

The nature of submillimetre and highly star-forming galaxies in the EAGLE simulation

Stuart McAlpine^{1,2*}, Ian Smail³, Richard G. Bower^{1,2}, A. M. Swinbank^{1,3}, James W. Trayford⁴, Tom Theuns², Maarten Baes^{1,5}, Peter Camps^{1,5}, Robert A. Crain^{1,6} and Joop Schaye^{1,4}

¹Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2a PO Box 64, FI-00014 University of Helsinki, Finland

²Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

³Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

⁴Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, the Netherlands

⁵Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281, B-9000 Gent, Belgium

⁶Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK

Accepted 2019 June 7. Received 2019 May 13; in original form 2019 January 7

ABSTRACT

We exploit EAGLE, a cosmological hydrodynamical simulation, to reproduce the selection of the observed submillimetre (submm) galaxy population by selecting the model galaxies at $z \geq 1$ with mock submm fluxes $S_{850\ \mu\text{m}} \geq 1$ mJy. We find a reasonable agreement between the model galaxies within this sample and the properties of the observed submm population, such as their star formation rates (SFRs) at $z < 3$, redshift distribution, and many integrated galaxy properties. We find that the median redshift of the $S_{850\ \mu\text{m}} \geq 1$ mJy model population is $z \approx 2.5$, and that they are massive galaxies ($M_* \sim 10^{11} M_\odot$) with high dust masses ($M_{\text{dust}} \sim 10^8 M_\odot$), gas fractions ($f_{\text{gas}} \approx 50$ per cent), and SFRs ($\dot{M}_* \approx 100 M_\odot \text{ yr}^{-1}$). In addition, we find that they have major and minor merger fractions similar to the general population, suggesting that mergers are not the sole driver of the high SFRs in the model submm galaxies. Instead, the $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies yield high SFRs primarily because they maintain a significant gas reservoir as a result of hosting an undermassive black hole relative to comparably massive galaxies. Not all ‘highly star-forming’ ($\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$) EAGLE galaxies have submm fluxes $S_{850\ \mu\text{m}} \geq 1$ mJy. We investigate the nature of these highly star-forming ‘Submm-Faint’ galaxies (i.e. $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ but $S_{850\ \mu\text{m}} < 1$ mJy) and find that they are similar to the model submm galaxies, being gas rich and hosting undermassive black holes. However, they are also typically at higher redshifts ($z > 4$) and are lower mass ($M_* \sim 10^{10} M_\odot$). These typically higher redshift galaxies show stronger evidence for having been triggered by major mergers, and critically, they are likely missed by most current submm surveys due to their higher dust temperatures and lower dust masses.

Key words: galaxies: active – galaxies: evolution – galaxies: formation – galaxies: high-redshift – galaxies: starburst.

1 INTRODUCTION

Submillimetre (submm) galaxies (SMGs) are a population of high-redshift galaxies ($z \approx 1\text{--}5$) that are inferred to have high star formation rates (SFRs; $\dot{M}_* \gtrsim 100 M_\odot \text{ yr}^{-1}$) and significant dust masses ($M_{\text{dust}} \gtrsim 10^8 M_\odot$). The luminous submm emission arises from the reprocessing of the ultraviolet (UV) light from young stars by obscuring dust (see Casey, Narayanan & Cooray 2014

for a review). SMGs are relatively rare, with a number density of $\sim 10^{-5} \text{ cMpc}^{-3}$ at $z \approx 2\text{--}3$ (e.g. Chapman et al. 2005; Simpson et al. 2014). However, this population is particularly interesting due to the apparently high levels of star formation and the many open questions that remain to be answered about their formation and evolution – for example: What triggers such extreme SFRs in SMGs and other strongly star-forming galaxies, mergers, or secular disc instabilities? How long can these galaxies form stars at such apparently high rates, and what terminates it? How do they evolve following the starburst episode? What are their descendants at $z \approx 0$? Many of these questions cannot be answered directly

* E-mail: stuart.mcalpine@helsinki.fi

by the observations; however, by combining the observational data with advanced numerical simulations, they can begin to be addressed.

Previous comparisons of models with observations of the high-redshift, high-SFR galaxy population have highlighted apparent issues with the current theories of galaxy formation (e.g. Baugh et al. 2005; Swinbank et al. 2008; Davé et al. 2010; Hayward et al. 2013). These models generally have difficulty in matching the observed SFRs or submm number counts of SMGs while maintaining agreement with other observations of the wider galaxy population. A number of solutions have been proposed for this discrepancy: including a top-heavy initial mass function (IMF), the unaccounted for blending of multiple galaxies into a single submm source in the observations, and the inability of hydrodynamical simulations to adequately resolve intense starbursts (leading to an underestimate in the SFR). However, recently the necessity of a radical solution to the discrepancy has perhaps weakened since higher resolution observational data has become available (e.g. Karim et al. 2013; Simpson et al. 2015; Cowie et al. 2018; Stach et al. 2018), which now resolves the blending of multiple SMGs into a single source (e.g. Cowley et al. 2015).

In this study we investigate the prevalence of submm and highly star-forming galaxies in the largest simulation of the EAGLE project (Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016). This simulation was calibrated to reproduce the observed galaxy stellar mass function, the sizes of galaxies, and the black hole–stellar mass relation at $z \approx 0.1$. Many other observed galaxy trends have also shown broad agreement with observations, both locally (e.g. Lagos et al. 2015; Schaye et al. 2015; Trayford et al. 2015; Bahé et al. 2016) and at higher redshift (e.g. Furlong et al. 2015; Rahmati et al. 2015; Furlong et al. 2017; McAlpine et al. 2017). 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 the SMG population. Observationally it is claimed that these two populations overlap strongly. We can test this in our model, by selecting the samples independently.

This paper is organized as follows. In Section 2 we provide a brief overview of the EAGLE simulation and discuss the sample selections. In Section 3.1 we compare our sample of submm-selected model galaxies to the observed submm population. We then examine the nature and triggering of the highly star-forming/submm model galaxies in Section 3.2 and their descendants at $z = 0$ in Section 3.3. Finally, in Section 4 we present a summary and our concluding remarks.

2 THE EAGLE SIMULATION

EAGLE (‘Evolution and Assembly of GaLaxies and their Environment’; Crain et al. 2015; Schaye et al. 2015)^{1,2} is a suite of cosmological smoothed particle hydrodynamics (SPH) simulations that cover a range of periodic volumes, numerical resolutions, and physical models. To incorporate the processes that operate below the simulation resolution a series of ‘subgrid’ prescriptions are implemented, namely radiative cooling and photoionization heating

(Wiersma, Schaye & Smith 2009a); star formation (Schaye & Dalla Vecchia 2008), stellar mass-loss (Wiersma et al. 2009b), and stellar feedback (Dalla Vecchia & Schaye 2012); and black hole growth via accretion and mergers, and black hole feedback (Springel, Di Matteo & Hernquist 2005; Schaye et al. 2015; Rosas-Guevara et al. 2016). The free parameters of these models are calibrated to reproduce the observed galaxy stellar mass function, galaxy sizes, and black hole mass–bulge mass relation at $z \approx 0.1$. A full description of the simulation and the calibration strategy can be found in Schaye et al. (2015) and Crain et al. (2015), respectively.

For this study we are interested in the most strongly star-forming ($\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$) and submm luminous (i.e. those with mock fluxes $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$) model galaxies, and therefore restrict our study to the largest simulation, Ref-L0100N1504, which contains the greatest number of these systems. This simulation is a cubic periodic volume 100 comoving megaparsec (cMpc) on each side, sampled by 1504^3 dark matter particles of mass $9.7 \times 10^6 M_\odot$ and an equal number of baryonic particles with an initial mass of $1.8 \times 10^6 M_\odot$. The subgrid parameters are those of the EAGLE reference model, described fully by Schaye et al. (2015). The simulation adopts a flat lambda cold dark matter (Λ CDM) cosmogony with parameters inferred from analysis of *Planck* data (Planck Collaboration I 2014): $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\sigma_8 = 0.8288$, $n_s = 0.9611$, and $H_0 = 67.77 \text{ km s}^{-1} \text{ Mpc}^{-1}$. A Chabrier (2003) stellar IMF is adopted. Unless otherwise stated, error estimates are from a bootstrap analysis.

2.1 Simulation output

The complete state of the simulation is stored at 29 intervals between redshifts $z = 20$ and $z = 0$, which we refer to as ‘snapshots’. In order to produce a halo and galaxy catalogue from these, the dark matter structure finding algorithm ‘friends of friends’ and the substructure finding algorithm SUBFIND (Springel et al. 2001; Dolag et al. 2009) are performed.

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 kiloparsec (pkpc), as per Schaye et al. (2015). The sizes of galaxies are quoted in physical units, unless stated otherwise.

2.1.1 Merger trees

Galaxies are tracked from their creation to the present day using a merger tree. Due to the hierarchical build-up, each galaxy today has many progenitors; therefore, the history of each galaxy is considered from the reference frame of its ‘main progenitor’, the branch of the galaxy’s full merger tree that contains the greatest total mass (see Qu et al. 2017 for full details).

Using these merger trees, the merger history of a galaxy can be established. We define the completion time of a galaxy merger with the main progenitor galaxy to be the cosmic time at the first simulation output where the two galaxies are now identified as a single bound system by the SUBFIND algorithm. Mergers are classified by their stellar mass ratio, $M_{*,1}/M_{*,2}$, where $M_{*,2}$ is the mass of the most massive member of the binary. They are considered major if $M_{*,1}/M_{*,2} \geq 1/4$, minor if $1/10 \leq M_{*,1}/M_{*,2} < 1/4$, and either major or minor if $M_{*,1}/M_{*,2} \geq 1/10$. To account for the stellar stripping that occurs during the later stages of the interaction, the

¹www.eaglesim.org

²The galaxy and halo catalogues of the simulation suite, as well as the particle data, are publicly available at <http://www.eaglesim.org/database.php> (McAlpine et al. 2016; The EAGLE team 2017).

stellar mass ratio is computed when the infalling galaxy had its maximum mass (e.g. Rodriguez-Gomez et al. 2015; Qu et al. 2017). Additionally, mergers are only considered ‘resolved’ when $M_{*,2} \geq 10^8 M_\odot$ (≈ 100 stellar particles), but this does not affect the results here, where we focus on much more massive galaxies.

2.1.2 Obtaining accurate star formation histories of EAGLE galaxies

The reported ‘instantaneous’ SFR of an EAGLE galaxy is computed from the current state of the galaxy’s associated gas particles. However, it is also possible to reverse engineer the SFR of a galaxy at a given time from the galaxy’s associated stellar particles at $z = 0$.³ By collectively binning the stellar particles in the galaxy at $z = 0$ by their birth time, weighting by their initial mass and dividing by the bin width, a robust SFR history is obtained, which is only limited in resolution by the total number of stellar particles sampled (i.e. the galaxy stellar mass at $z = 0$). For galaxies more massive than $M_*[z = 0] \geq 10^{10} M_\odot$ this can adequately resolve the SFR history down to intervals of ≈ 1 Myr, which achieves orders of magnitude better sampling than from the snapshot output. This method can be used to accurately study the SFR histories of individual galaxies (e.g. fig. 1 from McAlpine et al. 2017), or simply as a method to obtain the maximum SFR ever achieved by a galaxy throughout its lifetime.

2.1.3 Absolute calibration of SFRs

When comparing to the observed cosmic SFR density, Furlong et al. (2015) 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 typical local galaxies are systematically lower than previously estimated by ≈ 0.2 dex, yielding better agreement with the EAGLE prediction (see fig. 5 of Schaller et al. 2015).

As the observational data sets compared to in Section 3.1 utilize an earlier calibration, we *reduce* all observed SFRs by 0.2 dex. This serves to remove the known global systematic offset, making it simpler to focus on the trends with the observed submm population that are the topic of this paper.

2.1.4 Mock observables

The light emitted within a galaxy is subject to attenuation by dust in the interstellar medium (ISM). To accurately produce the observable properties of the model galaxies, we solve the three-dimensional radiative transfer problem in post-processing using information from the galaxy’s star-forming regions, stellar sources, and diffuse dust distribution using the radiative transfer code SKIRT (Baes et al. 2011; Camps & Baes 2015), as is detailed by Camps et al. (2018). This process infers the mock observables from the UV to the

submm wavebands.⁴ The free parameters in the radiative transfer model⁵ have been calibrated to reproduce far-infrared observables from the *Herschel* Reference Survey (Boselli et al. 2010) in the local Universe. Given the weaker observational constraints for these parameters at higher redshifts, the values obtained locally are applied to the galaxies at all redshifts; however, this is likely to be an oversimplification (see Camps et al. 2016 and Trayford et al. 2017 for more details).

For this analysis, we compare to the observed submm population using the SKIRT inferred observed-frame fluxes at $850 \mu\text{m}$ ($S_{850 \mu\text{m}}$, column SCUBA2.850 in the public database; McAlpine et al. 2016; Camps et al. 2018).

