

# What Influences the Efficiency of Galaxy Formation?



UNIVERSITY OF  
LIVERPOOL

by  
**Dominic Taylor** (201276356)

Supervisor: Prof Robert. A. Crain

A thesis submitted in partial fulfillment of the requirements  
for the degree of Master of Physics

in the  
Faculty of Science and Engineering  
Department of Physics

April 2022



# Declaration of Authorship

I, Dominic Taylor, declare that this thesis titled, ‘What Influences the Efficiency of Galaxy Formation?’ and the work presented in it are my own.  
I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---



UNIVERSITY OF LIVERPOOL

## *Abstract*

Faculty of Science and Engineering

Department of Physics

Master of Physics

### **What Influences the Efficiency of Galaxy Formation?**

by **Dominic Taylor**

An investigation into the second order influences on the efficiency of galaxy formation by exploring what physical processes of galaxies correlate with the stellar mass – halo mass relation at fixed halo mass, using the Evolution and Assembly of GaLaxies and their Environments (EAGLE) hydrodynamical simulations. Stellar and AGN feedback provide direct influence and morphological transformation depends on halo assembly history. It was found that feedback influences the ability of galaxies to exploit high birth densities where further studies at lower halo mass may be beneficial. Correlations with metallicity could be better untangled by studying its dependence on stellar feedback, and low and high mass early assembling haloes host more efficiently formed galaxies whereas between regimes late assemblers do. The scaling relation between regimes is sensitive to the energy fraction function and at high halo mass is sensitive to the modelling of AGN feedback as it is essential in reproducing observed morphological transformations, and both should be of prime focus to improve in future simulations so to reduce scatter in the relation.



# Contents

<b>Declaration of Authorship</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Galaxy Formation and Evolution . . . . .	1
1.2 Cosmological Framework . . . . .	5
<b>2 Cosmological Simulations</b>	<b>9</b>
2.1 Simulations . . . . .	9
2.2 EAGLE . . . . .	10
2.3 Data and Sample . . . . .	16
2.4 Aims of the Investigation . . . . .	18
<b>3 Preliminary Investigations of First Order Influences</b>	<b>20</b>
3.1 Stellar Feedback . . . . .	20
3.2 AGN Feedback . . . . .	23
3.3 Black Hole Mass . . . . .	25
<b>4 Stellar Mass - Halo Mass Relation</b>	<b>27</b>
4.1 Galaxy within a Spherical Aperture . . . . .	28
4.2 LOWESS . . . . .	29
4.3 Scaling Relation . . . . .	29
<b>5 Analysis and Results of the Second Order Influences</b>	<b>32</b>
5.1 Birth Density . . . . .	32
5.2 Metallicity . . . . .	36
5.3 Stellar Age . . . . .	38
5.4 Gas Density . . . . .	40
5.4.1 Within a 30kpc Spherical Aperture . . . . .	40
5.4.2 Total Halo . . . . .	43

5.5	Energy Fraction . . . . .	44
5.5.1	Dependence on Stellar Feedback Strength . . . . .	46
5.6	Gas Fraction . . . . .	48
5.6.1	Within a 30kpc Spherical Aperture . . . . .	48
5.6.2	Total Halo . . . . .	50
6	<b>Summary and Discussion</b>	<b>52</b>
<b>A</b>	<b>Computing Particle Quantities</b>	<b>56</b>
A.1	Wrapping the Simulation Box . . . . .	56
A.2	Particle Computed Stellar Mass versus Catalogued . . . . .	56
A.3	Running Median Fit of Third Variables . . . . .	57
<b>B</b>	<b>Poster Question and Answers</b>	<b>61</b>
	<b>Bibliography</b>	<b>63</b>

# List of Figures

1.1	A full sky map produced by the Wilkinson Microwave Anisotropy Probe (WMAP) of the cosmic microwave background radiation, residual energy from the Big Bang (Shu [35]). . . . .	2
1.2	A $100 \times 100 \times 20$ cMpc slice of the Ref-L100N1504 EAGLE simulation at $z = 0$ showing filaments, voids, walls and haloes of the cosmic web. Insets are of dimensions 10 cMpc and 60 ckpc on a side showing a galaxy with stellar mass of $3 \times 10^{10} M_{\odot}$ . Intensity shows gas density whereas colour corresponds to gas temperature such that $T_{\text{red}} > T_{\text{green}} > T_{\text{blue}}$ (Schaye et al. [33]). . . . .	3
1.3	Graph of the rotational velocity of galaxies with the predicted (dashed) curve from stars and gas, and observed (filled line) curve showing how the edges of galaxies have too large velocities to be bound to the galaxy by luminous matter alone (Price [31]). . . . .	6
2.1	Scaling relation between stellar mass and total halo mass with; stellar mass as a function of total halo mass (top panel) and; the ratio of stellar mass to total halo mass as a function of total halo mass (bottom panel). Light and dark shading show $1\sigma$ and $2\sigma$ errors, respectively (Moster et al. [28]). . . . .	12
2.2	Scaling relation of the ratio of stellar mass to total halo mass as a function of total halo mass, for redshifts $0 < z < 4$ . Points with error bars represent the observed relation from the $\Lambda$ CDM DUSTGRAIN- <i>pathfinder</i> simulations. Lines with shading represent the relation and $1\sigma$ error found by the mass function derived in Despali et al. [14] (Girelli et al. [20]). . . . .	13
2.3	Surface density maps of the stellar distributions of the GM-early (left column), reference model (centre column) and GM-late (right column) produced by EAGLE simulations with altered initial conditions. The top row shows surface densities face-on while the bottom row shows them edge-on. Overlaying contour lines indicate the star forming gas distribution, and specific SFR (sSFR) and stellar co-rotational kinetic energy fractions ( $\kappa_{co}$ ) are quoted for each variation (Davies et al. [13]). . . . .	16

2.4	Surface density map of the gas distribution of a 12.5 x 12.5 x 12.5 cMpc slice of the EAGLE Ref-L0025N0376 simulation at $z = 0$ , with yellow as the densest. . . . .	18
3.1	SMHM relation hexbin plot for central haloes in the EAGLE L0025N0376 simulation computed from particles within 30 kpc, normalised by the cosmic average baryon fraction, $\Omega_b/\Omega_0$ , and coloured by number density. Panels include the reference model (top left), altered initial conditions simulation to invoke weak stellar feedback (top right) and one to invoke strong stellar feedback (bottom left) to demonstrate the direct preferential influence of stellar feedback on galaxy formation efficiency at low $M_{200}$ . The blue dashed line included in each panel indicates the running median for the reference model for comparison against the running median for each variation (red dashed line). . . . .	22
3.2	SMHM relation hexbin plot for central haloes in the EAGLE L0050N0752 simulation computed from particles within 30 kpc, normalised by the cosmic average baryon fraction, $\Omega_b/\Omega_0$ , and coloured by number density. Panels include the reference model (top left), altered initial conditions for lower AGN feedback temperature (top right), higher AGN feedback temperature (middle left), no AGN feedback (middle right) and only AGN feedback (no stellar feedback; bottom left). The blue dashed line included in each panel indicates the reference model for comparison against the running median for each variation (red dashed lines) and the grey dashed line indicates efficiency unity. . . . .	24
3.3	SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation within 30kpc, taken from the catalogues and normalised by the cosmic average baryon fraction, $\Omega_b/\Omega_0$ . Hexbins are coloured by black hole mass. The black dashed line is the running median, the blue dot-dashed line indicates the reliability threshold (above-right ensures sufficient sampling in stellar and DM particles) and the grey dashed line shows efficiency unity. . . . .	26
4.1	Coordinate positions in the 32 <sup>nd</sup> subhalo of the EAGLE Ref-L0100N1504 simulation for all particles (left) and the same subhalo cut by a spherical aperture of 30 kpc from its centre of potential (right). . . . .	28

- 4.2 SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation computed from particles within 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by number density. The red dashed line is the running median, the blue dot-dashed line indicates the reliability threshold and the grey dashed line shows efficiency unity where all halo baryons would be converted into stars. . . . . 30
- 5.1 SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. . . . . 33
- 5.2 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line is the complete Spearman rank coefficient ( $\rho_c$ ) and the red dashed line is the reduced Spearman rank coefficient ( $\rho_r$ ). Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ). . . . . 34
- 5.3 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by mean metallicity (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. . . . . 36



- 5.8 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean gas mass-weighted Hydrogen number gas density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). . . . . 42
- 5.9 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in the total halo and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean gas mass-weighted total Hydrogen number gas density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ). . . . . 43
- 5.10 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by mean stellar mass-weighted energy fraction (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. . . . . 45
- 5.11 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted mean stellar mass-weighted energy fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ). . . . . 46

- 5.12 SMHM relation for centrals in the EAGLE L0025N0376 simulation computed from particles within 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Panels include (and are coloured by the mean stellar mass-weighted energy fraction  $f_{th}$  from) the reference model (top left), invoked weak stellar feedback (top right) and strong stellar feedback (bottom left). The black dashed line included in each panel indicates the running median for the reference model for comparison against the running median for each variation (red dashed line). Each panel includes a histogram of  $f_{th}$  for the corresponding simulation variation. . . . . 47
- 5.13 SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by ( $\log_{10}$  of the) gas fraction (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. . . . . 48
- 5.14 SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) gas fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot (\rho'_{stellar})$ . . . . . 49
- 5.15 SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in the total halo and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) gas fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot (\rho'_{stellar})$ . . . . . 51

A.1	Histogram of the difference between particle computed total stellar mass of central haloes and that from the catalogue of the EAGLE Ref-L0100N1504 simulation. . . . .	57
A.2	Hexbin plot of ( $\log_{10}$ of) mean stellar mass-weighted Hydrogen number birth density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	57
A.3	Hexbin plot of the mean metallicity for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	58
A.4	Hexbin plot of the mean stellar mass-weighted stellar ages (in Gyr) for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	58
A.5	Hexbin plot of the ( $\log_{10}$ of) mean gas mass-weighted Hydrogen number gas density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	58
A.6	Hexbin plot of the ( $\log_{10}$ of) mean gas mass-weighted Hydrogen number gas density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within the total halo, coloured by number density. The red dashed line shows the running median fit. . . . .	59
A.7	Hexbin plot of the mean stellar mass-weighted energy fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	59
A.8	Hexbin plot of the ( $\log_{10}$ of) gas fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit. . . . .	59
A.9	Hexbin plot of the ( $\log_{10}$ of) gas fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within the total halo, coloured by number density. The red dashed line shows the running median fit. . . . .	60

# List of Tables

2.1 Table of the relevant EAGLE simulation names, dimensions, numbers of dark matter particles and initial gas particles used in each. . . . .	17
--	----

# Chapter 1

## Introduction

### 1.1 Galaxy Formation and Evolution

The current cosmological paradigm indicates that the formation of galaxies occurs through the hierarchical merging model; the accretion of smaller systems onto more massive ones to form larger galactic systems with the addition of *in-situ* star formation. This was initially facilitated, according to inflation theory, by gravitational instabilities from the small density fluctuations present in the homogeneous early universe – residual energy of the Big Bang 13.8 billion years ago – connected through cosmic filaments and sheets.

Observations of the cosmic microwave background radiation (CMB) provide evidence for the homogeneous early universe, since the expanding universe has cooled over this duration, which was first identified by accident through the experiments on the first Telstar communication satellite. Excess radio noise was measured to be isotropic, the same in every direction. Subsequent investigations, such as by the Wilkinson Microwave Anisotropy Probe (WMAP), showed a very uniform distribution of radiation throughout the visible universe, however, with small colour differences indicative of the tiny fluctuations in the radiation intensity due to those

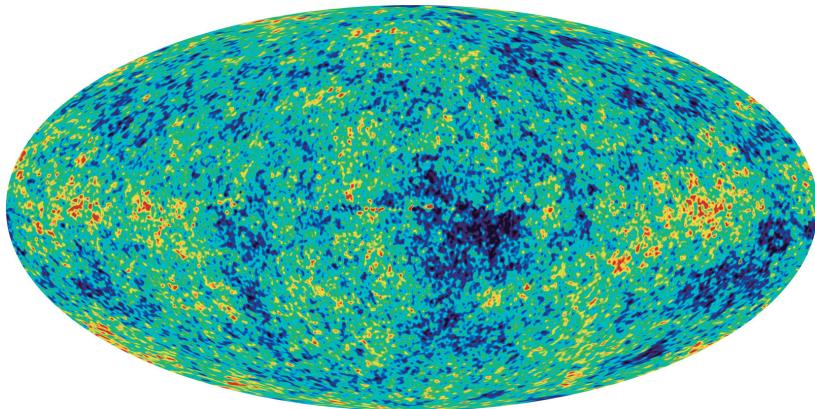


FIGURE 1.1: A full sky map produced by the Wilkinson Microwave Anisotropy Probe (WMAP) of the cosmic microwave background radiation, residual energy from the Big Bang (Shu [35]).

of the density of matter in the primordial universe (Shu [35]). An image of the full sky map produced by WMAP can be seen in Figure 1.1. Filaments, voids, walls and haloes, constitute what is known as the cosmic web - named after it's web-like appearance - where haloes of dark matter allow for the existence of large structures such as galaxies and satellite galaxies (Bond et al. [5]) which can be accreted onto the more massive systems. The cosmic web transports matter and energy between systems allowing for incredibly dynamic interactions in the universe. An image of a simulated cosmic web from the Evolution and Assembly of GaLaxies and their Environments (EAGLE) numerical simulations can be seen in Figure 1.2, showing large-scale structures forming within haloes of dark matter, connected by cosmic web filaments.

The baryon cycle is used to explain the dynamics involved in galaxy formation and evolution. This includes the interplay between the circumgalactic medium (CGM) – the reservoir of gas inside the halo region of a galaxy – and the ejection processes by which the gas is pushed out through intense interactions. Accreted gas from the surroundings is converted into stars by collapsing when the gravitational forces overcome the outward gas pressure. The interstellar medium (ISM), made up by dust and mainly

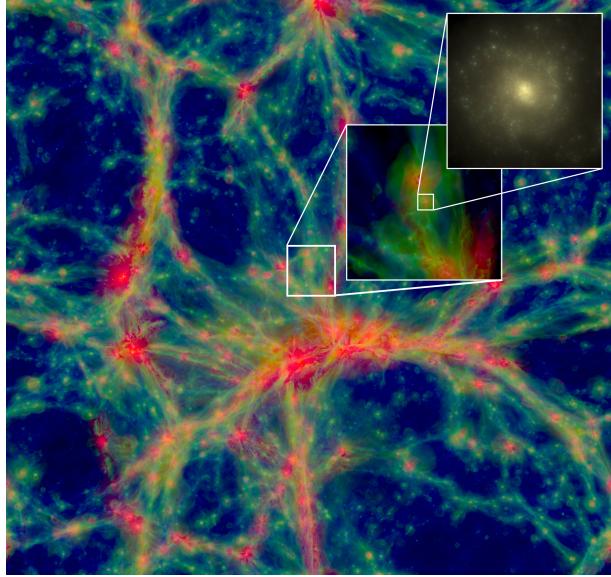


FIGURE 1.2: A  $100 \times 100 \times 20$  cMpc slice of the Ref-L100N1504 EAGLE simulation at  $z = 0$  showing filaments, voids, walls and haloes of the cosmic web. Insets are of dimensions 10 cMpc and 60 ckpc on a side showing a galaxy with stellar mass of  $3 \times 10^{10} M_{\odot}$ . Intensity shows gas density whereas colour corresponds to gas temperature such that  $T_{\text{red}} > T_{\text{green}} > T_{\text{blue}}$  (Schaye et al. [33]).

gas, feeds the process of star formation (SF) and is therefore an essential component in the evolution of galaxies. Gas in the ISM is comprised of a majority of atomic, ionized and molecular Hydrogen, however, it is enriched by the subsequent generations of stars as they produce heavier elements and expel this material into the medium between them after they die in events such as novae and supernovae. These events in combination with heating by stellar winds are referred to as stellar feedback; energy from star formation fed back into the ISM. Atomic Hydrogen (HI) is the least dense out of these three, with an average density of  $30 \text{ particles cm}^{-3}$ , ionized Hydrogen (HII) is denser with a density of  $10^4 \text{ cm}^{-3}$  and molecular Hydrogen ( $H_2$ ) is the most dense with between  $10^3$  and  $10^6 \text{ cm}^{-3}$ . Stars of mass above 8 solar masses ( $M_{\odot}$ ) have sufficiently high energy to emit photons capable of ionizing the gas around, producing HII.  $H_2$  forms from sufficiently cooled HI allowing the atoms to form covalent bonds and because  $H_2$  is comparably the most dense, one can conclude that the ISM is efficient in cooling which occurs through radiation from ions, atoms and

molecules – of which Carbon Monoxide (CO) is the most efficient coolant, especially in dark clouds due to its abundance and dipole moment (Dyson & Williams [16]). There exists a complex equilibrium between the accreted gas, conversion of that gas into stars, gas ejection and feedback processes. More is known about the stellar feedback from less massive stars than more massive stars because the heavier stars form within the densest parts of giant molecular gas clouds which prevents light from penetrating through, in order for them to be observed by optical telescopes. Through using radio telescopes, more massive stellar feedback can be observed, showing it to be similar to that of less massive stars. The effect of stellar winds and jets is to sweep and compress the surrounding material, causing the ISM to become turbulent. SF can be ignited through this compression of surrounding gas clouds but star forming regions can also be disrupted with the result on a galactic scale being detrimental to the overall star formation rate (SFR) (Pittard [30]).

Moreover, galactocentric supermassive black holes (SMBHs) can temporarily become active galactic nuclei (AGN) upon the sufficient supply of cold matter, which eject high energy from the centre of the galaxy outward of the disk plane and cause stars to further lose their mass which they also do through their evolution as planetary nebulae, novae and supernovae. AGN winds and jets provide a source of efficient feedback and gas outflows (Cimatti et al. [6]). At these sites, gas can move at nearly the speed of light in their violently dynamical activity and SMBHs contained in AGN grow in mass via the accretion of matter from neighbouring stars (Pittard [30]) and metal enriched gas from the halo which cools and is accreted onto the galactic disk (Cimatti et al. [6]). High energy processes of AGN feedback in combination with feedback associated with star formation cause disruptions in the nearby star forming regions and, as a consequence, inhibit SF. The details of how feedback influences SF are still not well understood but their relationships with macroscopic processes can be investigated.

