN luminosity (the shaded regions outline the $10^{\rm th}$–$90^{\rm th}$ percentile range), the difficulty in detecting low luminosity AGN ($L_{\rm bol} < 10^{43}$ erg s^{-1}), the relatively narrow range and therefore the need for accurate measurements of the stellar masses, and the need for large statistical samples of objects at multiple epochs will make this extremely challenging. It is therefore more plausible to find evidence for the rapid growth phase *indirectly* via the integrated BH accretion rate, i.e., the BH mass, as the three phases of BH evolution are also present within the BH mass–stellar mass relation (e.g., Crain et al., 2015; Schaye et al., 2015; Barber et al., 2016; Rosas-Guevara et al., 2016; Bower et al., 2017; McAlpine et al., 2017 for the case of EAGLE). The scatter in this relation is also predicted to change considerably with the mass of the galaxy host: galaxies below the critical halo virial temperature will host BHs with a small scatter around the seed mass; galaxies around the critical halo virial temperature will host a large dynamic range of BH masses, due to the rapid BH growth over this mass range; and BHs hosted in galaxies above the critical halo virial temperature return to a much smaller scatter due to the regulation from AGN feedback. Indeed, changing relationships between the mass of the galaxy host and that of the central BH across a range of stellar masses and morphologies have been found by empirical studies (e.g., Scott et al., 2013; Greene et al., 2016; Läsker et al., 2016; Martín-Navarro & Mezcua, 2018).

3.4.4 The dependence on the model

Three astrophysical prescriptions are crucial for forming the three phases of BH evolution investigated by this study: efficient stellar and AGN feedback, capable of regulating the gas inflow onto low- and high-mass galaxies, respectively, and the ability for BHs to grow rapidly when neither of these feedback processes are dominant. It is interesting to ask, then, to what extent the models that govern these processes influence the behaviour of BH growth in hydrodynamical simulations, and how ubiquitous the creation of these three phases may be.

Efficient stellar feedback using many different model implementations across a range of resolutions is found to restrict the growth of BHs within low-mass galaxies in hydrodynamical simulations (see Section 3.4.1). Interestingly, the ILLUSTRIS project (Vogelsberger et al., 2014), which is a cosmological hydrodynamical simulation that shares many similarities with the EAGLE project, shows no strong evidence of such behaviour (Sijacki et al., 2015). Unlike EAGLE, which models stellar feedback purely via the thermal injection of energy, ILLUSTRIS adopt a kinetic wind model that temporarily decouples the hydrodynamics. Kinetic injection schemes can be less efficient at disrupting early BH growth (see Dubois et al., 2015 for example). However, in the updated ILLUSTRIS-TNG model (Pillepich et al., 2018), where stellar feedback is now implemented partially thermally with a deliberate increased efficiency towards higher redshifts and in low-mass haloes, BH growth now appears limited in the familiar fashion below a critical mass (Pillepich et al., 2018; Weinberger et al., 2018). This phase of BH evolution is undoubtedly sensitive to the efficiency of the chosen stellar feedback model, however, efficient stellar feedback is crucial for replicating many of the observed properties of galaxies in hydrodynamical simulations, such as their sizes and star formation rates, and many hydrodynamical simulations have converged towards implementing a form of efficient stellar feedback as a result. Observations of BH activity (or lack thereof) in low-mass galaxies may therefore provide key insight for constraining stellar feedback models.

The choice of BH growth model may also have interesting implications. Many of the widely used and successful BH growth models that have faithfully replicated many of the observed properties of BHs in the local Universe are derived from the original Bondi prescription (Bondi & Hoyle, 1944), which is also the basis for the BH growth model within the EAGLE simulation (Rosas-Guevara et al., 2015). As Bondi-like accretion is proportional to the mass of the BH squared, BHs have the opportunity to grow at a rapid, non-linear rate if the conditions are favourable, hence the origin of such a short-lived rapid growth phase found by this study (see Figure 3.2). However, there are other BH growth models with alternate dependence on the mass of the BH, such as the gravitational torque-based prescription introduced by **?**, for which the accretion rate is proportional to the mass of the BH to a much lower power ($\frac{1}{6}$). In this regime, BHs do retain the capability to 'rapidly' grow, as is shown by Anglés-Alcázar et al. (2017), however at a sub-Bondi rate. This would presumably lengthen the duration of the rapid growth phase, yet once the BH becomes sufficiently massive it would still enter the AGN feedback regulated phase, and the three phases of BH evolution would theoretically remain distinct. Additionally, BH growth models can be sensitive to the resolution and scale over which the accretion rate is estimated (see Anglés-Alcázar et al., 2017), which may also impact the result. Observational measurements of the (changing) behaviour of the $M_{\rm BH}$–M_* relation around and beyond the critical transition mass will provide useful constraints between the different BH growth models.

Ultimately, to fully disentangle the direct influence of the stellar feedback and BH growth models on the three phases of BH evolution will require a parameter exploration coupled to a similar investigation as performed in this study.

3.5 Conclusions

We have investigated the rapid growth phase of BHs using the hydrodynamical cosmological EAGLE simulation. Our main conclusions are as follows:

- **The majority of massive BH life is spent in the AGN feedback regulated phase, at ≈60–90% of their lifetime.** The median duration of the rapid growth phase is only ≈ 1.4 Gyr, corresponding to ≈ 15% of their lifetime. The fraction of the present day BH mass accumulated during the rapid growth phase decreases with increasing BH mass (≈ 30% at $M_{\rm BH[z=0]} = 10^7$ $M_\odot$, decreasing to ≈ 5% at $M_{\rm BH[z=0]} = 10^9$ $M_\odot$). The remainder is acquired during the AGN feedback regulated phase, as no significant BH growth occurs during the stellar feedback regulated phase. See Figures 3.2 and 3.3.

- **BHs enter the rapid growth phase at a critical halo virial temperature ($T_{\rm vir} \approx 10^{5.6}$ K).** There is no fixed host galaxy stellar mass or halo mass at which the rapid growth phase begins. BHs initiating their rapid growth phase today do so in galaxies and haloes approximately an order of magnitude more massive than their high-redshift counterparts ($M_* \approx 10^{10.5}$ $M_\odot$ at $z \approx 0$ decreasing to $M_* \approx 10^9$ $M_\odot$ at $z \approx 6$ and $M_{200} \approx 10^{12.4}$ $M_\odot$ at $z \approx 0$ decreasing to $M_{200} \approx 10^{11.2}$ $M_\odot$ at $z \approx 6$). See Figure 3.5.

- **Galaxy–galaxy interactions are important for triggering the rapid growth phase.** Approximately 60% of the BHs initiating their rapid growth phase today ($z \approx 0$) do so within ±0.5 dynamical times of either a minor or major galaxy–galaxy merger ($\mu \geq \frac{1}{10}$) and ≈ 40% do so within ±0.5 dynamical times of a major merger ($\mu \geq \frac{1}{4}$). This is substantially higher than what is predicted from the background merger rate (≈ 10%). At higher redshifts the merger fractions remain high (≈ 60%), however the background merger rate has also substantially increased by these epochs (≈ 45%), making it difficult to directly disentangle the importance of mergers in triggering the rapid growth phase at high redshift. Minor mergers play much less of a role in triggering the rapid growth phase at all epochs. See Figures 3.6 and 3.7.

Chapter 4

The link between galaxy and black hole growth in the EAGLE simulation

4.1 Introduction

Substantial effort has been dedicated both observationally and theoretically to identifying the link between the growth of galaxies and their central supermassive black holes (BHs). However, the nature of this relationship remains poorly understood. Indirect evidence of a causal connection has been suggested empirically based on the *integrated* properties of galaxies and their BH counterparts. For example, galaxy bulge mass ($M_{*,\mathrm{bulge}}$) and the mass of the central BH (M_{BH}) exhibit a tight, approximately linear correlation for bulge masses in excess of $M_{*,\mathrm{bulge}} \sim 10^{10}\mathrm{M}_{\odot}$ (e.g, Magorrian et al., 1998; Kormendy & Ho, 2013; McConnell & Ma, 2013; Scott et al., 2013). However, at lower bulge mass, a steeper trend has been advocated (e.g, Scott et al., 2013; Greene et al., 2016). Additionally, the cosmic star formation rate (SFR) and black hole accretion rate (BHAR) densities broadly trace one another through time (e.g, Heckman et al., 2004; Aird et al., 2010; Madau & Dickinson, 2014).

A simple interpretation for these global relationships is that the growth rates that build these properties (i.e. the SFR of the galaxy and accretion rate of the BH) are proportional throughout their evolution, thus growing the two components in concert. More complex evolutionary scenarios have also been proposed. For example, a simple time-averaged relationship built from a common fuel reservoir of cold gas (Alexander & Hickox, 2012; Hickox et al., 2014), a rapid build up of galaxy and BH mass via merger induced starburst/quasar activity (e.g, Sanders et al., 1988; Di Matteo et al., 2005; Hopkins et al., 2008) or a mutual dependence on the mass or potential of the dark matter halo (Booth & Schaye, 2010, 2011; Bower et al., 2017). In these scenarios the SFR and BHAR do not necessarily trace each other directly and the observed correlations may only appear in massive galaxies due to an averaging of very different histories. Furthermore, Peng (2007) and Jahnke & Macciò (2011) go as far as to suggest there is no causal connection of any kind, with correlations only appearing as result of a random walk.

To test these scenarios, numerous observational studies have attempted to identify a direct link between the intrinsic growth rates of galaxies and their central BHs. Studies that investigate the mean SFR ($\langle \mathrm{SFR} \rangle$) as a function of BHAR consistently find no evidence for a correlation for moderate-luminosity sources ($L_{2-8\mathrm{keV}} \lesssim 10^{44}$ erg s^{-1}; e.g, Lutz et al., 2010; Harrison et al., 2012; Page et al., 2012; Mullaney et al., 2012a; Rosario et al., 2012; Stanley et al., 2015; Azadi et al., 2015). For high-luminosity sources ($L_{2-8\mathrm{keV}} > 10^{44}$ erg s) however, there has been significant disagreement as to if this relation becomes positively correlated (e.g. Lutz et al., 2010), negatively correlated (e.g. Page et al., 2012) or continues to remain uncorrelated (e.g, Harrison et al., 2012; Rosario et al., 2012; Stanley et al., 2015; Azadi et al., 2015). These disparities between various works at the high-luminosity end are likely due to small number statistics and sample variance (Harrison et al., 2012), and indeed, recent studies using large sample sizes confirm the extension of a flat trend to higher luminosities (Stanley et al., 2015; Azadi et al., 2015).

A flat trend for the $\langle \mathrm{SFR} \rangle$–BHAR relation could potentially be interpreted as re-

vealing an absence of a connection between SFR and BHAR. However, studies that have investigated the mean BHAR ($\langle$BHAR$\rangle$) as a function of SFR consistently find a *positive* relationship (e.g, Rafferty et al., 2011; Symeonidis et al., 2011; Mullaney et al., 2012b; Chen et al., 2013; Delvecchio et al., 2015). Within the paradigm of a linear $M_{\rm BH}$–$M_{\rm bulge,*}$ relation due to a universal co-evolution of BH and galaxy growth, both approaches are expected to produce a consistent, similarly linear result (see Section 4.3 for a derivation of why this is). Hickox et al. (2014) proposes a potential solution, suggesting that SFR and BHAR are connected *on average* over a period of 100 Myr, with a linear scaling. This relationship disappears when measured instantaneously owing to the rapid variability timescale of AGN, with respect to that of galactic star formation.

From a theoretical perspective, many simulations have focused on the growth of BHs in galaxy mergers (e.g, Di Matteo et al., 2005; Hopkins et al., 2005). Whilst both star formation and BH accretion are typically enhanced during the merger proper, the extent of the connection between SFR and BHAR pre- and post-merger event remains unclear. Neistein & Netzer (2014) demonstrate through the use of a semi-analytical model that the observed correlations between galaxies and their central BHs can be reproduced when BH growth occurs only during merger induced starbursts. This could explain the lack of a correlation between growth rates in low-luminosity systems whilst allowing for mutual enhancement during the merger events themselves. Thacker et al. (2014) investigate the impact of various feedback models on the SFR–BHAR parameter space in a set of equal mass merger simulations. They find a complex evolution for individual systems, even when averaged over 20 Myrs. Any correlation found is strongly dependent on the feedback model chosen, with the post-merger phase showing the strongest evidence for a positive connection. Using a high-resolution hydrodynamical merger suite, Volonteri et al. (2015a) find BHAR and galaxy-wide SFR to be typically temporally uncorrelated. They suggest in Volonteri et al. (2015b) that the observed discrepancy between the $\langle$SFR$\rangle$–BHAR and $\langle$BHAR$\rangle$–SFR relations seen observationally are a result of sampling two different projections

of the full bi-variate SFR–BHAR distribution whose build up is constructed from different behaviours between SFR and BHAR before, during and after the merger event.

It is now possible to extend these investigations to cosmological scale simulations. Using the semi-analytical code GALFORM, Gutcke et al. (2015) find a negative SFR–AGN luminosity correlation at low AGN luminosities, this then transitions to a strong positive correlation at high AGN luminosities. In the cosmological hydrodynamical simulation ILLUSTRIS, Sijacki et al. (2015) find a single trend in the SFR–BHAR relationship embedded in a large scatter, particularly in BHAR. Cosmological hydrodynamical simulations have the advantage of probing the entire galaxy population within a self consistent variety of environments with a diverse range of accretion and merger histories. Here we investigate to what extent galaxy and BH growth rates are connected within the "Evolution and Assembly of GaLaxies and their Environment" (EAGLE, Schaye et al., 2015; Crain et al., 2015) [1] [2] simulation. Adopting physical prescriptions for radiative cooling, star formation, stellar mass loss, BH accretion, BH mergers and both stellar and AGN feedback, EAGLE reproduces many observed properties of galaxies, BHs and the intergalactic medium with unprecedented fidelity (e.g, Schaye et al., 2015; Furlong et al., 2015b,a; Trayford et al., 2015; Schaller et al., 2015b; Lagos et al., 2015; Rahmati et al., 2015; Bahé et al., 2016; Crain et al., 2016; Rosas-Guevara et al., 2016; Segers et al., 2016; Trayford et al., 2016).

The chapter is organised as follows. In Section 4.2 we provide a brief overview of the EAGLE simulation suite, including the subgrid model prescription and simulation output. The results are presented in Section 5.3. We examine the EAGLE predictions of the ⟨SFR⟩–BHAR relationship for an AGN selected sample and the ⟨BHAR⟩–SFR relationship for a SFR selected sample in Section 4.4.1, finding good agreement with recent observational findings. To investigate why these trends might be different, we explore the effect of time-averaging each growth rate and examine potential sampling

[1] `www.eaglesim.org`

[2] Galaxy and halo catalogues of the simulation suite are publicly available at `http://www.eaglesim.org/database.php` (McAlpine et al., 2016).

biases in Section 4.4.2. Section 4.4.3 examines the influence of the host dark matter halo on both SFR and BHAR, finding that each exhibits a strong connection. Finally in Section 5.4, we discuss the changing relationship between SFR and BHAR as the halo grows and in Section 4.6, we present our conclusions.

4.2 The EAGLE simulation suite

A full overview of the EAGLE simulation suite can be found in Chapter 2.

4.2.1 Constructing histories of individual galaxies

In order to accurately trace the evolution of individual galaxies and their central BH for the analysis in Section 4.4.3, we require histories of a higher temporal resolution than is provided by the snipshot output. To do this, we follow galaxies and their central BH through cosmic time. BHARs are recorded at each timestep with a typical spacing of $\sim 10^3 - 10^4$ yr, yielding an 'instantaneous' rate. These can then be time-averaged over longer durations. Quoted 'instantaneous' SFRs are taken from the snapshot output, where they are computed based on the current star forming state of the gas contained within the galaxy. Time-averaged SFR histories are constructed from the stellar particles born within the main progenitor that reside in the galaxy at the present day. As these particles store both their birth time and initial mass, collectively they create a robust history of star formation for that galaxy. However, as these histories are sensitive in their resolution to the number of particles sampled, only galaxies containing more than 200 particles ($M_{*[z=0]} \approx 10^{8.5} M_{\odot}$) are considered for this study.

Figure 4.1 shows an example history of an individual galaxy's SFR (top panel) and accretion rate of the central BH (bottom panel) through cosmic time taken from the methods described above. We show each growth rate time-averaged over 5 Myr (blue and green lines) and 100 Myr (red and black lines) to highlight the large difference in

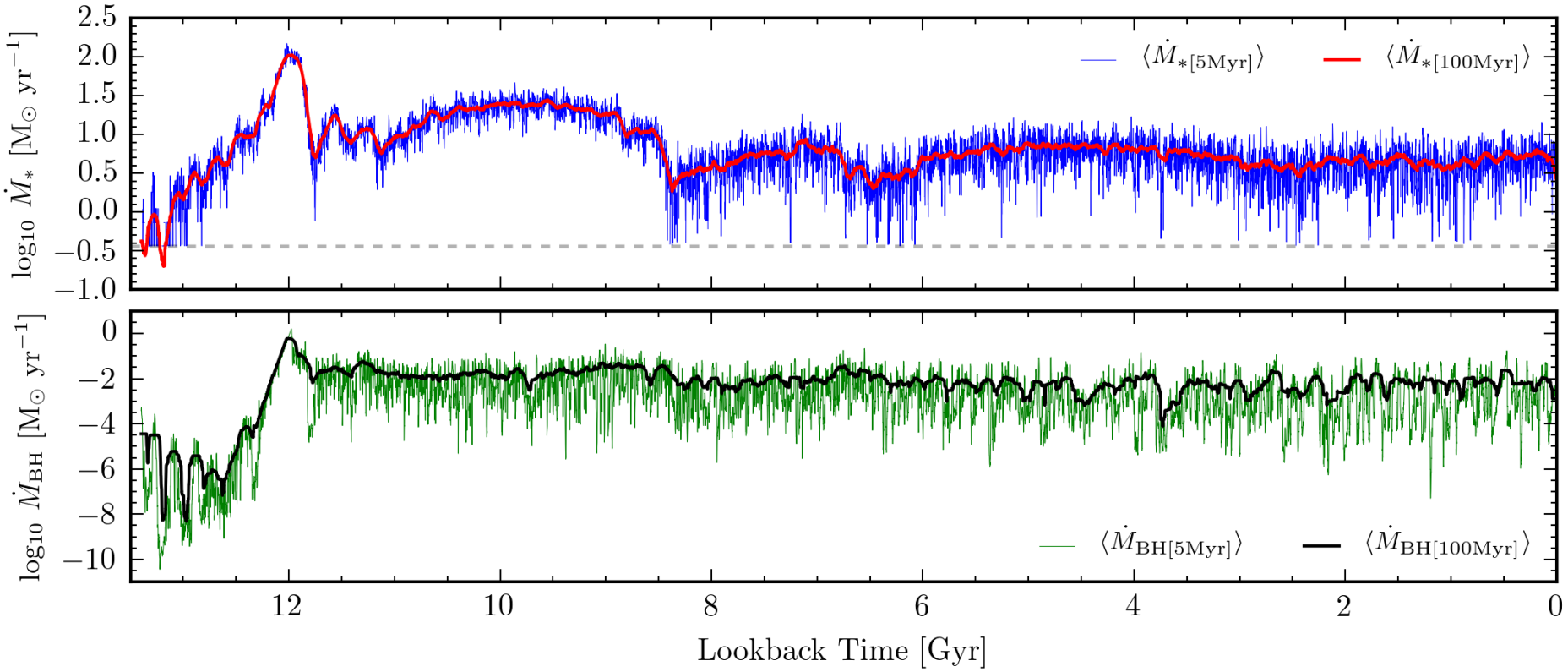

Figure 4.1: Galaxy (top panel) and BH (bottom panel) growth rates as a function of lookback time for an individual galaxy (`GalaxyID`=20324216, $M_{200[z=0]} = 10^{13.2}\mathrm{M}_\odot$, $M_{*[z=0]} = 10^{11.1}\mathrm{M}_\odot$, $M_{\mathrm{BH}[z=0]} = 10^{8.7}\mathrm{M}_\odot$). Blue (green) and red (black) lines show the SFR (BHAR) history averaged over 5 Myr and 100 Myr respectively. When averaged over short timescales, BHARs can vary by as much as ≈ 4 dex. SFRs however vary considerably less over the same time window (≈ 1 dex), and generally represent values in much closer agreement to their long term average rate. In this individual case the long term average trends of SFR and BHAR yield quite different evolutionary behaviours. As the both the particle mass and averaging time window is finite, the minimum possible SFR sampled for a 5 Myr timescale is shown as a dashed grey line.

variability scatter between the two timescales. This is particularly severe for BHAR, where values recorded over short timescales do not return a good approximation of the long term average rate, differing in value by as much as 4 dex. We have adopted 100 Myr as our long averaging duration as it reflects an estimate of the effective timescale for empirical indicators of star formation using the far-infrared (FIR, the tracer of star formation for the observational studies compared to in Section 4.4.1, see the discussions by Neistein & Netzer, 2014; Volonteri et al., 2015a). Although there are similar features between SFR and BHAR through time for this individual case (for example a common peak at a lookback time of 12 Gyr), globally the two histories are quite different.

4.2.2 The $M_{\rm BH}$–M_{200} relation

Figure 4.2 shows the $M_{\rm BH}$–M_{200} relation for central galaxies at $z = 0$. We have plotted $M_{\rm BH}$ as a function of halo mass rather than bulge or total stellar mass due to the crucial connection that M_{200} has with both SFR and BHAR (see Section 4.4.3 onwards). We note that the $M_{\rm BH}$–M_* relation also follows the same behaviour (see Figure 1 of Barber et al., 2016) and throughout this description M_{200} and M_* can be interchanged. The overlaid two dimensional bins are for the continued investigation in Section 4.4.3, where they are fully described.

The empirical relationship between BH mass and that of the classical bulge is well described by a single power law at high mass (e.g, Magorrian et al., 1998; Kormendy & Ho, 2013; McConnell & Ma, 2013 with gradient values of $\alpha \approx 1 - 1.3$ satisfying eq. (4.3.1)). Indeed, one of the calibration parameters of the simulation is to match the normalisation of this relationship. However, whereas traditionally this trend has been linearly extrapolated to lower-mass systems, EAGLE predicts a steepening of the trend. As a consequence, the relation between BH mass and the mass of the host galaxy or halo is not well described by a single power law. Interestingly, a steeper slope at intermediate masses is supported by recent observations of bulge (or

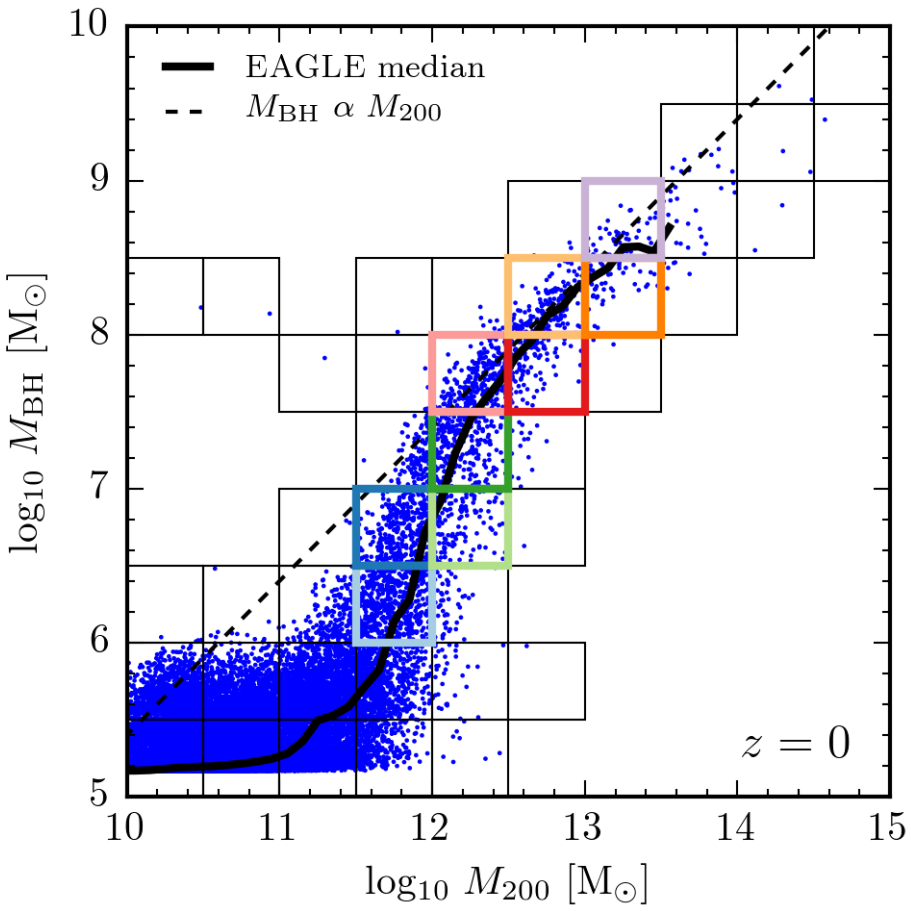

Figure 4.2: $M_{\rm BH}$–M_{200} relation for central galaxies at $z = 0$, where M_{200} is the halo mass. Each galaxy is represented by an individual blue point and the median trend is shown as a solid black line. The black dashed line shows a linear relationship, $M_{\rm BH} \propto M_{200}$, for reference. Overlaid two-dimensional bins are 0.5 dex on each side and contain at least one galaxy. Nine of these cells are used for the continued investigation in Section 5.4 and are outlined in colours that relate to the histories shown in Figures 4.9 and 4.10.

pseudobulge) systems (e.g, Scott et al., 2013; Greene et al., 2016). When total stellar mass is considered as the independent variable, Reines & Volonteri (2015) predict Seyfert-like systems yield and alternate $M_{\rm BH}$–M_* relationship to previously measured early-type systems. However, each trend is consistent with a linear relation, with Seyfert-like systems yielding a lower normalisation.