2.2 Sample selection

The primary motivation for this study is to infer the nature of the observed SMG and highly star-forming galaxy populations using analogues from the simulation. With the advent of ALMA the properties of the observed submm population are becoming better constrained, with the bulk of the population found at higher redshifts ($z > 1$; e.g. Simpson et al. 2014), as are the majority of highly star-forming sources found through other multiwavelength selection techniques ($\dot{M}_* \gtrsim 100 M_\odot \text{yr}^{-1}$; e.g. Madau & Dickinson 2014). For simplicity, in this study we therefore concentrate our analysis on the submm and highly star-forming model galaxies from the simulation above $z \geq 1$, where the bulk of the observed population lies and hence the vast majority of the observational constraints on these populations exist. We will return to the submm properties of low-redshift ($z < 1$) galaxies in a future study.

2.2.1 Sample 1: $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ galaxies

We select galaxies with mock $S_{850 \mu\text{m}}$ fluxes greater than 1 mJy. This cut broadly reflects the definitions in the literature for highly star-forming sources derived from the recent ALMA studies (e.g. Simpson et al. 2014; Cowie et al. 2018; Stach et al. 2018), and produces a selection of 62 $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ model galaxies in EAGLE at $z \geq 1$.

2.2.2 Sample 2: highly star-forming Submm-Faint galaxies

The observed submm population is synonymous with high SFRs ($\dot{M}_* \gtrsim 100 M_\odot \text{yr}^{-1}$), which is also true for the model galaxies in the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample (see Section 3.1.1). Yet it is interesting to ask if the reverse also holds – that is, are *all* ‘highly star-forming’ sources bright in the submm? To test this, we select the most highly star-forming galaxies in the simulation at $z \geq 1$, adopting a space density that is representative of the limits on highly star-forming populations derived from a variety of observational techniques. There is, however, considerable uncertainty in this quantity observationally. For example, surveys that select sources on the basis of their mid- or far-infrared luminosity (as a proxy for their SFR) have reported space densities of $\sim 10^{-5}$ to $2 \times 10^{-4} \text{ cMpc}^{-3}$ for galaxies with far-infrared luminosities of $\gtrsim 10^{12} L_\odot$ and inferred SFRs \gtrsim

⁴The rest-frame magnitudes and observer-frame fluxes for all EAGLE galaxies with stellar masses greater than $10^{8.5} M_\odot$ are publicly available (Camps et al. 2018).

⁵Notably for the dust distribution, f_{dust} , the fraction of the metallic gas locked up in dust (assumed to be 0.3), and T_{max} , the highest temperature at which gas contains dust (assumed to be 8000 K; see Camps et al. 2016).

³Note that only the stellar particles born within the main progenitor (see Section 2.1.1) galaxy are considered, so as to avoid combining the SFR histories of multiple progenitors.

$100 M_{\odot} \text{ yr}^{-1}$ at $z \approx 1.5\text{--}2.5$ (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). However, these studies suffer from a combination of active galactic nucleus (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 submillimetre surveys (e.g. Hodge et al. 2013; Scudder et al. 2018).

Given the range and uncertainties in the various estimates of the space density for highly star-forming galaxies from the observations, we have chosen to select an equivalent SFR limit that roughly corresponds to the space densities derived from the observations ($\sim 10^{-5}$ to $2 \times 10^{-4} \text{ cMpc}^{-3}$) and which allows us to isolate a sufficiently large sample ($\gtrsim 100$ galaxies) of highly star-forming galaxies to allow for a statistical analysis. Our adopted SFR limit is $\geq 80 M_{\odot} \text{ yr}^{-1}$, which corresponds to a space density of star-forming galaxies in EAGLE of $8 \times 10^{-5} \text{ cMpc}^{-3}$ at $z \geq 1$, and returns a sample of 84 galaxies.

We find, perhaps surprisingly, that only 32 per cent of the galaxies at $z \geq 1$ with a SFR $\geq 80 M_{\odot} \text{ yr}^{-1}$ have mock submm fluxes $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$, implying that a large number of highly star-forming galaxies in the Universe may not be detected by the current submm surveys. For this reason, we additionally investigate the nature of highly star-forming galaxies ($\dot{M}_{*} \geq 80 M_{\odot} \text{ yr}^{-1}$) that are ‘Submm-Faint’ ($S_{850 \mu\text{m}} < 1 \text{ mJy}$) to see how, if at all, these galaxies differ from the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ population. This will also reveal which subset of highly star-forming galaxies is selected by the submm surveys.

We note that we repeated the analysis of this study for highly star-forming $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ galaxies (i.e. adding the additional $\dot{M}_{*} \geq 80 M_{\odot} \text{ yr}^{-1}$ criterion to define the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample); however, the results were almost identical to those found for the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ only sample (given their naturally high SFRs, $\dot{M}_{*} \geq 50 M_{\odot} \text{ yr}^{-1}$). We remain with the pure submm-selected sample for this study as it most closely reflects that of the observations.

3 RESULTS

3.1 A comparison of simulated and observed SMGs

We begin with a comparison between the model galaxies in the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample and the results from a variety of observational studies of submm galaxies. Here, our aim is to establish if a submm selection applied to EAGLE serves to extract a population similar to that seen in the real Universe.

In the following, we will compare to the observed SMG population focusing on the first large ALMA-identified sample from the ALESS survey (Hodge et al. 2013). Subsequent analysis and study of the SMGs in this survey have provided a variety of empirical constraints on their natures, which we can compare to our model predictions (e.g. Simpson et al. 2014; da Cunha et al. 2015). We also make use of the earlier survey of Chapman et al. (2005) when discussing certain aspects, such as gas or black hole properties (e.g. Alexander et al. 2008; Bothwell et al. 2013).

3.1.1 Star formation rates

The SFRs inferred for the observed submm population are generally high ($\dot{M}_{*} \gtrsim 100 M_{\odot} \text{ yr}^{-1}$), which provides a good initial test for our model population. Fig. 1 shows the median SFR (\dot{M}_{*}) as a function

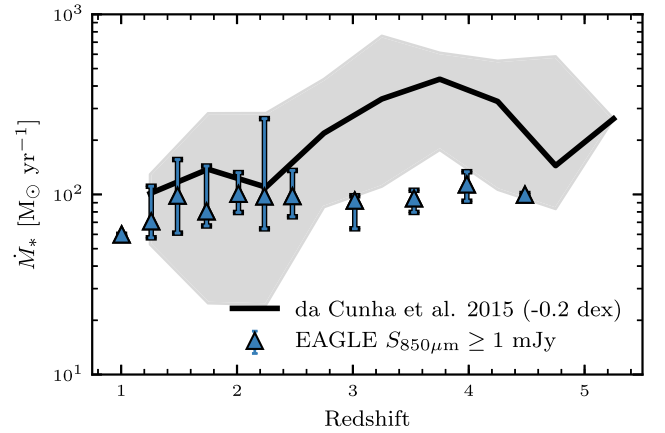


Figure 1. The median SFR as a function of redshift for both the EAGLE $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ galaxies at each simulation snapshot and the observed SMGs from da Cunha et al. (2015). A -0.2 dex SFR recalibration has been applied to the observational data at all redshifts (see Section 2.1.3). The error bars and the height of the shaded region outline the 10th–90th percentile range. The SFRs of the model galaxies in the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample are high ($\dot{M}_{*} > 50 M_{\odot} \text{ yr}^{-1}$, with a median SFR of $94 M_{\odot} \text{ yr}^{-1}$), and are in reasonable agreement with the SFRs inferred from the observations below $z \approx 3$ where the bulk of the population lies. However, the SFRs of the model galaxies in the high-redshift tail ($z \gtrsim 3$) are a factor of ≈ 3 lower than the observations.

of redshift for both the model galaxies in the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample and the observed SMGs from Hodge et al. (2013) as analysed by da Cunha et al. (2015).

Overall, the SFRs of the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ model galaxies are high ($\dot{M}_{*} \geq 50 M_{\odot} \text{ yr}^{-1}$), yielding a relatively constant median value between ≈ 70 and $100 M_{\odot} \text{ yr}^{-1}$ for all redshifts, and a maximum SFR of $294 M_{\odot} \text{ yr}^{-1}$ produced by a galaxy at $z = 2.2$. This confirms that a purely submm-selected sample of model galaxies at high redshifts does only return high-SFR sources. When compared to the SFRs inferred from the observations, there is good agreement at $z \lesssim 3$ where the bulk of the model and observed populations lie; however, a discrepancy appears above $z \gtrsim 3$, where the SFRs of the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ model galaxies are a factor of ≈ 3 lower than those inferred from observations. It is important to stress, however, that most of the observed and simulated submm galaxies lie below $z = 3$ (≈ 70 per cent; see next section), where there is good agreement between the SFRs from the two populations.

3.1.2 Redshift distribution

The spectroscopic redshifts for large surveys of observed SMGs suggest that this population is most abundant at redshift $z \approx 2\text{--}3$ (e.g. Chapman et al. 2005; Danielson et al. 2017). However, these studies are incomplete, and so here we compare the redshift distribution of the model galaxies in the $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ sample against the statistically corrected sample of observed SMGs using photometric estimates from Simpson et al. (2014).

Fig. 2 shows the volume number density of the model $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ galaxies at each simulation snapshot. It indicates that this population is most abundant at $z \approx 2.5$ (with a number density of $\sim 10^{-5} \text{ cMpc}^{-3}$), and is rarer towards both lower and higher redshifts ($\sim 10^{-6} \text{ cMpc}^{-3}$ at $z \approx 1$ and $z \approx 4$). To provide a measurement that is more easily comparable to the observations, Fig. 2 also shows the area number density (or counts) of the model $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$ galaxies. To compare, we overplot the redshift distribution

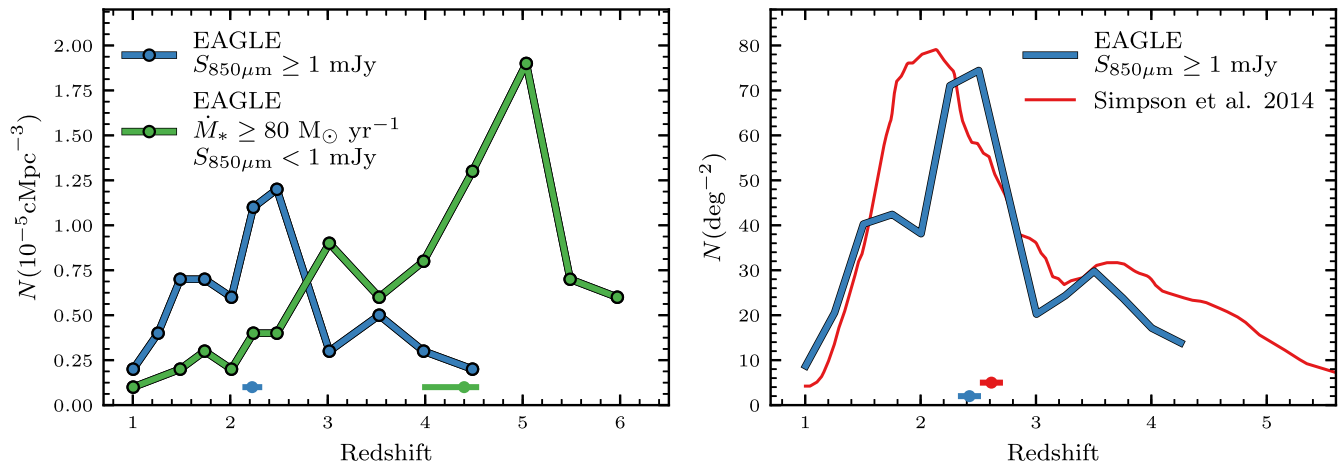


Figure 2. *Left-hand panel:* the number density of the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ and highly star-forming Submm-Faint model galaxies (in units of $10^{-5} \text{ cMpc}^{-3}$) at each simulation snapshot. The median redshift of each sample is indicated by an error bar along the lower axis ($z = 2.2^{+0.1}_{-0.1}$ and $z = 4.4^{+0.1}_{-0.4}$ for the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ and highly star-forming Submm-Faint population, respectively). The galaxies in the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample are most abundant at $z \approx 2.5$ and then become increasingly rare towards lower and higher redshifts. The galaxies within the highly star-forming Submm-Faint sample are predominantly at higher redshift ($z > 4$). *Right-hand panel:* the number of $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ model galaxies per square degree (median value of $z = 2.4^{+0.1}_{-0.1}$) compared to the statistically corrected observed sample of SMGs from Simpson et al. (2014) (median value of $z = 2.6^{+0.1}_{-0.1}$). The shape and the median value are similar for the two distributions. The position of the peak is slightly higher for the sample of $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ model galaxies ($z \approx 2.5$) compared to the observations ($z \approx 2$).

of observed SMGs from Simpson et al. (2014). Encouragingly, the model $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ galaxies yield a similar shape and median value ($z_{\text{EAGLE}} = 2.4^{+0.1}_{-0.1}$ compared to $z_{\text{obs}} = 2.6^{+0.1}_{-0.1}$) as the observations. Yet, whilst the overall behaviour between the two samples appears consistent (a peaked distribution that decays towards lower and higher redshifts), the location of the peak for the model $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ galaxies is slightly higher than is seen in these observations ($z_{\text{EAGLE}} \approx 2.5$ compared to $z_{\text{obs}} \approx 2$).