The ways in which galaxies form and continue to evolve can be partly understood through observing the Milky Way (MW) at the present day. Evidence for the hierarchical clustering of galaxies can be seen in observational data of the stars inside the stellar halo, where several major accretion events have already been discovered. These include; the ongoing accretion of the Sagittarius Dwarf Spheroidal Galaxy (Ibata et al. [24]), currently undergoing tidal disruption; Gaia-Enceladus (Helmi et al. [22]), proposed to have been accreted approximately 10 Gyr ago, contributing to the formation of the thick disk; and the recently discovered accretion of the satellite galaxy Heracles onto the Galactic bulge (Horta et al. [23]), approximately twice the mass of Gaia-Enceladus and occurring in the early history of the Milky Way, the halo of which it is postulated to have been a major building block.

## 1.2 Cosmological Framework

Swiss astronomer Fritz Zwicky in 1933 studied the redshift of interstellar clouds of gas and dust called nebulae, in particular those residing inside the Coma Cluster, to build upon the previous findings that these objects had much larger rotational velocities than needed to be bound by only the mass of luminous matter to the system. He subsequently postulated the idea of dark matter being a non-luminous massive entity giving rise to this observed effect. He was able to conclude that large velocities, velocity dispersions and observed Doppler effect could only be explained by the existence of a dark matter in a much greater density than luminous matter (Zwicky [41]). Similarly in 1970, American astronomer Vera Rubin together with Kent Ford observed the rotational velocities of stars in the Andromeda Galaxy. They found the stars towards the edge of the disk to rotate with larger velocities than expected from observations of planetary motion around the Sun, first identified by Johannes Kepler in the 1600s. They extended their investigations to a larger

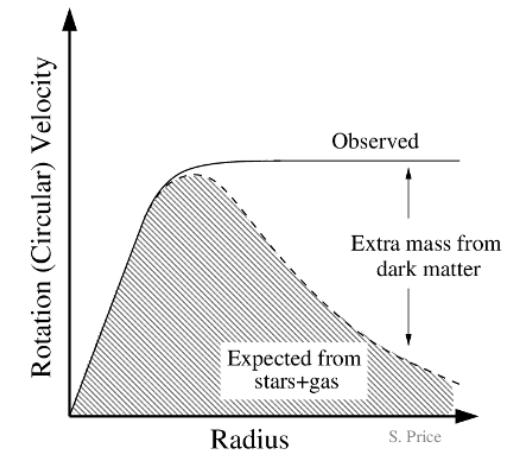


FIGURE 1.3: Graph of the rotational velocity of galaxies with the predicted (dashed) curve from stars and gas, and observed (filled line) curve showing how the edges of galaxies have too large velocities to be bound to the galaxy by luminous matter alone (Price [31]).

sample and found the same suggestions for the existence of dark matter (Bahcall [4]). The flattening effect of the rotation curve for galaxies is illustrated in Figure 1.3 and shows a general idea of the disparity between expected and observed effects.

In the late 1990s, it was found that not only is the universe expanding – first identified by Edwin P. Hubble in 1929 when he investigated the radial velocities of galaxies and their distances from earth by their redshifts (Bagla [3]) – but this expansion is accelerating with which the driving component was termed dark energy. The composition of the universe can be inferred through the CMB (Cimatti et al. [6]). Its acoustic peaks, interpreted from the fluctuations in the distribution of radiation, are stratified into different components (eg. baryonic, dark matter and radiation) and show that dark matter constitutes  $\sim 24\%$  of the energy of the universe and dark energy constitutes  $\sim 71\%$  meaning that there is about 95% of the universe which isn't understood. The current understanding of dark matter, however, is that it is incredibly weakly-interacting, massive and non-luminous, thus at astronomical scales it can only be detected indirectly through its gravitational interaction with other matter and, most noticeably light. It's interaction with light is demonstrated in gravitational

lensing, a phenomenon described by Einstein’s theory of general relativity, where a sufficiently large density of matter can bend the path of light causing it to distort and magnify much like a glass lens does.

Knowledge that the universe is expanding in combination with the information from the CMB has led to cosmology of the modern age. The Copernican Principle states that the earth does not occupy a position in the universe that is privileged over any other. As the universe is shown to be isotropic about earth (eg. galaxy counts on large scales are the same in every direction), it is also isotropic about every other point. This means that every object in the universe as a reference frame observes every other object, on large scales and excluding peculiar velocity, to recede away from it. By extension, as every position experiences the same, the universe must be homogeneous as well as isotropic; the Cosmological Principle (Taylor [39]).

The  $\Lambda$  cold dark matter ( $\Lambda$ -CDM) model of the universe describes the early universe as a hot, dense plasma of tightly coupled photons and matter with initial conditions established during the period of inflation. Specifically, the description of dark matter as ‘cold’ is in reference to the particle’s velocity, being slower than those candidates described as ‘hot’ such as neutrinos. One of the current favoured candidates is the weakly interacting massive particle (WIMP). The role of CDM is to enhance or negate the acoustic patterns of the density perturbations of the early universe in order to reproduce the observed amplitudes of the wave patterns of baryons and photons, and also aids in the gravitational collapse of over-densities to begin to form large scale structure observed in the present-day (redshift,  $z = 0$ ) universe. The form of dark energy, represented as the cosmological constant in the  $\Lambda$ -CDM model, causes the expansion of the universe over time due to the associated negative pressure and as this increasingly dominates over gravitational forces, expansion accelerates (NASA [29]). Independent parameters used in the model are generally described as follows, with values provided by the Plank Collaboration XVI (Akrami et al. [1]);

the scalar power law index ( $n_s$ ) is essential in characterising the strength of the CMB anisotropies where the spectrum is defined as a power law ( $\propto k^{n_s-1}$ ) and takes the value of (less than invariant,  $n_s = 1$ )  $n_s = 0.965$ ; the optical depth to reionisation ( $\tau$ ) can be interpreted as when emitting sources began to form and therefore a larger value implies an earlier onset of reionisation to facilitate this, taken as  $\tau = 0.054$ ; the matter density – the present-day fractional energy density of all matter – for the  $\Lambda$ -CDM model is  $\Omega_m = 0.315$ ; the amplitude of matter fluctuations ( $\sigma_8$ ) describes how the mean present-day density of matter in the universe is distributed at scales of  $8h^{-1}$  Mpc and is taken as  $\sigma_8 = 0.811$ ; and the Hubble constant describes the present-day expansion rate of spacetime and is taken as  $H_0 = 67.4 \text{ kms}^{-1}\text{Mpc}^{-1}$ . In combination, these parameters define the leading model for best describing the observable universe.

# Chapter 2

## Cosmological Simulations

### 2.1 Simulations

Observational astrophysics has been fundamental in the understanding of the formation and evolution of galaxies. By observing galaxies at high redshifts, their early histories can be witnessed as the light from them has travelled the distance with only a finite speed, and this distance has increased due to the expansion of the universe. However, before the existence of advanced cosmological simulations, the mechanisms involved in transitioning from a young galaxy to a modern age galaxy were not as easily understood; observers only had the capabilities of witnessing different galaxies at different respectful ages. NASA's forthcoming James Webb Space Telescope (JWST) will be able to observe galaxies at much higher redshifts than the current Hubble Space Telescope (HST) - which has already collected a vast amount of data since its launch in 1990 (Garner [19]) - and will provide groundbreaking observational insight into how galaxies form. The introduction of cosmological simulations, however, was revolutionary in unveiling galactic evolution. A system could be formed and allowed to evolve into one that agrees with the current observations of present-day galaxies. By rewinding the simulations, the methods of formation and evolution could

be better comprehended and with observations of galaxies at various redshifts, the theoretical models could be constrained to best reflect their real behaviours (Taylor [38]).

Initially in 1963, numerical simulations were applied in the formation of star clusters but were confined to the limit of only a few hundred particles. However, with the advancement of computing and understanding of the interactions between particles, the limit of particle number increased to more than 20,000 by 1981. Today, cosmological simulations can consist of billions of particles but require the aid of the world’s best supercomputers to run them, taking up to several months in the process. The larger the number of implemented particles, the better the system is at reflecting observations and so it should be noted that this is a fundamental constraint on the reliability of cosmological simulations (Taylor [38]). Moreover, the particles in this context should not be interpreted as representing actual particles such as the electron or proton, but rather represent sampling points of a fluid moving under its own gravity and governed by its own mass (Knebe [25]).

## 2.2 EAGLE

The Virgo Consortium’s EAGLE project was founded as a suite of hydrodynamical simulations following the formation of galaxies and supermassive black holes evolving in cosmologically representative volumes of a standard  $\Lambda$  cold dark matter model of the universe. Major improvements to the simulations consist of naturally evolving winds in the processes of stellar and AGN feedback, removing the need for predetermined speeds. The largest of the suite of simulations spans a box of 100 comoving Mpc (hereafter cMpc) on each side, large enough to contain up to 10,000 galaxies with masses similar to the Milky Way or even larger (EAGLE [17]). The suite has a

standard resolution of  $10^6 M_\odot$  gas particles, a smoothing length of 0.7 physical kpc (relative units (pkpc)), and adopts the philosophy of calibrating feedback efficiencies rather than predicting them from first principles which enables simpler feedback implementation. Feedback efficiencies are set by the small-scale microphysics which are not well understood and cannot be resolved in the simulations. Feedback associated with star formation is calibrated to reproduce the galaxy stellar mass function (GSMF) at  $z = 0$  while AGN feedback is calibrated to reproduce black hole (BH) scaling relations (Crain [9]). Stellar feedback now also takes into consideration the scale lengths of disc galaxies because prior models managed to reproduce the GSMF but produced over-compact galaxies, and failed to reproduce many galaxy properties. The  $100^3$  cMpc $^3$  simulation has a resolution of  $2 \times 10^6 M_\odot$  gas mass which has important consequences in the evolution of the simulated galaxies. It imposes a region in the temperature-density plane which cannot be tracked by the simulations, therefore, the gas is given an artificial pressure preventing it's collapse below about 1 kpc. This causes galactic gas and stellar disks to be too thick and so, to solve these problems, increased resolution and larger simulations are required and ongoing projects such as EAGLE-XL, C-EAGLE and EAGLE-2 are exploring such matters (Schaller [32]).

In 2010, Moster et al. [28] used unrelated high-resolution N-body simulations to investigate the relationship between galaxy stellar mass and the total halo mass of dark matter in which galaxies live. Masses of host haloes – those containing the more massive central galaxy – were determined by their virial mass while those of subhaloes – containing satellite galaxies – were determined by their maximum mass over their histories. The presented relationship was able to be described by a function requiring that the observed GSMF be reproduced. The stellar mass - total halo mass relation can be seen in Figure 2.1 where the top panel shows the stellar mass as a function of total halo mass while the bottom panel shows the ratio between the stellar mass and total halo mass (hereafter, stellar mass - total

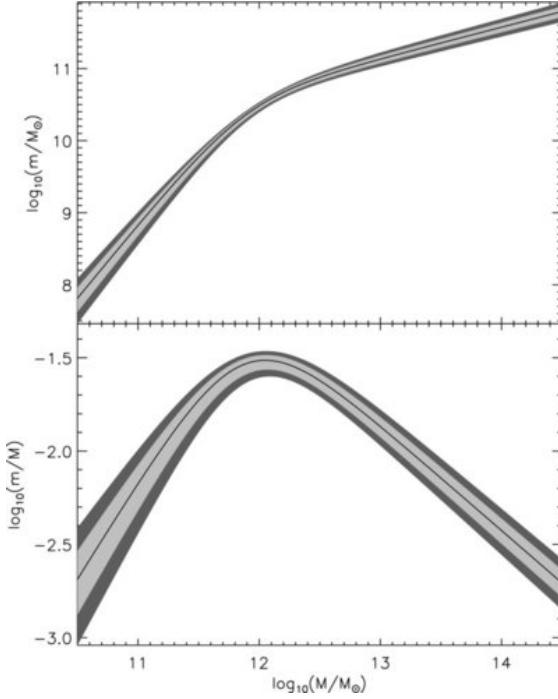


FIGURE 2.1: Scaling relation between stellar mass and total halo mass with; stellar mass as a function of total halo mass (top panel) and; the ratio of stellar mass to total halo mass as a function of total halo mass (bottom panel). Light and dark shading show  $1\sigma$  and  $2\sigma$  errors, respectively (Moster et al. [28]).

halo mass relation; SMHM) as a function of halo mass which is the relation of focus in the following chapters. They found the scaling relation for the ratio to peak at a total halo mass of  $10^{12} M_\odot$  indicating the most efficient mass of galaxies at converting gas into stars, to also decline steeply towards smaller halo mass and decline less steeply towards higher halo mass. They concluded that the interplay between different types of feedback influences the efficiency of SF where stellar feedback is dominant at low halo mass and AGN feedback dominates at high halo mass. One of the objectives of the investigation was to explore how the scaling relation varies with redshift. They found that the total halo mass, where the ratio peaks, increases with redshift and the normalisation decreases indicating that galaxies are less efficient in their SF at higher redshifts. This is illustrated in a plot of the SMHM spanning  $0 < z < 4$  by Girelli et al. [20] in Figure 2.2. Moreover,

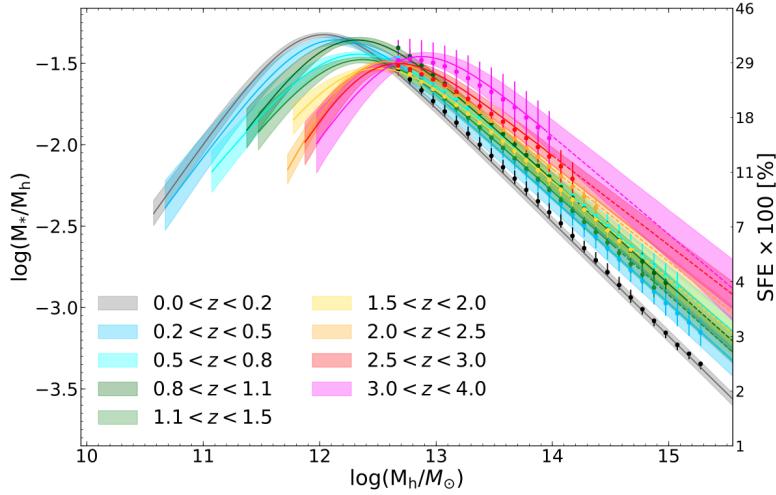


FIGURE 2.2: Scaling relation of the ratio of stellar mass to total halo mass as a function of total halo mass, for redshifts  $0 < z < 4$ . Points with error bars represent the observed relation from the  $\Lambda$ CDM DUSTGRAIN-pathfinder simulations. Lines with shading represent the relation and  $1\sigma$  error found by the mass function derived in Despali et al. [14] (Girelli et al. [20]).

clustering increases with total halo mass, and galaxy bias – the spatial distribution of galaxies as a function of the underlying dark matter density field (Coil [8]) – increases with redshift.

Matthee et al. [27] used the hydrodynamical EAGLE simulations to study the origin of the scatter in the SMHM relation, separating cause and effect by using both the baryonic simulation and the dark matter only (DMO) simulation. They determined that the scatter in stellar mass increases with redshift, decreases with halo mass although with a weak trend above  $10^{12} M_\odot$ , and correlates with the maximum circular velocity ( $V_{max}$ ) of the halo. When investigating influences on the scatter at fixed halo mass which explores the influences to the second order, they found that galaxies with larger stellar mass – therefore, more efficient – had a more concentrated halo, and suggested it was due to earlier formation times allowing for the galaxies to have a longer time for SF and accretion events to occur and/or due to higher binding energies causing less efficient feedback. In addition, they were able to rule out the influence of halo spin, sphericity, triaxiality,

substructure and environment on the scatter and point towards the possibility of more complex halo properties or non-linear/stochastic baryonic effects. Stellar and AGN feedback can heat dark matter haloes, providing stochastic density variations in gas distribution which may influence this (Freundlich et al. [18]). To complement findings by Moster et al. [28], it was concluded that the normalisation of the relation decreases with redshift and is relatively independent of halo mass except for being maximal at the peak, shifting this slightly to higher halo masses. In summary, their findings show that stellar mass correlates with  $V_{max}$ , however, the scatter does not correlate well with other halo properties.

Matthee et al. [27] revealed that scatter in the SMHM relation presented by EAGLE is driven essentially by assembly time – that is, how long it takes a dark matter halo to reach half its present-day mass – at low halo mass, and at large halo masses is driven by the mass of a galaxy’s central BH where a more massive BH correlates with less efficient formation at fixed stellar and halo mass. Subsequently, Davies et al. [12] investigated the mechanisms driving such effects. It was determined that AGN feedback through gas expulsion extends the cooling timescale of the CGM which hinders the fuelling of the ISM, allowing early assembling gas-poor haloes to preferentially host spheroidal, quenched galaxies and late assembling haloes to host disk galaxies with more gas. To lend credence to their suggestions, these correlations were also shown to be demonstrated in the IllustrisTNG hydrodynamical simulations which use majorly different implementations of stellar and AGN feedback, therefore, it is not a calibrated stratification but is purely natural. Such suggestions point towards a strong connection between assembly times of dark matter haloes and the quenching and morphological transformations, ie. the inhibiting of SF and transformations from disk gas-rich to spheroidal gas-poor galaxies. In comparisons between different haloes in simulations of large cosmic volumes yielding diverse environments, it is difficult to establish causal connections between galaxies in their CGM and the dark matter halo assembly histories

(Davies et al. [13]). In addition, there is an underlying bias in the  $\Lambda$ -CDM model which causes late assembling haloes of dark matter to live in less dense and less clustered environments. Halo assembly history also affects many properties of the galaxy populations within haloes. To account for correlations between environment and assembly history, the latter is required to be tracked to high detail for individual haloes (White [40]).