For EAGLE, BHs in massive systems ($M_{200} \gtrsim 10^{12.5}\mathrm{M}_\odot$) follow an approximately linear trend with halo mass (compare to the black dashed black line in Figure 4.2), but those hosted by haloes with mass $M_{200} \lesssim 10^{12.5}\mathrm{M}_\odot$ follow a much steeper relation and those in the lowest mass systems ($M_{200} \lesssim 10^{11.5}\mathrm{M}_\odot$) plateau at the seed BH mass.

Bower et al. (2017) argue that multiple physical processes drive the relation between $M_{\rm BH}$ and M_{200}. In low (high) mass systems stellar (AGN) feedback regulates the baryonic inflow to the galaxy, suppressing BH (continued stellar) growth. There is a critical transition halo mass ($M_{200} \sim 10^{12}\mathrm{M}_\odot$, hereafter $M_{\rm crit}$) separating these two regulatory regimes. Within $M_{\rm crit}$ haloes, neither feedback process is dominant, and as a result BHs grow at a highly non-linear rate. These phases create the flat, supra-linear and ∼linear regimes of BH growth seen in the integrated quantities of Figure 4.2 and have important consequences for the galaxy and BH growth rates investigated throughout this study.

4.2.3 Absolute calibration of SFRs

When comparing to the observed cosmic SFR density, Furlong et al. (2015b) found an almost constant -0.2 dex offset for redshifts $z \leq 3$. There is however continued uncertainty as to the absolute calibration of SFR indicators on which these observations rely. For example, Chang et al. (2015) find upon revisiting this calibration with the addition of WISE photometry to the full SDSS spectroscopic galaxy sample that the SFRs of local galaxies along the main-sequence are systematically lower than previously estimated by ≈ 0.2 dex, yielding good agreement with the EAGLE

prediction (see Figure 5 of Schaller et al., 2015a).

As the observational datasets compared to in Section 4.4.1 utilise an earlier calibration, we *reduce* all observed SFRs by 0.2 dex. The magnitude of this recalibration is shown as a red arrow in Figures 4.3 and 4.4. This serves to remove the known global systematic offset, making it simpler to focus on the trends with BHAR that are the topic of this paper.

4.3 Predictions from the integrated quantities

The major motivation for linking the growth of galaxies to the growth of their central BH has arisen empirically from the strong correlations seen in their integrated properties (the primary example being the tight $M_{\rm BH}$–$M_{*,\rm bulge}$ relation). We therefore require an evolutionary model that suitably fits this end point. In the simplest case, where BHs and their host galaxy grow in concert (or co-evolve), the relational form between their growth rates can be easily predicted. Given a functional form of the $M_{\rm BH}$–$M_{\rm bulge,*}$ relation described via

$$\log_{10} M_{\rm BH} = \alpha \log_{10} M_{\rm bulge,*} + \log_{10}\beta, \tag{4.3.1}$$

where α is the gradient of the slope and $\log_{10}\beta$ is the intercept, the predicted relation between the growth rates is simply found by differentiating with respect to time, i.e,

$$\log_{10}\dot{M}_{\rm BH} = \log_{10}\dot{M}_{\rm bulge,*} + \log_{10}\beta + \log_{10}\alpha + (\alpha - 1)\log_{10} M_{\rm bulge,*}. \tag{4.3.2}$$

In the trivial case where $\alpha = 1$ (i.e, a linear relation) this reduces to

$$\log_{10}\dot{M}_{\rm BH} = \log_{10}\dot{M}_{\rm bulge,*} + \log_{10}\beta. \tag{4.3.3}$$

Within this scenario, growth rates are directly proportional to one another scaled by the intercept, β, of the $M_{\rm BH}$–$M_{\rm bulge,*}$ relation. Therefore, if the functional form

between the growth rates is described via

$$\log_{10}\dot{M}_{\rm BH} = \gamma \log_{10}\dot{M}_{*} + \log_{10}\delta, \tag{4.3.4}$$

where again γ and δ are the slope and intercept values, we would expect $\delta = \beta$ and $\gamma = 1$ in the case where $\alpha = 1$.

Throughout this study we will test the hypothesis that there exists a broadly linear co-evolution between galaxies and their central BHs, a plausible scenario fitting the empirical $M_{\rm BH}$–$M_{*,\rm bulge}$ relation. We refer throughout to α as the gradient of the slope between the integrated properties ($M_{\rm BH}$–$M_{*/\rm bulge,*}$) and to γ as the slope between each growth rate ($\dot{M}_{\rm BH}$–$\dot{M}_{*}$), both in log space.

4.4 Results

4.4.1 Comparison to observations

We begin by comparing the predicted relationship between galaxy and BH growth rates to two recent observational studies using different selection criteria. Firstly, we explore the $\langle$SFR$\rangle$ versus BHAR relation for the *AGN selected* sample presented by Stanley et al. (2015). Secondly, we explore the $\langle$BHAR$\rangle$ versus SFR relation for the *SFR selected* sample presented by Delvecchio et al. (2015). Together they represent two of the largest sample sizes of their respective selection techniques, spanning multiple epochs. Large sample sizes such as these are key in overcoming the uncertainties inherent to low number statistics and in mitigating the potential redshift evolution biases that could be misinterpreted as an underlying trend.

As mentioned in Section 4.2.1, SFRs obtained via FIR tracers most likely probe the recent star formation history of a galaxy, rather than an instantaneous value. Therefore the comparative SFRs of EAGLE galaxies used in the analysis of Figures 4.3 and 4.4 are the time-averaged rate over the 100 Myr preceding the instantaneous

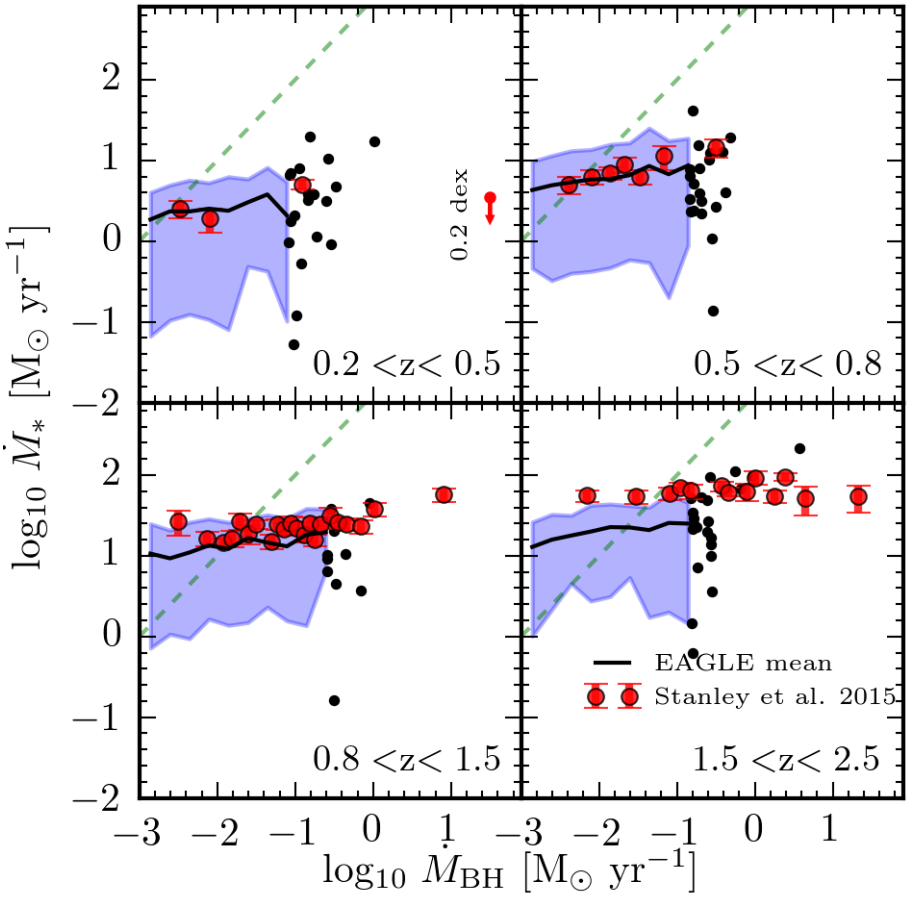

Figure 4.3: SFR ($\dot{M}_*$) as a function of BHAR ($\dot{M}_{\rm BH}$) for four continuous redshift ranges from $0.2 < z < 2.5$ for a BHAR selected sample. The recorded BHARs are instantaneous and the SFRs are time-averaged over the 100 Myr preceding the recorded BH accretion event. The *mean* SFR as a function of BHAR for central galaxies is shown as a black line, with the corresponding blue shaded region indicating the 10-90$^{\rm th}$ percentile range. Bins containing fewer than 10 objects have their galaxies represented individually as black solid circles. The linear relation $\dot{M}_{\rm BH}/\dot{M}_* = 10^{-3}$ is shown as a dashed green line and the data of Stanley et al. (2015) is represented as red circles. Fits to the EAGLE mean relations are tabulated in Table 4.1. The magnitude of the SFR recalibration applied to the data for all redshifts is indicated by a red arrow in the upper left panel (see Section 4.2.3). For each redshift range we find mean trends that are considerably flatter than a linear relation ($\gamma_{\rm S15} \ll 1$ in eq. (4.3.4)).

BHAR measurement. However, when performing the equivalent analysis using the instantaneous values of SFR at the time of the BHAR measurement we find no difference in the result, attesting to the stability of star formation as a process.

SFR as a function of BHAR for an AGN selected sample

The study of Stanley et al. (2015) consists of ≈ 2000 X-ray detected AGN spanning over three orders of magnitude in luminosity ($10^{42} < L_{2-8\mathrm{keV}} < 10^{45.5}$ erg s^{-1}) for the redshift range $z = 0.2 - 2.5$. To compare to the data we convert their quoted bolometric AGN luminosities (derived from X-ray luminosities) to BHARs via

$$\dot{M}_{\mathrm{BH}} = \frac{L_{\mathrm{AGN}}}{\epsilon_r c^2}, \tag{4.4.1}$$

where c is the speed of light and ϵ_r is the radiative efficiency of the accretion disk, which is assumed to be 0.1 (Shakura & Sunyaev, 1973). To emulate the sample selection of this study, we choose EAGLE galaxies based on a redshift independent detection limit of $\dot{M}_{\mathrm{BH}} = 10^{-3}$ M$_\odot$ yr^{-1}, corresponding to $L_{\mathrm{AGN}} \approx 10^{43}$ erg s^{-1} according to eq. (4.4.1) and is equivalent to $L_{2-8\mathrm{keV}} \approx 10^{42}$ erg s^{-1} using the conversion methods outlined in Section 3.2 of Stanley et al. (2015).

The results are presented in Figure 4.3, showing the *mean* SFR as a function of BHAR, represented by a solid black line. We see that for each redshift range the gradient of the relation remains shallow (note the same dynamic range is used for both axes), ranging in values $0.1 \leq \gamma_{\mathrm{S15}} \leq 0.2$ (see Table 4.1) and is in striking agreement with the "remarkably flat" relation reported by Stanley et al. (2015). This is considerably flatter than one would predict for a linear M_{BH}–$M_{\mathrm{bulge},*}$ relation from a co-evolution of growth, which we represent as a dashed green line[3] (see Section 4.3). The dynamic range of SFRs is modest, with a scatter of $\approx 1 - 1.5$ dex for all redshifts. The normalisations of $\langle \mathrm{SFR} \rangle$ in the three lowest redshift ranges ($0.2 < z < 1.5$) are in good agreement with the observational estimates (within

[3]Using eq. (4.3.3) with $\beta = 1000$ (McConnell & Ma, 2013).

$\approx$ 0.1 dex). However, the values in the highest redshift range ($1.5 < z < 2.5$) are systematically under predicted by $\approx$ 0.5 dex over and above the recalibration discussed in Section 4.2.3. We note that this highest bin is potentially subject to the largest systematic over estimate ($\approx$ 0.4 dex) due to the large fraction ($\approx$ 80%) of undetected FIR sources (included as upper limits) in the observations (see Stanley et al., 2015).

BHAR as a function of SFR for a SFR selected sample

The study of Delvecchio et al. (2015) consists of $\approx$ 8600 star-forming galaxies detected out to $z = 2.5$. The selection limits in SFR are redshift dependent, corresponding to 0.2, 1.0, 3.0, 8.0 and 25.0 $\mathrm{M_\odot\ yr^{-1}}$ for the five redshift ranges covered by this study from low to high respectively. We note that the data points from Delvecchio et al. (2015) are segregated also in stellar mass, however for simplicity we make no such distinction.

The comparison is shown in Figure 4.4, showing the *mean* BHAR as a function of SFR as the solid black line. Again, EAGLE shows a good consistency with the observational measurements (shown in red), only over predicting $\langle$BHAR$\rangle$ in the lowest redshift range ($0.0 < z < 0.5$). However, Delvecchio et al. (2015) mentions that the limited co-moving volume of this study at low redshift could potentially exclude the most luminous sources. The behaviour of the $\langle$BHAR$\rangle$–SFR relation is quite different from the $\langle$SFR$\rangle$–BHAR relation seen in Figure 4.3, adhering much closer to a linear trend. We see, uniformly, gradients close to unity ($1.0 \leq 1/\gamma_{\mathrm{D15}} \leq 1.2$, see Table 4.1) in good agreement with the linear M_{BH}–$M_{\mathrm{bulge},*}$ relation expected for a co-evolution of growth, shown as a dashed green line (note again the same dynamic range is used for both axes). An additional difference is the spread of values in the minimisation axis ($\dot{M}_{\mathrm{BH}}$ for this figure). The distribution of SFRs in Figure 4.4 span a relatively narrow dynamic range, whereas here, BHARs vary as much as $\approx$ 4 dex in the $10 - 90^{\mathrm{th}}$ percentile region. In fact, the dynamic range of BHARs is so large that the small fraction of galaxies whose values dominate the mean are able to pull

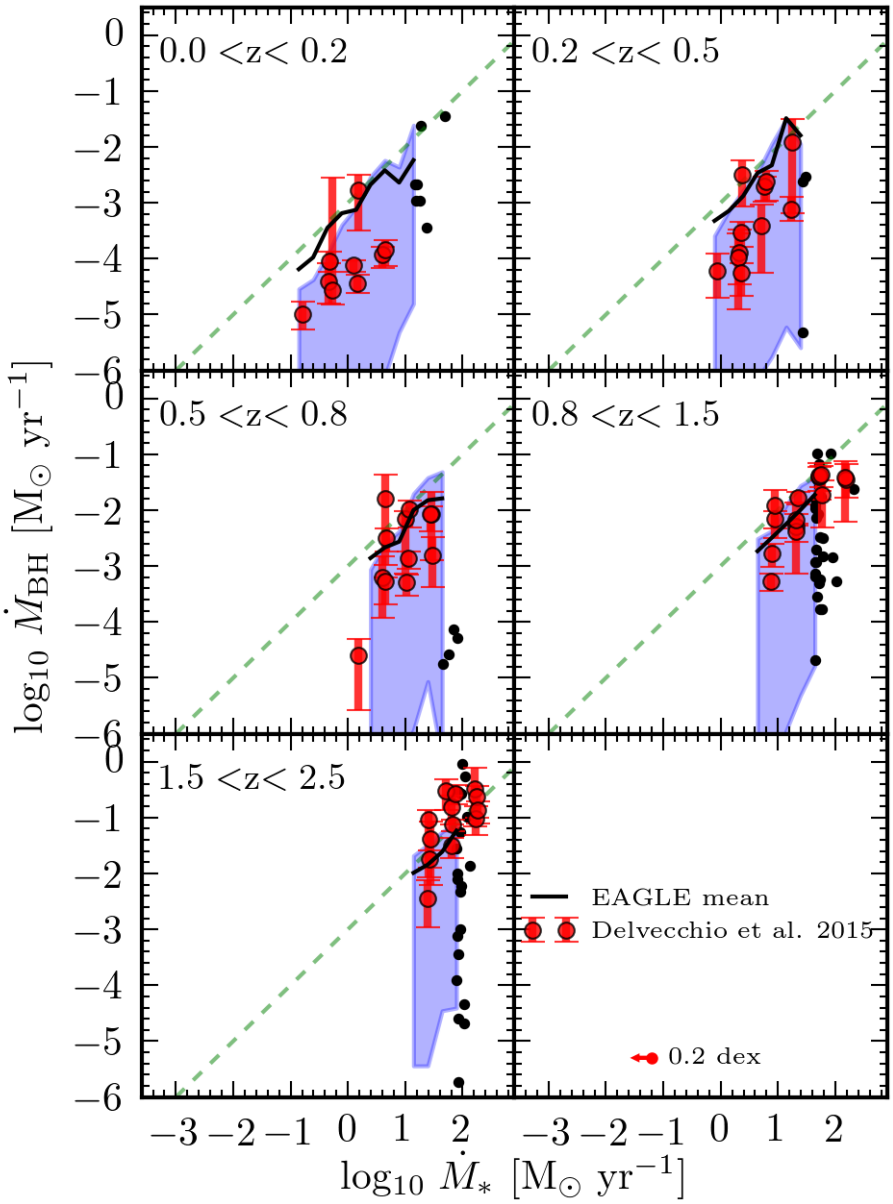

Figure 4.4: BHAR ($\dot{M}_{\rm BH}$) as a function of SFR ($\dot{M}_*$) for four continuous redshift ranges from $0.2 < z < 2.5$ for a SFR selected sample. The recorded BHARs are instantaneous and the SFRs are time-averaged over the 100 Myr preceding the recorded BH accretion event. The *mean* BHAR as a function of SFR for central galaxies is shown as a black line, with the corresponding blue shaded region indicating the 10-90$^{\rm th}$ percentile range. Bins containing fewer than 10 objects have their galaxies represented individually as black solid circles. The linear relation $\dot{M}_{\rm BH}/\dot{M}_* = 10^{-3}$ is shown as a dashed green line and the data of Delvecchio et al. (2015) is represented as red circles. Fits to the EAGLE mean relations are tabulated in Table 4.1. The magnitude of the SFR recalibration applied to the data for all redshifts is indicated by a red arrow in the lower right panel (see Section 4.2.3). For each redshift range we find gradients of the mean trend close to unity ($1/\gamma_{\rm D15} \approx 1$), in good agreement with a linear relation.

it outside this percentile range entirely in some places, suggesting the median to be a more suitable statistic to measure this trend.

Overall the agreement between EAGLE and the observations is excellent, particularly given that no information regarding this relation was considered during the calibration procedure. The difference in behaviour found empirically via alternate selection criteria is well reproduced by the simulation. We find, consistent with the Hickox et al. (2014) model and findings by Volonteri et al. (2015b), that ⟨SFR⟩–BHAR for an AGN selected sample exhibits a relatively flat trend ($\gamma_{\rm S15} \approx 0.15$), whilst that of ⟨BHAR⟩ with respect to SFR for a SFR selected sample is substantially steeper and close to unity ($1/\gamma_{\rm D15} \approx 1.1$). However, within the paradigm of a linear $M_{\rm BH}$–$M_{\rm bulge,*}$ relation created through co-evolution of growth these results are both predicted to be linear (i.e, $\gamma \approx 1$ for both, see Section 4.3). Therefore either the underlying relationship itself is fundamentally non-linear, or a fuller understanding of the two processes is required. In the next section we continue to examine potential reasons as to the cause of this difference.

4.4.2 Understanding the BHAR-SFR relationship

In this section we explore two potential reasons why the ⟨SFR⟩–BHAR and ⟨BHAR⟩–SFR trends are not each consistent with a linear relationship. We examine the hypothesis that (1) growth rates have an underlying linear connection only on average, which is masked when the unstable growth rate is observed instantaneously and (2) how selection biases due to the inability to probe the complete SFR–BHAR plane may play a role.

A time-averaged SFR–BHAR connection

Hickox et al. (2014) suggested that an underlying correlation held between a stable (galactic star formation) and unstable (BH accretion) process *on average* could be washed out if the unstable property is measured instantaneously. That is to say, if

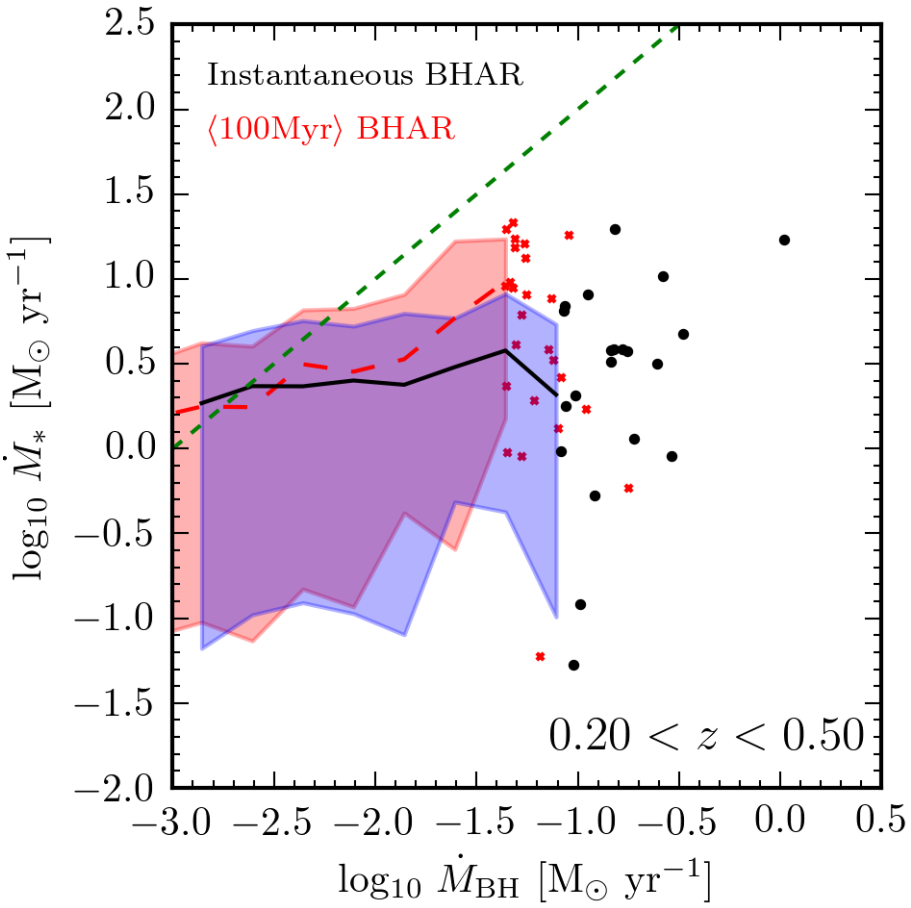

Figure 4.5: *blue region*: A replica of the upper left panel in Figure 4.3 (see caption for description of contents). Here BHARs are instantaneous and SFRs are the time-averaged rate over the 100 Myr preceding the BH event. *red region*: We repeat the analysis for the same central galaxies satisfying the Stanley et al. (2015) selection criteria (instantaneously, blue region), however now both SFR and BHAR are time-averaged over the same 100 Myr period. Fits to the time-averaged mean relations are shown in Table 4.1 (denoted with $\langle 100\ \mathrm{Myr}\rangle$). We find that even when both growth rates are time-averaged over 100 Myr, an AGN selected sample does not revert to a linear relationship between $\dot{M}_*$ and $\dot{M}_{\mathrm{BH}}$.

one could observe X-ray luminosities of AGN sources over prolonged periods, the underlying relationship between the two properties would begin to emerge. Indeed, with a simple model that assumes SFR and BHAR are connected on average with a linear scaling over a period of 100 Myr, Hickox et al. (2014) reproduce the empirical behaviours of both the ⟨SFR⟩–BHAR and ⟨BHAR⟩–SFR relationships whilst retaining a scenario consistent with a linear co-evolution between galaxies and their central BHs. While it is not possible to test observationally due to the length of these timescales, we are able to test this hypothesis using the simulation.

Figure 4.5 is similar to the upper left panel of Figure 4.3. The region in blue, with the black solid line, shows the original analysis of the ⟨SFR⟩–BHAR relation for the redshift range $0.2 < z < 0.5$ using instantaneous BHARs and SFRs that are time-averaged 100 Myr before the BH event. Overlaid in red, with the mean represented by a dashed line, is the same selection of galaxies (i.e, instantaneous $\dot{M}_{\rm BH} > 10^{-3}{\rm M}_{\odot}$ yr^{-1}) with each growth rate now time-averaged over 100 Myr. Interestingly, although the high BHARs shift systematically to lower values on average[4], both the dynamic range and slope of the mean remain broadly consistent with their instantaneous equivalents ($\gamma_{\rm S15} \rightarrow \gamma_{\rm S15\langle 100\ Myr\rangle} = 0.2 \rightarrow 0.3$, see Table 4.1). This behaviour remains for each redshift range explored by Stanley et al. (2015) (see Table 4.1). An alternate approach is to select galaxies in excess of $\dot{M}_{\rm BH} = 10^{-3}{\rm M}_{\odot}$ yr^{-1} on average (rather than instantaneously as done above) or indeed to prolong the averaging timescale to > 100 Myrs. However in each case, and for all redshift intervals, the gradient values remain well below unity ($0.30 < \gamma_{\rm S15\langle 100\ Myr\rangle} < 0.55$). This leads us to conclude that the average galaxy and BH growth rates for an AGN selected sample do not yield an underlying global linear relationship.