3.1.3 Galaxy and halo properties

Next, in Fig. 3 we compare a variety of properties of the model galaxies in the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample against those of the observed population. Each property is plotted as a function of redshift, and we show the halo mass (M_{halo}), the stellar mass (M_*), the central supermassive black hole mass (M_{BH}), the total gas fraction ($M_{\text{gas}}/M_{\text{gas+stars}}$), the stellar half-mass radius (HMR $_*$), the stellar velocity dispersion⁶ (σ_*), the dust mass (M_{dust}), the dust temperature (T_{dust}), and the stellar metallicity (Z_*). The values and associated errors for the properties of the model galaxies quoted below are the median and the 1σ uncertainties on the median, respectively (these values are also given in Table 1).

The model $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ galaxies reside in haloes of mass $(9.1^{+4.8}_{-0.2}) \times 10^{12} M_\odot$, have stellar masses of $(8.7^{+0.1}_{-2.6}) \times 10^{10} M_\odot$, and host black holes of mass $(15.4^{+9.6}_{-1.4}) \times 10^7 M_\odot$ at redshift $z \approx 2$. These values 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 ($37.4^{+3.2}_{-1.7}$ per cent at $z \approx 2$), stellar sizes ($2.5^{+0.4}_{-0.5}$ pkpc at $z \approx 2$), and velocity dispersions ($223^{+25}_{-17} \text{ km s}^{-1}$ at $z \approx 2$) yield reasonable agreements for a range of redshifts with the observations from Bothwell et al. (2013), Chen et al. (2015), and Bothwell et al.

⁶The one-dimensional velocity dispersion of the stars; see table B.1 from McAlpine et al. (2017).

(2013), respectively. The dust temperatures ($38.5^{+0.4}_{-0.9} \text{ K}$ at $z \approx 2$) match well 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 model [$(2.5^{+0.2}_{-0.5}) \times 10^8 M_\odot$ at $z \approx 2$] and the observations, with the observed submm galaxies apparently containing ≈ 3 – 4 times more dust than the model galaxies. Finally, the metallicities ($17.9^{+0.1}_{-1.8} \times 10^{-3}$ at $z \approx 2$) are potentially up to a factor of ≈ 2 higher than those estimated by Swinbank et al. (2004), but are consistent to within the error bars.

Although the observational uncertainties are typically large (often larger than the predicted scatter from the model), the model galaxies within the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample are in reasonable agreement with a variety of measurements from a variety of observations of submm-selected galaxies.

3.2 The nature of $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ galaxies and highly star-forming Submm-Faint galaxies

Given the broad agreement in the comparisons to the observed submm population above, we are encouraged to exploit the sample of $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ model galaxies in an attempt to answer the questions that the observations cannot easily address, such as the origin, evolution, and eventual fate of the submm population. In a parallel analysis, we investigate the nature of the highly star-forming galaxies that do not make it into the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample (i.e. $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ but $S_{850\mu\text{m}} < 1 \text{ mJy}$), to better understand what subset of high-SFR galaxies a submm selection extracts.

It is important to remind the reader for this section that the model galaxies in the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample are substantially more abundant than the model galaxies in the highly star-forming Submm-Faint sample at moderate redshift ($z \approx 2.5$), whereas the galaxies in the highly star-forming Submm-Faint sample are substantially more abundant than the galaxies in the $S_{850\mu\text{m}} \geq 1 \text{ mJy}$ sample at higher redshift ($z > 4$; see Fig. 2). Therefore, whilst we can compare the properties of the two model populations at

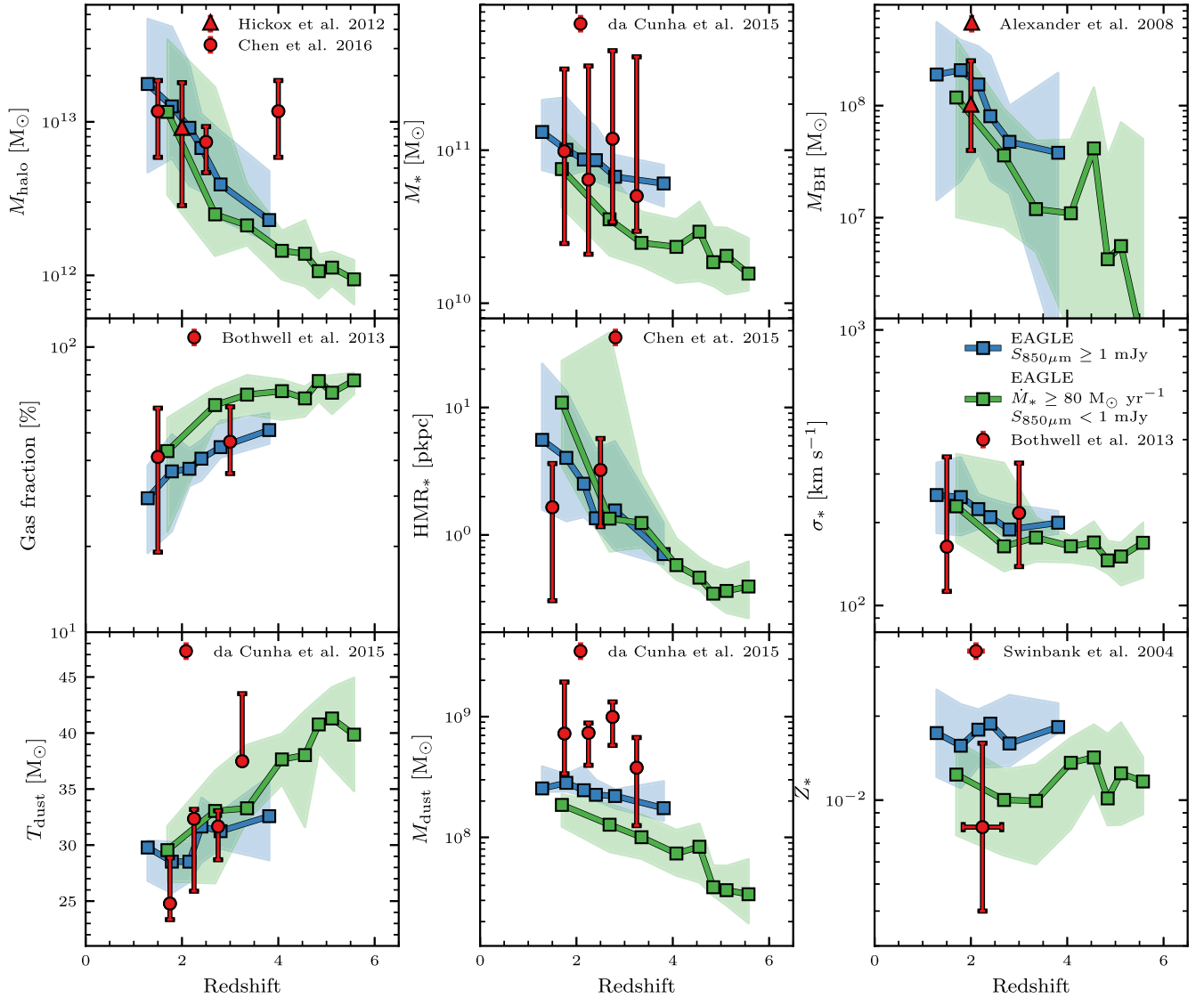


Figure 3. Various properties of the model galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy (blue) and highly star-forming Submm-Faint (green) samples, each plotted as a function of redshift. The median values of the EAGLE galaxies within each sample are represented by lines, with the shaded regions outlining the 10th–90th percentile ranges. It is ensured that each bin contains at least 10 galaxies. The red data points show estimates from observations of SMGs, and should therefore only be compared to the $S_{850\mu\text{m}} \geq 1$ mJy sample (shown in blue). *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 total gas fraction, $M_{\text{gas}}/M_{\text{gas+stars}}$, with observations from Bothwell et al. (2013). *Middle centre:* the stellar half-mass radius, with observations from Chen et al. (2015). *Middle right:* the stellar 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 stellar metallicity, with observations from Swinbank et al. (2004). We conclude that the $S_{850\mu\text{m}} \geq 1$ mJy galaxies in the EAGLE model are in reasonable agreement with the observed submm population across a range of observable properties. The EAGLE median values from this figure are tabulated in Tables 1 and 2.

a given redshift, the relative number density of galaxies between the two populations at the given redshift may be considerably different.

3.2.1 Comparison to the general population

We first investigate how the properties of the model galaxies in our two samples compare to the general model population. That is, what unique features do the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies exhibit that distinguish them from ‘typical’ model galaxies of their mass?

Fig. 4 shows the SFR, dust mass, total gas fraction, central supermassive black hole mass, stellar mass to halo mass ratio (M_*/M_{200}), and stellar metallicity of the galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples, each plotted as a function of either the stellar or halo mass. For clarity, we only show the model galaxies in the redshift range $z = 2$ –3, as these epochs contain the greatest number of sources between our two samples (see Fig. 2). We note that only the integrated properties of central galaxies are shown in this figure (the galaxies contained within both the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples are almost exclusively centrals at these times).

Table 1. The median values and the 1σ uncertainties (obtained via bootstrap resampling) for the integrated galaxy properties of the EAGLE $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies shown in Fig. 3. From left to right: the halo mass, stellar mass, central supermassive black hole mass, total gas fraction $M_{\text{gas}}/M_{\text{gas}+\text{stars}}$, stellar half-mass radius, stellar velocity dispersion, dust temperature, dust mass, and stellar metallicity. The quoted errors for the redshifts represent the bin width.

z	M_{halo} ($10^{12} M_{\odot}$)	M_{*} ($10^{10} M_{\odot}$)	M_{BH} ($10^7 M_{\odot}$)	f_{gas} (%)	HMR* (pkpc)	σ_{*} (km s^{-1})	T_{dust} (K)	M_{dust} ($10^8 M_{\odot}$)	Z_{*} (10^{-3})
1.3 ± 0.3	$17.6^{+4.7}_{-9.6}$	$13.2^{+3.8}_{-2.8}$	$19.0^{+11.7}_{-7.0}$	$29.5^{+4.3}_{-4.3}$	$5.6^{+1.4}_{-7.6}$	251^{+16}_{-13}	$29.8^{+1.5}_{-0.1}$	$2.5^{+0.1}_{-0.4}$	$17.4^{+3.3}_{-0.9}$
1.8 ± 0.2	$12.6^{+3.0}_{-2.8}$	$10.1^{+1.7}_{-3.1}$	$20.7^{+18.3}_{-1.8}$	$36.6^{+3.0}_{-1.4}$	$4.0^{+0.2}_{-3.2}$	247^{+24}_{-8}	$28.5^{+0.7}_{-1.0}$	$2.8^{+0.2}_{-0.3}$	$15.7^{+2.2}_{-3.7}$
2.2 ± 0.1	$9.1^{+4.8}_{-0.2}$	$8.7^{+0.1}_{-2.6}$	$15.4^{+9.5}_{-1.4}$	$37.4^{+3.2}_{-1.7}$	$2.5^{+0.4}_{-0.5}$	223^{+25}_{-17}	$28.5^{+0.4}_{-0.9}$	$2.5^{+0.2}_{-0.5}$	$17.9^{+0.1}_{-1.8}$
2.4 ± 0.1	$6.8^{+2.1}_{-0.7}$	$8.6^{+0.5}_{-2.2}$	$8.0^{+2.1}_{-5.5}$	$40.6^{+2.9}_{-2.6}$	$1.3^{+0.1}_{-0.5}$	209^{+5}_{-16}	$31.7^{+1.6}_{-1.3}$	$2.3^{+0.2}_{-0.1}$	$18.8^{+1.2}_{-0.9}$
2.8 ± 0.3	$3.9^{+0.6}_{-2.7}$	$6.7^{+0.5}_{-0.4}$	$4.8^{+1.5}_{-1.1}$	$44.6^{+0.3}_{-0.6}$	$1.6^{+0.2}_{-0.7}$	188^{+5}_{-14}	$31.2^{+1.1}_{-1.1}$	$2.2^{+0.3}_{-0.1}$	$16.0^{+1.3}_{-3.0}$
3.8 ± 0.7	$2.3^{+0.1}_{-0.3}$	$6.1^{+0.7}_{-0.1}$	$3.8^{+0.2}_{-7.9}$	$51.2^{+1.2}_{-3.7}$	$0.7^{+0.1}_{-0.1}$	199^{+8}_{-4}	$32.6^{+1.5}_{-0.8}$	$1.7^{+0.0}_{-0.3}$	$18.3^{+0.8}_{-0.4}$