Building on their findings, Davies et al. [13] aimed to investigate the influence of halo assembly history on the evolution of galaxy-CGM ecosystems to further examine the origins of the scatter in the SMHM relation, by disentangling the influences of assembly time and environment through the genetic modification technique (GM). Besides the unmodified “Organic” reference model in EAGLE, complementary initial condition simulations were also made to shift the assembly time of haloes resulting in an early assembler (GM-early) and a late assembler (GM-late) for which correlations with the evolution of the CGM could be investigated. Late assemblers had smaller total halo masses at  $z \approx 2$ , however, they reached similar present-day halo masses to those early assembling. Early assembling haloes were found to grow more massive central BHs by the present day. These more massive BHs inject more feedback energy into their surroundings, expelling more of the CGM (thus extending its cooling timescale) and prevent the replenishment of the ISM for allowing SF to persist, although earlier assembling haloes were found to have a higher peak SFR. Consequently, early assemblers form less stars from their gas and cause their hosted galaxies to be less efficient. Quenching of SF by the growing central BHs causes the galaxy to transform to a spheroidal, which was determined to occur at  $z \approx 1$ . A surface density map can be seen of the stellar and gas distribution in Figure 2.3 showing the morphological transformation dependence on assembly history where the GM-early galaxy can be seen to lack much more gaseous content causing it to be more spherical than both the reference model and GM-late, due to the SF quenching by the growing central BH.

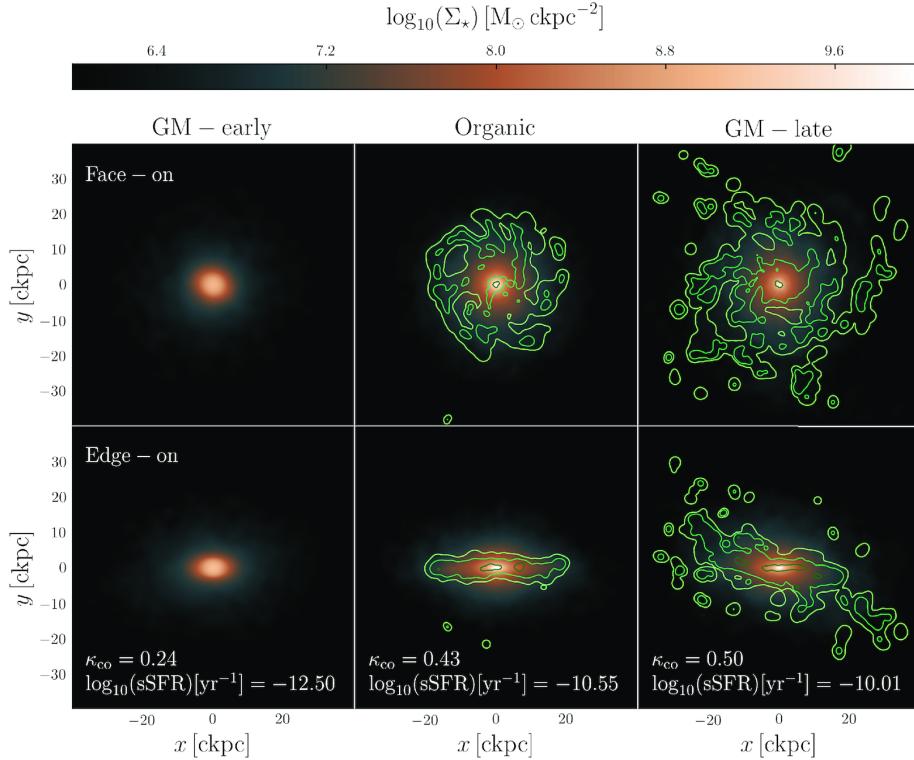


FIGURE 2.3: Surface density maps of the stellar distributions of the GM-early (left column), reference model (centre column) and GM-late (right column) produced by EAGLE simulations with altered initial conditions. The top row shows surface densities face-on while the bottom row shows them edge-on. Overlaying contour lines indicate the star forming gas distribution, and specific SFR (sSFR) and stellar co-rotational kinetic energy fractions ( $\kappa_{co}$ ) are quoted for each variation (Davies et al. [13]).

## 2.3 Data and Sample

Abundance matching (AM) is a technique used to relate central galaxies to their host haloes of dark matter through matching the GSMF from observations to the halo mass function obtained in a collisionless simulation. Dark matter and the stellar component of galaxies demonstrate collisionless behaviour as dark matter only interacts gravitationally and cannot collide with any other known particle or themselves and stars are sufficiently much more distant from one another than their sizes so that the chance of stellar collisions is extremely small (Taylor [38]). Scatter and evolution are accounted for in modern implementation of abundance matching and involve the assumption that the masses of satellite galaxies retain the mass

Simulation	Dimension [cMpc <sup>3</sup> ]	DM Particles	Initial Gas Particles
L0025N0376	$25^3$	$376^3$	$376^3$
L0050N0752	$50^3$	$752^3$	$752^3$
L0100N1504	$100^3$	$1504^3$	$1504^3$

TABLE 2.1: Table of the relevant EAGLE simulation names, dimensions, numbers of dark matter particles and initial gas particles used in each.

they had at the time they were central galaxies. In 2014, Schaye et al. [33] compared the results of the SMHM relation obtained in EAGLE simulations to those obtained by abundance matching to evaluate the precision. They determined that results obtained from EAGLE are in good agreement with those from abundance matching, although the maximum efficiency is reached at a slightly higher stellar mass through the latter. More appropriately, the peak of the SMHM relation describes the masses of galaxies and their haloes which are the “least inefficient”, as when the relation is normalised by the universal baryon fraction ( $\Omega_b/\Omega_m$ ) the peak corresponds to only  $\sim 10\%$  efficiency in EAGLE and  $\sim 25\%$  through AM, with disparity due to computational time constraints for EAGLE while also aiming to reproduce the observed GSMF at the function’s knee.

Total halo mass, referred to as  $M_{200}$  in the scaling relation, is defined as the total mass within the virial radius  $R_{200}$  from the dark matter particle corresponding to the Friends-of-Friends (FOF) halo of minimum gravitational potential, within which the mean internal density is  $200 \times \Omega_{crit}$ , the critical density of the universe (Schaye et al. [33]). This is used as a method for replicating observational techniques of measuring total halo masses and is the standard of measurement for this quantity. FOF is a technique used to group dark matter particles into their respective haloes, which essentially includes neighbouring particles within such a distance until the nearest particles lie too far away.

For further clarification, the simulations relevant to the investigation are described in Table 2.1 where the EAGLE simulation name, dimension, number of dark matter particles and the number of initial gas particles implemented

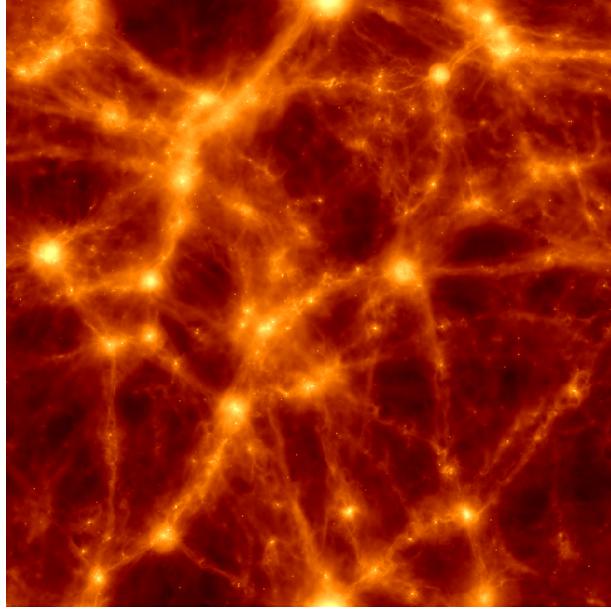


FIGURE 2.4: Surface density map of the gas distribution of a  $12.5 \times 12.5 \times 12.5$  cMpc slice of the EAGLE Ref-L0025N0376 simulation at  $z = 0$ , with yellow as the densest.

in each are listed, and a small slice of the EAGLE L0025N0376 reference model simulation can be seen as a surface density map of the gas distribution in Figure 2.4 where yellow regions are the densest. The reference model (Ref) for each particular dimension simulation corresponds to the evolution of the “Organic” system in which no initial conditions have been altered to produce a particular outcome. However, variations are essential in demonstrating how sensitive the evolution of galaxies is to the mechanisms employed, and will be used to discuss prior findings that will provide a solid footing for the current investigation.

## 2.4 Aims of the Investigation

In this investigation, I explore the second order influences on the efficiency of galaxy formation – variation in the fraction of stellar mass of the total halo mass at fixed halo mass – based on prior suggestions of the intimate connection between the assembly history of dark matter haloes and the

quenching and morphological transformations of hosted galaxies due to a direct dependence upon stellar and AGN feedback. The effect of assembly history is much more direct at low halo mass and so this regime will be explored as well as the high halo mass regime. The dependence of formation efficiency on galaxy properties is still not well understood, therefore, by investigating correlations at fixed total halo mass, more can be interpreted about which physical processes drive SF in galaxies.

My primary focus is on providing a demonstration of the first order influences on the formation efficiency in Chapter 3 as described in prior works, illustrating the significant role of stellar (Section 3.1) and AGN (Section 3.2) feedback in fostering SF in galaxies, supported by connections with black hole mass (Section 3.3). Subsequently, the determined stellar mass – halo mass relation will be discussed in Chapter 4. Then in Chapter 5, as a consequence, the interesting physical processes of birth density (Section 5.1), metallicity (Section 5.2), stellar age (Section 5.3), gas density (Section 5.4), energy fraction (Section 5.5) and gas fraction (Section 5.6) will be studied for their connections to each other and assembly history. As an extension, the dependence of efficiency on stellar feedback strength and energy fraction (Subsection 5.5.1) will also be briefly explored, and a summary and final discussion is provided in Chapter 6. The strong tidal stripping which satellite haloes are subjected to makes it harder to distinguish the relation between stellar mass and halo mass (Schaye et al. [33]), therefore, the focus is directed towards central galaxies only. For consistency, investigations into second order influences will be carried out at present-day redshift  $z = 0$ , therefore, any influences on efficiency by redshift can be ruled out and galaxy bias is at a constant minimum. I aim to place presented correlations onto a quantitative footing through the use of the Spearman rank correlation coefficient ( $\rho$ ) and p-value ( $p$ ) to establish the variables influencing the exhibited scatter in the SMHM.

# Chapter 3

## Preliminary Investigations of First Order Influences

To lay the foundations of the subsequent investigations into the second order effects, it is essential to present the first order influences on the efficiency of galaxy formation. Prior research by Moster et al. [28] concluded that the interplay between feedback processes provides a direct influence, where stellar feedback dominates at low halo masses whereas AGN dominates at high halo masses. Changing initial conditions and the mechanics through which feedback processes function is an invaluable tool for understanding their influences on galaxy formation efficiency. Data has been obtained for variations from the EAGLE Ref-L0025N0376 and Ref-L0050N0752 simulations and, therefore, do not contain the best resolution which the suite has to offer but nevertheless provide a key insight into first order influences.

### 3.1 Stellar Feedback

For the EAGLE Ref-L0025N0376 simulation, variations were carried out with altered conditions to invoke different strengths of stellar feedback.

These can be compared to the reference model to interpret how they influence efficiency at different regimes of total halo mass. The corresponding stellar mass - halo mass relations for these variations of this dimension simulation are illustrated together in Figure 3.1 where dashed lines indicate the running medians for the scaling relations – obtained through the use of the *statsmodels* non-parametric LOWESS function which is discussed in more detail in the following chapter in Section 4.2. Smearing at low halo mass occurs due to the implementation of stellar evolution and chemical enrichment in EAGLE and the top left panel shows the reference model only.

The top right panel shows a comparison between the reference model and the simulation with invoked weaker stellar feedback. Weaker stellar feedback causes the efficiency of low mass haloes to increase by a significant amount. Star forming regions are less disrupted by the less intense winds, novae and supernovae from stars and so are given the opportunity to enrich the galaxies with more stellar mass. The increase in efficiency can be interpreted from the increase in  $M_{\star}/M_{200}(\Omega_b/\Omega_0)$  of the running median of the weak feedback relation from that of the reference model relation. Taken at  $M_{200} = 10^{10.5} M_{\odot}$ , the increase is from  $10^{-2} \times 100 \approx 1\%$  efficient to  $10^{-1.5} \times 100 \approx 3\%$  efficient. The maximum  $M_{\star}/M_{200}(\Omega_b/\Omega_0)$  is also shifted to lower halo masses, indicating that weaker stellar feedback causes the maximum efficiency of converting gas into stars to occur in lower mass haloes, however, the effect is not very significant. At higher halo masses, weaker feedback is suggested to cause a tiny decrease in efficiency, however, with a very small sample at these masses it is difficult to conclusively establish this. The major effect of weaker stellar feedback is in inducing a change in formation efficiency in low mass haloes.

The bottom left panel shows a comparison between the reference model and the simulation with invoked stronger feedback from star formation. Stronger feedback causes a more dramatic change in efficiency as seen by

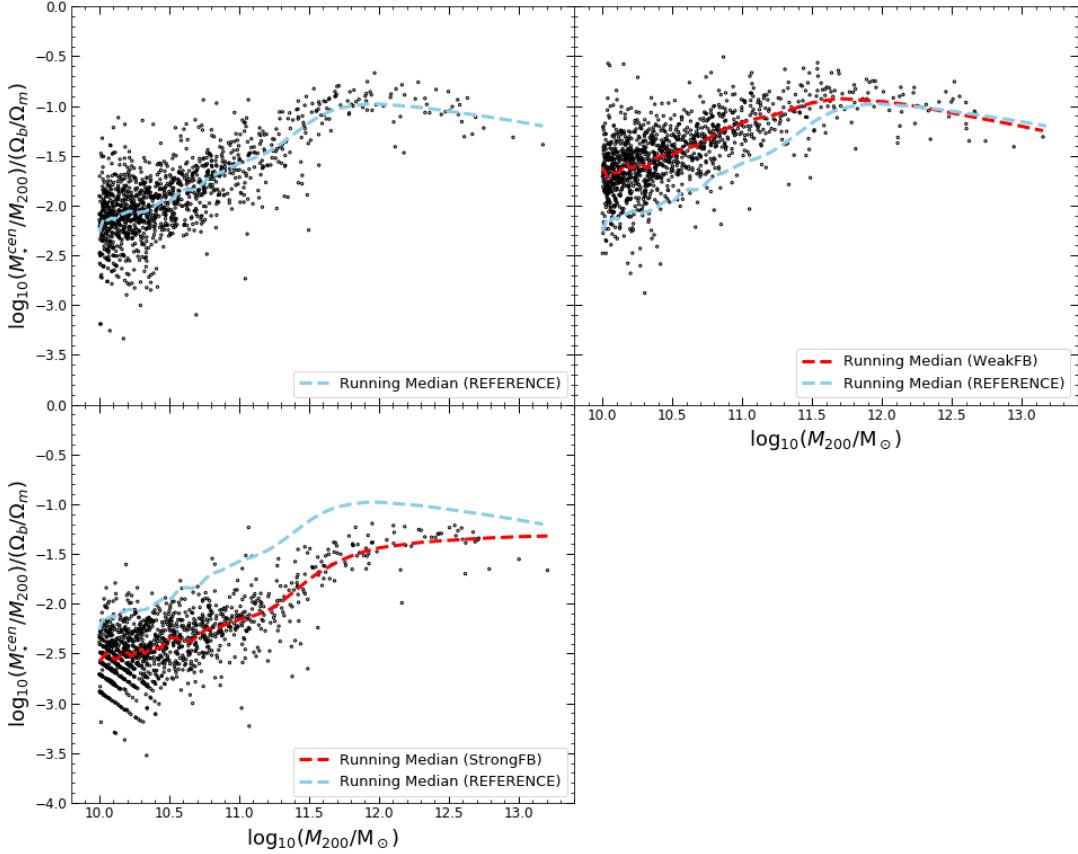


FIGURE 3.1: SMHM relation hexbin plot for central haloes in the EA-GLE L0025N0376 simulation computed from particles within 30 kpc, normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ , and coloured by number density. Panels include the reference model (top left), altered initial conditions simulation to invoke weak stellar feedback (top right) and one to invoke strong stellar feedback (bottom left) to demonstrate the direct preferential influence of stellar feedback on galaxy formation efficiency at low  $M_{200}$ . The blue dashed line included in each panel indicates the running median for the reference model for comparison against the running median for each variation (red dashed line).

the significant decrease in normalisation of the relation. At low  $M_{200}$ , efficiency is decreased by the onset of stronger feedback due to winds, novae and supernovae of stars disrupting their surrounding star forming regions. Again taken at  $M_{200} = 10^{10.5} M_\odot$ , the decrease in efficiency is from  $10^{-2} \times 100 \approx 1\%$  efficient to  $10^{-2.5} \times 100 \approx 0.3\%$  efficient. The maximum  $M_*/M_{200}(\Omega_b/\Omega_0)$  is also shifted, however to higher halo masses, where the peak is not displayed given the low resolution from the small sample obtained from L0025N0376. This indicates that stronger stellar feedback

causes the maximum efficiency of converting gas into stars to occur in higher mass haloes. It becomes immediately obvious that feedback due to SF is a dominant process in influencing the efficiency of galaxy formation in low mass haloes.

## 3.2 AGN Feedback

For the EAGLE Ref-L0050N0752 simulation, variations were carried out with altered conditions to change how AGN feedback functions in galaxies. These can be compared to the reference model to interpret how different temperatures of AGN heating influence the efficiency at different regimes, shown in Figure 3.2 where the relations are shown as hexbin plots to illustrate how dense the distribution is at low halo mass. The coloured dashed lines indicate the running medians for the scaling relations and the grey dashed line indicates efficiency unity where all halo baryons would be converted into stars.

The top left panel shows the reference model only. The top right panel shows a comparison between the reference model and the simulation with invoked lower AGN feedback temperature  $T_{AGN} = 10^8$  K compared to the reference AGN feedback temperature  $T_{AGN} = 10^{8.5}$  K. At low halo mass the scaling relation is unchanged, however, a lower AGN feedback temperature causes the efficiency to increase slightly at high halo mass.