Table 4.1: Slope (γ) and intercept (δ) of various relations satisfying eq. (4.3.4). S15 denotes the study of Stanley et al. (2015) and D15 denotes the study of Delvecchio et al. (2015). Columns 2 and 3 are the fitted values of the EAGLE mean relations investigated in Figures 4.3 and 4.4. Annotated with a * in columns 6 and 7 are the EAGLE *median* fits to the same datasets. Fits to the mean relation of the 100 Myr time-averaged growth rates from Section 4.4.2 for the same datasets are shown in columns 4 and 5. Errors on the individual mean/median EAGLE data points are taken from bootstrap re-sampling. Quoted fits and their associated errors were computed using the Python module LMFIT.

z	$\gamma_{\rm S15}$	$\log_{10}\delta_{\rm S15}$	$\gamma_{\rm S15}\langle 100{\rm Myr}\rangle$	$\log_{10}\delta_{\rm S15}\langle 100{\rm Myr}\rangle$	${\gamma_{\rm S15}}^*$	$\log_{10}{\delta_{\rm S15}}^*$
$0.20 < z < 0.50$	0.13 ± 0.04	0.67 ± 0.09	0.31 ± 0.05	1.18 ± 0.14	0.14 ± 0.02	0.46 ± 0.05
$0.50 < z < 0.80$	0.15 ± 0.02	1.07 ± 0.04	0.26 ± 0.01	1.36 ± 0.04	0.12 ± 0.03	0.83 ± 0.07
$0.80 < z < 1.50$	0.12 ± 0.03	1.32 ± 0.06	0.27 ± 0.02	1.63 ± 0.04	0.14 ± 0.03	1.23 ± 0.07
$1.50 < z < 2.50$	0.16 ± 0.02	1.60 ± 0.05	0.24 ± 0.02	1.72 ± 0.05	0.16 ± 0.02	1.51 ± 0.05
z	$1/\gamma_{\rm D15}$	$-\frac{\log_{10}\delta_{\rm D15}}{\gamma_{\rm D15}}$	$1/\gamma_{\rm D15}\langle 100{\rm Myr}\rangle$	$\log_{10}\delta_{\rm D15}\langle 100{\rm Myr}\rangle$	${1/\gamma_{\rm D15}}^*$	$-\frac{\log_{10}\delta_{\rm D15}}{\gamma_{\rm D15}}*$
$0.01 < z < 0.25$	1.07 ± 0.10	-3.21 ± 0.05	1.06 ± 0.06	-3.29 ± 0.04	2.68 ± 0.14	-5.52 ± 0.07
$0.25 < z < 0.50$	1.13 ± 0.11	-3.30 ± 0.05	0.94 ± 0.04	-3.07 ± 0.02	2.28 ± 0.20	-5.61 ± 0.10
$0.50 < z < 0.80$	1.06 ± 0.19	-3.35 ± 0.17	1.10 ± 0.03	-3.23 ± 0.02	2.21 ± 0.33	-5.99 ± 0.29
$0.80 < z < 1.50$	1.00 ± 0.01	-3.39 ± 0.15	1.27 ± 0.03	-3.44 ± 0.03	2.48 ± 0.21	-6.53 ± 0.23
$1.50 < z < 2.50$	0.96 ± 0.16	-3.17 ± 0.24	1.37 ± 0.04	-3.60 ± 0.05	1.61 ± 0.26	-5.18 ± 0.39

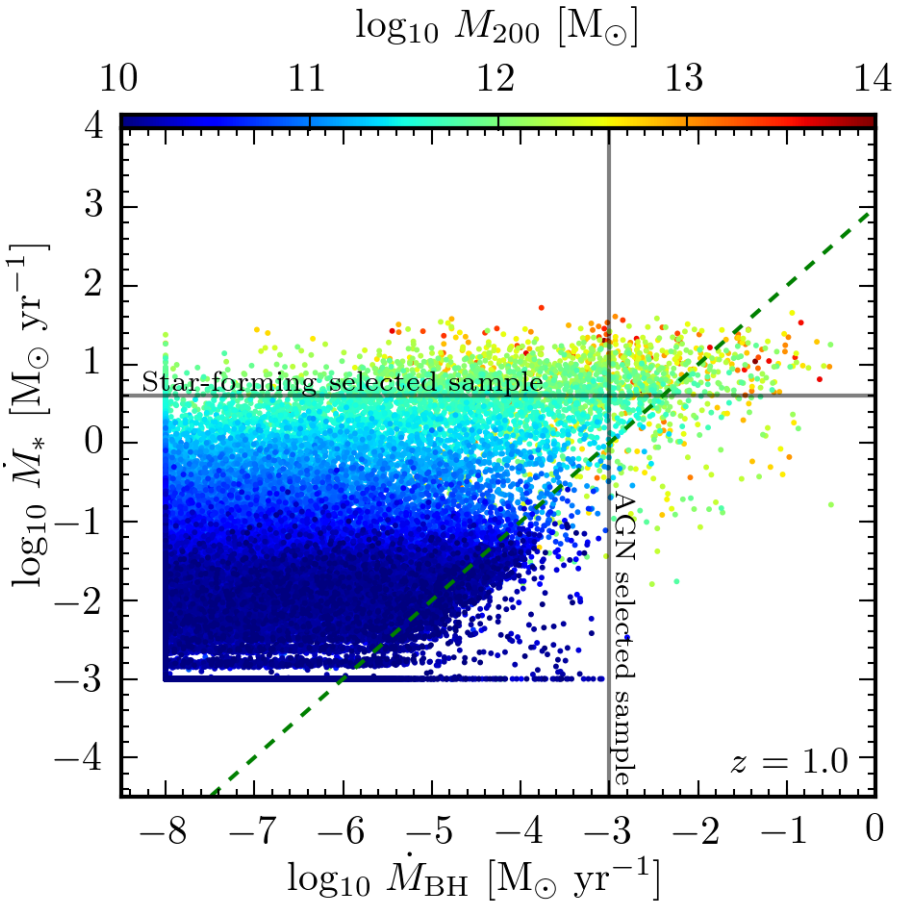

Figure 4.6: $\dot{M}_*$-$\dot{M}_{\rm BH}$ (SFR-BHAR) relation for central galaxies at $z = 1$. Growth rates are *instantaneous* and coloured by the mass of the halo (M_{200}). Values that are below 10^{-8} M$_\odot$ yr^{-1} for $\dot{M}_{\rm BH}$ and below 10^{-3} M$_\odot$ yr^{-1} for $\dot{M}_*$ are clipped to these limits. The approximate flux limits of Stanley et al. (2015) and Delvecchio et al. (2015) investigated in Section 4.4.1 are shown as vertical and horizontal solid lines respectively, highlighting the different regions of the full distribution that these surveys are able to probe. The dashed green line indicates the linear relation $\dot{M}_{\rm BH}/\dot{M}_* = 10^{-3}$.

Sampling different regions of the entire SFR-BHAR plane

As observational surveys are subject to various flux limitations, they can only sample particular regions of the full SFR–BHAR plane. If the global underlying relation is linear, and each property exhibits a moderate scatter ($\lesssim 1$ dex), each sub sample should also return a linear result. However, as the findings of previous sections do not to support an underlying linear relation, and because the scatter is large ($\gg 1$ dex), it is important to investigate the effect of this sampling.

Figure 4.6 shows the complete SFR–BHAR plane for all central galaxies at $z = 1$. In order to eliminate any potential bias incurred via redshift evolution in either growth rate we consider a discrete redshift rather than the continuous ranges of Section 4.4.1. Each data point represents the *instantaneous* state of a single galaxy and its central BH, coloured by the halo mass. Values that are below 10^{-8} $\mathrm{M}_{\odot}$ yr^{-1} for $\dot{M}_{\mathrm{BH}}$ and 10^{-3} $\mathrm{M}_{\odot}$ yr^{-1} for $\dot{M}_{*}$ are treated as 'zero' and are clipped to these values. The approximate flux limit of the AGN selected sample by Stanley et al. (2015) is shown as a solid vertical line and the approximate flux limit of the SFR selected sample by Delvecchio et al. (2015) is shown as a horizontal solid line.

It is further apparent that the global relationship between SFR and BHAR is not simply linear (reference with the dashed green line). Instead, a complicated relationship arises due to an amalgamation of three distinct behaviours of BH growth dependent on the mass of the host dark matter halo (see next section). It is therefore crucial to consider the particular region sampled before arriving at a particular conclusion. AGN selected samples, such as that of Stanley et al. (2015), currently probe a relatively limited region at the tip of the SFR–BHAR plane. With the exception of a few sources with rates $\dot{M}_{*} \ll 1\mathrm{M}_{\odot}$ yr^{-1}, galaxies satisfying this selection criteria are distributed over a relatively narrow range of SFRs. As such, each bin of BHAR yields a very similar value of $\langle \mathrm{SFR} \rangle$, creating an approximately flat trend. SFR selected samples, such as that of Delvecchio et al. (2015), sample a not too dissimilar

[4]The shift ($\approx$0.5 dex) to lower values in BHAR when averaging over 100 Myr arises due to the most luminous 'detections' commonly residing in peaks of the accretion rate history.

distribution of SFRs (this time due to the flux limit), however, the distribution of BHARs is much wider. This in turn yields a steeper relation. We note, that whilst the mean SFR provides a good proxy of the median SFR for an AGN selected study (compare columns 3 and 6 of Table 4.1), the mean BHAR for a SFR selected study is not a good proxy of the median value due to the distribution of BHARs having such a large scatter (compare columns 8 and 10 of Table 4.1). Although only the results from $z = 1$ have been shown here, when investigated we find the results remain true independent of redshift.

We therefore conclude that the different behaviour found for the $\langle$SFR$\rangle$–BHAR and $\langle$BHAR$\rangle$–SFR relations recovered by observational studies is due to sampling considerably different regions of the full (not universally linear) SFR–BHAR plane. We now continue to investigate the nature of this relationship in the EAGLE simulation and its evolution through time.

4.4.3 The connection to the host dark matter halo

The relationship between SFR and BHAR seen in Figure 4.6 is complicated, seemingly not adhering to a simple universal trend. However, there is evidence that each property has a link with the mass of the host dark matter halo, highlighted in the change of the data point colours, which transition smoothly from blue to red with increasing SFR, and systematically shift rightward in BHAR (with a large scatter) at high halo mass.

To examine this in more depth we sub categorise the $z = 1$ central galaxies into five continuous ranges of halo mass, showing the growth rates in Figure 4.7. Here we find, in fact, that the global make up of the SFR–BHAR plane in Figure 4.6 is resolved into a collection of two dimensional strips, wide in their dynamic range of BHAR ($\approx 4-6$ dex) yet generally much more compact in their SFR ($\approx 1-2$ dex). Each strip hosts a characteristic value of SFR (defined as the mode of the distribution, shown as a horizontal solid red line) that continuously increases with increasing halo mass.

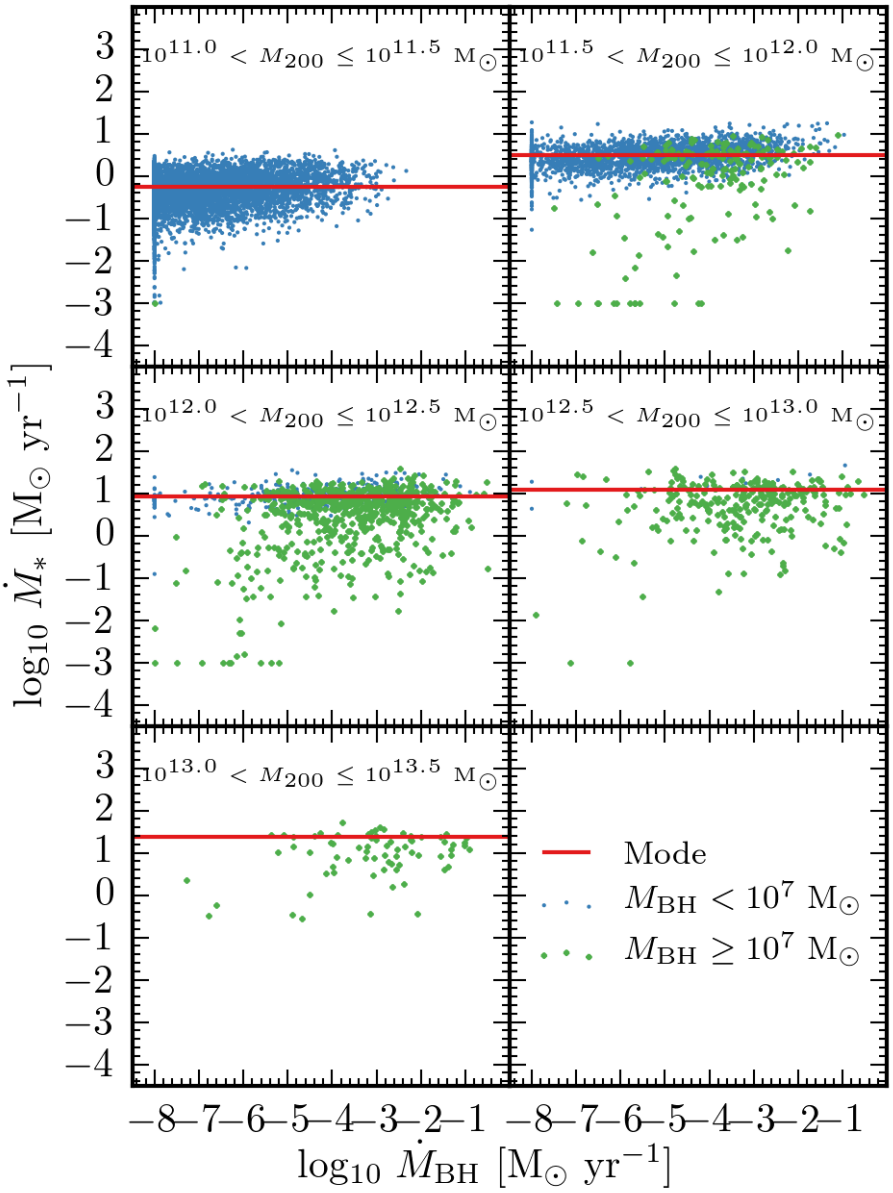

Figure 4.7: Growth rates at $z = 1$ (from Figure 4.6) for central galaxies separated into five continuous halo mass ranges. Points are coloured green if a galaxy hosts a massive BH ($M_{\rm BH} \geq 10^7 {\rm M}_\odot$) and blue otherwise. The characteristic SFR for a given halo mass bin (classified as the mode of the distribution) is shown as a horizontal solid red line. Each range of halo mass yields a relatively narrow distribution of SFRs ($1 - 2$ dex) and much wider distribution of BHARs (up to 8 dex). Larger halos are associated with larger characteristic SFRs and have a higher fraction of BHARs $> 10^{-4} {\rm M}_\odot$ yr^{-1}. Galaxies yielding SFRs far below the characteristic rate all contain massive BHs (green points) and have likely recently undergone a violent episode of AGN feedback reducing the current star-forming capability of the system.

This is in line with the 'star forming main sequence', where galaxies of increased stellar mass are seen to host larger SFRs (e.g, Elbaz et al., 2007). Interestingly, the rate of change with M_{200} for this characteristic SFR does not remain constant, initially increasing by $\Delta\dot{M}_* \approx 1$ dex between $10^{11.0} < M_{200} \leq 10^{12.0}\mathrm{M}_\odot$ and reducing to almost zero in the regime $10^{12.5} < M_{200} \leq 10^{13.5}\mathrm{M}_\odot$. This is potential evidence that SFRs in massive systems are not keeping pace with the increasing baryonic inflow rates for increasing halo mass at fixed redshift (e.g, Correa et al., 2015). BHARs show a less continuous behaviour however, broadly categorised by two rudimentary states: BHs residing in haloes below $\approx 10^{11.5}\mathrm{M}_\odot$ are typically accreting at a 'low' rate ($\dot{M}_{\mathrm{BH}} \ll 10^{-4}$ $\mathrm{M}_\odot$ yr^{-1}); BHs residing in haloes more massive than $\approx 10^{12.5}\mathrm{M}_\odot$ tend to be accreting at a 'high' rate ($\dot{M}_{\mathrm{BH}} > 10^{-4}$ $\mathrm{M}_\odot$ yr^{-1}). The fraction of galaxies with $\dot{M}_{\mathrm{BH}} \geq 10^{-4}\mathrm{M}_\odot$ yr^{-1} for a given halo mass bin is $\approx 3\%$, 21%, 55%, 70% and 78% from $10^{11.0} < M_{200} \leq 10^{11.5}\mathrm{M}_\odot$ to $10^{13.0} < M_{200} \leq 10^{13.5}\mathrm{M}_\odot$ respectively. Those in haloes between the mass range $10^{11.5} \sim 10^{12.5}\mathrm{M}_\odot$ are in an intermediate state.

A fraction of galaxies hosted by haloes with $M_{200} \gtrsim 10^{11.5}\mathrm{M}_\odot$ yield extremely low, or even zero SFRs. As all of these galaxies host massive BHs ($M_{\mathrm{BH}} \geq 10^7\mathrm{M}_\odot$, green dots), we are most likely seeing the effect of recent episodes of violent AGN feedback that have severely reduced the current star-forming capabilities of these systems. The cause, prevalence and impact of these feedback events will be the subject of a future paper.

We now investigate if the growth rate to halo connection evolves. To do this we return to the M_{BH}–M_{200} relation shown for $z = 0$ in Figure 4.2. The population is sub divided into two-dimensional bins, 0.5 dex on a side and outlined as squares. Here we investigate nine bins that lie along the median track through a continuous range spanning $10^{11.5} < M_{200} < 10^{13.5}\mathrm{M}_\odot$ in halo mass and $10^{6.0} < M_{\mathrm{BH}} < 10^{9.0}\mathrm{M}_\odot$ in BH mass, each outlined with a unique colour to reference their histories in Figures 4.8 to 4.10.

Figure 4.8 shows the evolution of the time-averaged SFR (top panel) and time-averaged BHAR (bottom panel) as a function of halo mass for all galaxies that come

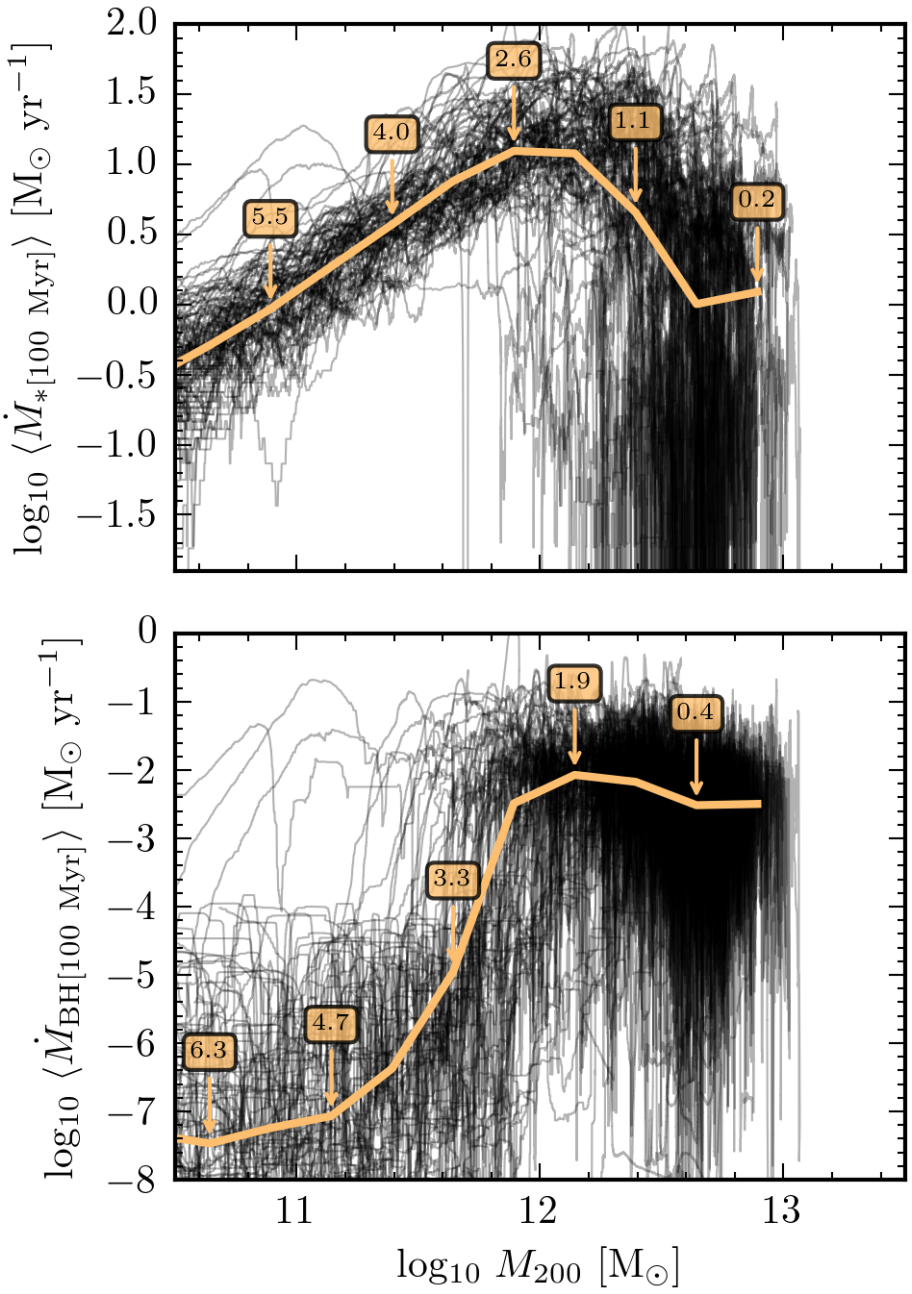

Figure 4.8: The evolutionary history of SFR (top panel) and BHAR (bottom panel) as a function of halo mass for all galaxies that come to reside in the $10^{12.5} < M_{200} < 10^{13.0}\mathrm{M}_\odot$, $10^{8.0} < M_{\mathrm{BH}} < 10^{8.5}\mathrm{M}_\odot$ two-dimensional bin of Figure 4.2 at $z = 0$ (outlined in yellow). Each black line is an individual history. The orange line shows the median trend, annotated with the median redshift at which these galaxies were hosted in haloes of that mass. For each panel, growth rates are time-averaged over 100 Myr as to overcome the noise induced when considering instantaneous rates. We see very different evolutionary behaviour for SFR and BHAR as the halo grows. SFRs initially rise and then decline, centred around $M_{200} \sim 10^{12}\mathrm{M}_\odot$. BHARs similarly transition from a low to high rate around this halo mass.

to reside in one of these two-dimensional bins at the present day (each solid black line is an individual history). We time-average both SFR and BHAR over 100 Myr in order to remove the inherent noise when considering instantaneous growth rates, and unveil the average trend. To eliminate galaxies that were previously classified as satellites of a more massive halo, we only consider central galaxies that have evolved monotonically in their halo mass (this excludes only $\approx 1\%$ of the $z = 0$ centrals population). We see, that although individual histories can be quite different, on average galaxies and their central BHs do follow a well defined path. The median SFR and BHAR of this population subset for a given halo mass are over-plotted in yellow, annotated by the median redshift at which they were hosted by haloes of that particular mass. As expected, an increasing halo mass corresponds to a decreasing redshift.

There is a striking difference in behaviour seen between the two growth rates as the halo grows. Initially the SFR increases steadily with halo mass. As the halo grows more massive than $\approx 10^{12}\mathrm{M}_{\odot}$ the SFRs begin to fluctuate between high and low values, yet overall there is a gradual decline of the median trend after this mass. Similarly, BHARs also change their behaviour around $\approx 10^{12}\mathrm{M}_{\odot}$, rapidly transitioning from a low ($\dot{M}_{\mathrm{BH}} \ll 10^{-4}\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$) to high ($\dot{M}_{\mathrm{BH}} > 10^{-4}\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$) rate. As with SFRs, BHARs decline a similar amount after the halo mass $\approx 10^{12}\mathrm{M}_{\odot}$ (note the many orders of magnitude difference in the scale of the growth rate axis between the two panels). We interpret therefore, given that the decline of SFR coincides with the peak of the rapid increase in BHAR, that AGN feedback is impeding the continued rise of SFRs in the most massive systems (see Figure 4.1 for an individual example of SFR reduction after the peak AGN activity at lookback time ≈ 12). We note that the decrease in halo mass accretion rate with declining redshift and the dependence of halo cooling rates on halo mass will play *additional* roles in shaping these histories. However, given the severity of the SFR reduction seen immediately after the BHAR peak, AGN feedback appears to be a dominant factor in hindering further galaxy growth.

Figure 4.9 extends this analysis to each of the highlighted two-dimensional bins in Figure 4.2, now showing only the median lines for clarity. Remarkably, the evolutionary behaviour is similar regardless of the final position in the $M_{\rm BH}$–M_{200} plane. The normalisation of each history is set by the evolving baryonic inflow rate at fixed halo mass. As this rate decreases with redshift (e.g, Correa et al., 2015), so does the normalisation of both the SFR and BHAR seen here (as each population reaches a particular halo mass at different times). We include also in the bottom panel of Figure 4.9 the median ratio between SFR and BHAR shown in the two panels above. This shows that galaxy growth is dominant over BH growth by many orders of magnitude in low mass haloes ($M_{200} \lesssim 10^{11.5}\mathrm{M}_{\odot}$). As BHARs settle to their 'high' rate in haloes of a mass above $M_{200} \sim 10^{12}$, the ratio between SFR and BHAR plateaus to an approximately constant value of $\sim 10^{3}$. Note the trends of both Figures 4.8 and 4.9 are not directly observable as they rely on median time-averaged growth rates in both SFR and BHAR of 100 Myr whilst also being binned by halo mass.