Table 2. The same as Table 1, but now for the EAGLE highly star-forming Submm-Faint galaxies.

z	M_{halo} ($10^{12} M_{\odot}$)	M_{*} ($10^{10} M_{\odot}$)	M_{BH} ($10^7 M_{\odot}$)	f_{gas} (%)	HMR* (pkpc)	σ_{*} (km s^{-1})	T_{dust} (K)	M_{dust} ($10^8 M_{\odot}$)	Z_{*} (10^{-3})
1.7 ± 0.6	$11.6^{+3.3}_{-2.4}$	$7.5^{+1.4}_{-3.5}$	$11.8^{+3.1}_{-9.4}$	$43.1^{+7.4}_{-5.4}$	$10.9^{+4.0}_{-2.3}$	228^{+11}_{-17}	$29.6^{+1.7}_{-1.0}$	$1.9^{+0.1}_{-0.1}$	$12.3^{+0.6}_{-0.4}$
2.7 ± 0.4	$2.5^{+0.5}_{-0.0}$	$3.5^{+0.3}_{-0.4}$	$3.6^{+0.6}_{-1.1}$	$62.6^{+1.1}_{-4.3}$	$1.3^{+0.5}_{-1.2}$	163^{+1}_{-123}	$33.0^{+1.5}_{-2.9}$	$1.3^{+0.1}_{-0.3}$	$10.0^{+0.4}_{-0.8}$
3.4 ± 0.3	$2.1^{+0.5}_{-0.0}$	$2.5^{+0.4}_{-1.2}$	$1.2^{+0.0}_{-2.5}$	$68.1^{+3.3}_{-5.0}$	$1.2^{+0.3}_{-0.8}$	176^{+25}_{-0}	$33.3^{+0.1}_{-1.9}$	$1.0^{+0.0}_{-0.0}$	$9.9^{+1.8}_{-2.1}$
4.1 ± 0.4	$1.4^{+0.2}_{-0.1}$	$2.3^{+0.3}_{-0.7}$	$1.1^{+0.5}_{-0.5}$	$70.0^{+6.0}_{-5.7}$	$0.6^{+0.1}_{-0.2}$	164^{+2}_{-7}	$37.6^{+0.7}_{-0.2}$	$0.7^{+0.0}_{-0.0}$	$13.6^{+3.2}_{-1.3}$
4.6 ± 0.0	$1.4^{+0.5}_{-0.2}$	$2.9^{+0.7}_{-0.5}$	$4.2^{+3.3}_{-1.1}$	$66.0^{+0.3}_{-2.5}$	$0.5^{+0.0}_{-0.1}$	169^{+7}_{-7}	$38.0^{+1.9}_{-1.7}$	$0.8^{+0.1}_{-0.1}$	$14.2^{+0.8}_{-0.6}$
4.8 ± 0.3	$1.1^{+0.0}_{-0.0}$	$1.9^{+0.1}_{-0.3}$	$0.4^{+0.1}_{-0.1}$	$76.0^{+5.0}_{-0.3}$	$0.3^{+0.0}_{-0.1}$	145^{+5}_{-9}	$40.8^{+0.1}_{-1.7}$	$0.4^{+0.0}_{-0.0}$	$10.1^{+0.4}_{-4.9}$
5.1 ± 0.0	$1.1^{+0.2}_{-0.0}$	$2.0^{+0.2}_{-0.2}$	$0.6^{+0.2}_{-0.7}$	$69.2^{+2.2}_{-4.5}$	$0.4^{+0.0}_{-0.1}$	150^{+3}_{-8}	$41.3^{+1.5}_{-0.2}$	$0.4^{+0.0}_{-0.0}$	$12.5^{+2.5}_{-1.9}$
5.6 ± 0.4	$0.9^{+0.1}_{-0.0}$	$1.6^{+0.2}_{-0.1}$	$0.1^{+0.1}_{-1.3}$	$76.4^{+4.3}_{-1.3}$	$0.4^{+0.1}_{-0.1}$	168^{+16}_{-5}	$39.9^{+1.2}_{-1.4}$	$0.3^{+0.1}_{-0.1}$	$11.7^{+2.2}_{-0.6}$

Upon first inspection, it is clear that the galaxies across both samples are among the most massive in the model universe, yet, as was indicated previously in Fig. 3, there is a preference for the $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies to be more massive (a median stellar mass of $8.4 \times 10^{10} M_{\odot}$ for the $S_{850\ \mu\text{m}} \geq 1$ mJy sample compared to $5.6 \times 10^{10} M_{\odot}$ for the highly star-forming Submm-Faint sample). By construction, each sample is probing the galaxies with the highest SFRs for their stellar mass, all lying well above the median trend, which is closely linked to their high dust masses and gas fractions. The two model populations also show evidence of hosting black holes that are undermassive for galaxies of their stellar masses, and the $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies have high stellar masses for their halo masses. Finally, the stellar metallicities of the model galaxies within each sample appear consistent with the median trend of the general population, although there is a hint that the galaxies from the highly star-forming Submm-Faint sample have lower metallicities than expected for their stellar mass.

We find that the model galaxies in both the $S_{850\ \mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples do not appear as ‘typical’ star-forming galaxies in the simulation; they are massive, unusually gas- and dust-rich galaxies that host undermassive black holes. A simple explanation is that the two model populations comprise the most massive galaxies which, for one reason or another, host an undermassive black hole. This is consistent with the results presented by Matthee & Schaye (2019), who find that the relative efficiency of black hole growth contributes to the scatter in the high-mass end of the star-forming sequence in the EAGLE simulation. Such an undermassive black hole would allow the model galaxies to retain a greater amount of gas in their centres due to a decreased cumulative amount of AGN feedback. In addition, having stellar masses that are high for their halo masses suggests that the $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies are undergoing a somewhat prolonged

period of high SFRs, whereas the highly star-forming Submm-Faint population have potentially initiated their starburst much more recently. This would be consistent with the highly star-forming Submm-Faint galaxies having lower stellar metallicities.

3.2.2 The evolution in the integrated properties for the $S_{850\ \mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies

Relative to the median trend of the general population, the $S_{850\ \mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies always have high SFRs, dust masses, and gas fractions for their stellar masses (see Fig. 4). However, the typical values of the integrated properties for the $S_{850\ \mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies do evolve as a function of redshift, as was shown in Fig. 3.

In general, the $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies at higher redshift are less massive, have higher gas fractions, are more compact, and have higher dust temperatures relative to their lower redshift counterparts. These trends also broadly represent the evolution of the properties of the galaxies in the highly star-forming Submm-Faint sample. Yet, although the trends are similar, the median values of the two populations are commonly offset from one another at a given redshift. For example, the halo, stellar, and black hole masses of the highly star-forming Submm-Faint galaxies are a factor of ≈ 2 – 4 times lower than those of the $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies (see Fig. 3); they are additionally more gas rich and more metal poor (both by up to a factor of ≈ 2), their dust temperatures are a few degrees higher, and their dust masses are a factor of ≈ 2 lower.

It appears, therefore, that the model galaxies that make up the two populations are different in several aspects, potentially suggesting two alternate evolutionary paths. The galaxies from the highly star-forming Submm-Faint subset are, on average, lower mass, more gas

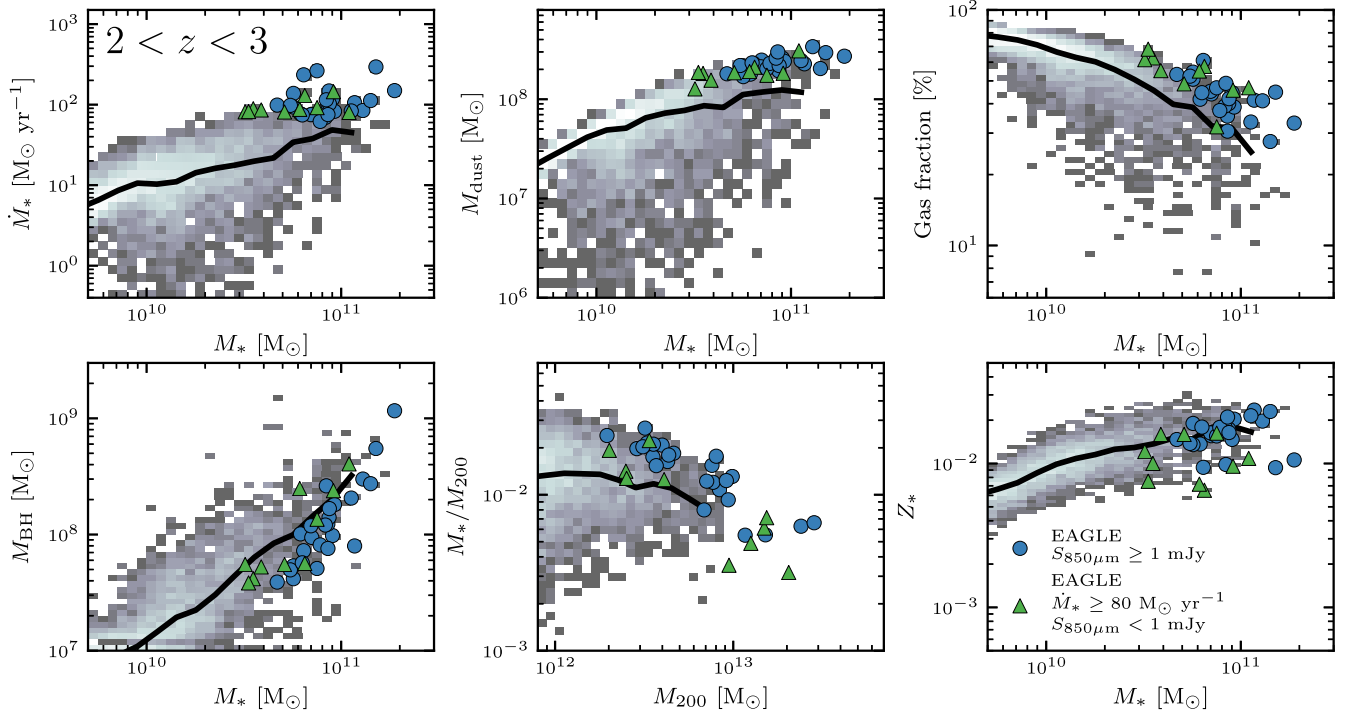


Figure 4. The properties at $z = 2\text{--}3$ of the model galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples (shown individually), plotted as a function of either the stellar or halo mass. We show the SFR (top left), dust mass (top centre), total gas fraction ($M_{\text{gas}}/M_{\text{gas+stars}}$, top right), central supermassive black hole mass (bottom left), stellar mass to halo mass ratio (bottom centre), and stellar metallicity (bottom right). The general population of all model galaxies in the simulation is represented by a two-dimensional histogram, coloured by the number of galaxies in each bin. The median trend for all model galaxies is indicated via a solid line. Both the $S_{850\mu\text{m}} \geq 1$ mJy galaxies and the highly star-forming Submm-Faint galaxies are typically massive ($M_* \approx 4 \times 10^{10}\text{--}2 \times 10^{11} M_\odot$), have high SFRs (by construction), and have the highest dust masses and gas fractions for their stellar masses, all relative to the median trends of the general population in the simulation. In addition, the galaxies within each sample show evidence of hosting undermassive black holes for their stellar masses and the $S_{850\mu\text{m}} \geq 1$ mJy galaxies have stellar masses that are high for their halo masses.

rich, more metal poor, contain less and slightly warmer dust, host lower mass black holes, and preferentially exist at higher redshifts (see Fig. 2) than the galaxies from the $S_{850\mu\text{m}} \geq 1$ mJy subset. Regardless of their classification, however, the galaxies from both samples are always the most gas rich and typically the most massive at their respective redshifts.

3.2.3 Merger fractions

Galaxy–galaxy mergers and interactions provide one potential triggering mechanism for starbursting galaxies: The induced tidal field disturbs any regular orbits of the gas, funnels material inwards towards the galaxy centre, and triggers star formation (e.g. Barnes & Hernquist 1991; Mihos & Hernquist 1996). It is reasonable to assume, therefore, that such interactions may be responsible for creating the observed submm and highly star-forming galaxy populations, which we test here.