The middle left panel shows a comparison between the reference and the simulation with a higher AGN feedback temperature  $T_{AGN} = 10^9$  K. Again, efficiency at low halo mass is unchanged, however, it is decreased slightly at high halo mass. The middle right panel shows the comparison between the reference and the simulation with AGN feedback disabled and so feedback only comes from SF. This still causes no change at low halo mass but is responsible for high mass haloes becoming even more efficient, to be roughly as efficient as those haloes able to convert the most of their gas into stars.

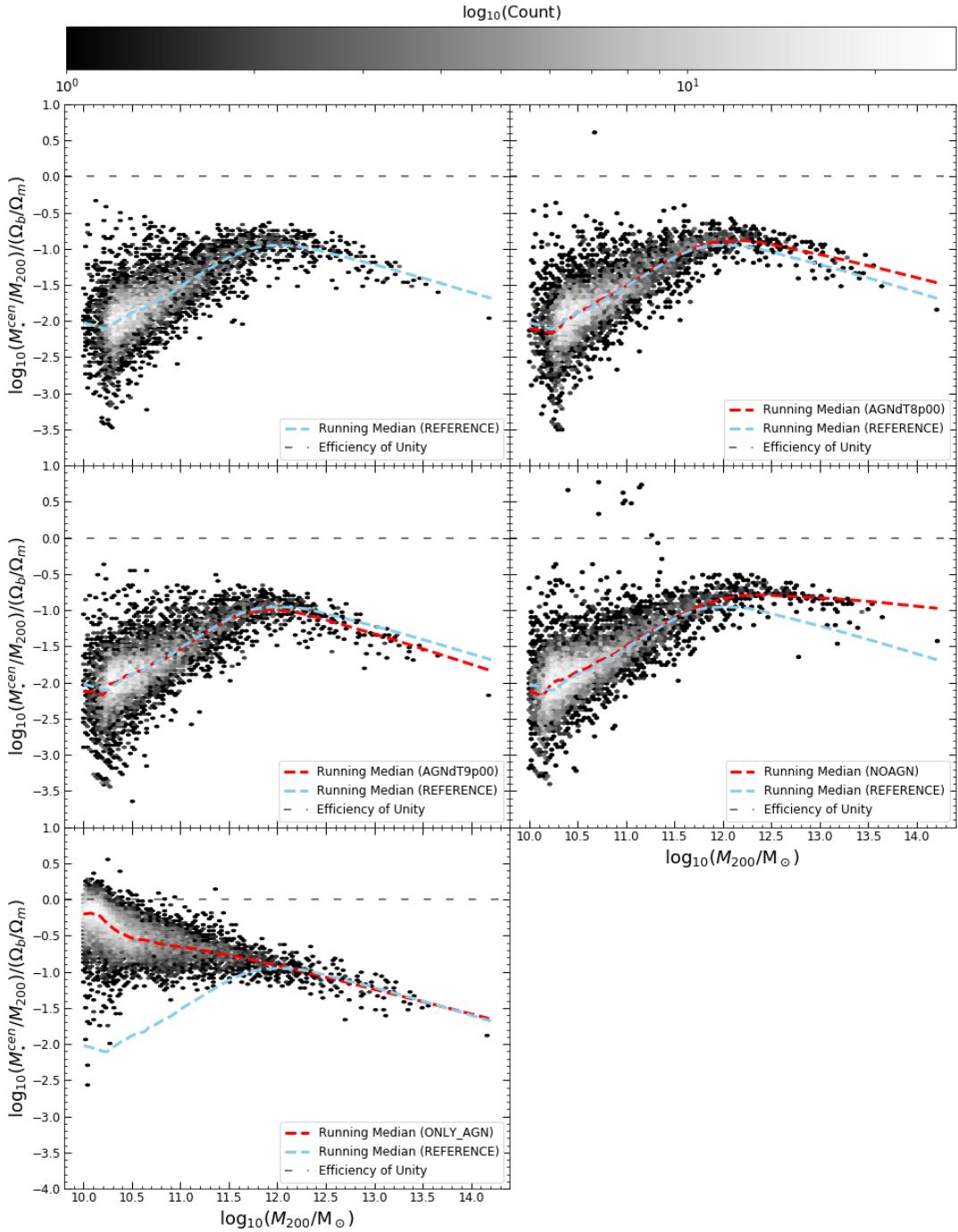


FIGURE 3.2: SMHM relation hexbin plot for central haloes in the EA-GLE L0050N0752 simulation computed from particles within 30 kpc, normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ , and coloured by number density. Panels include the reference model (top left), altered initial conditions for lower AGN feedback temperature (top right), higher AGN feedback temperature (middle left), no AGN feedback (middle right) and only AGN feedback (no stellar feedback; bottom left). The blue dashed line included in each panel indicates the reference model for comparison against the running median for each variation (red dashed lines) and the grey dashed line indicates efficiency unity.

The bottom left panel shows the reference against the simulation with only AGN feedback and with feedback from SF disabled. The efficiency at high halo mass is unchanged, however, the efficiency of low mass haloes is significantly increased above  $M_\star/M_{200}(\Omega_b/\Omega_0)$  of the reference model’s maximum efficiency, following the distribution of the relation at high halo mass and supporting results shown in the variations in the strength of feedback associated with star formation. The maximum efficiency increases from  $M_\star/M_{200}(\Omega_b/\Omega_0) \approx 10\%$  to a new peak at  $M_{200} = 10^{10} M_\odot$  of  $\approx 60\%$ , a much more significant efficiency increase than that which occurs due to invoking no AGN feedback. The SMHM relation is, therefore, more sensitive to stellar feedback than AGN feedback and one dominates over the other at specific regimes of total halo mass; at low halo mass, stellar feedback dominates and at high halo mass AGN feedback dominates.

### 3.3 Black Hole Mass

Davies et al. [13] investigated the influences of assembly history on formation efficiency and found that the mass of central BHs in early assembling haloes were able to grow earlier and quicker by the present day. With a more rapid growth in mass, central BHs exhibit stronger AGN feedback, which has been seen to have a direct influence on formation efficiency of high mass haloes. By extension, more massive central BHs cause a morphological transformation from disk to spherical galaxies due to the quenching of SF.

Figure 3.3 shows the SMHM relation for central galaxies in the EAGLE Ref-L0100N1504 simulation determined from catalogue data, again normalised by the cosmic average baryon fraction and coloured by BH mass,  $M_{BH}$  in units of solar mass. Although the hexbin plot is coloured by the sum of masses of all BHs contained in the halo, the significant majority of the mass will be due to the central BH and therefore the correlation is more

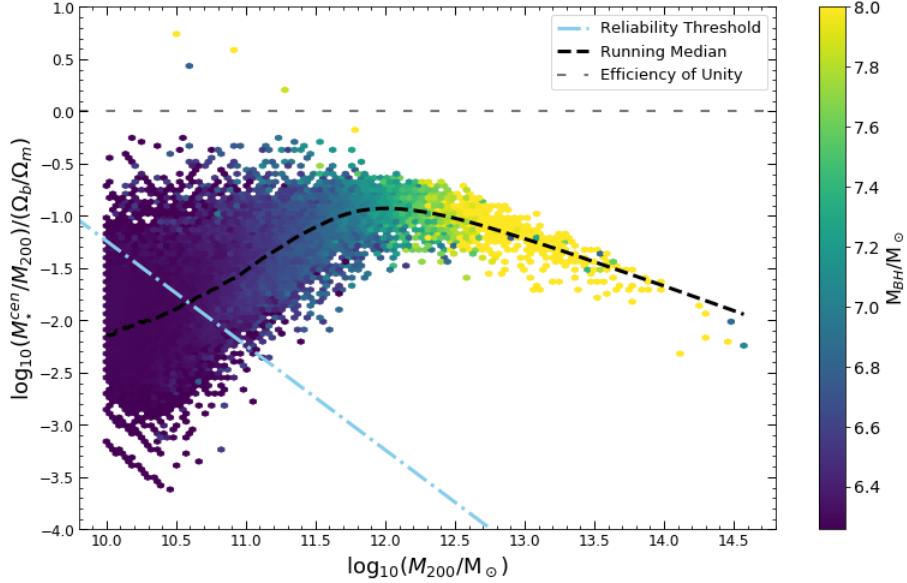


FIGURE 3.3: SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation within 30kpc, taken from the catalogues and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by black hole mass. The black dashed line is the running median, the blue dot-dashed line indicates the reliability threshold (above-right ensures sufficient sampling in stellar and DM particles) and the grey dashed line shows efficiency unity.

appropriately between the efficiency and central BH mass. This presents a direct influence on formation efficiency as high mass haloes have more massive central BHs, and efficiency is maximal in haloes with not too low mass and not too high mass central BHs. AGN feedback works to regulate the growth of massive galaxies (Crain et al. [10]) therefore by investigating other physical processes, interesting interplay with this type of feedback can be interpreted.

# Chapter 4

## Stellar Mass - Halo Mass

### Relation

Rather than being limited to those galaxy and halo quantities already stored in the EAGLE catalogues, many more characteristics of the galaxies and their host haloes could be found by computing quantities from the simulated particles themselves, that form them. This is made possible by wrapping the periodic simulation box by the use of the python module *pyread\_eagle*, a python-adapted function to be able to read particles in specified regions of the simulation (Davies [11]). Due to smearing of masses at low stellar and halo mass from the implementation of stellar evolution and chemical enrichment, only haloes of  $M_{200} > 10^{10} M_\odot$  are investigated and this selection, as well as reducing the sample to only centrals, is built into the wrapping of the simulation box which inevitably causes computation time to reduce. Another technique implemented into reducing computation time was to discard any regions of the simulation with particles that were far away from the relevant regions containing the central haloes and the code can be found in the GitHub repository linked in Appendix A.1. Stellar masses of all particles for each central halo were computed and a comparison between these and those stored in the catalogues can be found in Appendix A.2.

## 4.1 Galaxy within a Spherical Aperture

Because particles lying outside the specified cubic region of a halo, which are associated with those inside the region, are included when wrapping the simulation box, they need to be discarded through the use of a spherical aperture. This will also allow for a better comparison of the local effects of the physical processes which are to be investigated. This is carried out by cutting the selected region by a spherical aperture of a given radius away from a point of choice. An aperture is chosen around the centre of potential of the given halo, rather than around its centre of mass, because the centre of mass of two colliding galaxies would be situated in the middle of them whereas the centre of potential would be located in the more massive of the two objects and is a better representative of the centre of the system in general. A spherical aperture of 30 kpc from the centre of potential of each central halo in the sample was applied and an example of this can be seen in Figure 4.1 showing the arbitrary 32<sup>nd</sup> subhalo in cartesian coordinates. The relevant region changes from the asymmetrical shape (left panel) to the uniform sphere (right panel) of radius 30 kpc.

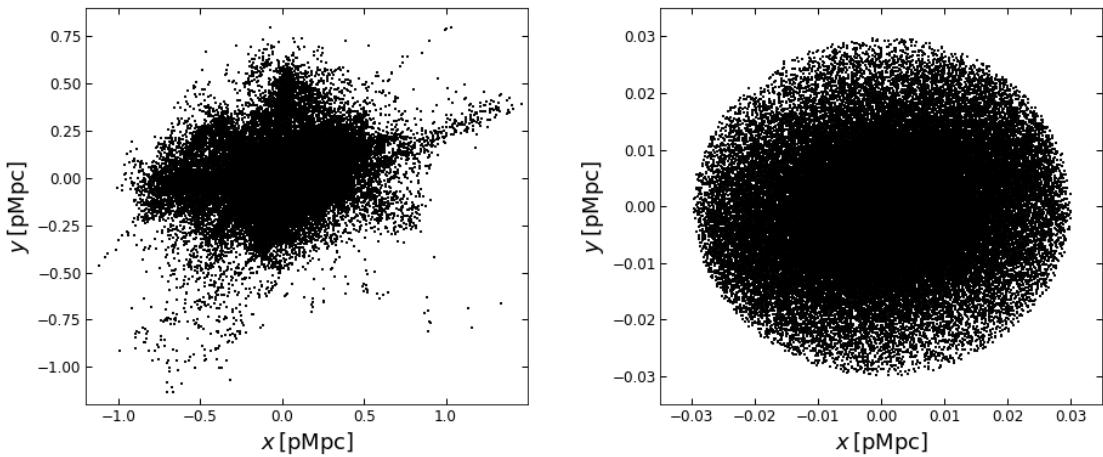


FIGURE 4.1: Coordinate positions in the 32<sup>nd</sup> subhalo of the EAGLE Ref-L0100N1504 simulation for all particles (left) and the same subhalo cut by a spherical aperture of 30 kpc from its centre of potential (right).

## 4.2 LOWESS

A more robust approach to providing a fit for the scaling relation rather than a median by simply binning the halo masses, is to use the Locally Weighted Scatterplot Smoothing (LOWESS) algorithm (Cleveland [7]). This fits a median value for each halo mass producing a smooth running median, through a weighted linear regression of a fraction of the data surrounding each datapoint. When colouring the scaling relation by a third variable as a function of halo mass – other than  $M_*/M_{200}(\Omega_b/\Omega_0)$  and  $M_{200}$  – to see how they correlate, the running median gives an interesting way to separate the vertical and horizontal dependence and correlations between  $M_*/M_{200}(\Omega_b/\Omega_0)$  and the quantity at fixed total halo mass can be inferred. The running median value of the third variable –  $M_{200}$  relation  $\tilde{z}(M_{200})$  is subtracted from the third variable  $z$ , leaving a median-subtracted value of that quantity,  $\Delta z = z(M_{200}) - \tilde{z}(M_{200})$ . By colouring the relation with  $\Delta z$ , only the secondary (vertical) dependence is shown. It will also be useful to compress the dynamic ranges of the colourbars showing the correlations and so the limits of them will not accurately represent the minimum and maximum values of the influencing third variables.

## 4.3 Scaling Relation

Figure 4.2 shows the SMHM relation obtained from the EAGLE Ref-L0100N1504 simulation with stellar mass computed from particles themselves, normalised by the cosmic average baryon fraction, as a hexbin plot coloured by number density to show how populated the low halo mass region is compared to high halo mass. The general trend shows that more massive galaxies are formed in more massive haloes of dark matter with a maximum efficiency of  $\sim 10\%$ . Interestingly, the stellar mass and total halo mass of the Milky Way galaxy places it at the peak of the SMHM relation. Therefore, this poses questions about whether or not the solar system is in

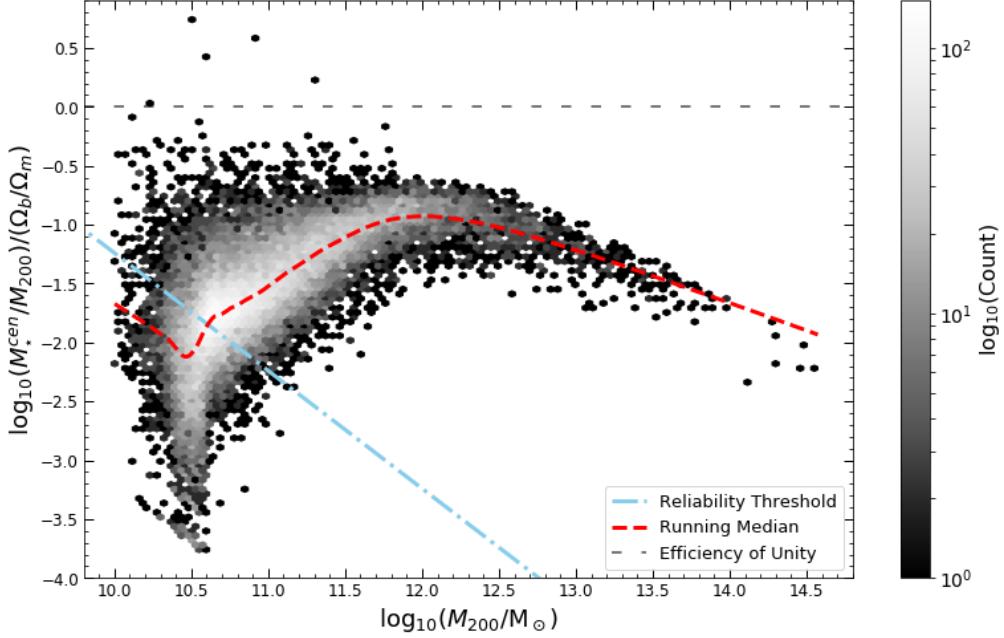


FIGURE 4.2: SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation computed from particles within 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by number density. The red dashed line is the running median, the blue dot-dashed line indicates the reliability threshold and the grey dashed line shows efficiency unity where all halo baryons would be converted into stars.

a special environment in the universe which may have facilitated the ability for intelligent life to exist on earth; a philosophical argument that has been discussed for hundreds of years. The dip in the running median, shown by the red dashed line, is not observed in the SMHM relation found from the catalogue data and, therefore, it is appropriate to be slightly circumspect about any correlations inferred for halo masses less than  $M_{200} \approx 11 M_\odot$ . A reliability threshold (blue dot-dashed line) is placed to indicate that above-right of it there is strong sampling in both stellar and DM particles and is defined as  $M_\star/M_{200}(\Omega_b/\Omega_0)$  corresponding to the mass of 50 star particles, found through the conversion from the mass of the dark matter particles. The grey dashed line indicates 100% efficiency in converting halo baryons to stars and the reason for a few galaxies having efficiency greater than unity could be due to the dynamical stripping of dark matter particles from the outer edge of the halo, causing the halo mass to decrease but the stellar

mass to remain constant. Stellar feedback influences the scaling relation at low halo mass and AGN feedback influences it at high halo mass. Defining scales of  $M_{200}$  where these each dominate over the other will make it easier to interpret correlations with other physical processes in each regime, therefore, the stellar scale and AGN scale are hereafter defined as halo masses of  $M_{200} = 10^{10.5} M_\odot$  and  $M_{200} = 10^{12} M_\odot$ , respectively.