4.5 Discussion

Throughout this investigation we have consistently found no evidence supporting a simple underlying relationship between the rate of a galaxy's star formation and the accretion rate of its central BH. Instead, a mutual dependence of each property upon the mass of the host halo yields a more complex connection. It is interesting to examine, then, how the relation between the SFR and BHAR evolves for individual objects. In the following discussion, we will provide a physical interpretation based on the Bower et al. (2017) (hereafter B16) model for BH growth (for a similar interpretation on the importance of SN feedback to BH growth see Dubois et al., 2015; Habouzit et al., 2016). However, we stress the simulation results are themselves independent of any physical interpretation.

Figure 4.10 equates the median trends of the SFR and BHAR histories shown in

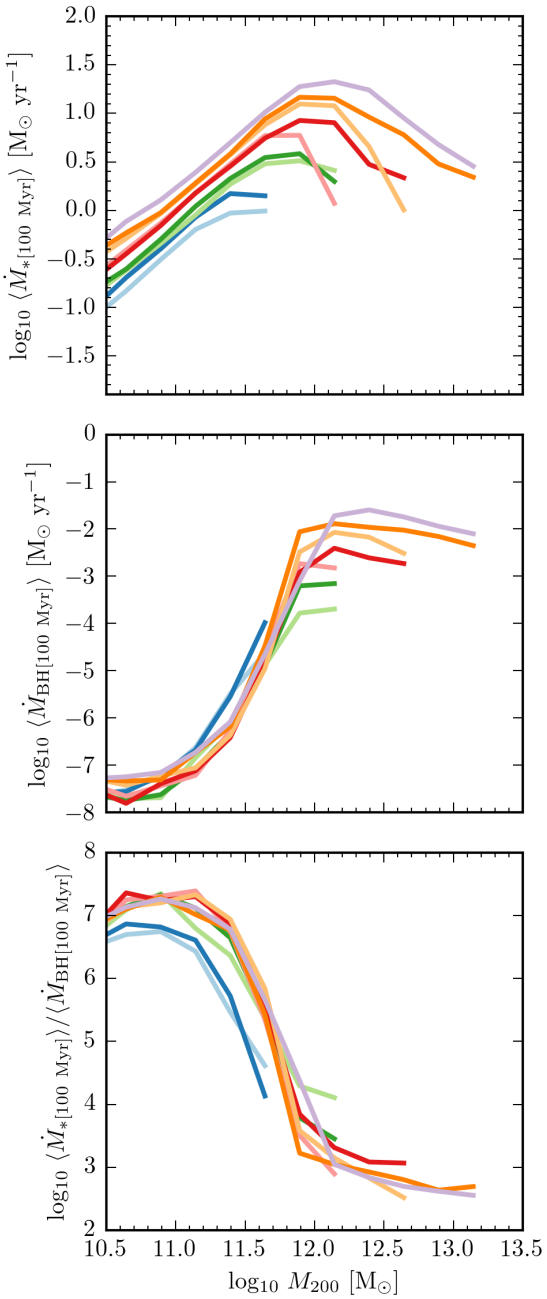

Figure 4.9: *top two panels*: A continuation of the analysis in Figure 4.8 to each of the nine chosen two-dimensional bins of Figure 4.2. The lines are the median track, with the colour corresponding to the outline in Figure 4.2. Regardless of where a galaxy is located on the M_{BH}–M_{200} plane at the present day, both the galaxy and its central BH evolve similarly, though different from each other. The change in normalisations between the histories is due to the declining baryonic inflow rates with decreasing redshift for a fixed halo mass. *bottom panel*: The median ratio between the SFR and BHAR from the two panels above. SFRs are initially dominant by many orders of magnitude in low mass haloes ($M_{200} \lesssim 10^{11.5}\mathrm{M}_\odot$), coming to plateau at an approximately constant value of $\sim 10^3$ in high mass haloes ($M_{200} \gtrsim 10^{12.5}\mathrm{M}_\odot$).

Figure 4.9. This specifies the 100 Myr average SFR as a function of the 100 Myr average BHAR in equal spacings of halo mass. Three distinct trends between SFR and BHAR emerge as the halo evolves: the *stellar feedback regulated* phase (shaded blue), the *non-linear BH growth* phase (shaded green) and the *AGN feedback regulated* phase (shaded red).

- *Region A - The stellar feedback regulated phase*: From the time of their seeding until they are hosted by haloes of mass $M_{200} \sim 10^{11.5}\mathrm{M}_{\odot}$ the BH accretion rates are negligible ($\dot{M}_{\mathrm{BH}} \leq 10^{-6}\mathrm{M}_{\odot}$ yr^{-1} on average). By contrast, SFRs increase steadily with halo mass. This behaviour produces the uncorrelated (yet causally connected) ∼vertical trend in region *A*, creating an imbalance of growth within these systems. As a result, BHs remain close to their seed mass whilst the halo/galaxy continues to grow around them (see the low-mass region of Figure 4.2).

  B16 interpret galaxies in this regime as being in a state of regulatory equilibrium. Energy injected by stars heats the ISM within the stellar vicinity, ejecting it, and causing it to rise buoyantly in the halo. This in turn creates an outflow of material balancing the freshly sourced fuel from the cosmic web, and as such prevents large gas densities from building up within the inner regions of these low-mass galaxies. Such low densities, coupled with the relatively low mass BHs living within these galaxies (BHAR $\propto M_{\mathrm{BH}}^2$), ensures that BHs fail to grow substantially.

- *Region B - The non-linear BH growth phase*: Both galaxies and BHs grow through the halo mass range $M_{200} \sim 10^{11.5} - 10^{12.0}\mathrm{M}_{\odot}$. However, whereas the SFRs continue to increase steadily with increasing halo mass, BHs rapidly transition to a non-linear phase of growth. This creates a highly non-linear *indirect* correlation between SFR and BHAR, connected through the host halo mass.

  The physical interpretation posited by B16 is that haloes that grow to the

transition mass, $M_{\rm crit}$, have become sufficiently massive to stall the regulatory outflow. Due to (what is now) the halos' hot coronae, heated gas ejected by stellar feedback loses the capability to rise buoyantly and therefore returns to the galaxy centre. Densities in the central regions of the galaxy are no longer kept low and a 'switch' to non-linear BH growth is triggered.

- *Region C - The AGN feedback regulated phase*: For haloes with masses above $M_{200} \sim 10^{12}{\rm M}_{\odot}$ SFRs and BHARs both decline on average, correlated with an approximately linear trend (compare to green dashed line, see also bottom panel of Figure 4.9).

  B16 argue that BHs in these haloes have become sufficiently massive (through their rapid non-linear growth) to efficiently regulate the gas inflow onto the galaxy themselves via AGN feedback. This again creates an equilibrium state, for which a fluctuating low level of (specific) BH accretion is maintained, keeping the outer halo hot and evaporating much of the new cold material trying to enter the system from the intergalactic medium.

Galaxies and their central BHs within the EAGLE simulation transition through multiple stages of growth as their host dark matter halo evolves, creating three distinct behaviours between SFR and BHAR. This is a stark contrast to a simple model where SFR and BHAR correlate globally via a linear relation, on average and for all halo masses. Whilst the underlying trend is only revealed when each growth rate is time-averaged (given the inherent noise of instantaneous growth rates), we only find an approximately linear correlation for the most massive systems ($M_{200} \gtrsim 10^{12.5}{\rm M}_{\odot}$).

In this paper we have emphasised the role of the halo and how its interaction with both SFR and BHAR shapes the growth rate relationship. However, additional factors may also contribute to the form this relationship takes. For example, Volonteri et al. (2015b) find using a suite of isolated merger simulations at fixed halo mass, that alternate behaviours between SFR and BHAR before, during and after the merger

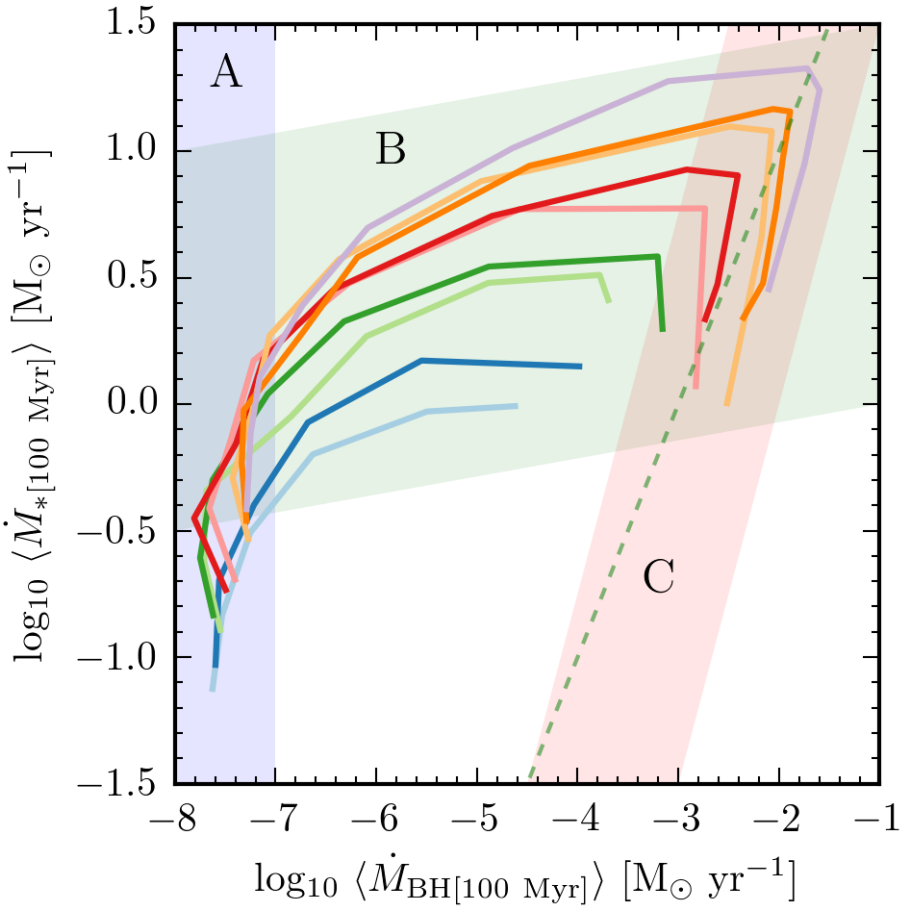

Figure 4.10: Each line shown here equates the median trends from the top two panels of Figure 4.9 to give the 100 Myr average SFR as a function of the 100 Myr average BHAR (in equal spacings of halo mass). Region A (shaded blue) corresponds to galaxies hosted by haloes with $M_{200} \lesssim M_{\mathrm{crit}}$. Galaxies in this regime increase their SFR with increasing halo mass, while BHARs remain negligible on average. As haloes reach $\sim M_{\mathrm{crit}}$ in region B (shaded green), SFRs continue to rise, however the BH growth increases by many orders of magnitude over this narrow halo mass range. For haloes in excess of $\gtrsim M_{\mathrm{crit}}$ shown in region C (shaded red), we see a reduction for both SFR and BHAR on average, yielding a approximately constant scaling between the two growth rates (compare to dashed green line which shows the linear relation $\dot{M}_{\mathrm{BH}}/\dot{M}_* = 10^{-3}$).

proper collectively contribute to form a complex two-dimensional plane. Additionally, Pontzen et al. (2017) reveal the particular importance differing merger histories can have on significantly altering the growth rate history of both that of the galaxy and the central BH. However, the global influence of mergers upon galaxy and BH growth rates in a full cosmological context remains open for debate, and will be the subject of a future paper.

4.6 Conclusions

We have investigated the relationship between the galaxy star formation rate (SFR) and the black hole accretion rate (BHAR) of the central black hole (BH) using the EAGLE cosmological hydrodynamical simulation. Our main conclusions are as follows:

- We compared EAGLE predictions to two recent observational studies in Figures 4.3 and 4.4. The simulation reproduces both the flat trend of the mean SFR ($\langle$SFR$\rangle$) as a function of BHAR found in the AGN selected study of Stanley et al. (2015) and the approximately linear trend of the mean BHAR ($\langle$BHAR$\rangle$) as a function of SFR found in the SFR selected study of Delvecchio et al. (2015).

- There is a moderate difference in the $\langle$SFR$\rangle$–BHAR relationship when time-averaging each growth rate over a 100 Myr period for an AGN selected study (Figure 4.5). However, this change was not found to be sufficient as to revert the trend to an underlying linear relationship as has been proposed by previous theoretical studies.

- Examining the complete $z = 1$ SFR–BHAR plane in Figure 4.6, we found no evidence for a simple universal global relationship between the two instantaneous growth rates. The difference between the trends found for the

⟨SFR⟩–BHAR and ⟨BHAR⟩–SFR relations from AGN and SFR selections respectively, is due to sampling different regions of this complex plane. The complexity of this plane results from both the rate of galactic star formation and the accretion rate of the central BH holding an evolving connection to the host dark matter halo (Figure 4.7).

- For a discrete redshift, the characteristic SFR of a halo increases smoothly with increasing halo mass (Figure 4.7). BHs in haloes of mass $M_{200} \lesssim 10^{11.5}\mathrm{M}_{\odot}$ accrete at a 'low' rate ($\dot{M}_{\mathrm{BH}} < 10^{-4}\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$). They then transition through haloes of mass $10^{11.5} \sim 10^{12.5}\mathrm{M}_{\odot}$ to a 'high' rate ($\dot{M}_{\mathrm{BH}} > 10^{-4}\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$) in haloes of mass $M_{200} \gtrsim 10^{12.5}\mathrm{M}_{\odot}$. However, the scatter in the BHAR at fixed halo mass is very large (up to ~ 6 dex). Galaxies with SFRs far below the characteristic SFR all contain massive BHs ($M_{\mathrm{BH}} \geq 10^{7}\mathrm{M}_{\odot}$).

- The median evolutionary trend for a galaxy's SFR and the accretion rate of its central BH, averaged over 100 Myr, are insensitive to the final properties of the system (Figure 4.9). By equating these trends together we found that the 100 Myr average SFR as a function of the 100 Myr average BHAR can be split into three regimes, separated by the halo mass (Figure 4.10). BHs hosted by haloes below the characteristic transition mass, M_{crit} (Bower et al., 2017M_{200} $\sim 10^{12}\mathrm{M}_{\odot}$), fail to grow effectively, yet the galaxy continues to grow with the halo. Once the halo reaches M_{crit} there is a non-linear 'switch' of BH growth that rapidly builds the mass of the BH. In the most massive haloes (M_{200} $> M_{\mathrm{crit}}$) both SFR and BHAR decline on average, with a roughly constant scaling of SFR/BHAR $\sim 10^{3}$.

Chapter 5

Strongly star-forming galaxies in the EAGLE simulation: triggering, descendants and submillimetre properties

5.1 Introduction

A consensus has been established towards the existence of a well defined approximately linear relationship between the star formation rate (SFR) and the stellar mass of star-forming galaxies. This so-called 'main-sequence' is apparent in both the low- ($z < 1$, e.g., Brinchmann et al., 2004; Salim et al., 2007) and high-redshift Universe ($z \approx 2$, e.g., Daddi et al., 2007), and is recovered independently by observations in many different wavebands; including the optical (e.g., Tasca et al., 2015), infrared (e.g., Elbaz et al., 2011; Rodighiero et al., 2014) and radio (e.g., Karim et al., 2011). The main-sequence has traditionally been parameterised in the form of a single power law, $\log_{10}\dot{M}_* = \alpha \log_{10} M_* + \beta$, with discovered slopes close to unity ($\alpha \approx 0.4$–1, e.g., Whitaker et al., 2014; Schreiber et al., 2015; Tomczak et al., 2016). However, studies have also suggested a second, shallower relation, at the high-mass end ($M_* \gtrsim 10^{10.5}$,

e.g., Lee et al., 2015; Tomczak et al., 2016). Regardless of the functional form chosen, the normalisation (β) of the main-sequence is observed to increase with increasing redshift (e.g., Speagle et al., 2014; Johnston et al., 2015; Schreiber et al., 2015; Tomczak et al., 2016). As the bulk of the star-forming population is seen to reside on or around this relation (with a scatter of around $\approx$ 0.2–0.4 dex, e.g., Rodighiero et al., 2011; Speagle et al., 2014; Schreiber et al., 2015), it has made the main-sequence synonymous with the 'typical' SFR of star-forming galaxies at a particular redshift.

Being able to replicate this well established relationship is therefore essential for hydrodynamical simulations of galaxies. Several simulations have recovered a tight relationship between the SFR and the stellar mass of star-forming galaxies (e.g,. Davé et al., 2011; Torrey et al., 2014), however, the observed normalisation of the main-sequence, and in particular the evolution of the normalisation, has been notoriously difficult to achieve. This discrepancy, whilst reduced, remains a tension in the current cosmological scale hydrodynamical simulations of today (e.g., Sparre et al., 2015; Furlong et al., 2015c).

Although the majority of star-forming galaxies occupy the locus of the main-sequence, there are significant outliers that yield uncharacteristically high SFRs for their stellar mass, commonly referred to as 'starbursting' galaxies. It has been suggested that main-sequence galaxies and starbursting galaxies are in fundamentally different star forming regimes: a smooth accretion of gas from the intergalactic medium forming a long lasting quasi-equilibrium state, and a more rapid, starburst mode, potentially created through major mergers or in relation to the density of the environment (e.g., Daddi et al., 2010; Genzel et al., 2010). However, such extreme starbursting objects are rare ($\approx$ 2% of the galaxy population) and are not thought to dominate the cosmic SFR density ($\lesssim$ 10%, Rodighiero et al., 2011).

One potential candidate of starbursting galaxies is the sub-millimeter (submm) galaxy (SMG) population. SMGs are a set of high-redshift galaxies with high SFRs and significant dust masses. The submm emission arises from the reprocessing of the UV light from young stars by the dust (see Casey et al., 2014 for a review). The

SMG population are relatively rare, with a number density of $\sim 10^{-5}\,\mathrm{h}^{-1}\,\mathrm{cMpc}^{-3}$ at $z \approx 2$–3 (e.g., Chapman et al., 2005; Simpson et al., 2014). However, this population remains particularly interesting due to the apparently high levels of star formation and the many open questions that remain unanswered about their formation and evolution. For example: What triggers such extreme star formation, mergers or disk instabilities? How long can these galaxies form stars at such apparently high rates? How do they evolve following the starburst episode?

In this study we investigate the prevalence of strongly star-forming galaxies in the largest simulation of the EAGLE project. This simulation was calibrated to reproduce the observed galaxy stellar mass function, galaxy sizes and black hole–stellar mass relation at $z \approx 0$ (Crain et al., 2015; Schaye et al., 2015). Many other observed galaxy trends both locally (e.g., Schaye et al., 2015; Bahé & McCarthy, 2015; Trayford et al., 2015) and at higher redshift (e.g., Furlong et al., 2015c; Rahmati et al., 2015) have also shown broad agreement with observations. Predictions from the simulation for the full model galaxy population are reasonably representative of the observed Universe, as a result this simulation presents an interesting testbed for more extreme populations, such as those galaxies with the highest SFRs, and their potential to match the SMG population. Observationally it is claimed that these two populations are strongly overlapping, however we can test this theory in the model.

This paper is organised as follows: In Section 5.2 we provide a brief overview of the EAGLE simulation and discuss our sample selections. In Section 5.3.1 we investigate what triggers the most strongly star-forming galaxies within the simulation an explore if star-forming galaxies are a fundamentally distinct population relative to the 'typically' star-forming galaxies. We then go on to investigate the submm properties of strongly star-forming galaxies in Section 5.3.2. Finally, in Section 5.4 we discuss our results and in Section 5.5 we present our concluding remarks.

5.2 The EAGLE simulation

A full overview of the EAGLE simulation suite can be found in Chapter 2. For this chapter we are interested in the most strongly star-forming galaxies ($\dot{M}_* \geq 80$ $\mathrm{M}_\odot\ \mathrm{yr}^{-1}$), and therefore restrict our study to the largest simulation, Ref-L0100N1504, which contains the greatest number of these objects.

The completion time of a galaxy–galaxy merger is defined as the cosmic time of the first simulation output where two galaxies that were previously identified as separate individually bound objects are now identified as a single bound object by the SUBFIND algorithm. Mergers are classified by their stellar mass ratio, $\mu = M_{*,1}/M_{*,2}$, where $M_{*,2}$ is the mass of the most massive member of the binary. They are considered major if $\mu \geq \frac{1}{4}$, minor if $\frac{1}{10} \leq \mu < \frac{1}{4}$ and either major or minor if $\mu \geq \frac{1}{10}$. To account for the stellar stripping that occurs during the later stages of the interaction, the stellar mass ratio is computed when the in-falling galaxy had its maximum mass (e.g., Rodriguez-Gomez et al., 2015; Qu et al., 2017). Additionally, mergers are only considered 'resolved' when $M_{*,2} \geq 10^8$ $\mathrm{M}_\odot$ (≈ 100 stellar particles).

Mock observables

The light observed from a galaxy is a combination of intrinsic and scattered light. Therefore to accurately compare the observables of model galaxies to the observations, we are required to solve the complete three-dimensional radiative transfer problem. For EAGLE, this is performed in post-processing using information from the star-forming regions, stellar sources, and the diffuse dust distribution of the model galaxies using the radiative transfer code SKIRT (Baes et al., 2011; Camps & Baes, 2015). This produces mock observables for the model galaxies from the ultraviolet to the submillimeter (submm) wavebands[1]. The radiative transfer post processing is only performed for the model galaxies within the 29 snapshot outputs.

[1]The rest-frame magnitudes and observer-frame fluxes for all EAGLE galaxies with stellar masses greater than $10^{8.5}$ $\mathrm{M}_\odot$ are publicly available (Camps et al., 2018).

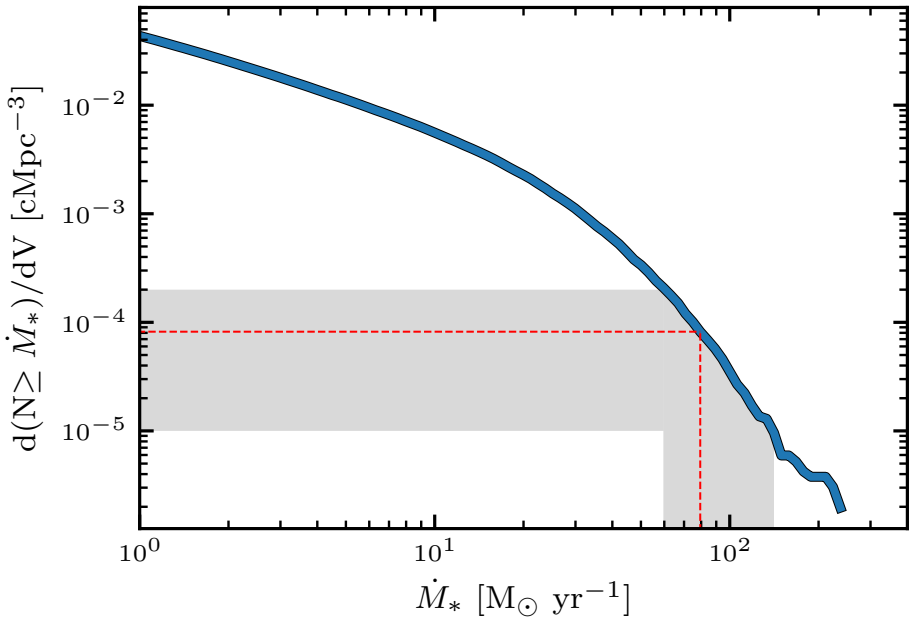

Figure 5.1: The cumulative number density of all star-forming galaxies in the EAGLE simulation ($0 < z < 20$), plotted as a function of the SFR ($\dot{M}_*$). The horizontal gray band highlights the range of space densities of strongly star-forming galaxies inferred from a variety of observations ($\sim 2 \times 10^{-4}$–10^{-5} cMpc^{-3}, see Section 5.2.1) and the vertical gray band highlights the cumulative SFRs of model galaxies these space densities correspond to ($\approx$ 60–140 M$_\odot$ yr^{-1}). Our selection criteria for strongly star-forming galaxies ($\dot{M}_* \geq 80$ M$_\odot$ yr^{-1}, corresponding to a space density in EAGLE of 8×10^{-5} cMpc^{-3}) is indicated by the dashed line.

For the analysis in Section 5.3.2 we compare to the observed submm population using the SKIRT inferred observed-frame fluxes at 450μm and 850μm ($S_{450\mu\mathrm{m}}$ and $S_{850\mu\mathrm{m}}$, columns SCUBA2_450 and SCUBA2_850 in the public database McAlpine et al., 2016; Camps et al., 2018).