In Fig. 5 we investigate the major merger fraction, defined as the fraction of galaxies that have completed or will complete a major merger ($M_{*,1}/M_{*,2} \geq 1/4$) within the previous/next dynamical time,⁷

⁷Equivalent to $|n_{\text{dyn}}| \leq 1$ from equation 1 in McAlpine et al. (2018), where n_{dyn} is the number of dynamical times to the nearest galaxy–galaxy merger with a stellar mass ratio $\geq 1/4$. The dynamical time, t_{dyn} , is defined as the free-fall time of the dark matter halo ($t_{\text{dyn}} \approx 1.6$ Gyr at $z = 0$, ≈ 0.5 Gyr at $z = 2$, and ≈ 0.2 Gyr at $z = 5$); see section 2.3 from McAlpine et al. (2018).

as a function of the stellar mass. The major merger fractions are shown separately for the $S_{850\mu\text{m}} \geq 1$ mJy model galaxies and the highly star-forming Submm-Faint model galaxies, with the merger fractions for all model galaxies in the simulation at a given redshift shown as our control. Due to their limited size, we only divide the two samples into bins of stellar mass and not also by redshift, realizing that the merger fraction of galaxies for a fixed stellar mass increases with increasing redshift (e.g. Rodriguez-Gomez et al. 2015; Qu et al. 2017). However, this effect is accounted for in part by expressing the merger fraction using a set number of dynamical times instead of using a fixed time interval. In addition, separating the galaxies by their stellar masses also acts to separate them by their redshift, as is indicated by the highlighted median redshift of the galaxies in each stellar mass bin (see also Fig. 3). Yet, there could remain a further dependence on the importance of mergers in creating the $S_{850\mu\text{m}} \geq 1$ mJy highly star-forming Submm-Faint populations with redshift (potentially similar to the discovered redshift dependence on mergers for triggering the rapid growth phase of black hole growth in the EAGLE simulation; see McAlpine et al. 2018).

We find no strong trend in the major merger fraction with stellar mass for either the $S_{850\mu\text{m}} \geq 1$ mJy or the highly star-forming Submm-Faint sample. The major merger fractions of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies are ≈ 35 per cent, but this is consistent with those of the general model population, suggesting that major mergers are not a *required* triggering mechanism for the model SMGs. What appears to be more important, therefore, is that they have

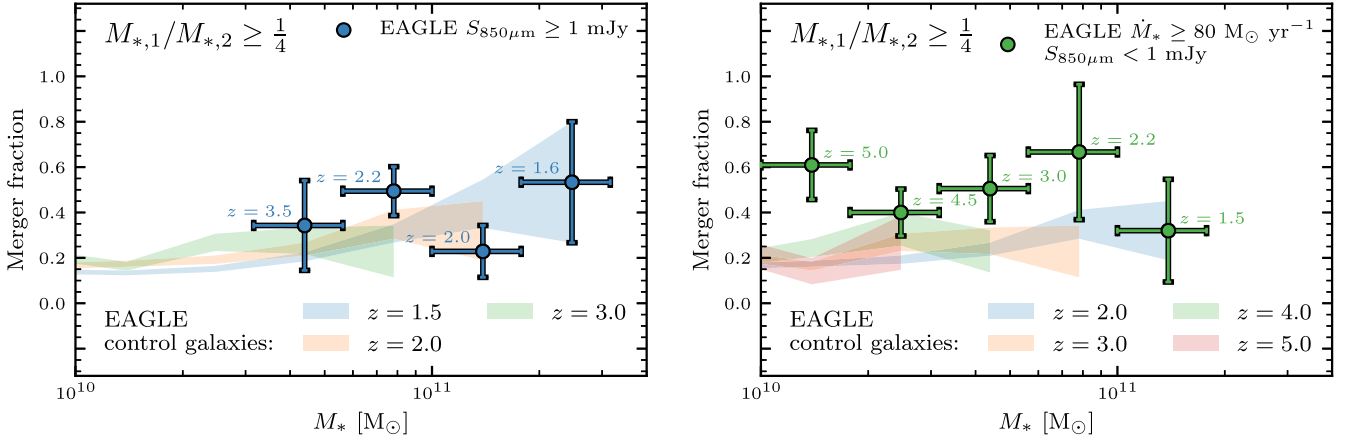


Figure 5. The major merger fraction for the two model samples, defined as the fraction of galaxies that have completed or will complete a major merger ($M_{*,1}/M_{*,2} \geq 1/4$) within the previous/next dynamical time, as a function of stellar mass. The merger fractions are shown separately for the model $S_{850\mu\text{m}} \geq 1$ mJy galaxies (left) and the model highly star-forming Submm-Faint galaxies (right), with the merger fractions of all the galaxies in the simulation at a given redshift shown as shaded regions, acting as our control. The vertical errors in the merger fractions and the height of the shaded regions represent the Poisson uncertainty. The errors on M_* indicate the extent of each stellar mass bin. Each data point is annotated with the median redshift of the galaxies in the given stellar mass bin. The merger fractions of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies are ≈ 35 per cent, but are consistent with the general model population, suggesting major mergers are not the primary trigger of the model SMG population. The highly star-forming Submm-Faint galaxies have merger fractions of ≈ 50 per cent, typically above the general population, (with the greatest excess coming at lower stellar masses $M_* \ll 10^{11} M_\odot$, where it is a factor of ≈ 3), suggesting major mergers are the main driver for triggering this galaxy subset.

a large gas reservoir and host an undermassive black hole (see Section 3.2.1). By contrast, the major merger fractions of the highly star-forming Submm-Faint galaxies are ≈ 50 per cent, typically higher than those of the general population (particularly at the lowest stellar masses, $M_* \sim 10^{10} M_\odot$, where the enhancement is a factor of ≈ 3), suggesting that major mergers are important in triggering this galaxy subset.

We note that whilst the merger fractions universally increase in the case of minor mergers ($1/10 \leq M_{*,1}/M_{*,2} < 1/4$) and either major or minor mergers ($M_{*,1}/M_{*,2} \geq 1/10$), the merger fractions of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies, while high, are always similar to those of the general population. An enhancement in the merger fraction above the control galaxies for the highly star-forming Submm-Faint population exists for all merger classifications; however, the enhancement is greatest in the case of major mergers.

3.2.4 The contribution to the cosmic SFR density

As the levels of star formation exhibited by the galaxies within the two samples are so high, it is then interesting to see, even though they are relatively rare, if the two populations can still contribute significantly to the cosmic SFR density (CSFRD) budget of the universe. The top left-hand panel of Fig. 6 shows the CSFRD (i.e. the total amount of star formation per unit volume) for galaxies with stellar masses $M_* \geq 10^{10} M_\odot$ in the redshift range $z = 1-7$. In addition to this, we show the star-forming density of the galaxies within each of the two samples, with their fractional contribution to the total CSFRD shown in the lower panel.

For galaxies with $M_* \geq 10^{10} M_\odot$, those in the $S_{850\mu\text{m}} \geq 1$ mJy sample contribute an average of 1–6 per cent towards the total CSFRD between the redshifts 1 and 4, with their greatest contribution coming at $z \approx 2.5$ (when the population is most abundant, see Fig. 2). At redshifts below $z \approx 2$, the highly star-forming Submm-Faint galaxies contribute less than 1 per cent towards the total CSFRD, yet this population becomes increasingly dominant, and eventually contributes >70 per cent towards the total CSFRD at $z >$

5. Therefore, the galaxies within the two samples are typically never a dominant source of star formation within the universe; the low numbers of highly star-forming sources are simply overwhelmed by the vastly increased abundance of lower star-forming galaxies. There is however an exception at the highest redshifts ($z > 5$), where the highly star-forming Submm-Faint galaxies become the primary contributor to the CSFRD.

If we restrict our consideration to only the highest mass galaxies ($M_* \geq 10^{11} M_\odot$, shown in the right-hand panel), the contribution to the CSFRD from both the $S_{850\mu\text{m}} \geq 1$ mJy and the highly star-forming Submm-Faint galaxies is much greater: ≈ 10 per cent of the CSFRD at $1 < z < 2$ is produced by the highly star-forming Submm-Faint population, and up to 100 per cent of the CSFRD density at $z \gtrsim 2$ is produced by $S_{850\mu\text{m}} \geq 1$ mJy galaxies.

3.3 The descendants of $S_{850\mu\text{m}} \geq 1$ mJy galaxies and highly star-forming Submm-Faint galaxies at $z = 0$

We conclude our analyses by investigating the descendants of the galaxies in both samples at $z = 0$. Here we are asking the following: What fraction of today’s galaxies have undergone a $S_{850\mu\text{m}} \geq 1$ mJy or highly star-forming Submm-Faint phase in their past? Do the galaxies in either of the two samples retain any signatures that could allow them to be identified in today’s parameter space? What fraction of the total *in situ* mass formed within the galaxies from the two samples was formed during high-SFR events?

Fig. 7 shows the $z = 0$ central supermassive black hole mass, stellar metallicity, alpha-enhancement ($[\alpha/\text{Fe}]_*$ ⁸), stellar mass to halo mass ratio, SFR, and total gas fraction, each plotted as a function of either the stellar or halo mass. The general population of model galaxies is shown as a two-dimensional histogram and

⁸We define alpha-enhancement following equation 1 from Segers et al. (2016), where we use $[\text{O}/\text{Fe}]_*$ as a proxy for $[\alpha/\text{Fe}]_*$.

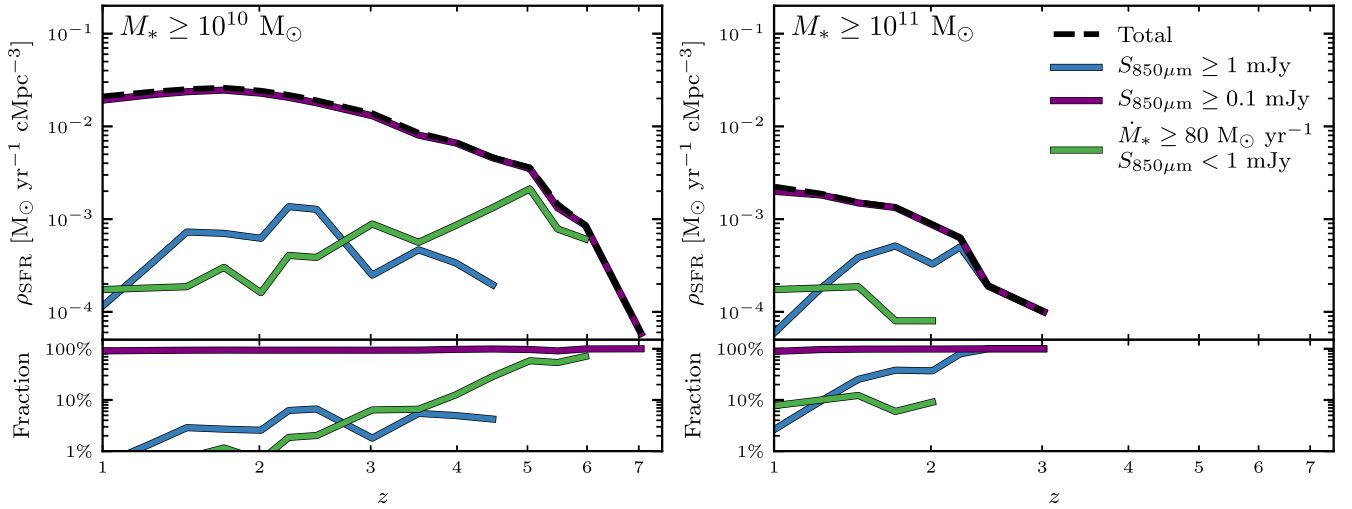


Figure 6. The cosmic SFR density (CSFRD) for galaxies with masses $M_* \geq 10^{10} M_\odot$ (left) and $M_* \geq 10^{11} M_\odot$ (right) in the redshift range $z = 1-7$. The star-forming density of the galaxies within the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples are also shown, with their fractional contribution to the total CSFRD shown in the lower panels. The additional $S_{850\mu\text{m}} \geq 0.1$ mJy sample of galaxies highlights what fraction of the total CSFRD we would expect to recover with a deeper submm survey. For the mass range $M_* \geq 10^{10} M_\odot$, neither the $S_{850\mu\text{m}} \geq 1$ mJy nor the highly star-forming Submm-Faint galaxies contribute more than ≈ 6 per cent to the total CSFRD, until $z \approx 4$, after which up to 70 per cent of the total CSFRD is produced in highly star-forming Submm-Faint galaxies. At higher masses ($M_* \geq 10^{11} M_\odot$), both samples contribute more to the total CSFRD: ≈ 10 per cent from the highly star-forming Submm-Faint galaxies in the redshift range $1 < z < 2$ and up to 100 per cent from the $S_{850\mu\text{m}} \geq 1$ mJy galaxies at redshifts $z > 2$.