# Chapter 5

## Analysis and Results of the Second Order Influences

### 5.1 Birth Density

Birth density  $\rho_{birth}$  is the density of gas particles at the instance of their conversion into star particles and is one of the essential components of the stellar energy fraction function implemented in the suite of simulations, which is explored in Section 5.5. The natural expectation would be that galaxies are more efficiently formed if they consist of a high density of gas particles as they convert to star particles but it has been shown that feedback significantly influences a galaxy’s SF ability. Figure 5.1 shows the SMHM relation coloured by ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density (left panel) and the median subtracted version (right panel). The Hydrogen number birth density  $n_{H,birth}$  is used as a convention and is related to the absolute density  $\rho_{birth}$  through  $n_{H,birth} \equiv X \rho_{birth} / m_H$  where  $X$  is the hydrogen abundance within 30 kpc for each subhalo in this instance (particle-by-particle due to its evolution through the simulation via stellar evolution) and  $m_H$  is the mass of the

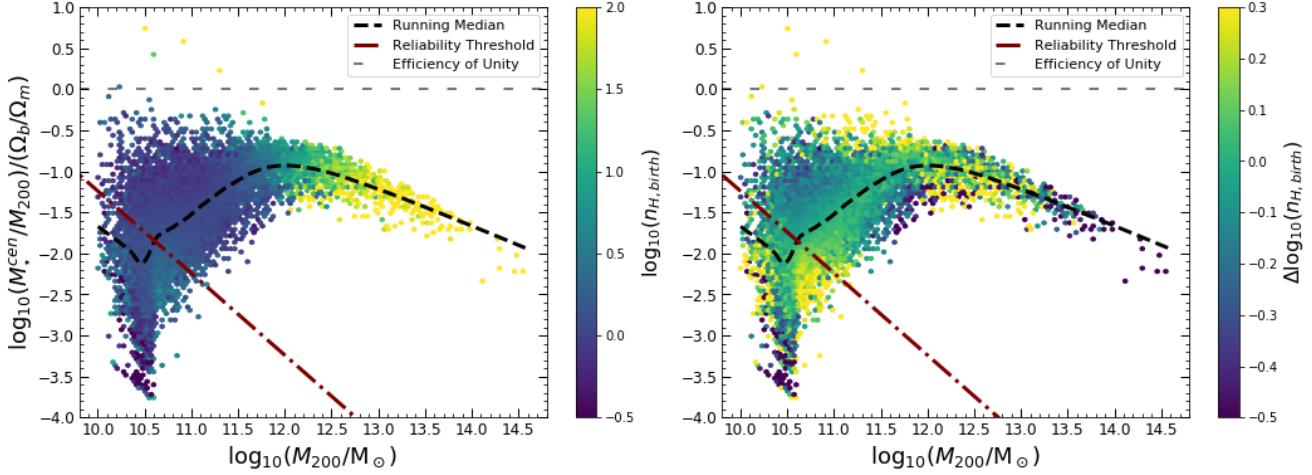


FIGURE 5.1: SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

hydrogen atom. The distribution of  $\log_{10}(n_{H,birth})(M_{200})$  with the running median fit obtained from the LOWESS algorithm can be seen in Appendix A.3. Analysis of the first order effects shows that higher mass haloes consist of higher birth densities meaning that stellar feedback does not significantly impact a galaxy’s ability to reap the benefits of a higher birth density environment. Exploring the second order influences indicates how galaxies depend upon birth density at similar halo masses. This immediately suggests that there are more complex dependencies on each type of feedback than would originally be suggested by first order influences.

To provide a statistical weighting for the observed second order correlations, the running Spearman rank correlation coefficient ( $\rho$ ) is found for the complete sample and shown as the solid black line in the bottom panel of Figure 5.2. The Spearman rank tests whether or not there is a monotonic relationship between the ratio of stellar mass to halo mass and the process of discussion, which in this case is birth density (Lund & Lund [26]). A correlation coefficient larger than zero suggests a positive correlation while

a coefficient below zero suggests a negative correlation and in combination with the obtained Spearman rank p-value ( $p$ ), the significance of the correlation can be inferred.  $p > 0.01$  indicates a correlation of low significance with which to be cautious when making any conclusions. Where significant, the mean correlation coefficient is quoted for haloes within a  $\pm 10^{0.05} M_{\odot}$  window around both the stellar and AGN scale halo mass. For birth density, correlation coefficients are insignificant at the AGN scale and so the mean is only provided for haloes at the stellar scale. At this scale,

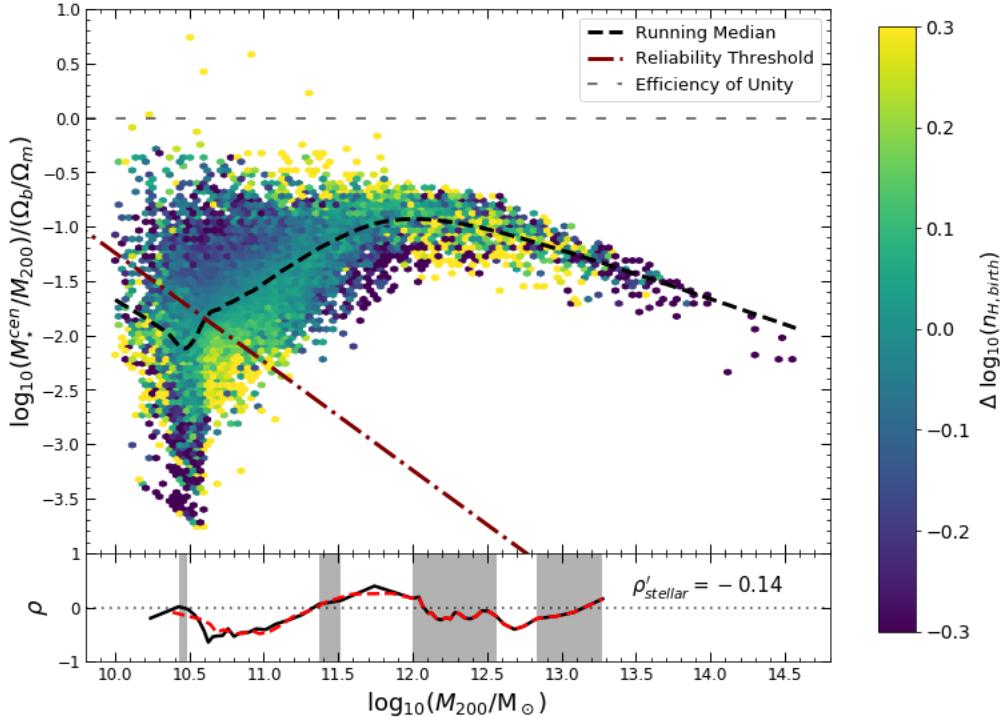


FIGURE 5.2: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line is the complete Spearman rank coefficient ( $\rho_c$ ) and the red dashed line is the reduced Spearman rank coefficient ( $\rho_r$ ). Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_{\odot}$  ( $\rho'_{stellar}$ ).

the correlation coefficient is quoted as  $\rho'_{stellar} = -0.14$ , suggesting a significant positive correlation between Hydrogen number birth density and  $M_*/M_{200}(\Omega_b/\Omega_0)$ ; galaxies are more efficient star formers if formed in low birth density environments. More massive haloes than that at the AGN scale, around  $M_{200} \approx 10^{12.7} M_\odot$ , also host galaxies that are more efficient if formed in low birth density environments. In contrast, between these two regimes, the correlation becomes positive suggesting that for halo masses where neither stellar nor AGN feedback dominate, galaxies are able to reap the benefits of their high birth density environments in order to catalyse SF.

To understand how the smearing at low stellar mass and halo mass may affect the interpreted correlations given from the statistical test, a separate reduced running Spearman rank correlation coefficient ( $\rho_r$ ) is retrieved for only the data above-right of the reliability threshold where there is strong sampling in both stellar and DM particles. This is plotted alongside that of the complete sample ( $\rho_c$ ) and is shown as a red dashed line. As expected,  $\rho_r$  agrees almost perfectly with  $\rho_c$  for  $M_{200} > 10^{11} M_\odot$  but provides key insight into the reliable correlation at smaller halo masses. There isn't as steep of a change in correlation at  $M_{200} \approx 10^{10.6} M_\odot$ , though the correlation is generally unchanged. However, where  $\rho_c$  decreases towards the minimum halo mass,  $\rho_r$  may be beginning to increase, suggesting that at even lower halo masses than are investigated, the correlation may flip as it does between the stellar and AGN scale and that the effect of stellar feedback may once again not inhibit a galaxy's ability to use its gas rich environment for forming stars. It would be of significant benefit to explore correlations at even lower halo masses in future investigations.

## 5.2 Metallicity

Another component of the energy fraction used in EAGLE simulations is the metallicity,  $\log_{10}(Z/Z_\odot)$  – the abundance of elements heavier than hydrogen and helium present in galaxies through their stars and the medium between, which is enriched via the subsequent generations of SF, normalised by solar metallicity. As well as being a possible influence on the efficiency, increased metallicity is also inherently a symptom of an increased SFR which Davies et al. [13] found to be larger in early assembling haloes than late assembling; early assembling haloes will host more generations of stars due to rapid bursts of SF, enriching the ISM with heavier and heavier elements. Figure 5.3 shows a clear correlation of larger mass haloes exhibiting larger present-day metallicities than low mass haloes and so to first order, galaxies are more efficiently formed with a metallicity slightly greater than the metallicity of the Sun,  $Z_\odot$ . The distribution of  $\log_{10}(Z/Z_\odot)(M_{200})$  with the running median fit can be seen in Appendix A.3.

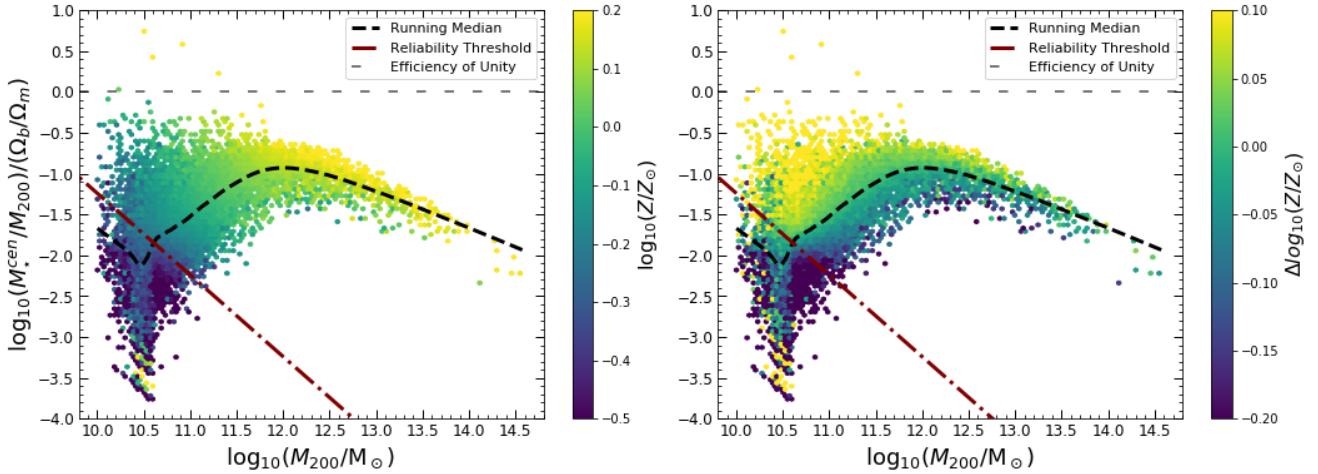


FIGURE 5.3: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by mean metallicity (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

At fixed halo mass, the correlation with metallicity is also quite clear and shows that for all halo masses, more efficient galaxies are more metal-rich as seen in Figure 5.4. There is a strong correlation for most of the relation, however at very low masses, there is a steep decline where the correlation coefficient decreases from a value at the AGN scale of  $\rho'_{AGN} = 0.76$  to  $\rho'_{stellar} = 0.31$  and is further confirmed by the reduced coefficient demonstrating the same change but more gradually over a slightly larger halo mass range. The correlation between metallicity and formation efficiency may be more strongly influenced by feedback from stars because of the intimate connection between the two. Metallicity is increased by a burst of

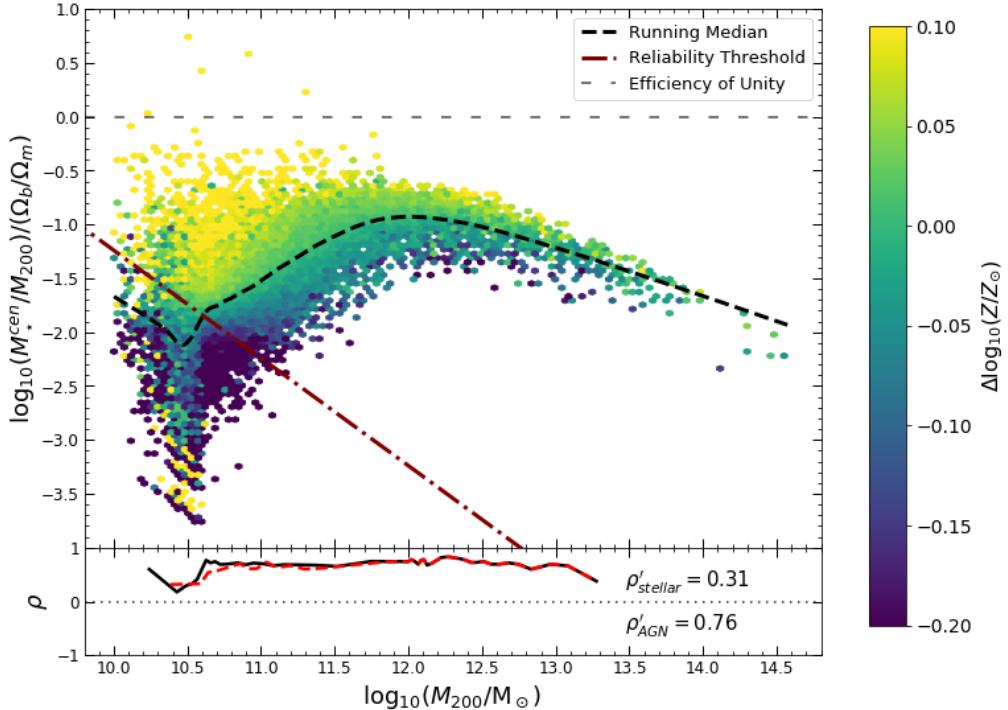


FIGURE 5.4: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted mean metallicity (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficients are quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ) and  $M_{200} = 10^{12} \pm 0.05 M_\odot$  ( $\rho'_{AGN}$ ).

SF which consequently drives stellar feedback and so disrupts the efficiency of galaxies in low mass haloes.

### 5.3 Stellar Age

Davies et al. [13] showed that earlier assembling haloes were able to start growing their central BH earlier causing a quenching of SF earlier than late assemblers. The haloes were also found to reach very similar masses by the present day. Galaxies of differing stellar masses may have reached the same total halo mass through either assembling earlier or later, however, because earlier assemblers will contain many more older stars than late assemblers, a correlation with stellar mass at fixed halo mass will unveil which sort of assembly history forms a galaxy more efficiently. Stellar age isn't the best proxy for assembly time however, and so correlations may only suggest a dependence rather than establish any cause-and-effect relationship. Figure 5.5 shows the SMHM coloured by the mean stellar mass-weighted stellar age in Gyr (left panel) and coloured by the median-subtracted version (right panel). The distribution of  $\text{Age}(M_{200})$  with the running median fit can be seen in Appendix A.3. The left panel indicates that to first order, low mass haloes and high mass haloes have averagely older stars and between regimes, haloes contain averagely younger stars. When differentiating the correlation into the secondary dependence only, things become more complex. Only through the use of statistical tests can any meaning be inferred.

The negative Spearman rank correlation coefficients for the complete sample at low and high halo masses that are significant, shown in the bottom panel of Figure 5.6, suggest that late assembling haloes are more efficient star formers at these regimes and so early formation suppresses SF. However,  $\rho_r$  indicates that there is no correlation at the stellar scale. For better

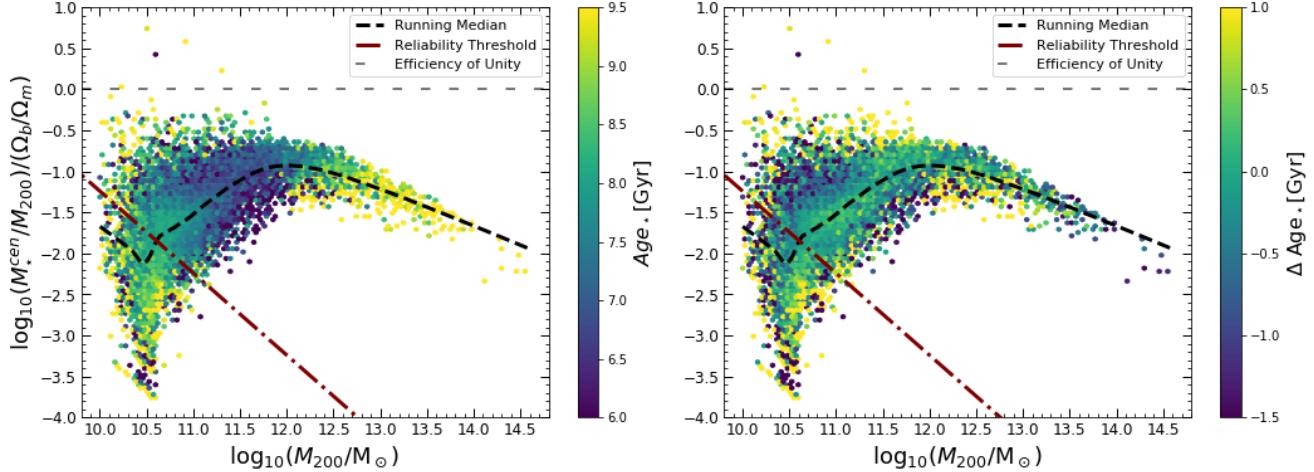


FIGURE 5.5: SMHM relation for centrals in the EAGLE RefL0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by mean stellar mass-weighted stellar age in Gyr (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

comparison with prior findings, investigations into stellar ages for the total halo may be beneficial. The stellar scale at low halo mass provides  $\rho'_{stellar} = -0.21$ , however, the correlation coefficient at the AGN scale is quoted as  $\rho'_{AGN} = 0.27$  suggesting that, between low and high regimes of halo mass, early assembling haloes host more efficiently formed galaxies. This is most likely because earlier forming haloes have a longer time in which they are able to produce stars and the energy fraction remains low. The AGN scale  $M_{200}$  describes the halo mass where AGN feedback begins to dominate over stellar feedback and so the correlation corresponding to a strong AGN feedback is more appropriately that of the high mass halo regime; negative. In other words, where AGN feedback is strong, more efficient galaxies are hosted by late assembling haloes. For halo masses greater than  $M_{200} \sim 10^{13} M_\odot$  however, haloes and their galaxies grow mostly through mergers and so the effects of assembly history on efficiency are not as strong.