The star formation histories of EAGLE galaxies from the individual stellar particles

5.2.1 Selecting strongly star-forming model galaxies

The reported 'instantaneous' SFRs of EAGLE galaxies are computed from the current state of the galaxy's associated gas particles. However, it is also possible to compute

the SFR of a galaxy at a given time by reverse engineering its SFR history from the stellar particles. By collectively binning the stellar particles of the galaxy at $z = 0$[2] by their birth time, weighting by their initial mass and dividing by the bin width, a SFR history is obtained, which is limited only by the total number of stellar particles used (i.e., the $\approx z = 0$ galaxy stellar mass). For galaxies more massive than 10^{10} $M_\odot$ this can adequately resolve the SFR history down to intervals of ≈ 1 Myr, which is an order of magnitude better than sampling from the snipshot output. This method can be used to accurately study the SFR histories of individual galaxies (for example see Figure 4.1), or simply as a robust method to obtain the maximum SFR ever achieved by a galaxy throughout its lifetime.

We wish to use EAGLE to study the properties of the most strongly star-forming galaxies ($\dot{M}_* \gtrsim 100$ $M_\odot$ yr^{-1}). For this we apply a selection to isolate the most strongly star-forming galaxies, adopting a space density which is representative of the limits on strongly star-forming populations derived from a variety of observational techniques. We take this approach as the various observational techniques used to estimate the space density of strongly star-forming galaxies at $z > 1$ suffer from a range of biases or shortcomings. For example, wide-field, narrow-band Hα surveys estimate space densities of $\sim 10^{-4}$ cMpc^{-3} at $z \geq 1.5$ (Sobral et al., 2013), however, many of these bright Hα sources are AGN (Sobral et al., 2016), while narrower field spectroscopic and grism surveys sample much smaller volumes and the grism surveys do not cover Hα beyond $z \approx 1.5$ (e.g., Atek et al., 2010; Brammer et al., 2012; Pirzkal et al., 2013). Equally, surveys which select sources on the basis of their mid- or far-infrared luminosity (as a proxy for their SFR) have reported space densities of $\sim 2 \times 10^{-4}$–10^{-5} cMpc^{-3} for galaxies with far-infrared luminosities of $\gtrsim 10^{12}$ $L_\odot$ and inferred SFRs of $\gtrsim 100$ $M_\odot$ yr^{-1} at $z \approx 1.5$–2.5 (e.g., Chapman et al., 2005; Magnelli et al., 2011; Casey et al., 2012; Gruppioni et al., 2013; Swinbank et al., 2014; Koprowski et al., 2017). Again, these studies suffer from a combination of

[2]Note that only the stellar particles born within the main-progenitor galaxy are considered, so as to avoid combining the SFR histories of multiple progenitors.

AGN contamination of the derived luminosities (a particular problem in the mid-infrared, e.g., Kirkpatrick et al., 2012, 2015; Del Moro et al., 2013), or blending and misidentification of the correct galaxy counterparts in low-resolution far-infrared and sub-millimeter surveys (e.g., Hodge et al., 2013; Casey et al., 2014; Simpson et al., 2014).

ALMA is starting to address some of these concerns, but has so far only undertaken surveys at sub-/millimeter wavelengths and therefore selects sources on a combination of dust temperature and far-infrared luminosity, rather than purely on the SFR. Nevertheless, current sub-/millimeter studies of blank fields, as well as follow-ups of wide-field single-dish sub-millimeter surveys yields space densities of ~ 1–5×10^{-5} cMpc^{-3} at equivalent 870 μm fluxes of 1–4 mJy, equivalent to inferred SFRs of 70–300 M$_\odot$ yr^{-1} (e.g., Swinbank et al., 2014; Aravena et al., 2016; Walter et al., 2016; Dunlop et al., 2017; Stach et al., 2018). Given the reliability of the identification of the counterparts to the sub-/millimeter sources provided by ALMA's sub-arcsecond spatial resolution, we will use these samples later in our comparison to the predicted properties from EAGLE.

Given the range and uncertainties in the various estimates of the space density for strongly star-forming galaxies from the observations, we have chosen to select an equivalent SFR limit which roughly corresponds to the space densities derived from the observations, $\sim 2 \times 10^{-4}$–10^{-5} cMpc^{-3}, and which allows us to isolate a sufficient sample ($\gtrsim 100$ galaxies) of strongly star-forming galaxies to allow for a statistical analysis (see Figure 5.1). Our adopted SFR limit is ≥ 80 M$_\odot$ yr^{-1}, which corresponds to a typical space density of star-forming galaxies in EAGLE of 8×10^{-5} cMpc^{-3} (indicated in Figure 5.1). Nevertheless, as we show in Section 5.3.2, the predicted properties of these strongly star-forming model galaxies are reasonably well-matched to the multiwavelength properties of the observed populations, supporting their use to address the questions which the observations either have difficulty answering, or cannot answer (e.g., triggering, evolutionary connections, etc).

We construct two samples: those galaxies that were identified to be strongly star-

forming at the times of the snapshot outputs (129 galaxies) and those galaxies that were identified to be strongly star-forming at the times of the snipshot outputs (962 galaxies). In total, 174 *unique* main-progenitor branches are identified as having hosted a strongly star-forming galaxy at least once along their history within the snipshot sample, compared to only 82 identified from the snapshot sample. The discrepancy is due to the temporal spacing of the output and their relative ability to capturing galaxies in a finite strongly star-forming phase. However, we confirm that the temporal spacing of the snipshots ($\approx$ 40–80 Myr) is sufficient to capture $> 95\%$ of the galaxies that were ever strongly star-forming (confirmed by the galaxy's maximum SFR derived from the stellar particles, see Section 5.2), giving us an almost fully complete sample of strongly star-forming galaxies in the simulation. Other than for Section 5.3.2, where we use the snapshot sample for the mock observable properties, the snipshot sample is used throughout this work.

Figure 5.2 demonstrates where, relative to the star forming 'main-sequence', strongly star-forming galaxies lie in the redshift range $2 < z < 3$. We define the main-sequence as the median SFR, for a given range in stellar mass, of all 'non-quenched' galaxies. For this redshift range that includes all galaxies with specific star-formation rates (sSFR $= \dot{M}_*/M_*$) greater than $10^{-9.6}$ yr^{-1}, however this value does evolve (see Matthee & Schaye, 2018). This so-called 'main-sequence' is plotted on the figure. Due to its shape, strongly star-forming galaxies become increasingly offset from the main-sequence as the stellar mass decreases (up to $\approx$ 0.75 dex). Additionally, as the characteristic SFR of main-sequence galaxies increases with increasing redshift (Furlong et al., 2015c; Matthee & Schaye, 2018), the offset between strongly star-forming galaxies and the main-sequence universally decreases for all stellar masses with increasing redshift.

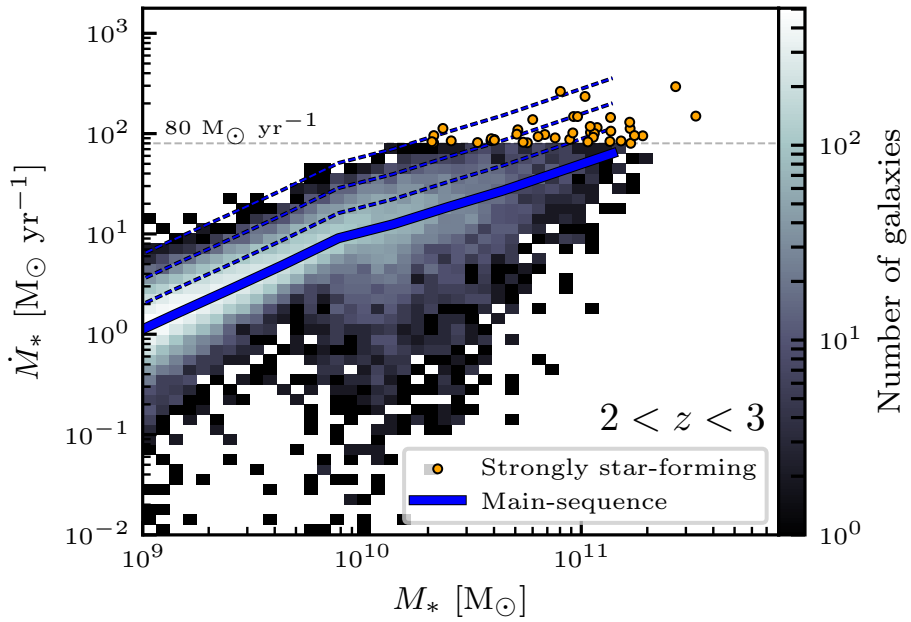

Figure 5.2: The SFR as a function of stellar mass for all galaxies in the redshift range $2 < z < 3$. The bulk of the population ($\dot{M}_* < 80$ $M_\odot$ yr^{-1}) is represented by a two-dimensional histogram, coloured by the number of galaxies in each bin. The galaxies meeting our strongly star-forming criteria ($\dot{M}_* \geq 80$ $M_\odot$ yr^{-1}) are highlighted individually as circles. The median value for all 'non-quenched' (sSFR $\geq 10^{-9.6}$ yr^{-1}) galaxies is shown as a solid line, which we refer to as the 'main-sequence'. The dashed lines show the median track offset +0.25, 0.50 and 0.75 dex in SFR, respectively. The most massive strongly star-forming galaxies ($M_* \approx 10^{11}$ $M_\odot$) lie up to $\approx$ 0.2 dex above the main-sequence, whereas less massive strongly star-forming galaxies ($M_* < 10^{11}$ $M_\odot$) lie $\approx$ 0.50–0.75 dex above the main-sequence.

5.3 Results

5.3.1 The nature of strongly star-forming galaxies

We begin by investigating the properties of the model galaxies within our sample around the time that they were strongly star forming. Our analysis focuses initially on the redshift range $2 < z < 3$, as this epoch contains the greatest number of strongly star-forming galaxies over the widest dynamic range of stellar masses within the simulation. It also brackets the peak of the cosmic SFR history (e.g., Madau & Dickinson, 2014) which is reproduced by EAGLE (Furlong et al., 2015c).

Figure 5.3 shows the evolution of the SFR and the stellar mass for the model galaxies within our sample that were strongly star-forming in the redshift range $2 < z < 3$. To provide a comparison, the equivalent evolution of the model galaxies on the main-sequence at these redshifts are also shown. Here, a galaxy is defined to reside on the main-sequence if the SFR is within ± 0.1 dex of the median value for all non-quenched galaxies for the given stellar mass (i.e., the line in Figure 5.2). The evolution is shown relative to the cosmic times when the model galaxies were strongly star-forming (or on the main-sequence), denoted t_0, and are tracked for 2.5 Gyr before and after this time. For each population, the galaxies are separated into three stellar mass bins at the time t_0: $M_*[t = t_0] = 0.2$–0.4, 0.4–0.9 and 1.0–2.0 $\times 10^{11}$ $\mathrm{M}_\odot$, with the bins containing 81, 155 and 71 strongly star-forming galaxies, respectively. These trends yield the typical characteristics of galaxies both before and after they were strongly star-forming, relative to the more 'typical' main-sequence population.

In the lowest stellar mass bin, there are strong differences in the evolutionary tracks between the two SFR classifications. By construction, the strongly star-forming and main-sequence populations are equivalent in their stellar masses at the time t_0, however, immediately following and preceding this time the tracks diverge. Strongly star-forming galaxies evolved from a set of galaxies that were on average less massive and less star-forming (both by ≈ 0.75 dex) than those on the main-sequence, yet

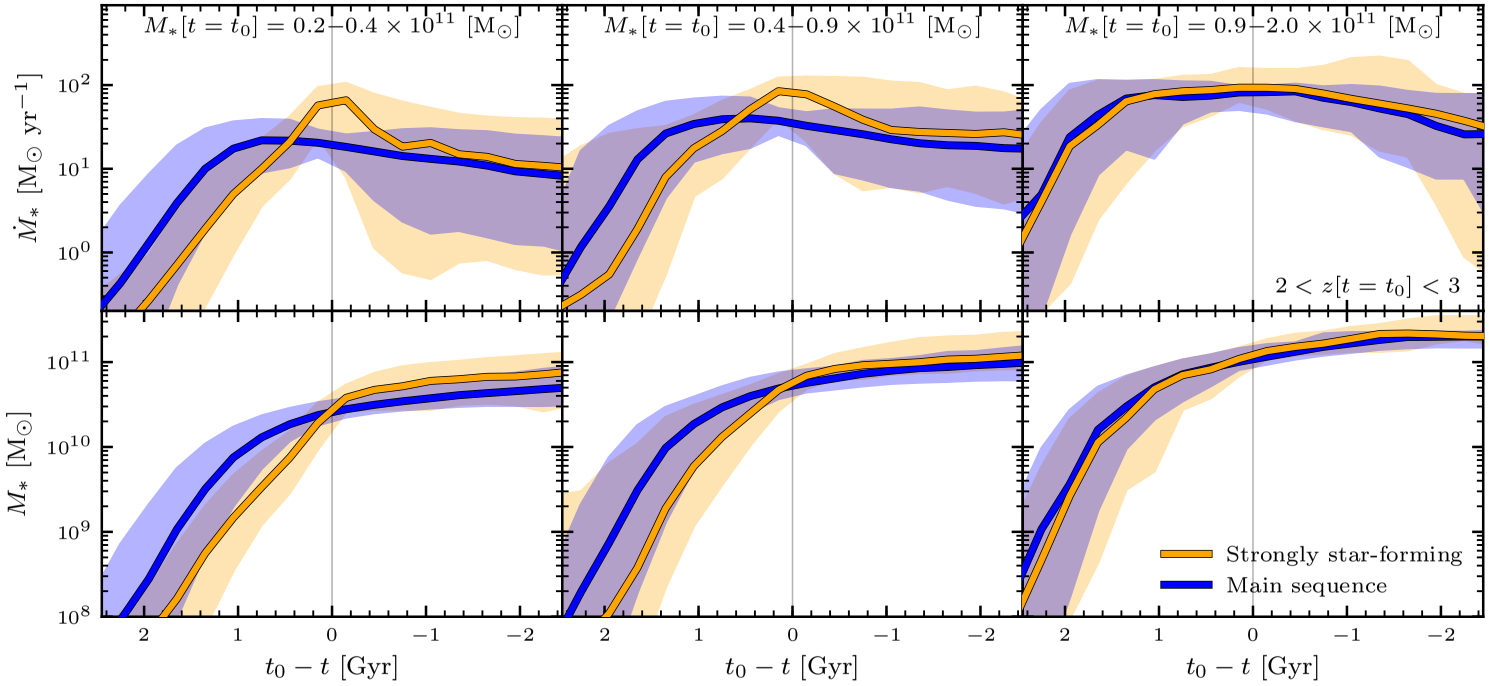

Figure 5.3: The evolution of the SFR (top) and the stellar mass (bottom) for the galaxies within our sample that were strongly star-forming in the redshift range $2 < z < 3$. For a comparison, the evolution of the galaxies on the main-sequence at this redshift are also shown. The evolution is measured relative to the cosmic times when the galaxies were strongly star-forming (or on the main-sequence), denoted t_0 (indicated by a vertical line). The galaxies are divided into three stellar mass bins at the time t_0: $M_*[t = t_0] = 0.2$–0.4, 0.4–0.9 and 0.9–2.0 $\times 10^{11}$ M$_\odot$ from the left to right panels, respectively. The solid lines are the median values, and the shaded regions outline the $10^{\text{th}} - 90^{\text{th}}$ percentile ranges. Strongly star-forming galaxies, whilst having higher SFRs and equivalent stellar masses relative to main-sequence galaxies at the time t_0 (by construction), evolved from lower mass galaxies with lower SFRs, and evolve into galaxies that are more massive and marginally more star forming. The size of the offsets between the two population lessens as the stellar mass increases. This is due to the main-sequence galaxies falling closer to the strongly star-forming selection at the highest stellar masses (see Figure 5.2). The evolution of further properties for the same set of galaxies is shown in Figure 5.3.

tend to evolve into galaxies that are more massive (by ≈ 0.25 dex) and marginally more star-forming (by ≈ 0.1 dex). These underlying behaviours are also present in the two more massive stellar mass bins; however, as the main-sequence galaxies fall closer to the strongly star-forming selection at the highest stellar masses (see Figure 5.2), the relative size of these offsets reduce. In fact, the offsets between the strongly star-forming and main-sequence population in the highest stellar mass bin are only marginal, suggesting that the strongly star-forming galaxies and 'typically' star-forming galaxies share the same star-formation process at this stellar mass.

To further explore the differences in the two populations we show the evolution of the black hole mass to stellar mass ratio ($M_{\rm BH}/M_*$), the total gas fraction ($M_{\rm gas}/M_{\rm gas+stars}$) and the stellar mass to halo mass ratio (M_*/M_{200}) in Figure 5.4. For this figure, we only show the evolution of the model galaxies in the lowest stellar mass bin of Figure 5.3 ($M_*[t = t_0] = 0.2$–0.4×10^{11} M$_\odot$), for clarity, however the properties evolve with increasing stellar mass in exactly the same manner as in Figure 5.3, whereby the relative size of the offsets between the two populations reduce.

The typical evolutionary pathway for the black holes hosted by the strongly star-forming model galaxies is very different than for those hosted by the main-sequence model galaxies. Both populations initially evolve towards the time t_0 with a declining trend, created because the masses of the black holes are remaining approximately constant as their host galaxies grow around them. This is due to the almost complete suppression of black hole growth in lower mass haloes ($M_{200} \lesssim 10^{12}$ M$_\odot$) from efficient stellar feedback (Bower et al., 2017, Chapter 3). On average, and we note there is a large scatter in the black hole properties of both populations, at ≈ 1.5 Gyr before t_0 for the main-sequence galaxies and at ≈ 0.5 Gyr before t_0 for the strongly star-forming galaxies, the hosted black holes then transition into their rapid growth phase, a phase of accelerated black hole growth that initiates within haloes of a critical mass ($M_{200} \sim 10^{12}$ M$_\odot$, see Chapter 3). As this phase has initiated later for the strongly star-forming population, their black holes at the time t_0 are considerably

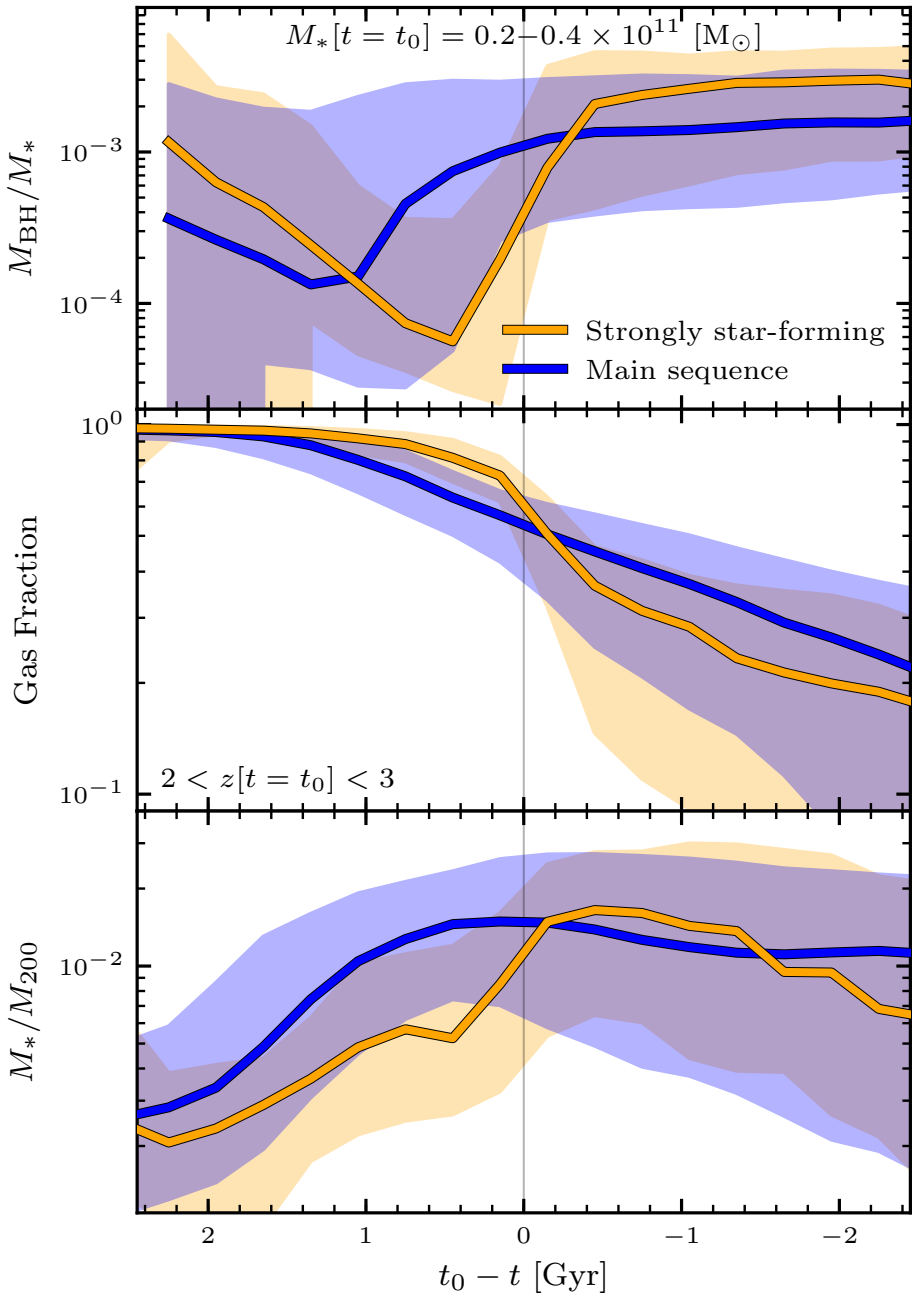

Figure 5.4: The format of this figure is identical to that of Figure 5.3, now including the evolution of: the black hole mass to stellar mass ratio ($M_{\rm BH}/M_*$, 1$^{\rm st}$ row), the total gas fraction ($M_{\rm gas}/M_{\rm gas+stars}$, 2$^{\rm nd}$ row) and the stellar mass to halo mass ratio (M_*/M_{200}, 3$^{\rm rd}$ row). For clarity, we only show the evolution of the galaxies in the lowest of the three stellar mass bins introduced in Figure 5.3 ($M_*[t = t_0] = 0.2$–0.4×10^{11} M$_\odot$). At the time t_0, the black holes hosted by the strongly star-forming galaxies are a factor of ≈ 5 more undermassive for their stellar mass relative to the main-sequence galaxies. The gas fractions of strongly star-forming galaxies are up to a factor of ≈ 3 higher than the main-sequence population before t_0, and up to a factor of ≈ 2 lower after t_0.

undermassive for their stellar mass relative to the main-sequence population (by a factor of ≈ 5). In fact, the majority of the black holes hosted by the strongly star-forming galaxies are still within their rapid growth phase at the time t_0 (and will therefore be luminous AGN, see Chapter 3), whereas the majority of the black holes hosted by the main-sequence galaxies have apparently completed this phase in their evolution. Additionally, as the mass accumulated during the rapid growth phase for the black holes hosted by the strongly star-forming galaxies is typically greater than for those hosted by the main-sequence galaxies, their black holes become (and remain) relatively overmassive for their stellar mass. Therefore it would appear that the black holes hosted by the strongly star-forming population are undergoing a more recent, and what will turn out to be a more vigorous, rapid growth phase, relative to their main-sequence counterparts.

The gas fractions of the strongly star-forming galaxies are up to a factor of ≈ 3 higher than the main-sequence population before the time t_0. This is inextricably linked to the evolution of the black hole, as the time at which the gas fractions begin to decline for the two populations is the same as when the hosted black holes initiate their rapid growth phase (and start to produce disruptive amounts of AGN feedback). As the decline happens earlier for the strongly star-forming population, an offset develops. After t_0, the gas fractions of the once strongly star-forming galaxies now develop to be up to a factor of ≈ 2 *lower* than the galaxies coming from the main-sequence. There are two causes to this effect: (1) the once strongly star-forming galaxies have had a higher fraction of their baryons converted into stars and into growing the black hole, and (2) the overmassive black holes that now reside within the once strongly star-forming galaxies produce an increased amount of AGN feedback, relative to those hosted by the galaxies once on the main-sequence.

Finally, at the time t_0 the stellar mass to halo mass ratio for the strongly star-forming galaxies is lower by a factor of ≈ 2 relative to the galaxies on the main-sequence (again linked to their higher gas fractions). After t_0, strongly star-forming galaxies then briefly become overmassive for their halo mass (due to the injected mass from

the starburst), and then decline to become undermassive for their halo mass (as their sSFRs are reduced due to their overmassive black holes and lower gas fractions), both relative to the main-sequence population. It is clear, then, that the evolution of strongly star-forming galaxies are distinct from the equivalent evolution of the 'typically' star-forming population.

Although only shown here for the galaxies in the redshift range $2 < z < 3$, these underlying behaviours are found to be universal throughout the entire redshift range covered by our strongly star-forming sample ($1 < z < 6$). As the characteristic SFR for 'typical' star-forming galaxies increases with increasing redshift (Furlong et al., 2015c; Matthee & Schaye, 2018), the evolution of strongly star-forming galaxies and main-sequence galaxies become less offset from each other with increasing redshift (in a similar manner to what was seen for an increasing stellar mass at fixed redshift in Figure 5.3). However, the majority of strongly star-forming galaxies, which are at lower stellar masses ($M_* \lesssim 10^{11}$ M$_\odot$), remain distinct from the main sequence population at all times, and the higher mass strongly star-forming galaxies ($M_* \gtrsim 10^{11}$ M$_\odot$) remain broadly consistent with the typically star-forming population.