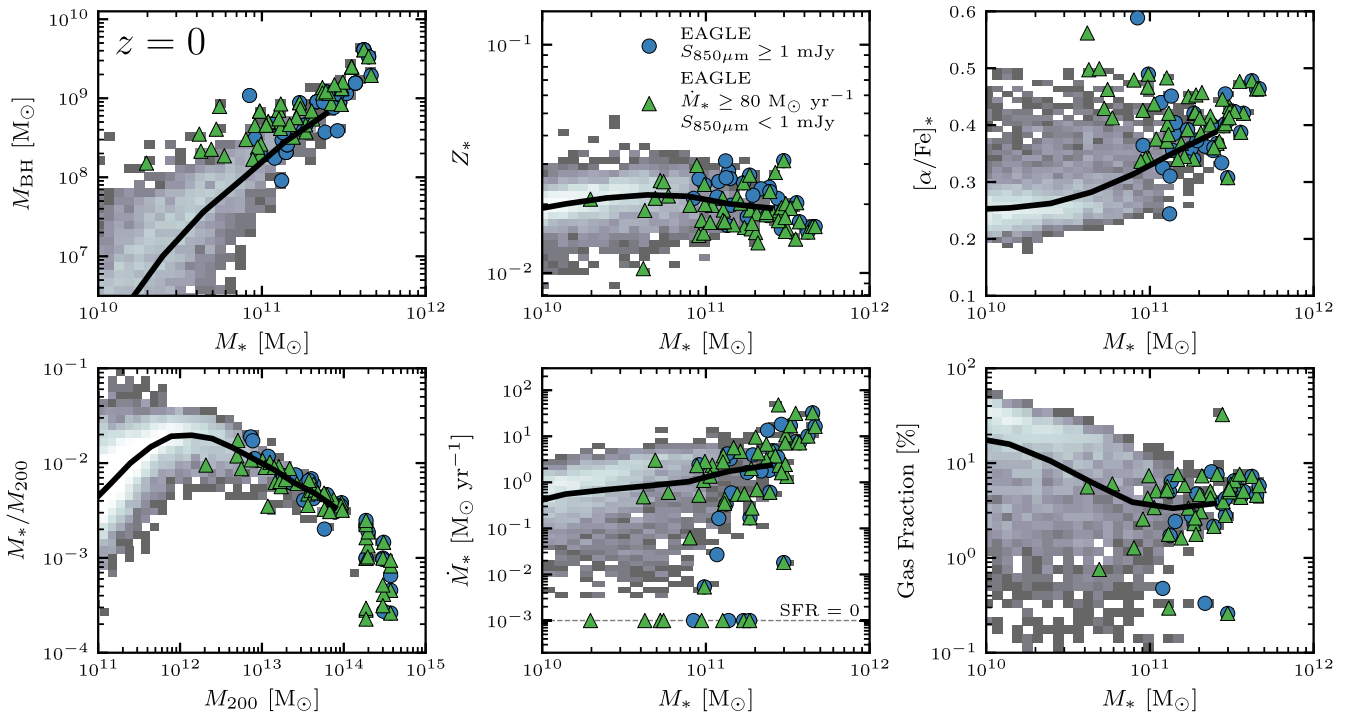


Figure 7. The properties at $z = 0$ of the descendants of the model galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples plotted as a function of either the stellar or halo mass. We plot the central supermassive black hole mass (top left), stellar metallicity (top centre), alpha-enhancement (top right), stellar mass to halo mass ratio (bottom left), SFR (bottom centre), and total gas fraction ($M_{\text{gas}}/M_{\text{gas+stars}}$, bottom right). The total model galaxy population is represented as a two-dimensional histogram, coloured by the number of galaxies in each bin. The median trend for all galaxies is indicated via a solid line. We assign galaxies with a zero SFR to a value of $M_* = 10^{-3} M_\odot \text{ yr}^{-1}$ (indicated by a horizontal dashed line on the figure). In all panels, both central and satellite galaxies are shown (except in the lower left-hand panel, where the general population only shows central galaxies for clarity). Descendants of galaxies that were once $S_{850\mu\text{m}} \geq 1$ mJy systems are not obviously distinguishable from the general population at a given stellar mass in any of the integrated properties we have explored (i.e. they follow the median trends). However, the once highly star-forming Submm-Faint galaxies have, on average, overmassive black holes, lower metallicities, and are more α -enhanced for their stellar masses at $z = 0$.

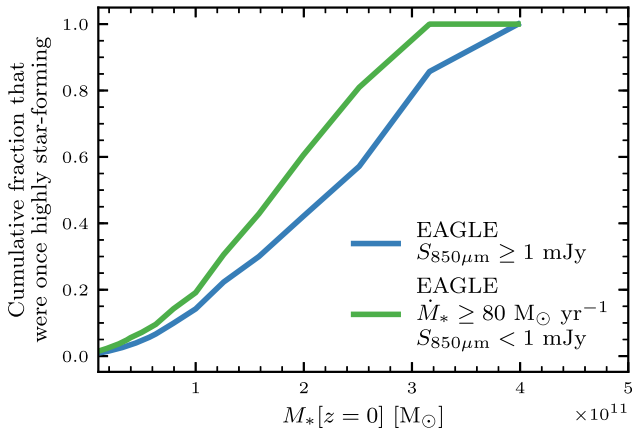


Figure 8. The cumulative fraction of galaxies at $z = 0$ that were ever once $S_{850\mu\text{m}} \geq 1$ mJy or highly star forming and submm faint, as a function of the present day stellar mass. The descendants of these galaxies are rare; only ≈ 20 per cent of the galaxies at $z = 0$ with stellar masses greater than $M_*[z = 0] \geq 10^{11} M_\odot$ were once highly star forming. However, this fraction rapidly increases to 100 per cent for stellar masses $M_*[z = 0] \geq 3 \times 10^{11} M_\odot$.

the descendants of the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint model galaxies are highlighted individually.

3.3.1 The fraction of today's galaxies that were once highly star forming

The clearest distinguishing feature in Fig. 7 for the descendants of both $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies, which was also true at the times when they were selected (see Fig. 4), is that they are typically massive. The median masses are $M_* = 2.1 \times 10^{11} M_\odot$ and $M_{200} = 6.9 \times 10^{13} M_\odot$ and $M_* = 1.8 \times 10^{11} M_\odot$ and $M_{200} = 6.5 \times 10^{13} M_\odot$ for the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples, respectively. Yet, not all massive galaxies today were once highly star forming. To quantify this more clearly, we show Fig. 8, which presents the cumulative fraction of the galaxies at $z = 0$, as a function of the $z = 0$ stellar mass, that were ever either $S_{850\mu\text{m}} \geq 1$ mJy or highly star forming and submm faint in their past (i.e. the cumulative fraction of today's population that are the descendants of the galaxies within either sample). We see a rapid decline in the cumulative fraction for both classifications as the stellar mass decreases, from 100 per cent of galaxies above $3 \times 10^{11} M_\odot$ to just ≈ 20 per cent of the galaxies above $1 \times 10^{11} M_\odot$.⁹ We note that these fractions serve as a lower limit, as it is unlikely that all instances of galaxies in a high-SFR phase can be captured by the temporal spacings of the simulation snapshots. However, by considering the galaxies today that were once highly star forming (i.e. if they were ever above $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ in their past, obtained by examining accurate SFR histories; see Section 2.1.2), we find the temporal spacing of the snapshots is sufficient to capture >80 per cent of high-SFR events, making for an almost complete sample.

Therefore, with the exception of the galaxies with the highest stellar masses ($M_* \gtrsim 1-3 \times 10^{11} M_\odot$), it is rare for a given galaxy today to have been either $S_{850\mu\text{m}} \geq 1$ mJy or highly star forming

⁹We note that a given galaxy at $z = 0$ could have been *both* $S_{850\mu\text{m}} \geq 1$ mJy and highly star forming and submm faint at different points in its history. Thus, the sum of the cumulative fractions between the two samples can exceed 1.

and submm faint in its past, particularly at lower stellar masses ($M_* < 10^{11} M_\odot$).

3.3.2 The descendants in today's parameter space

Now we ask what characteristics (at a fixed mass) may be imprinted upon the descendants of the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming submm-faint galaxies that could potentially distinguish them from the general population of galaxies at $z = 0$.

For many of the integrated properties shown in Fig. 7, and for the majority of the integrated properties we have explored, we do not find that the descendants of the galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples are distinguishable from the general population at $z = 0$ (i.e. they lie on or around the median trend for galaxies of their stellar/halo masses). For example, they have the expected stellar masses for their halo masses, and the expected sizes, SFRs, and total gas fractions for their stellar masses. In fact, for the $S_{850\mu\text{m}} \geq 1$ mJy population, no property appears obviously offset away from the median trends of the general population at a given stellar/halo mass. However, some properties do begin to isolate the descendants of the highly star-forming Submm-Faint sample. They typically host black holes overmassive for their stellar masses, with the most extreme examples occurring at lower stellar masses ($M_* \ll 10^{11} M_\odot$), suggesting a strong link between high-redshift starbursting galaxies and black hole growth. In addition, their descendants are typically metal poor and alpha-enhanced, the combination of which would suggest they have undergone a rapid stellar mass build-up followed by a rapid quenching of continued star formation in their past, which again likely points towards a period of rapid black hole growth (and its associated AGN feedback; see also Segers et al. 2016).

Whilst only a minority of the galaxies in the two samples were classified as satellites at the times that they were highly star forming (10 per cent of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies and 3 per cent of the highly star-forming Submm-Faint galaxies), 35 per cent of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies and 32 per cent of the highly star-forming Submm-Faint galaxies evolve to become satellites by $z = 0$. This would suggest a link between the galaxies within the two samples and the environment. In fact, many of the descendants (particularly from the highly star-forming Submm-Faint sample) reside as satellites in today's most massive haloes. For example, in the three most massive haloes from the simulation at $z = 0$, 25 of their galaxies were once either $S_{850\mu\text{m}} \geq 1$ mJy or highly star forming and submm faint in their past. The act of galaxies infalling into these massive haloes is the cause for a number of the descendants having 'low' stellar masses (due to stellar stripping), zero SFRs (due to ram pressure stripping), and excessively overmassive black holes for their stellar mass (again as a result of the stellar stripping; see Barber et al. 2016 for a study of these sources).

3.3.3 The contribution to the in situ stellar mass budget from high-SFR events

For a galaxy to be in one of our two samples it must achieve an intrinsically high SFR or submm flux (which also indicates a high SFR; see Fig. 1). We are therefore capturing galaxies at a moment in time when they are among the most highly star-forming systems at their respective redshifts. However, without knowledge of the typical duration of the highly star-forming phase, it is unclear how much this star-forming episode will contribute to the total *in situ* stellar mass budget of the galaxy. That is, is such a phase typically

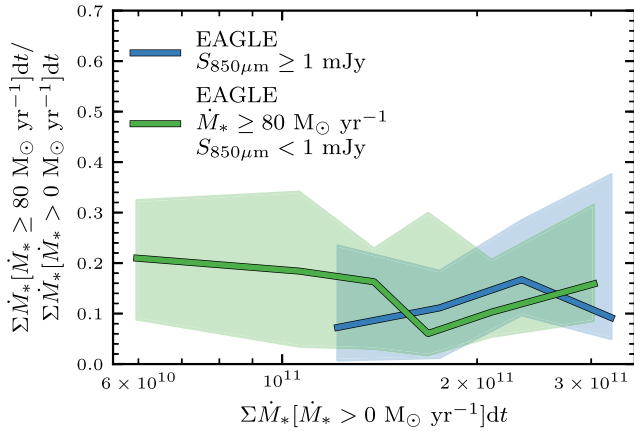


Figure 9. For the descendants of the model galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint samples at $z = 0$, the fraction of the total initial stellar mass formed *in situ* within the galaxy at a rate above $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ as a function of the total initial mass that formed *in situ* within the galaxy. On average, the fractions in both samples are low (10–20 per cent, implying a typical starbursting duration of ≈ 100 –400 Myr), indicating that high-SFR events are not the main contributor to the *in situ* production of stars for the galaxies in both samples. Instead, the majority of stars produced *in situ* within these galaxies are formed at lower SFRs, over longer periods of time.

rapid and short-lived, and so may contribute relatively little to the total *in situ* stellar mass budget, or does a significant fraction of the total stellar mass that is grown within the galaxy form during this phase?

Fig. 9 shows, for the descendants of both the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies at $z = 0$, the fraction of the total *initial* stellar mass that was *formed in situ within the galaxy* at a rate above $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$, as a function of the total initial stellar mass that was formed *in situ* within the galaxy (both of these values are computed by integrating the galaxy’s entire star formation history; see Section 2.1.2). For both samples, periods of high SFR are never the main contributor to the production of stars within these galaxies, typically contributing an average of 10–20 per cent to the total *in situ* initial stellar mass. To put this into context, for an average galaxy that has formed $10^{11} M_\odot$ worth of *in situ* stars by $z = 0$, ≈ 20 per cent of these were formed during periods of high SFRs, implying a typical duration of ≈ 250 Myr for the high-SFR phase(s) (combined between one or multiple events). Therefore, the majority of the *in situ* stellar mass formed within the descendants of the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies occurred at lower SFRs, and over longer periods of time.