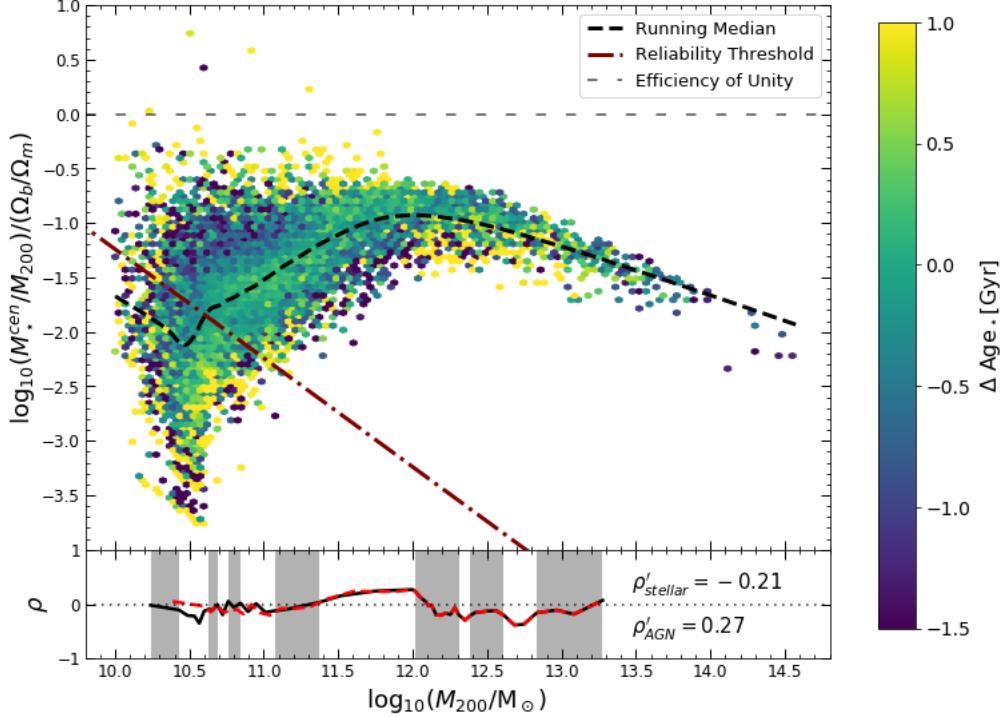


FIGURE 5.6: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted mean stellar mass-weighted stellar age in Gyr (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficients are quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ) and  $M_{200} = 10^{12} \pm 0.05 M_\odot$  ( $\rho'_{AGN}$ ).

## 5.4 Gas Density

### 5.4.1 Within a 30kpc Spherical Aperture

The growing mass of the central BH corresponds to a morphological transformation from a gaseous, disk-like galaxy to one that is spherical with much less gas and so it is immediately obvious that to first order, increasing BH mass should correlate with lower gas density  $\rho_{gas}$  where BH mass grows with increasing total halo mass. This is illustrated in the left panel of

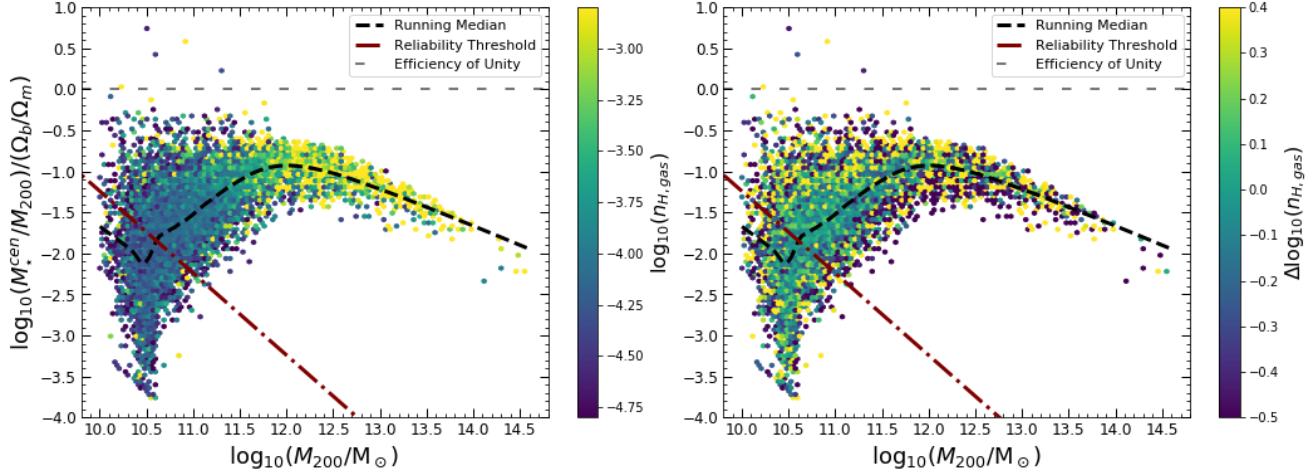


FIGURE 5.7: SMHM relation for centrals in the EAGLE RefL0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by ( $\log_{10}$  of) the mean gas mass-weighted Hydrogen number gas density (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

Figure 5.7 where there is a clear indication that high mass haloes contain a higher density of gas - represented as ( $\log_{10}$  of) the mean gas mass-weighted Hydrogen number gas density  $n_{H,gas}$  within 30 kpc in the same way that birth density is - than low mass at the present day. On the other hand, correlations with  $\Delta \log_{10}(n_{H,gas})$  are extremely more difficult to visually identify where the distribution of  $\log_{10}(n_{H,gas})(M_{200})$  with the running median fit can be seen in Appendix A.3.

On further inspection into the second order correlations with  $\Delta \log(n_{H,gas})$  using the Spearman rank coefficients, for particles within a 30 kpc spherical aperture, it can be seen in Figure 5.8 that correlations only exist with significance for limited ranges of halo mass. Statistically speaking, it is difficult to identify significant correlations across the scaling relation for this quantity, when excluding everything further than 30 kpc from the halo centre of potential. Near the AGN scale at  $M_{200} \approx 10^{12.1} M_\odot$  and at a higher halo mass of  $M_{200} \approx 10^{11.3} M_\odot$ , the correlation is significantly positive, however, given the size of the windows, it is difficult to generalise

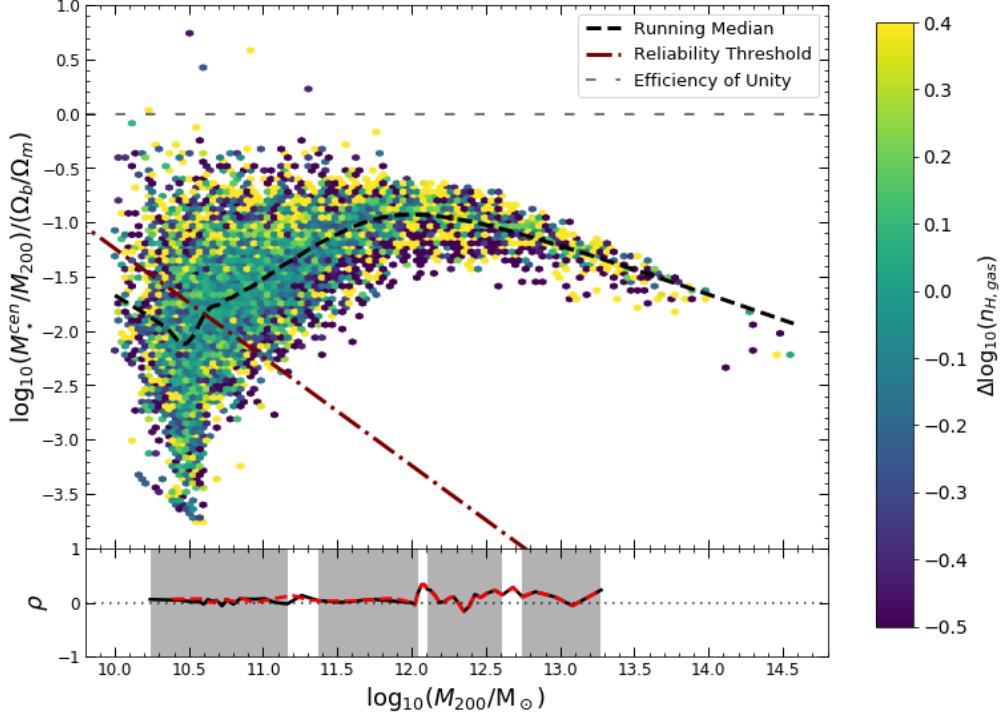


FIGURE 5.8: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean gas mass-weighted Hydrogen number gas density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ).

these results to the entire sample. Between stellar and AGN scales at  $M_{200} \approx 10^{11.3} M_\odot$ , the correlation coefficient is also significant but suggests there isn't a strong positive or negative correlation. In other words, where neither stellar nor AGN feedback dominate, there is no stratification of the present gas density by stellar mass. Another characteristic describing the amount of gas inside the galaxies is the gas fraction, which may provide further insight into the connection with the amount of gas. Dependencies on this quantity are explored in Section 5.6. In order to better compare with the results from Davies et al. [13] where it was found that early assembling galaxies became spheroidal by the present day, the gas density without any spherical aperture to include the entire halo can be investigated to correlate

with the SMHM relation at fixed total halo mass.

### 5.4.2 Total Halo

With the inclusion of the stellar mass and gas density further from the centre of potential than 30 kpc, more can be understood about the correlation of the gas density and the historic SF efficiency. Figure 5.9 shows the same SMHM relation, however, with stellar masses computed for the total halo rather than only particles within 30 kpc and is also coloured by

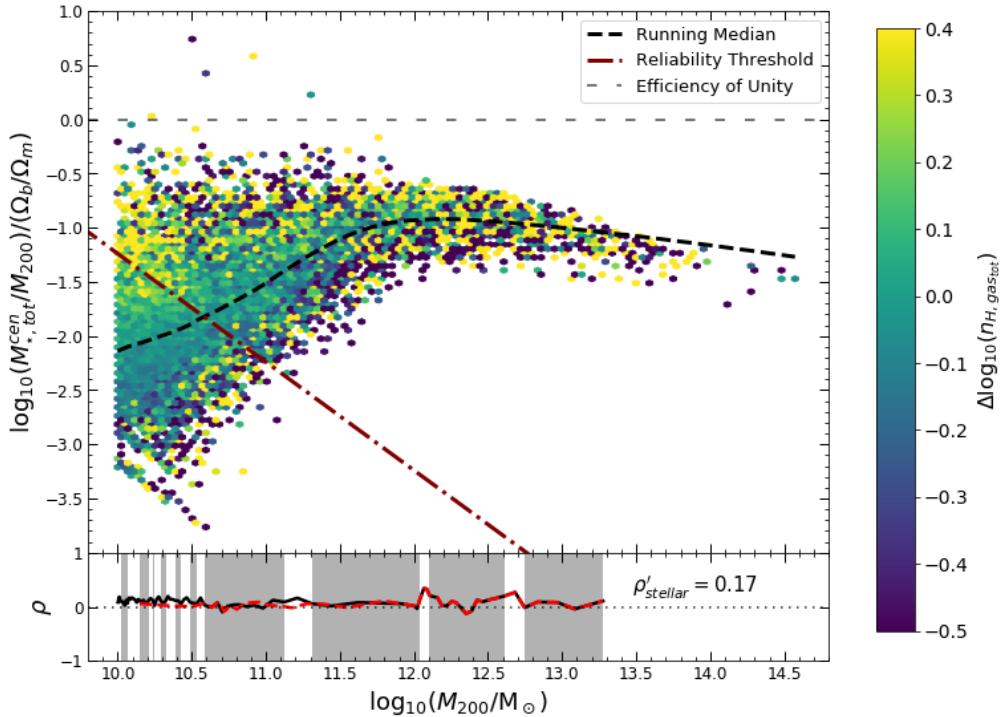


FIGURE 5.9: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in the total halo and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) mean gas mass-weighted total Hydrogen number gas density (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ).

$\Delta \log(n_{H,gas})$  with the distribution of the (total)  $\log_{10}(n_{H,gas})(M_{200})$  with the running median fit shown in Appendix A.3. The same message is retrieved from interpretations as that of the aperture-cut relation though at first glance it seems to present slightly more information about the correlation at low halo mass, where  $\rho'_{stellar} = 0.17$  suggesting a significant but small positive correlation. However,  $\rho_r$  provides a better indication of the correlation at this regime and shows there to be none present. In combination with results from the aperture-cut relation, this suggests that galaxies can have arbitrary gas densities which have no dependence on their formation histories.

## 5.5 Energy Fraction

Energy fraction  $f_{th}$  is the fraction of energy available from stellar feedback to couple to the ISM and is, therefore, a feature of the EAGLE simulations which has been manually controlled, being dependent upon naturally occurring processes such as the previously mentioned galaxy birth density and metallicity. Due to its dependence on stellar feedback, the significance of correlations at the low halo mass is of more interest in interpreting results from. To first order, it can be seen generally in Figure 5.10 that  $f_{th}$  grows with halo mass, although it may also increase towards low halo masses where the distribution of  $f_{th}(M_{200})$  with the running median fit can be seen in Appendix A.3. Interestingly, formation efficiency is seen to be maximum for haloes with  $f_{th} \approx 1$ , suggesting that efficiency is hindered by both weak and strong stellar feedback coupling with the ISM and so a compromise is found between the two.

For low mass haloes where feedback associated with SF dominates, efficiency was seen to be maximal in metal-rich galaxies in low birth density environments. Figure 5.11 shows significant correlation between the stellar and AGN scales suggesting that low mass haloes host galaxies which are

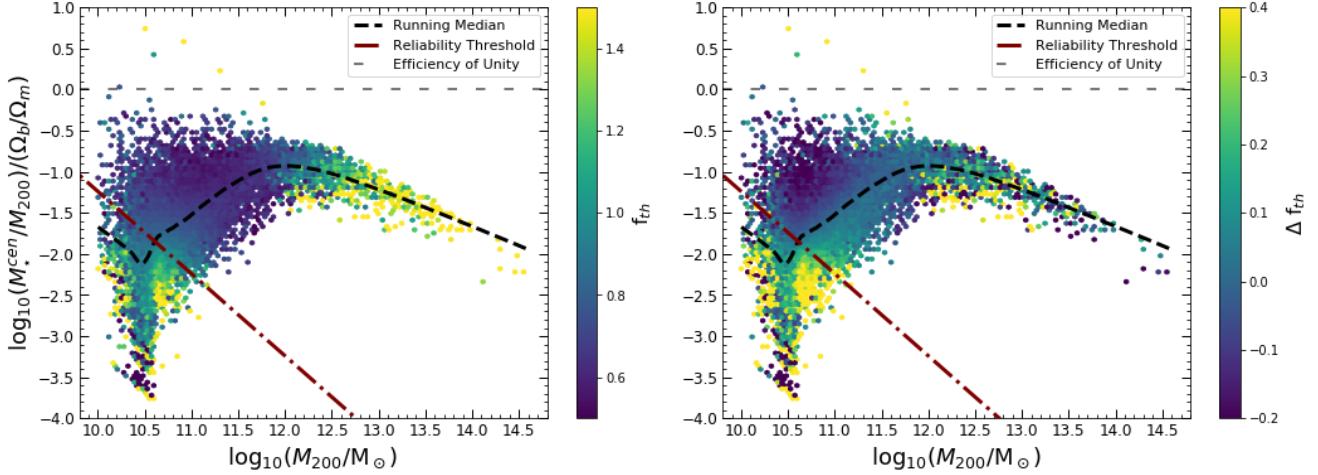


FIGURE 5.10: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by mean stellar mass-weighted energy fraction (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

also more efficient if formed in environments with slightly low  $f_{th}$  and so less energy from stellar winds and supernovae had an influence on the ISM, as expected. However, the correlation is not as negative for low mass haloes suggesting that the efficiency of formation of galaxies is independent of the  $f_{th}$  and must, therefore, depend on different processes. In other words, other factors may cause efficiency to correlate with birth density and metallicity, but not with  $f_{th}$  at this regime, for example additional variables in the energy fraction function. The determination of the correlation coefficient of  $\rho'_{stellar} = -0.24$  accounts for  $0.05 M_\odot$  either side of the stellar scale causing it to be influenced by the minimum coefficient achieved at slightly higher halo mass, and so does not accurately represent the observed correlation at this scale. Nevertheless, between regimes,  $f_{th}$  influences the efficiency significantly meaning that the simulations are quite sensitive to the input function governing this process.

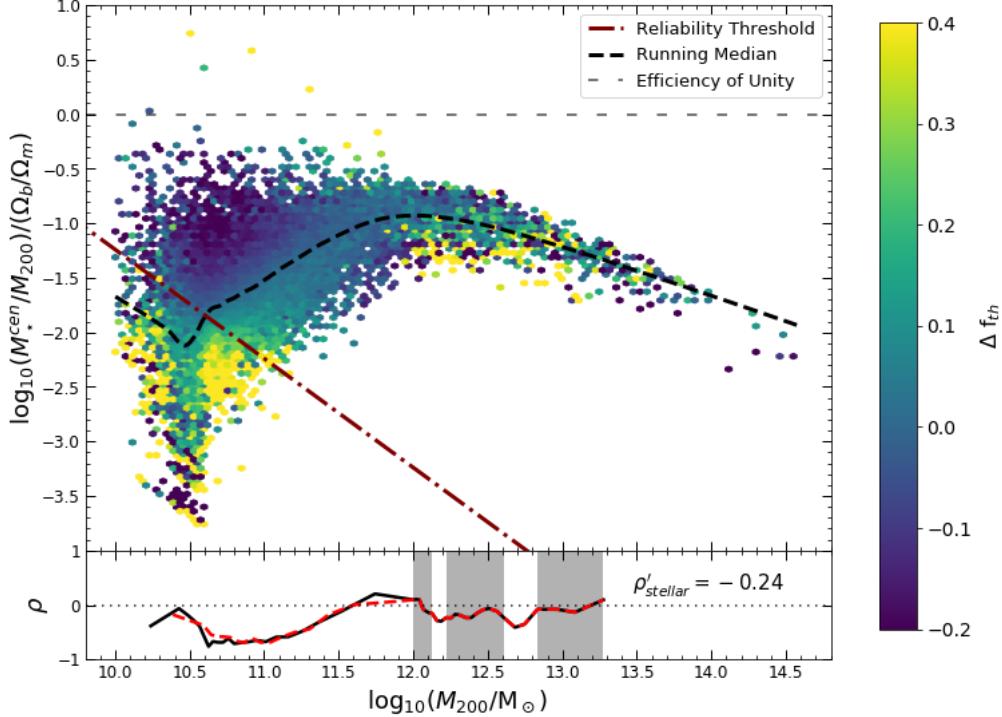


FIGURE 5.11: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation computed from particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted mean stellar mass-weighted energy fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar}$ ).