The merger rates of strongly star-forming galaxies

In addition to the intrinsic differences in their galaxy properties, relative to the general population, interactions may also play an important role in triggering the most strongly star-forming galaxies. To investigate this, we include Figure 5.5, which shows the merger fraction for the galaxies within our sample that were strongly star-forming in the redshift range $2 < z < 3$. Here, the merger fraction is defined as the fraction of galaxies that have undergone or will undergo a merger within the previous/next dynamical time[3] The galaxies are separated into the same three stellar mass bins as Figure 5.3. The merger fraction is computed separately for major mergers ($\mu \geq \frac{1}{4}$), minor mergers ($\frac{1}{10} \leq \mu < \frac{1}{4}$) and major or minor mergers

[3]Equivalent to $|n_{\rm dyn}| \leq 1$ from eq. (3.2.1). The dynamical time, $t_{\rm dyn}$, is defined as the free fall time of the dark matter halo ($t_{\rm dyn} \approx 1.6$ Gyr at $z = 0$, ≈ 0.5 Gyr at $z = 2$ and ≈ 0.2 Gyr at $z = 5$), see Section 3.2.3.

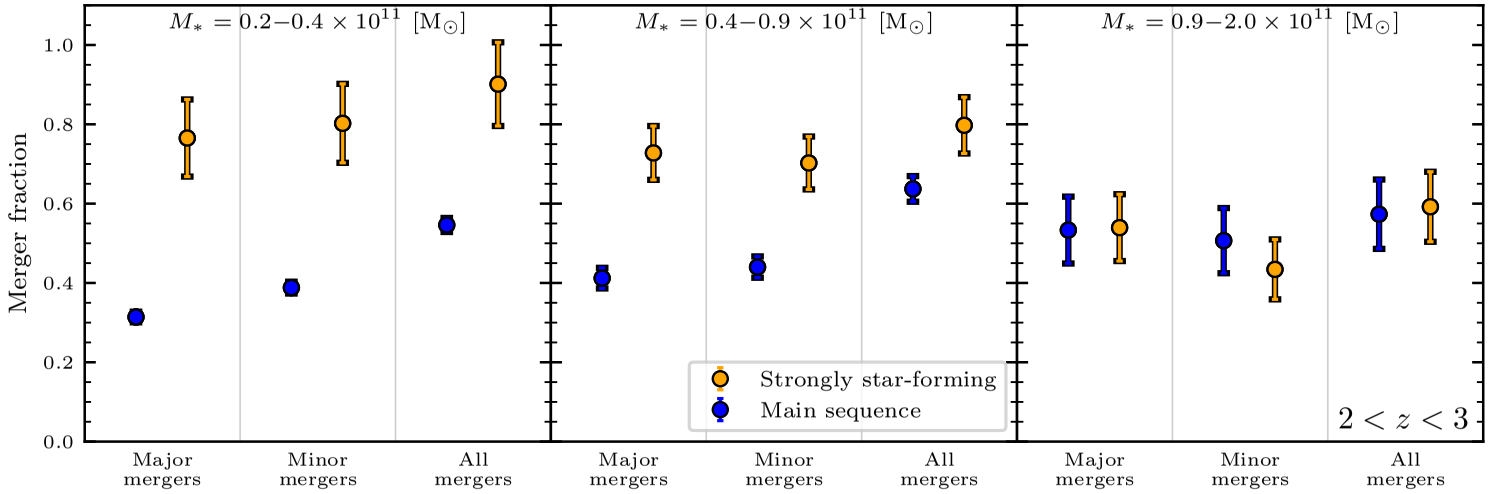

Figure 5.5: The merger fraction, defined as the fraction of galaxies that have undergone or will undergo a merger within the previous/next dynamical time, for the galaxies within our sample that were strongly star-forming in the redshift range $2 < z < 3$. As with Figure 5.3, a comparison to the galaxies on the main-sequence at this redshift is also shown. The galaxies are divided into the same three stellar mass bins as Figure 5.3, increasing in mass from the left to right panels. The merger fraction is computed separately for major mergers ($\mu \geq \frac{1}{4}$), minor mergers ($\frac{1}{10} \leq \mu < \frac{1}{4}$) and major or minor mergers ($\mu \geq \frac{1}{10}$) from left to right in each panel, respectively. The errors are the Poisson uncertainties. For each merger classification, the merger fraction of strongly star-forming galaxies decreases with increasing stellar mass, whilst the merger fraction of the main-sequence galaxies marginally increases. Strongly star-forming galaxies have merger fractions that are a factor of ≈ 2 greater than the main-sequence population at lower stellar masses, however, at higher stellar masses this excess is essentially erased. These enhancements suggest that mergers are playing a role in triggering strongly star-forming galaxies, at least at lower stellar masses.

($\mu \geq \frac{1}{10}$). Again, for comparison, we show the equivalent merger fractions of the main-sequence galaxies at this redshift. These fractions indicate the importance of the recent merger history for driving strongly star-forming galaxies, relative to those on the main-sequence.

For each merger classification, the merger fraction of the strongly star-forming population reduces with increasing stellar mass, whist the merger fraction of the main-sequence galaxies marginally increases; $90 \pm 11\%$ ($55 \pm 2\%$) for $M_* = 0.2$–0.4 $\times 10^{11}$ $M_\odot$, $80 \pm 7\%$ ($64 \pm 3\%$) for $M_* = 0.4$–0.9 $\times 10^{11}$ $M_\odot$ and $60 \pm 9\%$ ($57 \pm 8\%$) for $M_* = 0.9$–2.0 $\times 10^{11}$ $M_\odot$ for the strongly star-forming (main-sequence) galaxies, for $\mu \geq \frac{1}{10}$. The errors quoted are the Poisson uncertainties. Strongly star-forming galaxies have merger fractions that are approximately a factor of 2 greater than the main-sequence population in the lowest stellar mass bin, reaffirming that they are potentially derived from two distinct star-forming populations. However, analogous the trends with the stellar mass in Figure 5.3, this enhancement is all but erased in the highest stellar mass bin, again suggesting that strongly star-forming galaxies at these masses are no different in their star formation process from the 'typical' star-forming population.

This underlying behaviour is present at all redshifts covered by our strongly star-forming sample ($1 < z < 6$). However, analogous to the redshift trends reported in the previous section, the enhancement in the merger fractions between the strongly star-forming galaxies and those on the main-sequence decreases with increasing redshift. This is again due to the increase in the characteristic SFR of main-sequence galaxies with increasing redshift, but *also* due to the universal increase in the background merger rate with increasing redshift (e.g., Rodriguez-Gomez et al., 2015; Qu et al., 2017).

These enhancements yield a clear indication that mergers are indeed an important factor in producing strongly star-forming galaxies, at least at lower stellar masses. We return to this discussion in Section 5.4.1.

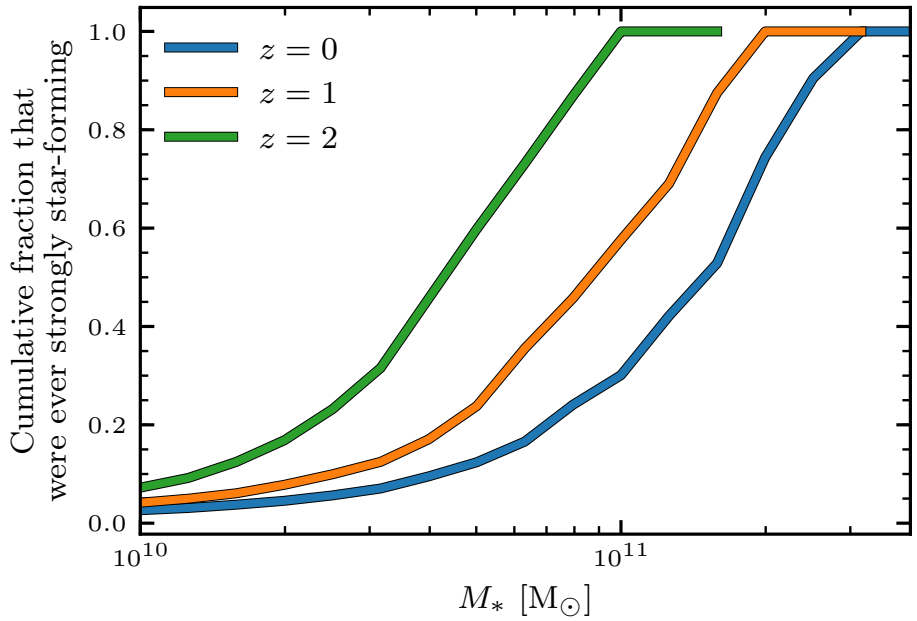

Figure 5.6: The cumulative fraction of all model galaxies at redshifts $z = 0$, 1 and 2, as a function of the stellar mass, that have ever been strongly star-forming ($\dot{M}_* \geq 80$ M$_\odot$ yr^{-1}) up until the considered redshift. Only the most massive galaxies ($M_* \gtrsim 10^{11}$ M$_\odot$) at each redshift have all experienced a strongly star-forming phase at some point in their history. The vast majority of galaxies have never had such an event. The cumulative fraction at fixed stellar mass decreases with decreasing redshift. Thus the majority ($\approx 70\%$) of massive galaxies ($M_* > 10^{11}$ M$_\odot$) at low redshift have grown either via more moderate periods of SFR activity, or by galaxy mergers.

Is there a ubiquitous strongly star-forming phase of galaxy evolution?

Until now, our analysis has focused on the properties of galaxies in the redshift range $2 < z < 3$ that have been captured in a strongly star-forming phase. It is interesting to ask, then, how ubiquitous such a phase of evolution is. That is, do all galaxies (or all galaxies above a critical mass) strongly star-form at some point during their lifetimes?

Figure 5.6 shows, for all galaxies at redshifts $z = 0, 1$ and 2, the cumulative fraction above a given stellar mass that have *ever*[4] been strongly star-forming up until the

[4]That is, if the galaxy has ever recorded a SFR ≥ 80 M$_\odot$ yr^{-1} along its main progenitor branch up until the considered redshift.

considered redshift. To ensure we capture even the shortest strongly star-forming durations (that could potentially be missed by the temporal spacings of the snipshots), we compute the maximum SFR for each galaxy using the individual stellar particle histories (see Section 5.2). There are two results that stand out from this figure: (1) the vast majority of galaxies at $z = 0$ have never undergone a strongly star-forming phase, and (2) the cumulative fraction at fixed stellar mass is a strongly epoch dependent quantity. All galaxies at $z = 2$ with stellar masses greater than 1×10^{11} $\mathrm{M}_{\odot}$ have undergone a strongly star-forming phase, this however decreases to $\approx 60\%$ at $z = 1$, and to $\approx 30\%$ at $z = 0$. Therefore the majority of massive galaxies ($M_* > 10^{11}$ $\mathrm{M}_{\odot}$) at the present day have accumulated their mass either through more moderate levels of star-formation, or galaxy mergers. This is consistent with the results from Qu et al. (2017), who show that the most massive galaxies in the EAGLE simulation today have grown primarily via mergers.

Therefore only the most massive galaxies ($M_* \gtrsim 10^{11}$ $\mathrm{M}_{\odot}$) at each epoch are likely to have previously undergone a strongly star-forming phase. At lower stellar masses such a phase of galaxy evolution becomes increasingly rare, particularly at lower redshift.

The descendants of strongly star-forming galaxies

We conclude this section by examining where the descendants of the galaxies within our sample lie relative to the global population at $z = 0$. Figure 5.7 shows the sSFR, the central supermassive black hole mass and the stellar half mass radius of all galaxies at $z = 0$, each plotted as a function of the stellar mass. The descendants of the galaxies within our sample (i.e., those that were ever once strongly star-forming) are indicated individually, the remainder of the $z = 0$ population is shown as a two-dimensional histogram.

Upon first inspection, the descendants of the strongly star-forming galaxies are, for the most part, among the most massive central and satellite galaxies today. There

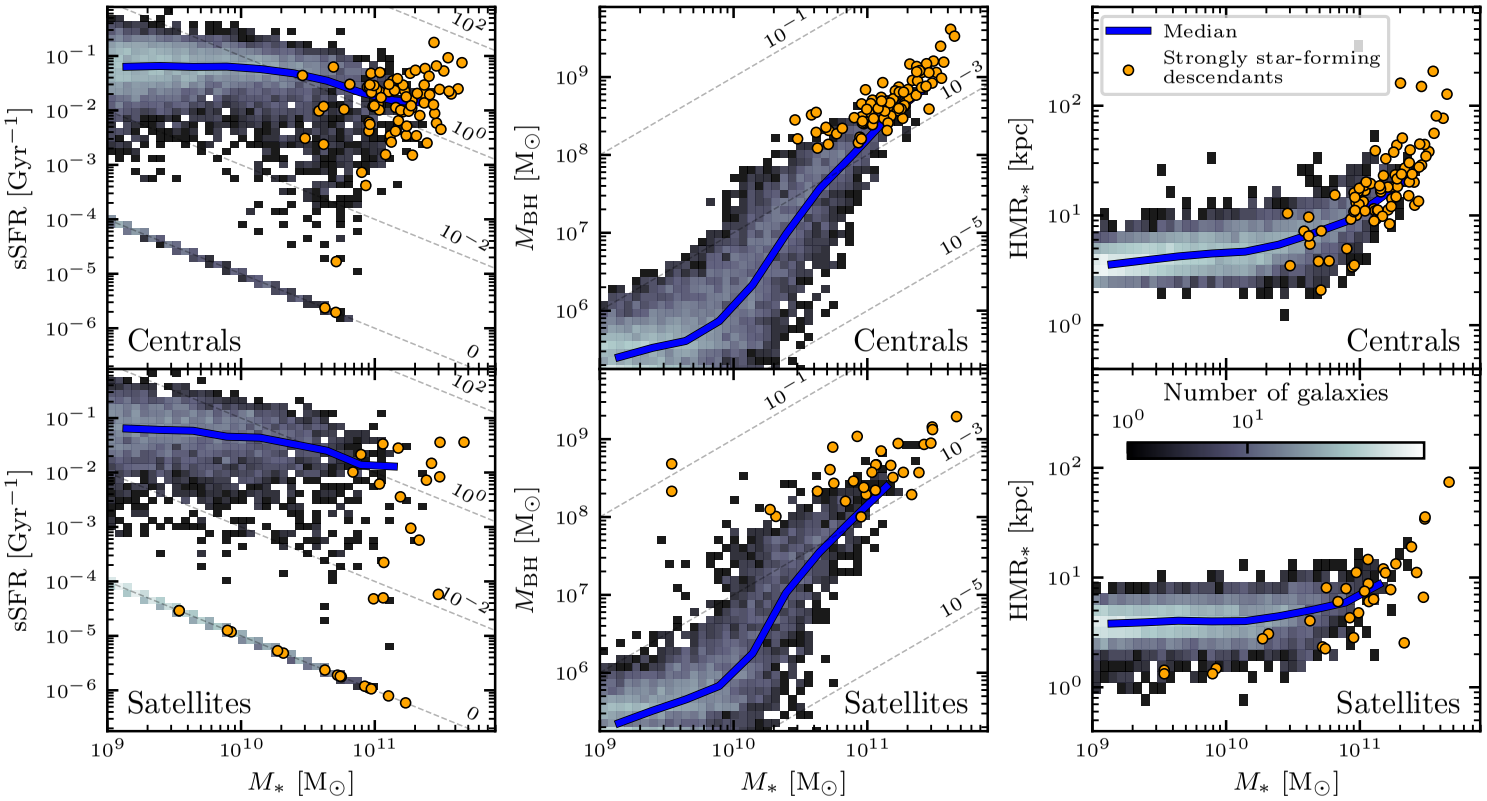

Figure 5.7: The $z = 0$ parameter space of the descendants of the galaxies contained within our star-forming sample. The descendants are divided into those that become either a central (top) or satellite (bottom) galaxy. The two-dimensional histograms show the $z = 0$ population, coloured by the number of galaxies in each bin. These histograms exclude the strongly star-forming descendants, which are indicated individually. The median relation for *all* galaxies is shown as a solid line. Each property is plotted as a function of the total stellar mass. *Left panels*: the sSFR ($\dot{M}_*/M_*$). The dashed lines indicate values of a constant SFR, including non star-forming galaxies ($\dot{M}_* = 0$). *Middle panels*: the central supermassive black hole mass. The dashed lines indicate values of a constant $M_{\rm BH}/M_*$ ratio. *Right panels*: the stellar half mass radius (HMR$_*$). The descendants of the most strongly star-forming galaxies evolve to be among the most massive galaxies today. However, not all of the most massive galaxies today had a high star-forming event in their past (see also Figure 5.6). There is little to distinguish the strongly star-forming descendants and the general population at $z = 0$ in the regions where the stellar masses overlap. An exception is in the middle panels, where the descendants of strongly star-forming galaxies have, on average, overmassive black holes for their stellar mass compared to the general population.

are a few exceptions in the satellite population, where a small number of low mass galaxies have previously been strongly star-forming. However, this subset of galaxies have deceptively low stellar masses as they are currently undergoing the process of being ram-pressure- and stellar-stripped as they in-fall into more massive haloes; seen from their total lack of star-forming gas ($\dot{M}_* = 0$ M$_\odot$ yr^{-1}), their extreme over massive black holes (see Barber et al., 2016 for a study of these compact objects), and their compact stellar sizes. The vast majority of the strongly star-forming descendants are actively star-forming today[5], however no more or less so (in the region of overlap) than the general population. The galaxies that were once strongly star-forming host, on average, more massive black holes for a fixed stellar mass than the general population, which was also alluded to in Figure 5.4. In terms of their stellar sizes, the strongly star-forming descendants are no more or less compact than the general population.

Therefore the descendants of strongly star-forming galaxies are not easily differentiable from the general galaxy population at $z = 0$. A possible exception is with the black hole masses, where the descendants of strongly star-forming galaxies are found to host the most massive black holes for a given stellar mass.

5.3.2 The submillimeter properties of strongly star-forming galaxies

Current surveys in the far-infrared and submm consistently report galaxy populations at $z \gtrsim 1$ with space densities of $\sim 10^{-4}$–10^{-5} cMpc^{-3} and inferred SFRs in excess of 100 M$_\odot$ yr^{-1}, which would easily meet our strongly star forming criteria. We now wish to compare the EAGLE strongly star-forming galaxies with these observations. We note, for this section we transition to using the strongly star-forming sample produced by the snapshot output, as the radiative transfer post-processing is only

[5]However, this result is inevitably skewed by the fact that the EAGLE model fails to completely quench star formation in the most massive galaxies, resulting in massive galaxies with bluer colours (Trayford et al., 2015) and higher gas fractions (Schaye et al., 2015) than is observed.

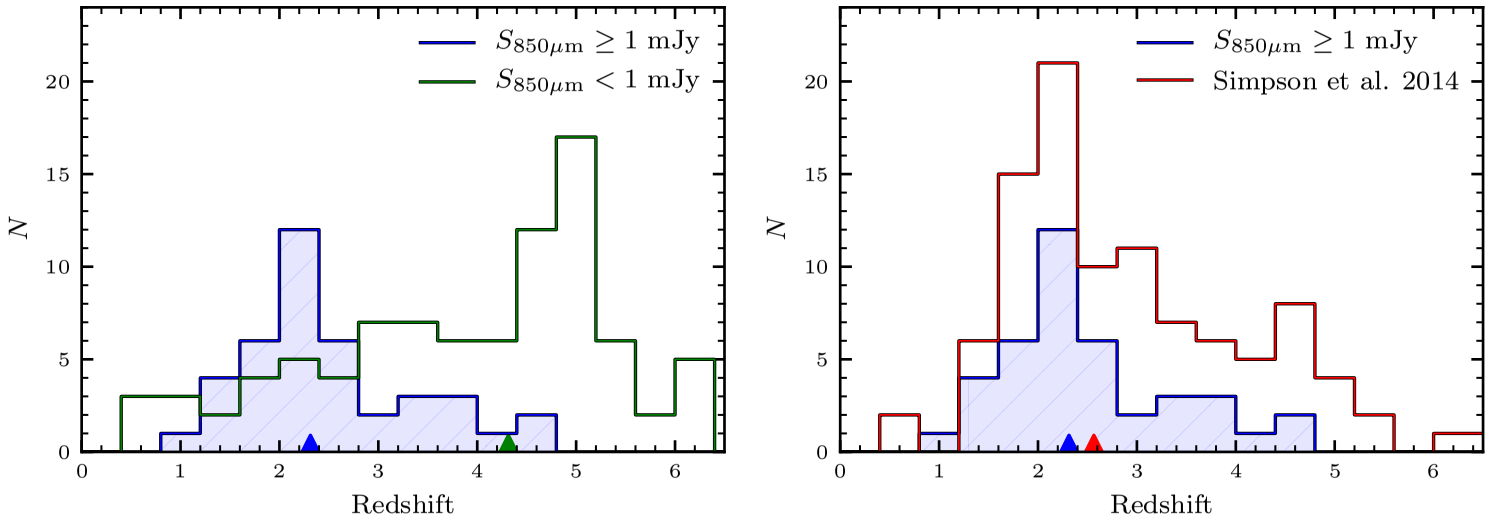

Figure 5.8: *Left panel*: the redshift distribution of the Submm-Bright and Submm-Faint samples. The median redshift of each sample is indicated by an arrow on the lower axis ($z = 2.32\pm^{0.14}_{0.04}$ and $z = 4.26\pm^{0.25}_{0.31}$ for the Submm-Bright and Submm-Faint population, respectively). Strongly star-forming Submm-Bright galaxies are rare at low redshift ($z \lesssim 1$), most abundant at $z \approx 2$, and become increasingly rare towards higher redshift ($z \gtrsim 3$). The number of strongly star-forming Submm-Faint galaxies increases steadily with increasing redshift until $z \approx 5$, where the counts start to rapidly decline. *Right panel*: a comparison between the Submm-Bright sample and the statistically corrected ALESS sample from Simpson et al. (2014) (the median redshift of the ALESS sample is $z = 2.56\pm^{0.25}_{0.16}$). The distributions are similar both in the position of the peak and the overall behaviour. The quoted errors on the medians are the 1σ uncertainties from bootstrap resampling.

computed for the model galaxies at these times (see Section 5.2).

We define the model galaxies with $S_{850\mu\mathrm{m}}$ greater than 1 mJy to be 'Submm-Bright' and those with $S_{850\mu\mathrm{m}}$ less than 1 mJy to be 'Submm-Faint'. This cut broadly reflects the definitions in the literature for strongly star-forming sources derived from the observations (e.g., Simpson et al., 2014; Cowie et al., 2018). Perhaps surprisingly, we find that the majority of the model galaxies within the strongly star-forming sample are Submm-Faint (89/129, 69%), as opposed to Submm-Bright (40/129, 31%). Figure 5.8 shows the redshift distributions of the galaxies within the Submm-Bright and Submm-Faint subsets. Interestingly, whilst all these model galaxies are strongly star-forming, those that are Submm-Bright preferentially exist at intermediate redshifts ($z \approx 2$) and those that are Submm-Faint preferentially exist at higher redshifts ($z \approx 5$).

As the strongly star-forming Submm-Faint galaxies within our sample would likely not be detected by current submm surveys, we only compare directly to the Submm-Bright subset in the next section. We return to the properties of the Submm-Faint strongly star-forming galaxies, and what differentiates them from the galaxies within the Submm-Bright subset, in Section 5.3.2.

A comparison to the observed submillimeter population

Figure 5.8 compares the redshift distributions of the model galaxies within the Submm-Bright sample to the statistically corrected ALESS sample of submm galaxies from Simpson et al. (2014). The model galaxies show a similar distribution to the observations, peaking around a redshift of $z \approx 2.5$, declining rapidly towards lower redshift, and declining more gradually towards higher redshift, agreeing well with one another.