3.4 The hidden population of highly star-forming galaxies with faint submm fluxes

There is increasing empirical evidence that SMGs at higher redshifts ($z \geq 3$) have higher dust temperatures than their lower redshift ($z \approx 2$) counterparts (Cooke et al. 2018; Schreiber et al. 2018). An increasing dust temperature with increasing redshift is also seen in the model galaxies from the $S_{850\mu\text{m}} \geq 1$ mJy sample ($T_{\text{dust}} = 29.8^{+1.5}_{-0.1}$ K at $z \approx 1$ to $32.6^{+1.5}_{-0.8}$ K at $z \approx 4$; see Fig. 3). This could be linked to physical differences in their galaxy properties, such as their SFR surface densities, which would be consistent with the trend of an approximately constant SFR (see Fig. 1) combined with a decreasing galaxy size with increasing redshift (see Fig. 3), as found for the $S_{850\mu\text{m}} \geq 1$ mJy model galaxies. However, the model

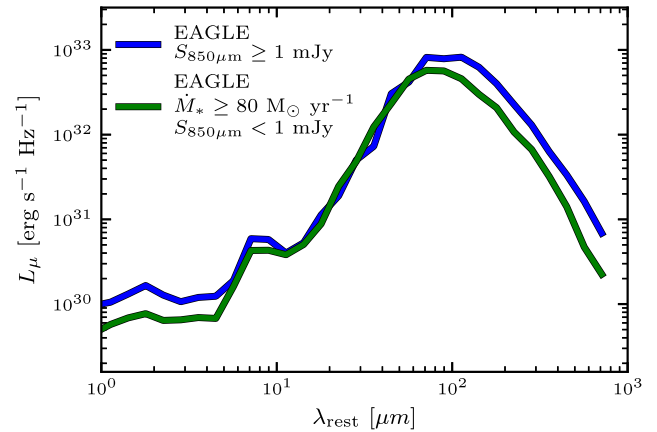


Figure 10. Composite broad-band spectral energy distributions for the $S_{850\mu\text{m}} \geq 1$ mJy and highly star-forming Submm-Faint galaxies. The leftward shifting of the peak in the dust emission for the highly star-forming Submm-Faint sample signifies hotter dust temperatures than the $S_{850\mu\text{m}} \geq 1$ mJy galaxies.

galaxies from the highly star-forming Submm-Faint sample have dust temperatures that are systematically higher than those from the $S_{850\mu\text{m}} \geq 1$ mJy sample at all redshifts ($T_{\text{dust}} = 29.6^{+1.7}_{-1.0}$ K at $z \approx 2$ to $39.9^{+1.2}_{-1.4}$ K at $z \approx 6$; see Fig. 3), which, in combination with their lower dust masses (see Fig. 3), is why they do not make it into the $S_{850\mu\text{m}} \geq 1$ mJy sample.

The reason an increased dust temperature may exclude the highly star-forming Submm-Faint galaxies from the $S_{850\mu\text{m}} \geq 1$ mJy sample becomes clear when looking at the spectral energy distributions (SEDs) of the galaxies. Fig. 10 shows the composite broad-band SED for the two populations of model galaxies, which clearly reveals a shift to higher temperatures in the dust emission peak for the highly star-forming Submm-Faint galaxies relative to the $S_{850\mu\text{m}} \geq 1$ mJy galaxies, confirming the hotter dust component, and a corresponding reduction in the the submm brightness. In addition, the lower dust masses of the highly star-forming Submm-Faint population further reduce the output at longer wavelengths. The existence of a population of ‘hot’, dusty ultraluminous galaxies at high redshift has been suggested previously in the literature (e.g. Blain et al. 2004; Chapman 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 higher characteristic dust temperatures than the detected submm population.

To uncover the model population of highly star-forming Submm-Faint galaxies observationally, one therefore has to observe deeper in the submm/mm. One way to achieve this is through gravitationally lensed sources, or, in recent years, ALMA has begun to explore the fainter submm/mm population ($S_{850\mu\text{m}} < 1$ mJy) by overcoming the high confusion limit inherent to single-dish telescopes (e.g. Hatsukade et al. 2013; Fujimoto et al. 2016). Both the lensing studies (e.g. Hsu et al. 2016) and the studies of deep ALMA imaging in the *Hubble* Ultra Deep Field (Dunlop et al. 2017) and GOODS-S field (Franco et al. 2018; Hatsukade et al. 2018) estimate that up to ≈ 90 per cent of star formation in the Universe is locked up in dust-obscured systems (i.e. those with $S_{1.1\text{mm}}/S_{850\mu\text{m}} \gtrsim 0.1$ mJy). This is similarly true for the EAGLE model population, where $\gtrsim 90$ per cent of the CSFRD in galaxies with masses greater than $M_* \geq 10^{10} M_\odot$ is recovered when using a reduced flux limit of $S_{850\mu\text{m}} \geq 0.1$ mJy (see Fig. 6). However, reducing the flux limit would not only reveal the

highly star-forming Submm-Faint population, it would primarily uncover the majority of the general star-forming population at each epoch. Thus, other than being able to independently constrain the SFR of the galaxy, or preferentially selecting those at $z \gtrsim 5$, it will remain observationally extremely challenging to uncover the highly star-forming Submm-Faint population as found by EAGLE.

4 SUMMARY AND CONCLUSIONS

In this study, we have investigated the nature of the model galaxies in the EAGLE simulation with the highest SFRs throughout cosmic time. Our analysis focused on two model galaxy samples. The first sample comprised the simulated galaxies that yielded the highest mock submm fluxes ($S_{850\mu\text{m}} \geq 1$ mJy), which acted as analogues to the observed submm population. The second sample contained the simulated galaxies that were defined as highly star forming ($\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$), but did not make it into the $S_{850\mu\text{m}} \geq 1$ mJy sample (i.e. they are also $S_{850\mu\text{m}} < 1$ mJy, or ‘Submm-Faint’). This second sample revealed what subset of the model galaxies with high SFRs the $S_{850\mu\text{m}} \geq 1$ mJy sample is (or more importantly is not) selecting.

In Section 3.1 we began with a comparison to the observed submm population, finding some encouraging similarities when compared against the $S_{850\mu\text{m}} \geq 1$ mJy model galaxy sample. First, a purely submm-based selection returned the simulated galaxies with the highest SFRs ($\dot{M}_* \approx 50\text{--}300 M_\odot \text{ yr}^{-1}$, with a median SFR of $94 M_\odot \text{ yr}^{-1}$), agreeing well with the inferred SFRs of observed submm sources at $z \lesssim 3$. However, the model $S_{850\mu\text{m}} \geq 1$ mJy galaxies in the tail-end of the redshift distribution ($z > 3$) had SFRs that were a factor of ≈ 3 lower than those inferred by the observations, yet we note that less than a fifth of the $S_{850\mu\text{m}} \geq 1$ mJy model galaxies lie at these redshifts. Additionally, the redshift distribution of the galaxies in the $S_{850\mu\text{m}} \geq 1$ mJy sample reasonably reproduced the shape and median value of the observed distribution. Finally, many of the integrated properties for the $S_{850\mu\text{m}} \geq 1$ mJy model galaxies broadly agreed with the observed values of submm sources. In combination, these results give us the confidence to use the $S_{850\mu\text{m}} \geq 1$ mJy sample in an attempt to answer the questions relating to the origin, evolution, and eventual fate of the observed submm population.

The $S_{850\mu\text{m}} \geq 1$ mJy galaxies were found to be typically massive ($M_* \sim 10^{11} M_\odot$), star-forming ($\dot{M}_* \approx 100 M_\odot \text{ yr}^{-1}$) galaxies at $z \approx 2\text{--}3$ with considerable dust masses ($M_{\text{dust}} \sim 10^8 M_\odot$) and gas fractions ($f_{\text{gas}} \approx 50$ per cent) that host black holes that are undermassive for galaxies of their stellar masses. The $z = 0$ descendants of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies were again massive ($M_* \geq 10^{11} M_\odot$), with all of the galaxies at $z = 0$ with stellar masses above $M_* \geq 3 \times 10^{11} M_\odot$ having had a $S_{850\mu\text{m}} \geq 1$ mJy phase in their past. However, at a given stellar mass, the $S_{850\mu\text{m}} \geq 1$ mJy galaxies were not obviously distinguishable from the general population of model galaxies. As the black holes of the once $S_{850\mu\text{m}} \geq 1$ mJy galaxies do not remain undermassive by $z = 0$, and come to lie on or around the median value for their stellar mass by the present day, they must have experienced a vigorous period of black hole growth (delayed after the starbursting phase) which grew the black hole and lowered the gas and dust content via AGN feedback. The typical life cycle of model SMGs, and their connection to black hole growth, will be the subject of future work.

The galaxies from the $S_{850\mu\text{m}} \geq 1$ mJy sample had major merger ($M_{*,1}/M_{*,2} \geq 1/4$) fractions of ≈ 35 per cent, but this is similar to those of the general population, suggesting that major mergers are not the primary driver of the model submm population. This

result would conflict with the findings of many observational studies that have suggested ≈ 100 per cent major merger fractions for submm galaxies (e.g. Ivison et al. 2007; Tacconi et al. 2008; Engel et al. 2010; Alaghband-Zadeh et al. 2012; Chen et al. 2015). However, the concept of galaxies ‘in the state of a merger’ is poorly defined; observations must deduce this state from a galaxy’s asymmetry/disturbance or relative distance to a close companion, whereas we define it as being within a fixed number of dynamical times from coalescence. Furthermore, major mergers may play a greater role in triggering the most extreme ($S_{850\mu\text{m}} \gtrsim 5$ mJy) SMGs, which are those primarily studied observationally (however, our current simulation volume is too small to fully explore this regime). Regardless of the definition, a mass/redshift-matched control sample is essential to begin to infer the ‘importance’ of mergers as triggering mechanisms, yet, given the natural abundance of mergers within a hierarchical formation of structure, even then disentangling their true importance can remain challenging. As the merger fractions of the $S_{850\mu\text{m}} \geq 1$ mJy galaxies were so similar to the merger fractions of the general population, we would argue that mergers and interactions may be sufficient, but not necessary, to create a submm starburst. Instead, the factor that appears essential is a significant gas reservoir due to an undermassive black hole.

The duration of the ‘submm phase’ for the $S_{850\mu\text{m}} \geq 1$ mJy galaxies was found to be relatively brief (100–400 Myr) and contributed an average of $\approx 10\text{--}20$ per cent towards the total *in situ* stellar mass production of the model galaxies. The importance of the submm phase for building up the stellar mass within a galaxy is an interesting question that is often raised when discussing the nature of submm sources. Given the levels of star formation that are inferred for the observed submm population (100s to 1000s of $M_\odot \text{ yr}^{-1}$), the stellar mass of a massive galaxy ($M_* \gtrsim 10^{11} M_\odot$) could be formed in its totality in just a few hundred million years. This has spawned the suggestion that submm galaxies are the progenitors of today’s most massive spheroidal galaxies (e.g. Lilly et al. 1999; Swinbank et al. 2006; Fu et al. 2013; Simpson et al. 2014). Indeed, Swinbank et al. (2006) demonstrated that by adopting a submm-burst lifetime of 300 Myr, coupled with the observed properties of SMGs at $z \approx 2$, the stellar populations of submm galaxies could evolve on to the scaling relations of the most massive elliptical galaxies in the local Universe (see also Simpson et al. 2014). The alternate extreme has also been suggested, whereby the submm phase makes no significant contribution (≈ 2 per cent) to the eventual stellar mass of the SMG descendants at $z = 0$ (González et al. 2011). However, our results suggest the correct answer lies somewhere between these two extremes.

In contrast to the $S_{850\mu\text{m}} \geq 1$ mJy galaxies, the highly star-forming Submm-Faint galaxies were found to be a set of lower mass ($M_* \sim 10^{10} M_\odot$) higher redshift ($z > 4$) galaxies (again with high dust masses, high gas fractions, and undermassive black holes) that show stronger evidence of being driven primarily by major mergers. A similar fraction of the *in situ* stellar mass budget is built within the starbursting phase compared with the $S_{850\mu\text{m}} \geq 1$ mJy galaxies ($\approx 10\text{--}20$ per cent). Their descendants at $z = 0$, whilst also massive ($M_* \gtrsim 10^{11} M_\odot$), were typically metal-poor, alpha-enhanced, and hosted overmassive black holes for galaxies of their stellar masses, potentially making them identifiable in today’s parameter space.