### 5.5.1 Dependence on Stellar Feedback Strength

To investigate further the correlations between  $f_{th}$  and formation efficiency and how they depend upon feedback strength, the variations from the Ref-L0025N0376 simulation could be coloured by their respective (absolute) mean stellar mass-weighted energy fractions which have markedly different distributions. Histograms of those for each variation are shown in each panel of Figure 5.12 with identical axis limits. This provides the SMHM relation for the reference model (top left), the altered initial conditions simulation to invoke weaker stellar feedback (top right) and stronger stellar

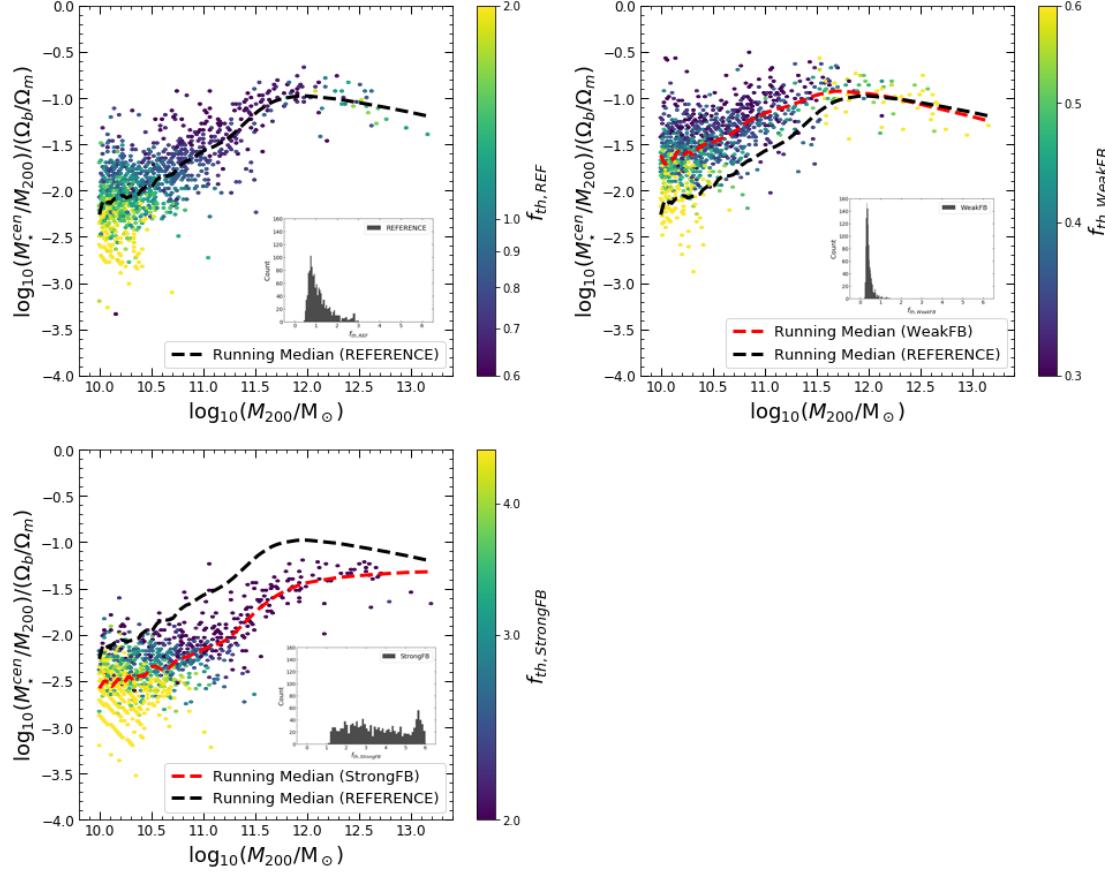


FIGURE 5.12: SMHM relation for centrals in the EAGLE L0025N0376 simulation computed from particles within 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Panels include (and are coloured by) the mean stellar mass-weighted energy fraction  $f_{th}$  from) the reference model (top left), invoked weak stellar feedback (top right) and strong stellar feedback (bottom left). The black dashed line included in each panel indicates the running median for the reference model for comparison against the running median for each variation (red dashed line). Each panel includes a histogram of  $f_{th}$  for the corresponding simulation variation.

feedback (bottom left). Therefore, to first order  $f_{th}$  the peaks of the relation show that the effect of weaker stellar feedback causes galaxies to be more efficiently formed with slightly higher  $f_{th}$ , and for strong stellar feedback with lower  $f_{th}$  though better sampling through the use of variations from the higher resolution simulations such as Ref-L0100N1504 would give a clearer indication. The maximum efficiency is achieved by a compromise between feedback strength and energy available from star formation feedback to couple to the ISM.

## 5.6 Gas Fraction

### 5.6.1 Within a 30kpc Spherical Aperture

The gas fraction  $f_{gas}$  is the fraction at  $z = 0$  of the gas mass out of the total mass of a galaxy and so does not indicate the environments that galaxies have formed in but rather shows a correlation between the formation efficiency and the present-day gaseous content of the galaxies. Naturally, this would be expected to complement the findings by Davies et al. [13] which showed the morphological transformation dependence upon the assembly history, however, they investigated the total halo as opposed to using a 30 kpc spherical aperture. Therefore, the local effects on  $f_{gas}$  are shown using the aperture and the total halo for comparison. Because of the gaussian-like number distribution of  $f_{gas}$ , the  $\log_{10}$  of  $f_{gas}$  is used to colour the SMHM relation as shown in the left panel of Figure 5.13 with the distribution of  $\log_{10}(f_{gas})(M_{200})$  and the running median fit shown in Appendix A.3. To first order, more efficiently formed galaxies have slightly

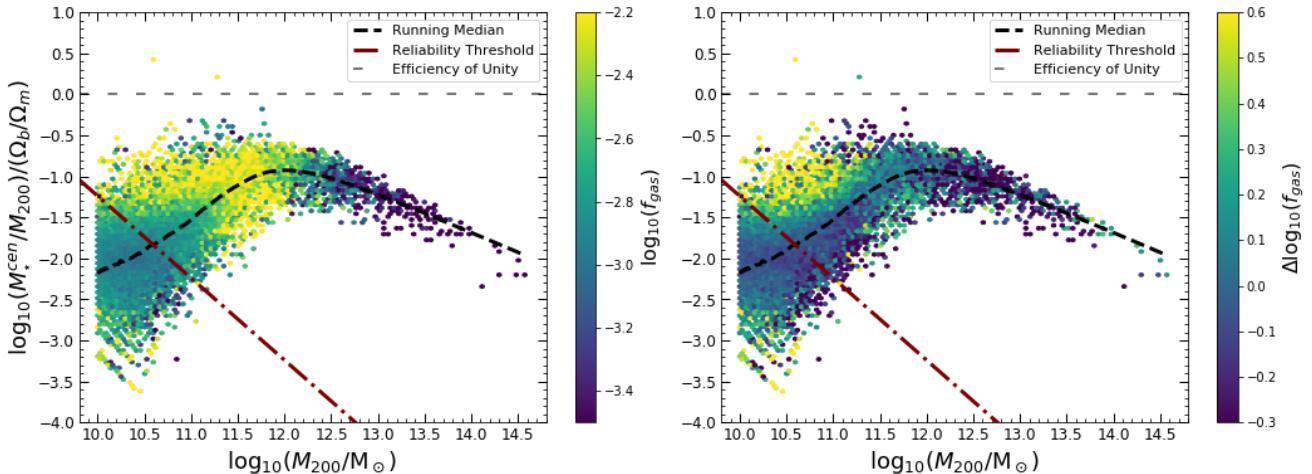


FIGURE 5.13: SMHM relation for centrals in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . Hexbins are coloured by ( $\log_{10}$  of the) gas fraction (left panel) and the median subtracted version (right panel). The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity.

higher  $\log_{10}(f_{\text{gas}})$ , low mass haloes have averagely lower gas fractions and high mass haloes have even lower gas fractions but as expected the correlations become more complex at the second order.

By colouring with the ( $\log_{10}$  of the) median-subtracted gas fraction  $\Delta \log_{10}(f_{\text{gas}})$  as shown in Figure 5.14, correlations at fixed total halo mass can be investigated.  $\rho_c$  and  $\rho_r$  show that at low halo mass, more efficiently formed galaxies are those with a more gaseous component at the present day with a fairly strong correlation coefficient of  $\rho'_{\text{stellar}} = 0.2$ . This complements the findings from the correlations with stellar age where it was

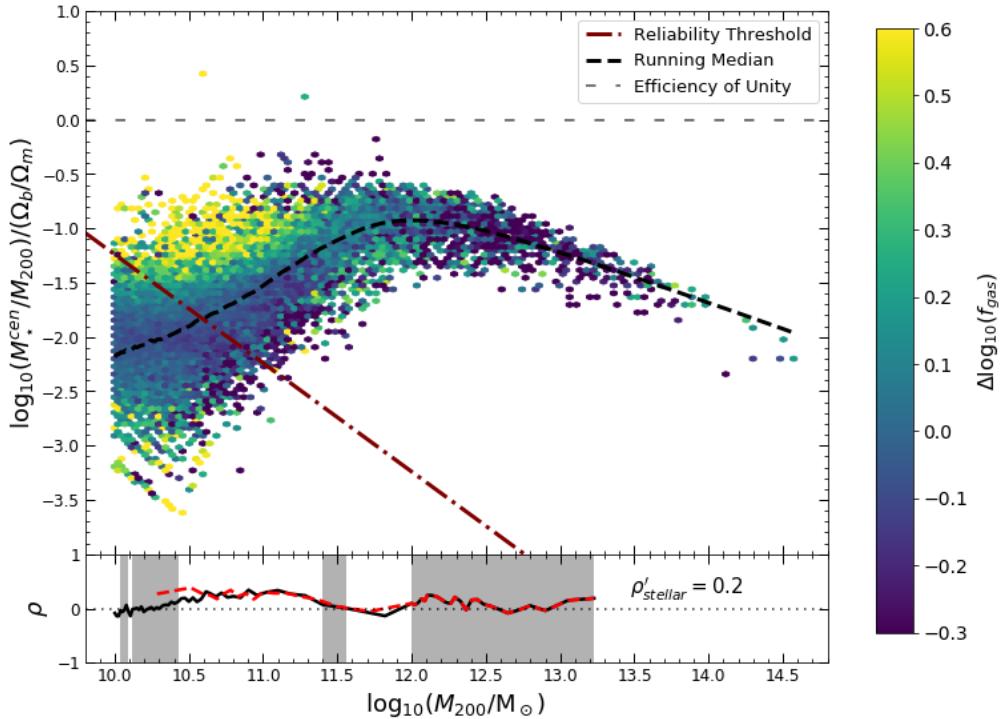


FIGURE 5.14: SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in 30 kpc and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) gas fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{\text{stellar}}$ ).

seen that late assembling galaxies were more efficiently formed at this halo mass which retain their disc by  $z = 0$  as opposed to early assemblers which have their SF quenched at earlier redshift. At the AGN scale, the correlation diminishes to nothing where there is no stratification by gas fraction, and so the amount of gas present today is independent of the efficiency of the galaxy in producing stars. At this halo mass, earlier assembling haloes were seen to be more efficient formers and so a negative correlation would be expected with  $\log_{10}(f_{gas})$ . This therefore shows that the amount of gas within 30 kpc from the halo centre of potential is independent of the assembly history for massive haloes. The effect on the whole galaxy may be better observed by investigating the total halo.

### 5.6.2 Total Halo

Figure 5.15 shows the same SMHM however with the inclusion of particles inside the total halo and so is coloured by  $\Delta \log_{10}(f_{gas,tot})$ . The distribution of  $\log_{10}(f_{gas,tot})(M_{200})$  and the running median fit is shown in Appendix A.3 and is markedly different to that of the aperture-cut for high mass haloes, highlighting the different correlations observed with the median-subtracted value past the stellar scale. At low halo mass,  $\rho_c$  shows a negative correlation as  $\rho'_{stellar} = -0.19$  which suggests that more efficiently formed galaxies actually have smaller amounts of gas across the total halo at the present day. However,  $\rho_r$  provides a more reliable perspective on the correlation at this regime which supports the findings from the aperture-cut correlations with  $f_{gas}$ . Most notably, the correlation is seen to diminish to zero at lower halo masses than in the aperture-cut relation and becomes negative between the stellar and AGN scale which agrees with interpretations from correlations with stellar age. Over the total halo, there is less gas at the present day in earlier assemblers which host more efficiently formed galaxies. Massive efficiently formed galaxies are built through preferential *ex-situ* mass acquisition which dominates the stellar

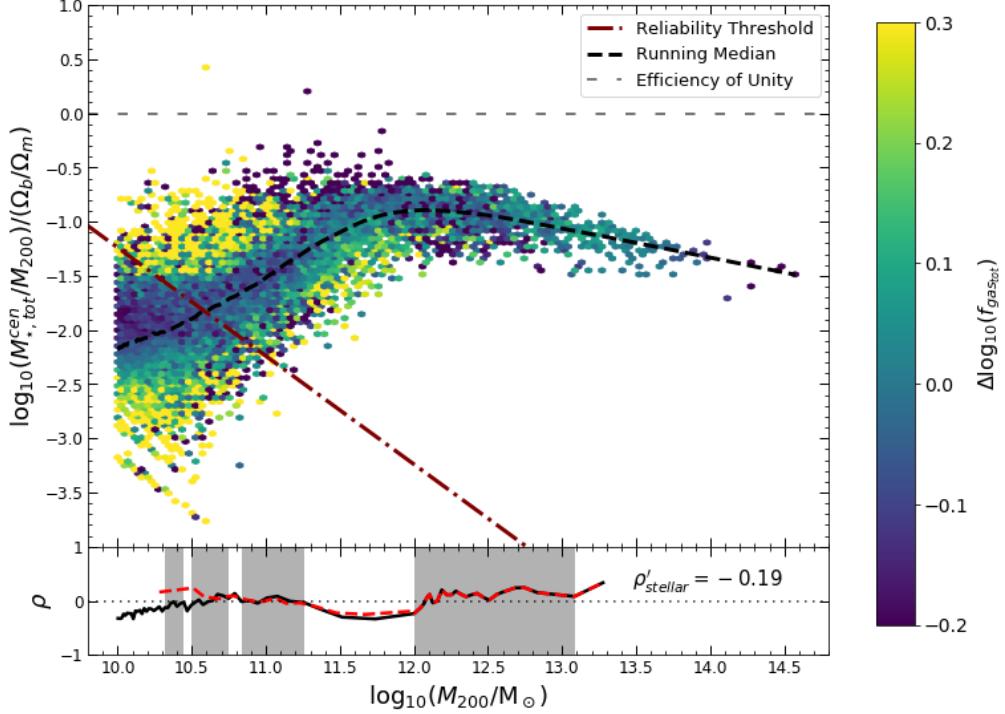


FIGURE 5.15: SMHM relation for central haloes in the EAGLE Ref-L0100N1504 simulation taken from the catalogues for particles in the total halo and normalised by the cosmic average baryon fraction,  $\Omega_b/\Omega_0$ . The black dashed line is the running median, the maroon dot-dashed line the reliability threshold and the grey dashed line is efficiency unity. Hexbins are coloured by the median-subtracted ( $\log_{10}$  of the) gas fraction (top panel) with the running Spearman rank coefficient  $\rho$  in the bottom panel. The solid black line indicates  $\rho_c$  and the red dashed line indicates  $\rho_r$ . Shade regions correspond to  $\rho_c$  values of low significance ( $p > 0.01$ ). As significant, the mean correlation coefficient is quoted for haloes within  $M_{200} = 10^{10.5} \pm 0.05 M_\odot$  ( $\rho'_{stellar} = -0.19$ ).

mass budget. These are able to become spheroidal because *ex-situ* stars are extended further from the disc plane and the role of AGN feedback is to suppress *in-situ* SF, resulting in a more highly dispersed distribution of stars (Dubois et al. [15]).

# Chapter 6

## Summary and Discussion

The efficiency of galaxy formation described in the stellar mass – halo mass relation exhibits significant scatter with a direct dependence on the feedback associated with star formation and AGN activity. In addition, the quenching and morphological transformations of galaxies are intimately connected to the assembly histories of their host haloes. In this investigation, further physical processes in galaxies such as the birth density, metallicity, stellar age, gas density, energy fraction and gas fraction have been studied to explore their correlations with formation efficiency at fixed halo mass because they trace changes in morphology and determine the mechanisms by which galaxies evolve, unveiling second order influences on the scatter in the scaling relation.

At low and high halo mass, galaxies are more efficient formers if formed in low birth density environments because of the disruption to SF by stellar and AGN feedback, respectively. Between regimes of halo mass, where feedback from star formation and AGN is weak, galaxies are able to reap the benefits of the high density of gas at the instance of its conversion into stars. The direct influence of feedback on formation efficiency is demonstrated here through its influence on a galaxy’s ability to use the fuel from its surroundings. With only the inclusion of reliable sampling in stellar and DM particles, the Spearman rank correlation coefficient suggests that, for

halo masses smaller than investigated, a galaxy may be more efficient if formed in high birth density environments as it is between the stellar and AGN scale. Future investigations into the influence of birth density on formation efficiency would benefit by reducing the minimum halo mass which in turn would benefit from better resolution simulations allowing for stronger reliability at these masses.