We then compare the halo and galaxy properties of the galaxies within the Submm-Bright sample against a variety of independent submm selected observations in Figure 5.9. Each property is plotted as a function of redshift, and shows: the halo

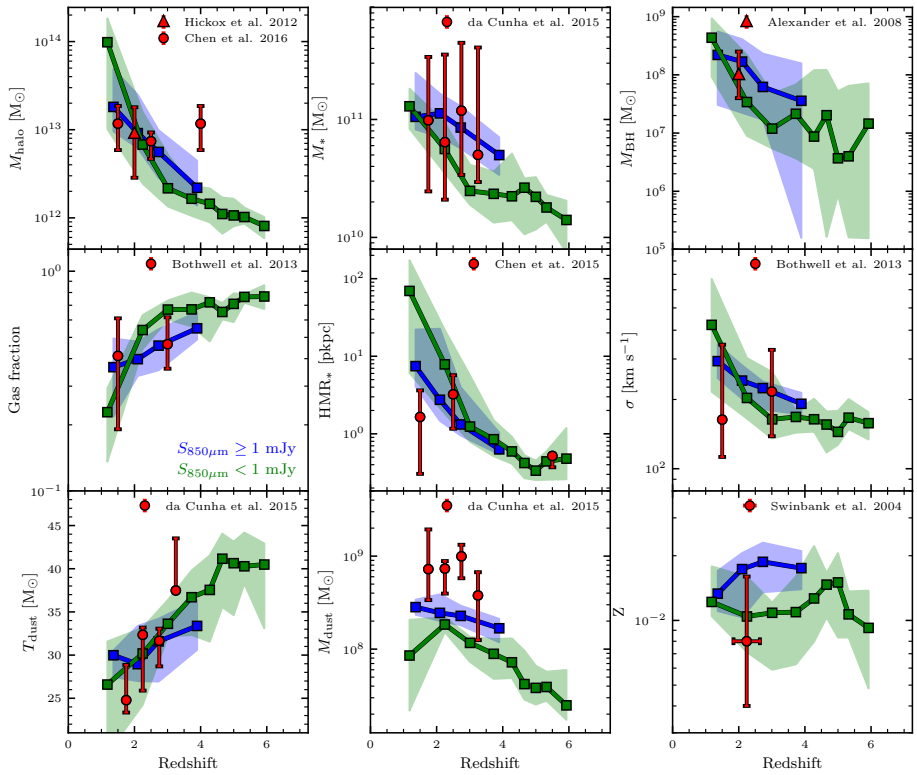

Figure 5.9: The halo and galaxy properties of the galaxies within the Submm-Bright and Submm-Faint samples, each plotted as a function of redshift. The lines represent the median values, with the shaded regions outlining the $10^{\text{th}} - 90^{\text{th}}$ percentile ranges. Each bin is ensured to contain at least 10 galaxies. The data points are observations of submm galaxies, and should therefore only be compared to the Submm-Bright sample. *Top left*: the halo mass, with observations from Hickox et al. (2012) and Chen et al. (2016). *Top centre*: the total stellar mass, with observations from da Cunha et al. (2015). *Top right*: the central supermassive black hole mass, with observations from Alexander et al. (2008). *Middle left*: the gas fraction, $M_{\text{gas}}/M_{\text{gas+stars}}$, with observations from Bothwell et al. (2013). *Middle centre*: the stellar half mass radius (HMR_*), with observations from Chen et al. (2015). *Middle right*: the velocity dispersion, with observations from Bothwell et al. (2013). *Bottom left*: the dust temperature, with observations from da Cunha et al. (2015). *Bottom centre*: the dust mass, with observations from da Cunha et al. (2015). *Bottom right*: the metallicity, with observations from Swinbank et al. (2014). We conclude that the strongly star-forming Submm-Bright galaxies in the EAGLE model are in reasonable agreement to the observed submm population across a range of observable properties. This suggests that the physical processes which create Submm-Bright model galaxies may also apply to submm galaxies.

mass, the stellar mass, the hosted central supermassive black hole mass, the total gas fraction ($M_{\rm gas}/M_{\rm gas+stars}$), the stellar half mass radius ($\rm HMR_*$), the velocity dispersion (σ), the dust mass, the dust temperature and the metallicity. The values and associated errors for the properties of the Submm-Bright model galaxies quoted below are the median and the 1σ uncertainties on the median (from bootstrap resampling), respectively.

The model galaxies reside in halos of mass $9.3\,\pm^{0.7}_{1.2}\times10^{12}$ $\rm M_\odot$, have stellar masses of $1.3\,\pm^{0.2}_{0.3}\times10^{11}$ $\rm M_\odot$ and host black holes of mass $1.8\,\pm^{0.9}_{0.2}\times10^{8}$ $\rm M_\odot$ at redshift $z\approx2$. These agree well with the observed clustering halo mass estimates from Hickox et al. (2012) and Chen et al. (2016), the stellar mass estimates from da Cunha et al. (2015) and the black hole mass estimates from Alexander et al. (2008), respectively. The total gas fractions ($39\pm^{3}_{3}\%$ at $z\approx2$), stellar sizes ($100\pm^{11}_{8}$ pkpc at $z\approx2$) and velocity dispersions ($236\pm^{10}_{10}$ km $\rm s^{-1}$ at $z\approx2$) yield reasonable agreements for a range of redshifts to the observations from Bothwell et al. (2013), Chen et al. (2015) and Bothwell et al. (2013), respectively. For the dust properties; the dust temperatures ($30\pm^{1}_{1}$ K at $z\approx2$) match reasonably well to the observations from da Cunha et al. (2015) for a range of redshifts, however, there is a systematic discrepancy at all redshifts between the dust masses (again from da Cunha et al., 2015) predicted by the EAGLE model ($2.4\,\pm^{0.3}_{0.2}\times10^{8}$ $\rm M_\odot$ at $z\approx2$) and the observations, with the observed submm galaxies containing $\approx$ 3–4 times more dust than the model galaxies. Finally, the metallicities ($0.018\pm^{0.001}_{0.003}$ at $z\approx2$) are potentially up to a factor of ≈2 greater than those estimated from Swinbank et al. (2014).

Although the observational uncertainties are typically large (often larger than the predicted scatter from the model), the model galaxies within the Submm-Bright sample yield a reasonable agreement to a variety of independent measurements from a variety submm selected observations.

The evolution of model submillimeter galaxies

As the majority of the observed submm populations are contained to within the redshift range $2 \lesssim z \lesssim 3$, it is challenging then to derive any evolutionary trends. However, we can predict the evolution of these objects using the Submm-Bright model galaxies.

The host haloes are predicted to decrease in their mass by approximately an order of magnitude between the redshifts $z \approx 1$ and $z \approx 4$ (from $2.4 \pm^{2.1}_{0.6} \times 10^{13}$ $M_{\odot}$ at $z \approx 1$ to $2.2 \pm^{0.3}_{0.3} \times 10^{12}$ $M_{\odot}$ at $z \approx 4$). Both the stellar and black hole masses are also predicted to decrease in mass with increasing redshift (from $1.3 \pm^{0.5}_{0.3} \times 10^{11}$ $M_{\odot}$ at $z \approx 1$ to $5.2 \pm^{0.8}_{0.3} \times 10^{10}$ $M_{\odot}$ at $z \approx 4$ for the galaxy masses and from $2.4 \pm^{1.5}_{0.3} \times 10^{8}$ $M_{\odot}$ at $z \approx 1$ to $3.7 \pm^{0.6}_{0.1} \times 10^{7}$ $M_{\odot}$ at $z \approx 4$ for the black hole masses). The total gas fractions are predicted to increase with increasing redshift (from $33 \pm^{5}_{6}\%$ at $z \approx 1$ to $55 \pm^{2}_{3}\%$ at $z \approx 4$). Both the stellar sizes and velocity dispersions are predicted to decrease with increasing redshift (from $11.4 \pm^{10}_{4}$ pkpc at $z \approx 1$ to $0.6 \pm^{0.1}_{0.0}$ pkpc at $z \approx 4$ for the stellar sizes and from $321 \pm^{50}_{50}$ km s^{-1} at $z \approx 1$ to $192 \pm^{8}_{4}$ km s^{-1} at $z \approx 4$ for the velocity dispersions). For the dust properties; the dust masses are predicted to decrease with increasing redshift (from $2.8 \pm^{0.2}_{0.2} \times 10^{8}$ $M_{\odot}$ at $z \approx 1$ to $1.7 \pm^{0.1}_{0.2} \times 10^{8}$ $M_{\odot}$ at $z \approx 4$) and the dust temperatures are predicted to increase with increasing redshift (from $30 \pm^{0.3}_{1.9}$ K at $z \approx 1$ to $33 \pm^{1.2}_{0.8}$ K at $z \approx 4$). Finally, the metallicities remain approximately constant at all redshifts (from $0.0140 \pm^{0.001}_{0.001}$ at $z \approx 1$ to $0.0175 \pm^{0.001}_{0.001}$ at $z \approx 4$).

In general, higher redshift Submm-Bright galaxies are predicted to be less massive, have increased gas fractions, be more compact and have warmer dust temperatures than their lower redshift counterparts.

The differences between submillimeter bright and submillimeter faint strongly star-forming galaxies

We now turn to the properties of the galaxies within the Submm-Faint subset, to see how, if at all, they differ from the galaxies within the Submm-Bright subset. Figure 5.9 additionally shows the halo and galaxy properties of the Submm-Faint population as a function of redshift.

The general behavioural trends discovered in the previous section for the model galaxies within the Submm-Bright subset remain applicable to the model galaxies within the Submm-Faint subset. However, the two populations are commonly offset in their absolute values for a fixed redshift. For example; the host halo, stellar and black hole masses of the Submm-Faint galaxies at $z \gtrsim 2$ are lower than the Submm-Bright galaxies by up to a factor of a few, the galaxies within the Submm-Faint subset are noticeably more gas rich and more metal poor at $z \gtrsim 2$ than the Submm-Bright population (both by up to a factor of ≈ 2) and the dust temperatures of the Submm-Faint galaxies are hotter, and their dust masses lower, than the Submm-Bright galaxies at $z \gtrsim 2$. At lower redshifts ($z \lesssim 2$), many of the properties for the two model populations come to overlap, however, we note that there is a large scatter in the properties for this redshift range, as the number of model galaxies becomes increasingly limited (see Figure 5.8).

It appears therefore, that the galaxies from the two model populations are fundamentally distinct. The galaxies from the Submm-Faint subset are, on average; hotter, lower mass, more gas rich, more metal poor, contain less dust and host lower mass black holes than the galaxies from the Submm-Bright subset. We discuss these subsets further in Section 5.4.1.

5.4 Discussion

5.4.1 What triggers strongly star-forming galaxies?

In Section 5.3 we compared the evolution and merger fractions of strongly star-forming model galaxies against the evolution and merger fractions of the more 'typical' ('main-sequence') star-forming model galaxies in an attempt to ascertain the physics surrounding the formation of this active population. To first order, two simple explanations exist, either (1) the star-formation process in strongly star-forming galaxies is the same as those on the main-sequence and the spread in the SFRs for galaxies of a given mass is purely down to the short-term variability in the stochastic star formation process, or (2) the star-formation process in strongly star-forming galaxies is fundamentally different to those operating on the main-sequence. However, we note that such a simplified view will be complicated if the star-formation process that drives the main-sequence varies with redshift; such that a different star-formation process operates in high redshift high SFR main-sequence galaxies compared to the typical $z \approx 0$ star-forming galaxies.

For the most massive strongly star-forming galaxies ($M_* \gtrsim 10^{11}$ M$_\odot$), scenario #1 likely applies. Their SFRs are minimally offset (if at all) from the main-sequence (< 0.25 dex, see Figure 5.2) which is consistent with the predicted scatter from SFR variability (see below), and their evolutionary tracks (see Figures 5.3 and 5.4) and their merger fractions (see Figure 5.5) are almost identical to those of the galaxies on the main-sequence. Or, to reiterate, strongly star-forming galaxies are simply the typical star-forming galaxies in this regime.

We therefore focus the remainder of our discussion on the origin of less massive strongly star-forming galaxies ($M_* \lesssim 10^{11}$ M$_\odot$), which make up the majority of our sample (79%). From this point forward we will refer to these objects simply as 'strongly star-forming galaxies' and not 'lower mass strongly star-forming galaxies', for clarity.

Are strongly star-forming galaxies fundamentally different?

Matthee & Schaye (2018) have shown that star formation burstiness on short ($\lesssim$ 1 Gyr) timescales creates ≈ 0.25 dex of scatter around the main-sequence in the EAGLE simulation. It is therefore not feasible that strongly star-forming galaxies are simply main-sequence galaxies captured at a high SFR, as they are too far offset from the main-sequence to be explained by short term variability alone, even at higher redshifts (they are offset by ≈ 0.75 dex at $2 < z < 3$, see Figure 5.2).

Furthermore, in Section 5.3.1 we investigated the properties of galaxies in the redshift range $2 < z < 3$ both before and after they were strongly star-forming. We found, relative to an equivalent population of galaxies on the main-sequence, that strongly star-forming galaxies evolved from a set of galaxies that were less massive (see Figure 5.3), more gas rich, and whose undermassive black holes had much more recently entered their rapid growth phase (see Figure 5.4). The fact that the strongly star-forming galaxies were so offset in their properties from their main-sequence equivalents both before and after they were strongly star forming further reinforced the fact that strongly star-forming galaxies are not simply main-sequence galaxies captured at a high SFR, and they are, in fact, two fundamentally distinct galaxy populations. These trends were present for the entirety of the redshift range covered by our sample ($1 < z < 6$).

We would conclude, therefore, that strongly star-forming galaxies ($M_* \lesssim 10^{11}$ M$_\odot$) are fundamentally distinct in their star formation process from the main-sequence population, at least for $z < 6$.

The role of the black holes

The evolutionary state of the black holes in strongly star-forming galaxies is likely the key to their higher rates of star formation. Strongly star-forming galaxies were found to host undermassive black holes still within their rapid growth phase, whereas the equivalent galaxies on the main-sequence hosted black holes that had completed

this phase in their evolution. This resulted in strongly star-forming galaxies being more gas rich relative to their main-sequence counterparts, as the hosted black holes were considerably less massive and had yet to meaningfully impact their host galaxy's gas reservoirs through AGN feedback, allowing for higher SFRs. This is in agreement with Matthee & Schaye (2018), who found that the relative efficiency of black hole growth is a source of scatter around the main-sequence. Despite being relatively underdeveloped at the time of the starburst, the black holes hosted by strongly star-forming galaxies went on to become noticeably *overmassive* relative to the main-sequence population (see Figure 5.4) and evolved into the most massive black holes for a given stellar mass at the present day (see Figure 5.7). This is a strong indication that vigorous black hole growth is directly linked with starbursting systems.

Are mergers required?

In Section 5.3.1 we found that 80–100% of strongly star-forming galaxies in the redshift range $2 < z < 3$ had undergone, or were about to undergo, either a minor or merger within the next/previous dynamical time; approximately a factor of two greater than the equivalent fraction found for the main-sequence population. This combination of such a high intrinsic fraction coupled with the vast excess over the expectation value from the galaxies on the main-sequence would strongly suggest that interactions are important for producing strongly star forming events in this regime. These intrinsically high fractions were generic for all redshifts covered by our sample ($1 < z < 6$), however, the excess above the main-sequence population was reduced slightly at higher redshifts due to the increase in the background merger rate with increasing redshift (e.g, Rodriguez-Gomez et al., 2015; Qu et al., 2017).

It would therefore appear that mergers are ubiquitously present around the strongly star forming phase in galaxies with masses $M_* \lesssim 10^{11}$ $\mathrm{M}_\odot$.

A minimum mass to sustain a starburst

Although not explicitly explored by this study, a minimum mass scale is also likely required to *sustain* a strongly star forming event. In galaxies below a critical mass ($\ll L_*$), stellar feedback is efficient at driving an outflow and balances the rate of cosmic inflow (e.g., White & Frenk, 1991; Finlator & Davé, 2008; Bouché et al., 2010; Schaye et al., 2010). Therefore strongly star forming events in these systems (for example those induced by an interaction) would likely 'burn out' after a short period of time, and these galaxies would quickly return to their quasi-equilibrium state. This critical mass scale is also linked to the black hole, as the growth of black holes is seriously hindered whilst stellar feedback remains efficient at driving an outflow (e.g., Dubois et al., 2015; Anglés-Alcázar et al., 2017; Bower et al., 2017; Habouzit et al., 2017). This critical mass scale today is $M_* \sim 10^{10}$ $\mathrm{M}_\odot$, however it does decrease with decreasing redshift (see McAlpine et al., 2018). Therefore the galaxies that are capable of sustaining a starburst are also those that host black holes that are primed to enter their rapid growth phase.

Creating strongly star-forming galaxies

By combining these results we would argue that strongly star-forming galaxies are the product of a sequence of well timed events.

1. The galaxy must host an underdeveloped black hole (one that has not yet entered its rapid growth phase); this ensures that the galaxy maintains a gas rich reservoir and perhaps more crucially that the black hole is not immediately too massive as to efficiently shutdown the triggered starburst event via AGN feedback.

2. A merger is required to trigger the initial starburst.

3. The galaxy must exceed the critical mass to sustain the starburst.

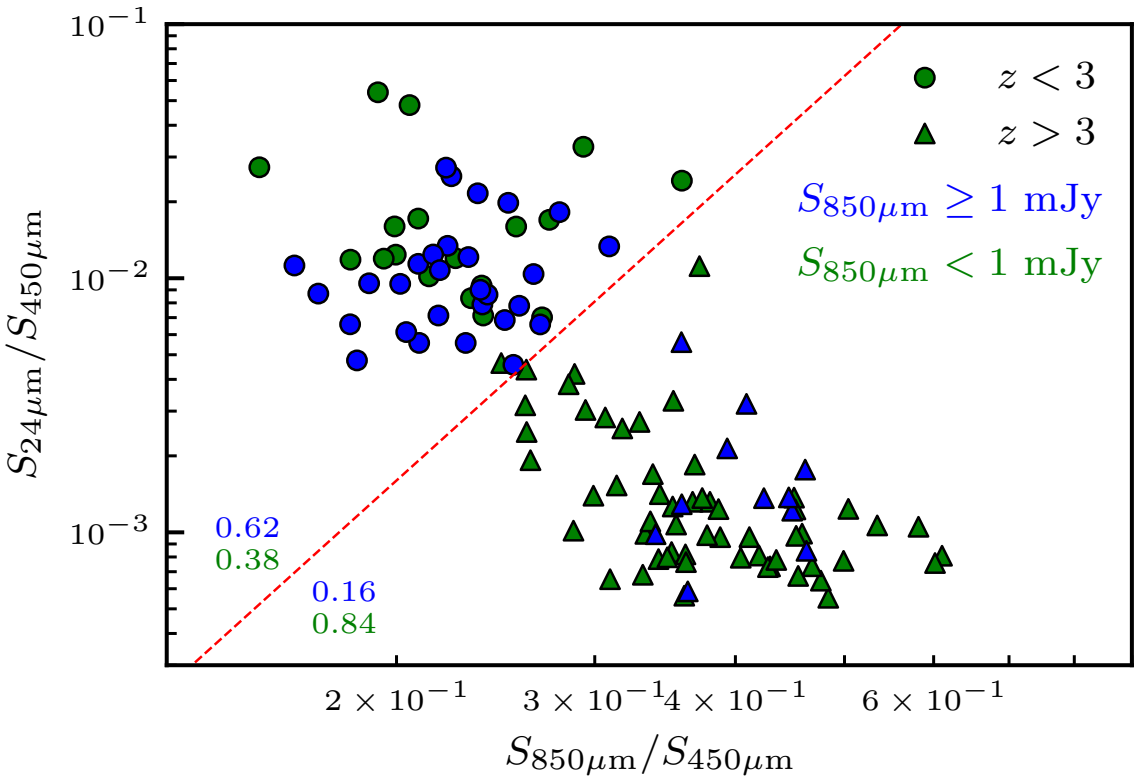

Figure 5.10: The colour–colour relation of $S_{24\mu m}/S_{450\mu m}$ versus $S_{850\mu m}/S_{450\mu m}$ for all the strongly star-forming galaxies within our sample with $S_{450\mu m} \geq 1$ mJy. Low- and high-redshift sources are indicated by different symbols, and those that are Submm-Bright and Submm-Faint are segregated by the colour of their symbols. A simple colour cut, indicated by a diagonal dashed line, isolates the high-redshift sources. The vast majority (84%) of galaxies to the lower right of this cut will be Submm-Faint strongly star-forming galaxies.

If these three conditions are met, the galaxy can produce, and sustain, a strongly star-forming event.

Identifying strongly star-forming galaxies

In Section 5.3.2 we found that the majority of strongly star-forming model galaxies were faint at submm wavelengths ($S_{850\mu m} \leq 1$ mJy), and would likely remain undetected by current widefield submm surveys. The existence of a population of 'hot', dusty ultraluminous galaxies at high-redshift has been suggested previously in the literature (e.g., Chapman et al., 2004; Blain et al., 2004; Casey et al., 2009).

These galaxies exhibit similar radio and optical characteristics to the high redshift submm-selected population, however, they are faint in the submm, suggesting a hidden population with hotter characteristic dust temperatures than the detected submm population. Additionally, studies that stack $S_{870\mu\mathrm{m}}$ imaging of high-redshift sources report characteristic dust temperatures that are considerably warmer than their lower-redshift counterparts (Cooke et al., 2018; Schreiber et al., 2018). However, conclusive evidence of a high redshift population of strongly star-forming submm faint galaxies, as found by EAGLE, has so far remained elusive.

Here we suggest a potential method to observationally identify the strongly star-forming Submm-Faint population at high redshift. Figure 5.10 shows the ratio of $S_{24\mu\mathrm{m}}$ to $S_{450\mu\mathrm{m}}$ fluxes against the ratio of $S_{850\mu\mathrm{m}}$ to $S_{450\mu\mathrm{m}}$ fluxes for the model galaxies within out sample. We impose a cut of $S_{450\mu\mathrm{m}} \geq 1$ mJy to emulate a sensitive $S_{450\mu\mathrm{m}}$ selected sample at the flux levels achievable with ALMA, however this only excludes a few sources from our original sample. We indicate low- and high-redshift strongly star-forming sources separately, whilst also segregating between the Submm-Faint and Submm-Bright galaxies. We find that imposing a simple colour cut (indicated by a dashed diagonal line) isolates the high-redshift typically Submm-Faint strongly star-forming galaxies from their low-redshift and typically Submm-Bright counterparts. The vast majority of the sources to the lower right of this diagram are comprised of the Submm-Faint high redshift strongly star-forming galaxies (84%).

5.5 Conclusions

We have investigated the triggering, descendants and submillimeter properties of strongly star-forming ($\dot{M}_* \geq 80$ $\mathrm{M}_\odot$ yr^{-1}) galaxies within the EAGLE hydrodynamical cosmological simulation. Our main conclusions are as follows:

- **Strongly star-forming galaxies and 'typically' star-forming 'main-sequence' galaxies do not follow the same evolutionary pathway**.

With the exception of the most massive strongly-star forming galaxies ($M_* \gtrsim 10^{11}$ M$_\odot$), which are the typical star-forming population for their stellar mass, the majority of strongly-star forming galaxies ($M_* \lesssim 10^{11}$ M$_\odot$) evolve from a set of galaxies that are less massive (see Figure 5.3), have higher gas fractions and host less-massive black holes that have entered their rapid growth phase more recently (see Figure 5.4), all relative to the 'typical' star-forming population.

- **Interactions are important for triggering a strongly star-forming phase in lower mass galaxies ($M_* \lesssim 10^{11}$ M$_\odot$).** The merger fractions of strongly star-forming galaxies are ubiquitously high at all redshifts. For lower-mass strongly star-forming galaxies the merger fractions are close to 100% and are a factor of ≈ 2 greater than the less active star-forming population (see Figure 5.5).

- **The descendants of galaxies which experienced a strongly star-forming phase are among the most massive galaxies today**. However, not all of the most massive ($M_* \gtrsim 10^{11}$ M$_\odot$) central and satellite galaxies today had strongly star-forming progenitors (see Figure 5.7). To experience a strongly star forming phase is rare; only $\approx 30\%$ of galaxies more massive than $M_* \geq 10^{11}$ M$_\odot$ today have ever been strongly star forming ($\dot{M}_* \geq 80$ M$_\odot$ yr^{-1}) in their past (see Figure 5.6). There is little to distinguish the strongly star-forming descendants and the general population at $z = 0$, however, strongly star-forming galaxies have, on average, overmassive black holes for their stellar mass compared to the general population (see Figure 5.7).

- **Not all strongly star-forming galaxies are submillimetre bright**. Only $\approx 30\%$ of strongly star-forming galaxies ($\dot{M}_* \geq 80$ M$_\odot$ yr^{-1}) are 'bright' at submillimeter wavelengths ($S_{850\mu\mathrm{m}} \geq 1$ mJy). The majority of strongly star-forming galaxies would therefore likely be missed by current submillimeter surveys. However, the population of 'warm', high-redshift, Submm-Faint strongly

star-forming sources (see Figure 5.8) could potentially be identified using a combination of deep mid-infrared and submm observations (see Figure 5.10).

- **Strongly star-forming galaxies are produced through a culmination of three coinciding events:**

  1. The galaxy must host an underdeveloped black hole (one that has not yet entered its rapid growth phase); this ensures that the galaxy maintains a gas rich reservoir and perhaps more crucially that the black hole is not immediately too massive as to efficiently shutdown the triggered starburst event via AGN feedback.
  2. A merger is required to trigger the initial starburst.
  3. The galaxy must exceed the critical mass to sustain the starburst.

Chapter 6

Conclusions

The accuracy of hydrodynamical simulations has dramatically increased over the course of the past 25 years. This is due in large part to an increased understanding of the physical prescriptions that are used to govern the astrophysical processes. Even though these remain unresolved, the continued development of state-of-the-art 'subgrid' models that faithfully capture the coarse-grained effects of unresolved scales have provided a better understanding of 'calibration' versus 'validation'. Because of this, hydrodynamical simulations have been able to broadly reproduce many of the observed properties of galaxies within the Universe; most notably the cosmic star formation history (e.g., Springel & Hernquist, 2003; Schaye et al., 2010) and the galaxy stellar mass function (e.g., Crain et al., 2009; Oppenheimer et al., 2010; Puchwein & Springel, 2013). Additionally, the advancements in the computational efficiency of the gravity and hydrodynamics solvers, and in the developments of computer architectures, have also been key. Following this, many large collaborations have now managed to produce accurate hydrodynamical simulations on cosmological scales; such as the ILLUSTRIS (Vogelsberger et al., 2014) and ILLUSTRIS-TNG (Pillepich et al., 2018) simulations, the EAGLE (Schaye et al., 2015; Crain et al., 2015) simulation, the HORIZON-AGN (Dubois et al., 2016) simulation and the ROMULUS (Tremmel et al., 2017) simulation. These collectively represent the state-of-the-art in cosmological hydrodynamical simulations today.