To summarize, whilst collectively the galaxies in the two samples have the highest SFRs in the simulation and are the progenitors of the most massive galaxies at $z \approx 0$ ($M_* > 10^{11} M_\odot$), they each preferentially select galaxies in two different regimes. The $S_{850\mu\text{m}} \geq 1$ mJy galaxies are massive dust- and gas-rich star-forming galaxies at $z \approx 2\text{--}3$ with undermassive black holes. These requirements

therefore largely preclude the existence of $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies at higher redshifts (given how rare such massive galaxies would be at these times), as they predominantly need a suitably massive, but also suitably enriched, gas reservoir. By contrast, the highly star-forming Submm-Faint galaxies are high-redshift, lower mass, gas- and dust-rich star-forming galaxies. Galaxies primarily enter this sample through a starburst triggered via a major merger; however, it will also include galaxies at the massive end of the star-forming main sequence at high redshift (recognizing that the typical SFR for a galaxy of fixed mass increases with increasing redshift; e.g. Furlong et al. 2015, for the case of EAGLE). Both scenarios predominantly select $M_* \sim 10^{10} M_\odot$ galaxies, thus largely excluding them from the $S_{850\ \mu\text{m}} \geq 1$ mJy sample. These high-redshift galaxies are likely missed by many submm surveys due to their lower masses and higher dust temperatures.

We have used the EAGLE cosmological hydrodynamical simulation of galaxy formation to investigate the nature of the model galaxies with the highest mock submm fluxes ($S_{850\ \mu\text{m}} \geq 1$ mJy). In addition, we investigated the nature of the ‘highly star-forming’ model galaxies that were also ‘Submm-Faint’ (i.e. $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ but $S_{850\ \mu\text{m}} < 1$ mJy). Here we report our main conclusions.

4.1 $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies

(i) **The $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies broadly reproduce the properties of the observed submm population.** They have high SFRs ($\dot{M}_* \approx 50\text{--}300 M_\odot \text{ yr}^{-1}$; see Fig. 1), broadly reproduce the shape and median value of the observed redshift distribution (see Fig. 2), and reproduce a variety of integrated galaxy and halo properties from the current observations of the submm population (see Fig. 3).

(ii) **The integrated properties of the $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies evolve with redshift.** At a given redshift, the $S_{850\ \mu\text{m}} \geq 1$ mJy model population comprises massive ($M_* \sim 10^{11} M_\odot$) gas- ($f_{\text{gas}} \approx 50$ per cent) and dust-rich ($M_{\text{dust}} \sim 10^8 M_\odot$) starbursting galaxies ($\dot{M}_* \approx 100 M_\odot \text{ yr}^{-1}$) that host undermassive black holes (see Fig. 4). With decreasing redshift, the $S_{850\ \mu\text{m}} \geq 1$ mJy population have higher halo, stellar, and black hole masses, they become increasingly gas poor, their sizes, velocity dispersions, and dust masses increase, their metallicities remain approximately constant, and their dust temperatures decrease (see Fig. 3).

(iii) **$S_{850\ \mu\text{m}} \geq 1$ mJy galaxies have major merger ($M_{*,1}/M_{*,2} \geq 1/4$) fractions of ≈ 35 per cent, similar to the general population.** This would suggest that major mergers are not the primary trigger of the EAGLE submm population (see Fig. 5). Instead, what is critical is that there is an adequate gas reservoir present due to a lower than average black hole mass.

(iv) **$S_{850\ \mu\text{m}} \geq 1$ mJy galaxies are significant contributors (50–100 per cent) to the total cosmic SFR density in massive galaxies ($M_* \geq 10^{11} M_\odot$), particularly above $z \approx 1.5$.** However, only $\approx 1\text{--}6$ per cent of the total cosmic SFR density for galaxies with stellar masses $M_* \geq 10^{10} M_\odot$ is produced by $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies in the redshift range $1 < z < 4$ (see Fig. 6).

(v) **The majority of star production in $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies occurs outside the high-SFR event(s).** On average, 10–20 per cent of the stars that are born within the $S_{850\ \mu\text{m}} \geq 1$ mJy model galaxies do so at high SFRs ($\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$), which implies a typical starbursting duration of $\approx 100\text{--}400$ Myr (combined between one or more events). Therefore, the majority of the stellar mass build-up in $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies occurs at lower SFRs, over longer periods of time (see Fig. 9).

(vi) **The descendants of the $S_{850\ \mu\text{m}} \geq 1$ mJy population at $z = 0$ are massive ($M_* > 10^{11} M_\odot$).** However, not all massive galaxies today were once $S_{850\ \mu\text{m}} \geq 1$ mJy (see Fig. 8). At a given stellar mass, the descendants of $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies have gas fractions, dust masses, black hole masses, metallicities, and alpha-enhancements (and all other integrated properties that we have explored) that are similar to the median trend for all galaxies (see Fig. 7). This suggests that it would be difficult to identify the descendants of $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies using any of these criteria in today’s parameter space. A moderate fraction of the once $S_{850\ \mu\text{m}} \geq 1$ mJy galaxies evolve to become satellites by the present day (35 per cent, up from 10 per cent at the time they were selected), suggesting the environment may play a role in their formation.

4.2 Highly star-forming Submm-Faint galaxies (i.e. $\dot{M}_* \geq 80 M_\odot \text{ yr}^{-1}$ but $S_{850\ \mu\text{m}} < 1$ mJy)

(i) **Highly star-forming Submm-Faint galaxies predominantly exist at higher redshift ($z > 4$; see Fig. 2).** Similar to the $S_{850\ \mu\text{m}} \geq 1$ mJy population, highly star-forming Submm-Faint galaxies have high SFRs, dusts masses, and gas fractions, and host undermassive black holes (see Fig. 4). The integrated properties of the highly star-forming Submm-Faint galaxies evolve with redshift in a similar manner to the $S_{850\ \mu\text{m}} \geq 1$ mJy galaxy population (see Fig. 3).

(ii) **Highly star-forming Submm-Faint galaxies have major merger ($M_{*,1}/M_{*,2} \geq 1/4$) fractions typically greater than the general population.** This would suggest that major mergers are more important for triggering this galaxy subset, particularly at lower stellar masses ($M_* \ll 10^{11} M_\odot$), where the enhancement in the major merger fraction is a factor of ≈ 3 greater than those of the general population (see Fig. 5).

(iii) **Highly star-forming Submm-Faint galaxies contribute up to 70 per cent towards the total cosmic SFR density in galaxies with stellar masses $M_* \geq 10^{10}$ at $z \gtrsim 4$.** However, below $z \approx 4$, they contribute only 1–4 per cent (see Fig. 6).

(iv) **The descendants of highly star-forming Submm-Faint galaxies at $z = 0$ have overmassive black holes, are metal poor, and are alpha-enhanced for their stellar masses, relative to the median trends of the general population (see Fig. 7).** In addition, many of the once highly star-forming Submm-Faint galaxies evolve to become satellites of the most massive haloes within the simulation by the present day (32 per cent of the descendants evolved to become satellites, up from 3 per cent at the time they were highly star forming), suggesting the environment may play a role in their formation.

(v) **Highly star-forming Submm-Faint galaxies are faint in the submm wavebands due to their higher dust temperatures and lower dust masses.** This means that a large fraction of the highly star-forming galaxies within the Universe could be potentially missed by the current submm surveys (see Figs 3 and 10), and also suggests that current observational works may be underestimating the total contribution to the cosmic SFR density above $z \gtrsim 3$ (as they would be missed in both the UV and submm bands).

ACKNOWLEDGEMENTS

SM thanks Michelle Furlong and Jonathan Davies for their many contributions to this work.

This work was supported by the Science and Technology Facilities Council (grant number ST/P000541/1) and the Academy of

Finland (grant number 314238). RAC is a Royal Society University Research Fellow. IRS acknowledges support from the (0:content-type="word-wrap")(0:funding-source)ERC(0:funding-source)(0:content-type="word-wrap") Advanced Grant DUSTYGAL (321334).

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure.

REFERENCES

- Alaghband-Zadeh S. et al., 2012, *MNRAS*, 424, 2232
- Alexander D. M. et al., 2008, *AJ*, 135, 1968
- Baes M., Verstappen J., De Looze I., Fritz J., Saffly W., Vidal Pérez E., Stalevski M., Valcke S., 2011, *ApJS*, 196, 22
- 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. E., Hernquist L. E., 1991, *ApJ*, 370, L65
- Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, *MNRAS*, 356, 1191
- Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, *ApJ*, 611, 52
- Boselli A. et al., 2010, *PASP*, 122, 261
- Bothwell M. S. et al., 2013, *MNRAS*, 429, 3047
- Camps P., Baes M., 2015, *Astron. Comput.*, 9, 20
- Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, *MNRAS*, 462, 1057
- 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.-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
- Cowie L. L., González-López J., Barger A. J., Bauer F. E., Hsu L.-Y., Wang W.-H., 2018, *ApJ*, 865, 106
- Cowley W. I., Lacey C. G., Baugh C. M., Cole S., 2015, *MNRAS*, 446, 1784
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- da Cunha E. et al., 2015, *ApJ*, 806, 110
- Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140
- Danielson A. L. R. et al., 2017, *ApJ*, 840, 78
- Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Kereš D., Weinberg D. H., 2010, *MNRAS*, 404, 1355
- Del Moro A. et al., 2013, *A&A*, 549, A59
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Dunlop J. S. et al., 2017, *MNRAS*, 466, 861
- Engel H. et al., 2010, *ApJ*, 724, 233
- Franco M. et al., 2018, *A&A*, 620, A152
- Fu H. et al., 2013, *Nature*, 498, 338
- Fujimoto S., Ouchi M., Ono Y., Shibuya T., Ishigaki M., Nagai H., Momose R., 2016, *ApJS*, 222, 1
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Furlong M. et al., 2017, *MNRAS*, 465, 722
- González J. E., Lacey C. G., Baugh C. M., Frenk C. S., 2011, *MNRAS*, 413, 749
- Grupponi C. et al., 2013, *MNRAS*, 432, 23
- Hatsukade B., Ohta K., Seko A., Yabe K., Akiyama M., 2013, *ApJ*, 769, L27
- Hatsukade B. et al., 2018, *PASJ*, 70, 105
- Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, *MNRAS*, 428, 2529
- Hickox R. C. et al., 2012, *MNRAS*, 421, 284
- Hodge J. A. et al., 2013, *ApJ*, 768, 91
- Hsu L.-Y., Cowie L. L., Chen C.-C., Barger A. J., Wang W.-H., 2016, *ApJ*, 829, 25
- Ivison R. J. et al., 2007, *MNRAS*, 380, 199
- Karim A. et al., 2013, *MNRAS*, 432, 2
- 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
- Lagos C. d. P. et al., 2015, *MNRAS*, 452, 3815
- Lilly S. J., Eales S. A., Gear W. K. P., Hammer F., Le Fèvre O., Crampton D., Bond J. R., Dunne L., 1999, *ApJ*, 518, 641
- Madau P., Dickinson M., 2014, *ARA&A*, 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
- Matthee J., Schaye J., 2019, *MNRAS*, 484, 915
- McAlpine S. et al., 2016, *Astron. Comput.*, 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., Schaye J., Theuns T., 2018, *MNRAS*, 481, 3118
- Mihos J. C., Hernquist L., 1996, *ApJ*, 464, 641
- Planck Collaboration I, 2014, *A&A*, 571, A1
- Qu Y. et al., 2017, *MNRAS*, 464, 1659
- Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, *MNRAS*, 452, 2034
- Rodriguez-Gomez V. et al., 2015, *MNRAS*, 449, 49
- Rosar-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, *MNRAS*, 462, 190
- Schaller M. et al., 2015, *MNRAS*, 451, 1247
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Franco M., 2018, *A&A*, 609, A30
- Scudder J. M., Oliver S., Hurley P. D., Wardlow J. L., Wang L., Farrah D., 2018, *MNRAS*, 480, 4124
- Segers M. C., Crain R. A., Schaye J., Bower R. G., Furlong M., Schaller M., Theuns T., 2016, *MNRAS*, 456, 1235
- Simpson J. M. et al., 2014, *ApJ*, 788, 125
- Simpson J. M. et al., 2015, *ApJ*, 807, 128
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, 361, 776
- Stach S. M. et al., 2018, *ApJ*, 860, 161
- Swinbank A. M., Smail I., Chapman S. C., Blain A. W., Ivison R. J., Keel W. C., 2004, *ApJ*, 617, 64
- Swinbank A. M., Chapman S. C., Smail I., Lindner C., Borys C., Blain A. W., Ivison R. J., Lewis G. F., 2006, *MNRAS*, 371, 465
- Swinbank A. M. et al., 2008, *MNRAS*, 391, 420
- Swinbank A. M. et al., 2014, *MNRAS*, 438, 1267
- Tacconi L. J. et al., 2008, *ApJ*, 680, 246
- The EAGLE team, 2017, preprint ([arXiv:1706.09899](https://arxiv.org/abs/1706.09899))
- Trayford J. W. et al., 2015, *MNRAS*, 452, 2879
- Trayford J. W. et al., 2017, *MNRAS*, 470, 771
- 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

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.