There exists a strong positive correlation between metallicity and formation efficiency for all halo masses, where increased metallicity is both an ingredient via the energy fraction and a symptom of a burst of SF. Because an increase in the SFR allows galaxies in early assembling haloes to host more generations of stars and enrich the ISM with heavier elements, it becomes difficult to establish the role of metallicity in influencing the efficiency of formation. Further investigations would be required to disentangle cause and effect from the exhibited correlations. It would be expected that the increase in rate of stellar feedback from the increase in SFR would influence the correlation more strongly and so using the variations of EAGLE simulations which have their stellar and AGN feedback altered could help to understand if this is the case.

Stellar age provides a relatively weak proxy for the different assembly histories where early assemblers have their SF quenched earlier and so contain averagely older stars than late assemblers which retain their disks of star forming gas. Nevertheless, second order investigations suggest that at both low and high halo mass, late assembling haloes host more efficiently formed galaxies because SF is quenched in early assemblers via the rapid growth in mass of the central BH and increased rate of AGN feedback. For halo masses when stellar and AGN feedback are weak, early assembling haloes host more efficiently formed galaxies because they are able to start forming stars earlier and the energy fraction remains low.

Within a spherical aperture of 30 kpc, there is no strong correlation between formation efficiency and gas density. Galaxies can have arbitrary

gas densities with no dependence upon their assembly histories. Investigations into the gas density and stellar mass for particles inside the total halo support these findings, however, the significance of the Spearman rank correlation is low and so it is difficult to confirm such results.

Energy fraction describes the fraction of energy available from stellar feedback to couple to the ISM. At both low and high halo mass, galaxies are more efficiently formed if formed in environments with low  $f_{th}$  although less-so in low halo mass. At this regime, other factors in the  $f_{th}$  function may more strongly influence efficiency, however, between the stellar and AGN scales there is a strong dependence of efficiency on  $f_{th}$ . It is clear that the stellar mass – halo mass relation is significantly sensitive to the energy fraction function implemented in the EAGLE simulations. A manually engineered function of  $f_{th}$  dependent on metallicity and birth density is highly susceptible to the limits of the current understanding of galaxy formation and evolution. This highlights a component that is essential to improve in order for the scaling relation obtained from cosmological simulations to reproduce that which is observed. A brief investigation into how efficiency depends on both  $f_{th}$  and stellar feedback strength indicates that when stellar feedback is weak, galaxies are more efficient with high  $f_{th}$  and when strong, are more efficient with low  $f_{th}$ . Therefore, a compromise is met to maximise efficiency, between the stellar feedback strength and the energy available from stellar feedback to influence the ISM as neither can be low while the other is.

In low mass haloes, more efficiently formed galaxies have a smaller fraction of gas at the present day within 30 kpc, in agreement with galaxies in late assembling haloes being more efficient which retain their gas content. However, at the AGN scale, the amount of gas is independent of the formation efficiency and by extension is independent of the assembly history, contrasting against findings by Davies et al. [13] which found galaxies in early assembling haloes to have their SF quenched and which were seen in Section 5.3 to be more efficient at this halo mass, but they found galaxies to

have less gas when investigating the total halo. Only when accounting for all particles inside the halo is it found that there is less gas at the present day in early assembling galaxies at this scale. Massive efficiently formed galaxies are built through preferential *ex-situ* mass acquisition which dominates the stellar mass budget. These are able to become spheroidal because *ex-situ* stars are extended further from the disc plane and the role of AGN feedback is to suppress *in-situ* SF. Models of central BHs and AGN feedback are therefore essential in the morphological transformations of galaxies and determine the relation at high halo mass.

Thus, the models implemented in the EAGLE simulations to describe the behaviour of feedback associated with star formation and central BHs in their AGN states is therefore a crucial component in determining the stellar mass for the scaling relation and should be a theme of focus in future investigations which hope to successfully reproduce the observed SMHM relation with minimal scatter. To reduce scatter at low halo mass, benefits could be found by improving the resolution to understand the influence of birth density on the stellar mass and further investigating the influences of metallicity and its dependence on feedback associated with star formation. Between regimes, scatter could be reduced by improving the energy fraction function implemented in EAGLE as the relation is significantly sensitive to what is not well known. At high halo mass, the determined relation would immediately benefit from improved resolution but also the models implemented of the central BHs and activity from AGN as they play key roles in determining the evolutionary paths of galaxies through their morphological transformations.

# Appendix A

## Computing Particle Quantities

### A.1 Wrapping the Simulation Box

For assessment purposes, the following public GitHub repository contains the code (written in Python) used to wrap the periodic EAGLE Ref-L0100N1504 simulation box to obtain quantities from the particles themselves rather than the catalogues, along with an example code used to plot the EAGLE Ref-L0100N1504 simulation stellar mass - halo mass relation coloured by ( $\log_{10}$  of the) Hydrogen number birth density and the median-subtracted version with the running complete and reduced Spearman rank correlation coefficients:

[https://github.com/dom-taylor1/MPhys\\_EAGLE](https://github.com/dom-taylor1/MPhys_EAGLE)

### A.2 Particle Computed Stellar Mass versus Catalogued

The maximum difference observed for computing stellar masses from the EAGLE Ref-L0025N0376 was magnitudes smaller than that observed for

Ref-L0100N1504, however, the majority of the error is within  $\sim 1\%$ , providing sufficient precision for computing quantities from particles themselves.

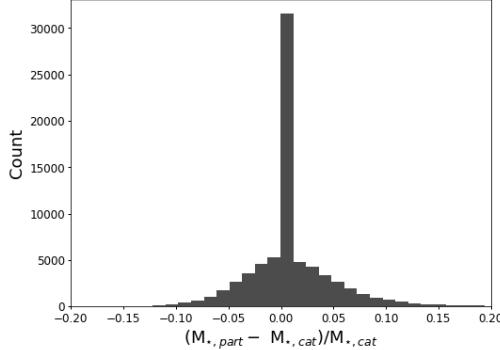


FIGURE A.1: Histogram of the difference between particle computed total stellar mass of central haloes and that from the catalogue of the EAGLE Ref-L0100N1504 simulation.

### A.3 Running Median Fit of Third Variables

The following plots provide the running medians of the physical processes investigated to correlate with the SMHM relation, which make use of the fraction 0.2 of the total data in order to fit a value for every halo mass, are iterated a total of 3 times and use only weighted regression.

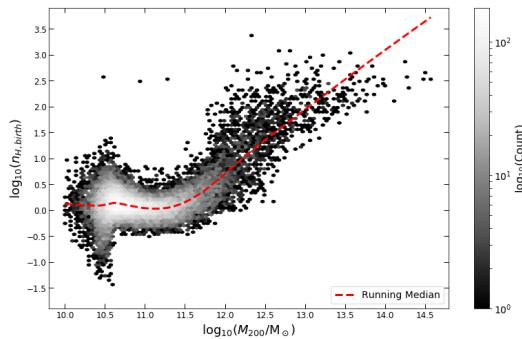


FIGURE A.2: Hexbin plot of ( $\log_{10}$  of the) mean stellar mass-weighted Hydrogen number birth density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

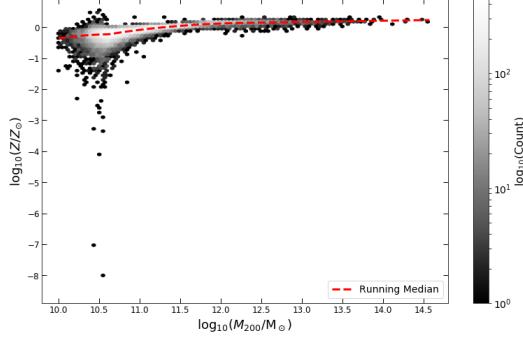


FIGURE A.3: Hexbin plot of the mean metallicity for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

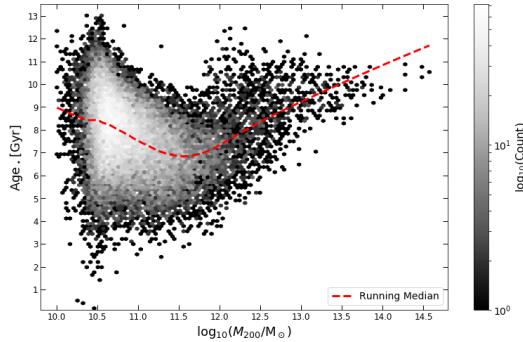


FIGURE A.4: Hexbin plot of the mean stellar mass-weighted stellar ages (in Gyr) for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

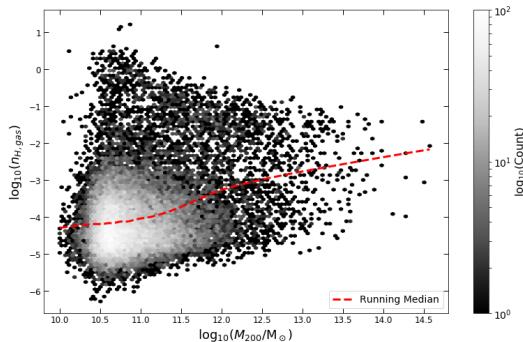


FIGURE A.5: Hexbin plot of the ( $\log_{10}$  of) mean gas mass-weighted Hydrogen number gas density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

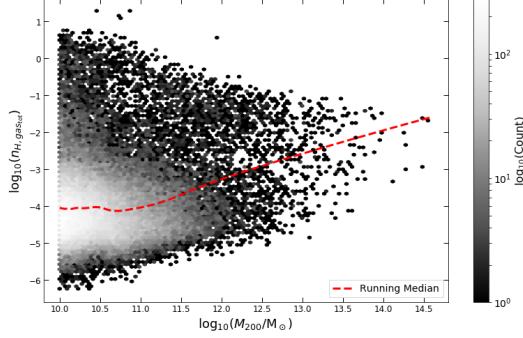


FIGURE A.6: Hexbin plot of the ( $\log_{10}$  of) mean gas mass-weighted Hydrogen number gas density for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within the total halo, coloured by number density. The red dashed line shows the running median fit.

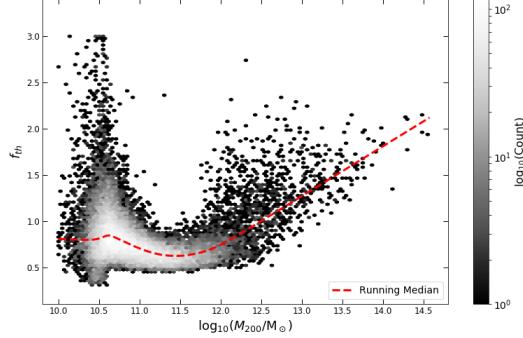


FIGURE A.7: Hexbin plot of the mean stellar mass-weighted energy fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

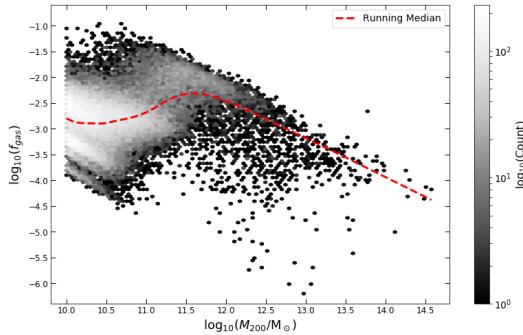


FIGURE A.8: Hexbin plot of the ( $\log_{10}$  of) gas fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within 30 kpc, coloured by number density. The red dashed line shows the running median fit.

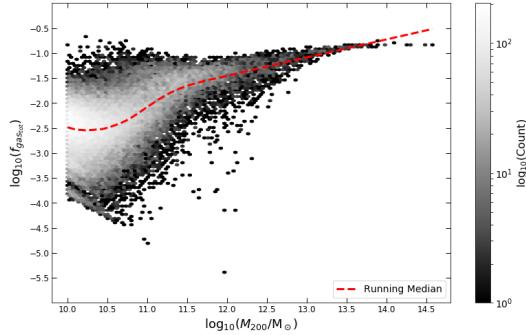


FIGURE A.9: Hexbin plot of the ( $\log_{10}$  of) gas fraction for central haloes in the EAGLE Ref-L0100N1504 simulation, for particles within the total halo, coloured by number density. The red dashed line shows the running median fit.

## Appendix B

### Poster Question and Answers

**1) Is it possible to vary the composition (hydrogen/helium/metals etc) of the ‘Gas’ in your simulations? Would it be of interest to do this?**

Yes, the EAGLE simulations record the abundances of different elements in the gas of galaxies. Carbon, for example, is a good tracer of H<sub>2</sub> in metal-poor galaxies (Glover & Clark [21]) which provides insight into the cooling efficiency of the ISM and so this could be investigated to determine what causes metal-poor galaxies to be less efficient formers for every halo mass.

**2) You conclude that black hole behaviour is important in some of the simulations and suggest it for future work. What kinds of black hole behaviour are there, or what kinds of behaviour can/should be simulated?**

Black holes cause powerful energetic events like winds and relativistic jets of radio, X-ray or ultraviolet radiation when they ignite AGN activity via the sufficient accretion of cold matter from the surrounding stars, dust and gas. Emission from AGN is not currently taken into account in post-processing of the simulations causing a discrepancy between observed luminosity functions and those in EAGLE (Baes et al. [2]). A delicate balance between the accretion and feedback from AGN is required to quench star formation

in low stellar mass galaxies and keep it quenched in those of higher stellar mass. For more information, see page 4.

# Bibliography

- [1] Akrami Y., Arroja F., Ashdown M., Aumont J., Baccigalupi C., Ballardini M., Banday A. J., Barreiro R. B., Bartolo N., et al. 2020, *Astronomy & Astrophysics*, 641, A10
- [2] Baes M., Trčka A., Camps P., Trayford J., Katsianis A., Marchetti L., Theuns T., Vaccari M., Vandenbroucke B., 2020, *Monthly Notices of the Royal Astronomical Society*, 494, 2912–2924
- [3] Bagla J. S., , 2009, Hubble, Hubble’s law and the expanding universe
- [4] Bahcall N. A., 2017, *Proceedings of the National Academy of Sciences*, 114, 2099
- [5] Bond J. R., Kofman L., Pogosyan D., 1996, *Nature*, 380, 603–606
- [6] Cimatti A., Fraternali F., Nipoti C., , 2019, *Introduction to Galaxy Formation and Evolution. From Primordial Gas to Present-Day Galaxies*
- [7] Cleveland W., 1979, *Robust Locally Weighted Regression and Smoothing Scatterplots*. American Statistical Association
- [8] Coil A., , Large Scale Structure of the Universe
- [9] Crain R., , The Eagle simulations
- [10] Crain R. A., Schaye J., Bower R. G., Furlong M., Schaller M., Theuns T., Dalla Vecchia C., Frenk C. S., McCarthy I. G., Helly J. C., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1937–1961
- [11] Davies J., , 2020, Examples of common EAGLE tasks

- 
- [12] Davies J. J., Crain R. A., Oppenheimer B. D., Schaye J., 2019, Monthly Notices of the Royal Astronomical Society, 491, 4462–4480
  - [13] Davies J. J., Crain R. A., Pontzen A., 2021, Monthly Notices of the Royal Astronomical Society, 501, 236
  - [14] Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2486
  - [15] Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, Monthly Notices of the Royal Astronomical Society, 463, 3948–3964
  - [16] Dyson J. E., Williams D. A., , The physics of the interstellar medium, 2 edn
  - [17] EAGLE, 2021
  - [18] Freundlich J., El-Zant A., Combes F., 2016, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I., eds, SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics How baryonic feedback processes can affect dark matter halos: a stochastic model. pp 153–156
  - [19] Garner R., , 2018, About the Hubble Space Telescope
  - [20] Girelli G., Pozzetti L., Bolzonella M., Giocoli C., Marulli F., Baldi M., 2020, Astronomy Astrophysics, 634, A135
  - [21] Glover S. C. O., Clark P. C., 2016, Monthly Notices of the Royal Astronomical Society, 456, 3596–3609
  - [22] Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85–88
  - [23] Horta D., Schiavon R. P., Mackereth J. T., Pfeffer J., Mason A. C., Kisku S., Fragkoudi F., Allende Prieto C., Cunha K., Hasselquist S.,

- et al. 2020, Monthly Notices of the Royal Astronomical Society, 500, 1385–1403
- [24] Ibata Gilmore Irwin 1994, Nature, 370, 1
- [25] Knebe A., , 2007, How to simulate the Universe in a computer
- [26] Lund A., Lund M., , Spearman’s Rank-Order Correlation - A guide to when to use it, what it does and what the assumptions are.
- [27] Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2016, Monthly Notices of the Royal Astronomical Society, 465, 2381–2396
- [28] Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, The Astrophysical Journal, 710, 903–923
- [29] NASA, 2021
- [30] Pittard J., , 2021
- [31] Price S., , Galaxy Rotation Curve
- [32] Schaller M., 2017, Eagle simulation limitations
- [33] Schaye J., Crain R. A., Bower R. G., Furlong M., Schaller M., Theuns T., Dalla Vecchia C., Frenk C. S., McCarthy I. G., Helly J. C., et al. 2014, Monthly Notices of the Royal Astronomical Society, 446, 521–554
- [34] Scolnic D. M., Jones D. O., Rest A., Pan Y. C., Chornock R., Foley R. J., Huber M. E., Kessler R., Narayan G., Riess A. G., et al. 2018, The Astrophysical Journal, 859, 101
- [35] Shu F. H., , 2021, Cosmic microwave background
- [36] Steigman G., Kneller J., Zentner A., 2001, The Astrophysical Journal
- [37] Susskind, Lesson 2 : Matter-dominated and radiation-dominated universes

- [38] Taylor D., 2019, Physical processes in computer simulations of galaxy formation and recent advances, Unpublished
- [39] Taylor D., 2020, Relativity and Cosmology Written Assignment, Unpublished
- [40] White S., 2017, Assembly bias
- [41] Zwicky F., 2008, General Relativity and Gravitation, 41, 207