These simulations have strived towards the same goal, to recreate a realistic representation of our Universe, however they have concentrated on different aspects of galaxy evolution, making them a complimentary set of simulation suites. For example, the ROMULUS simulation suite has focused primarily on an improved set of supermassive black hole physics; such as physically motivated models for supermassive black hole formation, dynamics and the accretion of rotationally supported gas. These simulations are optimal at higher resolutions (in order to resolve the higher density regions where supermassive black hole are likely to form), which limits their ability to be performed in much larger volumes ($\ll (100\ \mathrm{cMpc})^3$). HORIZON-AGN is similarly focused towards the impact of suppermassive black holes upon their host galaxies, through the combination of a relatively lower resolution cosmological volume ($(100\ \mathrm{cMpc})^3$) and a series of high resolution zoom-in simulations. Both the ILLUSTRIS and EAGLE simulation suites have adopted a broader scope, aiming to reproduce key observables of the global population (such as the galaxy stellar mass function and cosmic star-formation history) in a single, large ($(100\ \mathrm{cMpc})^3$), relatively high resolution simulation.

Each of these simulation suites, whilst adopting alternate subgrid model calibration strategies and different hydrodynamical solvers, have been shown to broadly reproduce many of the observed properties of low- and high-redshift galaxies; however, each simulation remains subject to a number of shortcomings. For example, the EAGLE simulation systematically falls short of the observed normalisation of the cosmic star-formation history by ≈ 0.2 dex throughout cosmic time, it also fails to produce enough galaxies at the 'knee' of the galaxy stellar mass function. Interestingly, whilst the ILLUSTRIS simulation well reproduces both the shape and normalisation of the cosmic star-formation history, it overproduces both the amount of galaxies below and above the knee of the stellar mass function. In addition, both the ILLUSTRIS and EAGLE simulations have baryon fractions that are too high when compared to observations. Each of these results would suggest that the feedback produced from star-formation and black hole accretion remains inefficient in various

regimes and that their modelling needs to be improved for the next generation, either through an improved calibration strategy or an implementation of additional physics.

Even with these shortcomings this generation of cosmological hydrodynamical simulations have been incredibly successful in reproducing many aspects of the cosmos. In the case of the EAGLE simulation, which is the simulation suite this thesis is based upon; it has successfully reproduced the observed $z = 0$ Tully-Fisher relation, specific star formation rates and the column density distribution of intergalactic C IV and O VI (Schaye et al., 2015), many of the H I and H_2 properties of galaxies (Bahé et al., 2016; Lagos et al., 2015), the column density distribution of intergalactic metals (Schaye et al., 2015), galaxy rotation curves (Schaller et al., 2015b), the $z = 0$ luminosity function and colour-magnitude diagram (Trayford et al., 2015), the evolution of the galaxy stellar mass function (Furlong et al., 2015b), the high-redshift H I column density distribution (Rahmati et al., 2015) and the AGN luminosity function (Rosas-Guevara et al., 2016), to name but a few. The collective successes of the simulations make them an incredibly useful tool in furthering our understanding of galaxy evolution.

As the fidelity of the EAGLE model galaxy population has been established, we have used the simulation suite as a tool to answer the questions that observations cannot directly address; such as to the origin and evolution of particular populations of galaxies. For this thesis, we have added to the analysis of the EAGLE simulation suite by investigating the rarer and more unusual galaxy populations that reside with our Universe; those that host actively accreting supermassive black holes, and those yielding unusually high star formation rates for their stellar mass. Each of these populations represents only the tail end of the full underlying galaxy population distribution, yet, due to their extreme, and therefore easily detectable luminosities, they are some of the most well observed objects within our Universe. By using the EAGLE simulation suite, we have attempted to shed some light on the formation and evolution of such extreme objects, and to see how representative these populations

are in relation to the more 'typical' galaxies within our Universe.

6.1 Thesis results & Future work

In this thesis we have investigated the complex evolutionary pathway taken by the black holes within the EAGLE simulation, and how such an evolution influences the properties of the galaxy hosts around them. The black hole evolution was separated into three distinct phases, each relating to the mass of the host dark matter halo (see Chapter 3 for an investigation of these three phases of black hole evolution). As a reminder:

1. *The stellar feedback regulated phase.* The growth of central black holes within low-mass haloes ($M_{\rm halo} \ll 10^{12}$ M$_\odot$) is almost completely suppressed. This is due to the efficiency of stellar feedback during this regime, from its ability to generate turbulence within the gas and drive an effective outflow. The combination of these processes ensures that the central gas densities within these galaxies remain low, resulting in an extremely limited growth of the central black hole.

2. *The rapid black hole growth phase.* As haloes grow ($M_{\rm halo} \sim 10^{12}$ M$_\odot$), the stellar feedback loses its ability to drive an effective outflow, and begins to stall. This gives the first opportunity for a high gas density to build up within the galaxy centre, and with it the first meaningful period of black hole growth. As Bondi-like accretion is proportional to the mass of the black hole squared, the response from the black hole is highly non-linear, and the black holes hosted by these systems quickly become massive ($M_{\rm BH} \gtrsim 10^7$ M$_\odot$).

3. *The AGN feedback regulated phase.* Directly following the burst of rapid growth, the central black hole gradually becomes able to regulate the gas inflow onto the halo itself via efficient AGN feedback. Therefore in massive haloes, $M_{\rm halo}$

$\gtrsim 10^{12}$ $M_{\odot}$, a regulatory equilibrium is once again restored, and the specific growth rates of these black holes dramatically decrease.

These three phases of black hole evolution were found to have a noticeable knock-on effect to the properties of their galaxy hosts. For example, in Chapter 4 we discovered multiple relationships between the star-formation rate of galaxies and the accretion rate of their central black holes depending on the current phase of black hole evolution. We argued that the multiple relationships created a complex two-dimensional plane between the star-formation rate and the black hole accretion rate for the galaxy population, which is ultimately responsible for the alternate empirical trends found between the two properties depending on the initial selection method used. Additionally, in Chapter 5 we found that only the galaxies that hosted black holes that had not yet entered their rapid growth phase (and also experienced a triggering galaxy–galaxy merger) could go on to become the most strongly star-forming galaxies for their stellar mass. When considering the interplay between black holes and their host galaxies, it will always be essential to additionally consider the current phase of black hole evolution.

6.1.1 Future work

The appeal of cosmological hydrodynamical simulations is in their predictive power. We have explored in this thesis the nature of supermassive black holes and how they may share a complex relationship with their host galaxies, and also how short-lived starbursting galaxies may be formed purely from a merger driven scenario, making them fundamentally distinct from the general population of star-forming galaxies. Studies such as these, and others like them, act as a 'proof-of-concept', in a first step towards understanding how such exotic populations may be formed. The onus moving forward will be to combine the predictions from simulations like EAGLE with both new and existing observations.

As an example related to this work, we have presented multiple potential methods as to how the rapid growth phase of black holes could be discovered observationally.

- The most direct method would be to obtain a large sample of accurate black hole and stellar mass measurements, in order to construct the black hole mass–stellar mass relation (making sure to bracket the critical transition stellar mass). Even in the local Universe, the tell-tale imprints of the multiple phases of black hole evolution are predicted to be embedded within this relation. Additionally, if the relation was to be made at multiple epochs, the evolution of the critical transition mass is also predicted to be seen from this relation. However, obtaining accurate black hole (and stellar) mass measurements, particularly in lower mass systems and at higher redshifts, is notoriously difficult.

- An alternate, and potentially more obtainable method, would be to look for evidence of the three phases of black hole evolution *indirectly* through the central gas densities of the galaxy hosts. In systems below the critical transition mass, the gas densities within the centres of galaxies are predicted to remain low (due to efficient stellar feedback), the central gas densities are then predicted to rise as the mass of the galaxy approaches the critical transition mass (as the stellar feedback begins to stall), and finally the central gas densities are then again predicted to reduce in systems above the critical mass (due to the AGN feedback). By examining the spatial gas distributions of galaxies over a wide dynamic range of stellar masses with instruments such as ALMA, we could begin to constrain many black hole evolutionary scenarios.

- One final prediction regarding the rapid growth phase was the prevalence of galaxy–galaxy mergers around the onset of this phase. We found that galaxy interactions became increasingly important in triggering this phase of black hole evolution as the redshift decreased, and therefore predicted that a substantially higher fraction of AGN in galaxies around the critical transition mass today should be in the state of a merger relative to a control population

of similar mass galaxies. The AGN merger fraction of critical mass galaxies was predicted to remain approximately constant at higher redshifts, however the merger rate of the control population was also predicted to increase. We note, that it is unclear how AGN variability would effect this observation, and it will be the subject of future work.

We can take what we have learned from the successes and failures of the current generation of cosmological hydrodynamical simulations to feed into the improvements for the next generation. Work on the next generation of simulation suites has already begun, and will include:

- By identifying the shortcomings of the current generation, by seeing what properties of the galaxy population cannot be accurately reproduced, we can reflect and improve upon the subgrid model prescriptions. This can be achieved, in part, by performing an improved calibration strategy of the subgrid parameters, in an attempt to provide a better fit to the observational data. A second option is to revisit to physics that go into these models, from first principles. For example, obtaining the correct balance of AGN feedback in both low- and high-mass galaxies is challenging, and has traditionally yielded an excess in the baryon fractions of high-mass galaxies, leading to an overproduction of stars. Therefore improvements to the AGN feedback model (or models) will be one of the critical elements going forward in producing a more accurate galaxy population.

- The current generation of simulations are still considerably too small to replicate the volumes of upcoming (and current) observational surveys. Therefore to compete with semi-analytic models to provide mock catalogues for these surveys, the hydrodynamical simulations will be required to increase their volume by at least an order of magnitude. This will be a tremendous computing challenge, and is not currently feasible (if we wish to remain at the current resolution). The transition to such large volumes will only be possible

in the advent of new, more computationally efficient, simulation codes, such as the parallel task-based code SWIFT (Schaller et al., 2018), which promises speed-ups of an order of magnitude over previous codes.

- Being able to go to much higher resolutions is also an attractive prospect. Although this sacrifices the investigation of galaxy populations on cosmological scales, these simulations can provide invaluable incite into individual processes that operate on much smaller scales. For example; the AURIGA (Grand et al., 2017) and APOSTLE (Sawala et al., 2016) projects take advantage of such high resolutions to investigate the satellite populations of Milky Way analogues within the local group, providing valuable incite to the 'missing satellites' and 'to big to fail' problems; Pontzen et al. (2017) employ a 'genetic modification approach' to investigate the importance of mergers and AGN to quench star formation at high redshift; and Costa et al. (2018) investigate radiatively-driven outflows of high-redshift quasar host galaxies to see how they may also quench star formation. By using these dedicated simulations to train the physical prescriptions that go into the full scale cosmological simulations will give us the clearest picture of galaxy evolution moving forward.

Despite the current limitations, there is still much to be learned from this generation of cosmological hydrodynamical simulations. The EAGLE simulation has proven to be an invaluable tool in furthering our understanding of galaxy evolution.

Bibliography

Adler R. J., Casey B., Jacob O. C., 1995, American Journal of Physics, 63, 620

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

Aguirre A., Hernquist L., Schaye J., Katz N., Weinberg D. H., Gardner J., 2001, ApJ, 561, 521

Aird J., et al., 2010, MNRAS, 401, 2531

Alexander D. M., Hickox R. C., 2012, New Astron. Rev., 56, 93

Alexander D. M., et al., 2008, AJ, 135, 1968

Alpher R. A., Herman R. C., 1948a, Physical Review, 74, 1737

Alpher R. A., Herman R., 1948b, Nature, 162, 774

Alpher R. A., Bethe H., Gamow G., 1948, Physical Review, 73, 803

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

Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046

Aravena M., et al., 2016, ApJ, 833, 68

Atek H., et al., 2010, ApJ, 723, 104

Azadi M., et al., 2015, ApJ, 806, 187

Baes M., Verstappen J., De Looze I., Fritz J., Saftly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, ApJS, 196, 22

Bahé Y. M., McCarthy I. G., 2015, MNRAS, 447, 969

Bahé Y. M., et al., 2016, MNRAS, 456, 1115

Barber C., Schaye J., Bower R. G., Crain R. A., Schaller M., Theuns T., 2016, MNRAS, 460, 1147

Barnes J., Hut P., 1986, Nature, 324, 446

Bennett C. L., et al., 2013, ApJS, 208, 20

Benson A. J., 2010, Phys. Rep., 495, 33

Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

Binney J., Tremaine S., 1987, Galactic dynamics

Binney J., Gerhard O., Silk J., 2001, MNRAS, 321, 471

Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, ApJ, 611, 52

Bondi H., Hoyle F., 1944, MNRAS, 104, 273

Booth C. M., Schaye J., 2009, MNRAS, 398, 53

Booth C. M., Schaye J., 2010, MNRAS, 405, L1

Booth C. M., Schaye J., 2011, MNRAS, 413, 1158

Bothwell M. S., et al., 2013, MNRAS, 429, 3047

Bouché N., et al., 2010, ApJ, 718, 1001

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. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, MNRAS, 407, 2017

Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, MNRAS, 465, 32

Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40

Brammer G. B., et al., 2012, ApJS, 200, 13

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Camps P., Baes M., 2015, Astronomy and Computing, 9, 20

Camps P., et al., 2018, ApJS, 234, 20

Casey C. M., et al., 2009, MNRAS, 399, 121

Casey C. M., et al., 2012, ApJ, 761, 140

Casey C. M., Narayanan D., Cooray A., 2014, Phys. Rep., 541, 45

Chabrier G., 2003, PASP, 115, 763

Chang Y.-Y., van der Wel A., da Cunha E., Rix H.-W., 2015, ApJS, 219, 8

Chapman S. C., Smail I., Blain A. W., Ivison R. J., 2004, ApJ, 614, 671

Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772

Chen C.-T. J., et al., 2013, ApJ, 773, 3

Chen C.-C., et al., 2015, ApJ, 799, 194

Chen C.-C., et al., 2016, ApJ, 831, 91

Cooke E. A., et al., 2018, ApJ, 861, 100

Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, MNRAS, 450, 1521

Costa T., Rosdahl J., Sijacki D., Haehnelt M. G., 2018, MNRAS, 479, 2079

Cowie L. L., Gonzalez-Lopez J., Barger A. J., Bauer F. E., Hsu L.-Y., Wang W.-H., 2018, preprint, (arXiv:1805.09424)

Crain R. A., et al., 2009, MNRAS, 399, 1773

Crain R. A., et al., 2015, MNRAS, 450, 1937

Crain R. A., et al., 2016, preprint, (arXiv:1604.06803)

Croton D. J., et al., 2006, MNRAS, 365, 11

Cullen L., Dehnen W., 2010, MNRAS, 408, 669

Daddi E., et al., 2007, ApJ, 670, 156

Daddi E., et al., 2010, ApJ, 714, L118

Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140

Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 11

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Dekel A., Silk J., 1986, ApJ, 303, 39

Del Moro A., et al., 2013, A&A, 549, A59

Delvecchio I., et al., 2015, MNRAS, 449, 373

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497

Dubois Y., Teyssier R., 2008, A&A, 477, 79

Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., Teyssier R., 2015, MNRAS, 452, 1502

Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948

Dunlop J. S., et al., 2017, MNRAS, 466, 861

Durier F., Dalla Vecchia C., 2012, MNRAS, 419, 465

Efstathiou G., Sutherland W. J., Maddox S. J., 1990, Nature, 348, 705

Elbaz D., et al., 2007, A&A, 468, 33

Elbaz D., et al., 2011, A&A, 533, A119

Fabian A. C., 2012, ARA&A, 50, 455

Fakhouri O., Ma C.-P., 2008, MNRAS, 386, 577

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, PASP, 110, 761

Finlator K., Davé R., 2008, MNRAS, 385, 2181

Frenk C. S., White S. D. M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Fukugita M., Hogan C. J., Peebles P. J. E., 1998, ApJ, 503, 518

Furlong M., et al., 2015a, preprint, (`arXiv:1510.05645`)

Furlong M., et al., 2015b, MNRAS, 450, 4486

Furlong M., et al., 2015c, MNRAS, 450, 4486

Gamow G., 1948, Nature, 162, 680

Genel S., Genzel R., Bouché N., Naab T., Sternberg A., 2009, ApJ, 701, 2002

Genel S., et al., 2014, MNRAS, 445, 175

Genzel R., et al., 2010, MNRAS, 407, 2091

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Grand R. J. J., et al., 2017, MNRAS, 467, 179

Greene J. E., et al., 2016, ApJ, 826, L32

Gruppioni C., et al., 2013, MNRAS, 432, 23

Gutcke T. A., Fanidakis N., Macciò A. V., Lacey C., 2015, MNRAS, 451, 3759

Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays. (arXiv:astro-ph/0106018)

Habouzit M., Volonteri M., Dubois Y., 2016, preprint, (arXiv:1605.09394)

Habouzit M., Volonteri M., Dubois Y., 2017, MNRAS, 468, 3935

Harrison C. M., et al., 2012, ApJ, 760, L15

Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109

Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, MNRAS, 338, 913

Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, MNRAS, 396, 535

Hickox R. C., et al., 2012, MNRAS, 421, 284

Hickox R. C., Mullaney J. R., Alexander D. M., Chen C.-T. J., Civano F. M., Goulding A. D., Hainline K. N., 2014, ApJ, 782, 9

Hodge J. A., et al., 2013, ApJ, 768, 91

Hopkins P. F., 2013, MNRAS, 428, 2840

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005, ApJ, 630, 705

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Jahnke K., Macciò A. V., 2011, ApJ, 734, 92

Jenkins A., 2010, MNRAS, 403, 1859

Jenkins A., 2013, MNRAS, 434, 2094

Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115

Johnston R., Vaccari M., Jarvis M., Smith M., Giovannoli E., Häußler B., Prescott M., 2015, MNRAS, 453, 2540

Karim A., et al., 2011, ApJ, 730, 61

Katz N., Weinberg D. H., Hernquist L., 1996, ApJS, 105, 19

Keller B. W., Wadsley J., Couchman H. M. P., 2016, MNRAS, 463, 1431

Kennicutt Jr. R. C., 1998, ARA&A, 36, 189

Khachatryan V., et al., 2015, European Physical Journal C, 75, 235

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kirkpatrick A., et al., 2012, ApJ, 759, 139

Kirkpatrick A., Pope A., Sajina A., Roebuck E., Yan L., Armus L., Díaz-Santos T., Stierwalt S., 2015, ApJ, 814, 9

Koprowski M. P., Dunlop J. S., Michałowski M. J., Coppin K. E. K., Geach J. E., McLure R. J., Scott D., van der Werf P. P., 2017, MNRAS, 471, 4155

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Lagos C. d. P., et al., 2015, MNRAS, 452, 3815

Läsker R., Greene J. E., Seth A., van de Ven G., Braatz J. A., Henkel C., Lo K. Y., 2016, ApJ, 825, 3

Laureijs R., et al., 2011, preprint, (arXiv:1110.3193)

Lee N., et al., 2015, ApJ, 801, 80

Lucy L. B., 1977, AJ, 82, 1013

Lutz D., et al., 2010, ApJ, 712, 1287

Madau P., Dickinson M., 2014, Annual Review of Astron and Astrophys, 52, 415

Magnelli B., Elbaz D., Chary R. R., Dickinson M., Le Borgne D., Frayer D. T., Willmer C. N. A., 2011, A&A, 528, A35

Magorrian J., et al., 1998, AJ, 115, 2285

Martín-Navarro I., Mezcua M., 2018, ApJ, 855, L20

Massey R., Kitching T., Richard J., 2010, Reports on Progress in Physics, 73, 086901

Matthee J., Schaye J., 2018, preprint, (arXiv:1805.05956)

McAlpine S., et al., 2016, Astronomy and Computing, 15, 72

McAlpine S., Bower R. G., Harrison C. M., Crain R. A., Schaller M., Schaye J., Theuns T., 2017, MNRAS, 468, 3395

McAlpine S., Bower R. G., Rosario D. J., Crain R. A., Schaller M., Schaye J., Theuns T., 2018, preprint, (arXiv:1805.08293)

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Monaghan J. J., 1992, ARA&A, 30, 543

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Mullaney J. R., et al., 2012a, MNRAS, 419, 95

Mullaney J. R., et al., 2012b, ApJ, 753, L30

Navarro J. F., Eke V. R., Frenk C. S., 1996, MNRAS, 283, L72

Neistein E., Netzer H., 2014, MNRAS, 437, 3373

Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325

Page M. J., et al., 2012, Nature, 485, 213

Peng C. Y., 2007, ApJ, 671, 1098

Perlmutter S., et al., 1999, ApJ, 517, 565

Pillepich A., et al., 2018, MNRAS, 473, 4077

Pirzkal N., et al., 2013, ApJ, 772, 48

Planck Collaboration et al., 2014, Astronomy and Astrophysics, 571, A1

Pontzen A., Tremmel M., Roth N., Peiris H. V., Saintonge A., Volonteri M., Quinn T., Governato F., 2017, MNRAS, 465, 547

Portinari L., Chiosi C., Bressan A., 1998, A&A, 334, 505

Price D. J., 2008, Journal of Computational Physics, 227, 10040

Price D. J., 2012, Journal of Computational Physics, 231, 759

Puchwein E., Springel V., 2013, MNRAS, 428, 2966

Qu Y., et al., 2017, MNRAS, 464, 1659

Rafferty D. A., Brandt W. N., Alexander D. M., Xue Y. Q., Bauer F. E., Lehmer B. D., Luo B., Papovich C., 2011, ApJ, 742, 3

Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034

Rees M. J., 1966, Nature, 211, 468

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Reines A. E., Volonteri M., 2015, ApJ, 813, 82

Riess A. G., et al., 1998, AJ, 116, 1009

Rodighiero G., et al., 2011, ApJ, 739, L40

Rodighiero G., et al., 2014, MNRAS, 443, 19

Rodriguez-Gomez V., et al., 2015, MNRAS, 449, 49

Rosario D. J., et al., 2012, A&A, 545, A45

Rosas-Guevara Y. M., et al., 2015, MNRAS, 454, 1038

Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, MNRAS, 462, 190

Rubin V. C., Ford Jr. W. K., Thonnard N., 1980, ApJ, 238, 471

Rugh S. E., Zinkernagel H., 2000, ArXiv High Energy Physics - Theory e-prints,

Salim S., et al., 2007, ApJS, 173, 267

Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74

Sawala T., et al., 2016, MNRAS, 457, 1931

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2008, MNRAS, 389, 1137

Schaller M., et al., 2015a, MNRAS, 451, 1247

Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015b, MNRAS, 454, 2277

Schaller M., Gonnet P., Chalk A. B. G., Draper P. W., 2018, SWIFT: SPH With Inter-dependent Fine-grained Tasking, Astrophysics Source Code Library (ascl:1805.020)

Schaye J., 2004, ApJ, 609, 667

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210

Schaye J., et al., 2010, MNRAS, 402, 1536

Schaye J., et al., 2015, MNRAS, 446, 521

Schreiber C., et al., 2015, A&A, 575, A74

Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Franco M., 2018, A&A, 609, A30

Scott N., Graham A. W., Schombert J., 2013, ApJ, 768, 76

Segers M. C., Crain R. A., Schaye J., Bower R. G., Furlong M., Schaller M., Theuns T., 2016, MNRAS, 456, 1235

Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Simpson J. M., et al., 2014, ApJ, 788, 125

Smoot G. F., et al., 1992, ApJ, 396, L1

Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, MNRAS, 428, 1128

Sobral D., Kohn S. A., Best P. N., Smail I., Harrison C. M., Stott J., Calhau J., Matthee J., 2016, MNRAS, 457, 1739

Sparre M., et al., 2015, MNRAS, 447, 3548

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15

Springel V., 2005, MNRAS, 364, 1105

Springel V., Hernquist L., 2003, MNRAS, 339, 312

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V., Di Matteo T., Hernquist L., 2005a, MNRAS, 361, 776

Springel V., et al., 2005b, Nature, 435, 629

Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137

Springel V., et al., 2008, MNRAS, 391, 1685

Stach S. M., et al., 2018, preprint, (`arXiv:1805.05362`)

Stanley F., Harrison C. M., Alexander D. M., Swinbank A. M., Aird J. A., Del Moro A., Hickox R. C., Mullaney J. R., 2015, MNRAS, 453, 591

Steigman G., 2007, Annual Review of Nuclear and Particle Science, 57, 463

Stinson G. S., Bailin J., Couchman H., Wadsley J., Shen S., Nickerson S., Brook C., Quinn T., 2010, MNRAS, 408, 812

Swinbank A. M., et al., 2014, MNRAS, 438, 1267

Symeonidis M., et al., 2011, MNRAS, 417, 2239

Tasca L. A. M., et al., 2015, A&A, 581, A54

Teyssier R., 2002, A&A, 385, 337

Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068

Thacker R. J., MacMackin C., Wurster J., Hobbs A., 2014, MNRAS, 443, 1125

Tomczak A. R., et al., 2016, ApJ, 817, 118

Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, MNRAS, 438, 1985

Trayford J. W., et al., 2015, MNRAS, 452, 2879

Trayford J. W., Theuns T., Bower R. G., Crain R. A., Lagos C. d. P., Schaller M., Schaye J., 2016, MNRAS, 460, 3925

Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769

Vogelsberger M., et al., 2014, MNRAS, 444, 1518

Volonteri M., Capelo P. R., Netzer H., Bellovary J., Dotti M., Governato F., 2015a, MNRAS, 449, 1470

Volonteri M., Capelo P. R., Netzer H., Bellovary J., Dotti M., Governato F., 2015b, MNRAS, 452, L6

Wadsley J. W., Keller B. W., Quinn T. R., 2017, MNRAS, 471, 2357

Walter F., et al., 2016, ApJ, 833, 67

Weinberger R., et al., 2018, MNRAS, 479, 4056

Wendland H., 1995, Advances in Computational Mathematics, 4, 389

Whitaker K. E., et al., 2014, ApJ, 795, 104

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

Zwicky F., 1937, ApJ, 86, 217

da Cunha E., et al., 2015, ApJ, 806, 110