

THE INFLUENCES OF DARK MATTER HALO
FORMATION HISTORY ON BLACK HOLE
GROWTH, CIRCUMGALACTIC GAS EXPULSION
AND GALAXY EVOLUTION

by

Jonathan J. Davies

A thesis submitted in partial fulfillment of the requirements of
Liverpool John Moores University
for the degree of
Doctor of Philosophy

April 2020

Declaration

The work presented in this thesis was carried out at the Astrophysics Research Institute, Liverpool John Moores University. Unless otherwise stated, it is the original work of the author.

While registered as a candidate for the degree of Doctor of Philosophy, for which submission is now made, the author has not been registered as a candidate for any other award. This thesis has not been submitted in whole, or in part, for any other degree.

Jonathan Davies
Astrophysics Research Institute
Liverpool John Moores University
ic2, Liverpool Science Park
146 Brownlow Hill
Liverpool
L3 5RF
UK

APRIL 2020

Abstract

Modern models predict that galaxies do not evolve in isolation, but exist in haloes of dark matter and gas, embedded in the large-scale “cosmic web”. To build a self-consistent model for how galaxies form and evolve, it is vital to understand how galaxies regulate their formation through interaction with the “circumgalactic medium” that surrounds them, and at present the nature and impact of these interactions is poorly-constrained by both theory and observations. This thesis presents new insights into how galaxies affect their surroundings through feedback processes, the impact of the properties of the CGM on galaxy evolution, and how the properties of galaxy-CGM ecosystems are fundamentally connected to the assembly histories of their dark matter haloes.

I begin by examining the origin of scatter in the relationship between the gas fraction and mass of dark matter haloes hosting present-day $\sim L^*$ central galaxies in the EAGLE simulations. The scatter is uncorrelated with the accretion rate of the central galaxy’s black hole (BH), but correlates strongly and negatively with the BH’s mass, implicating differences in the expulsion of gas by active galactic nucleus (AGN) feedback, throughout the assembly of the halo, as the main cause of scatter. Haloes whose central galaxies host undermassive BHs also tend to retain a higher gas fraction, and exhibit elevated star formation rates (SFRs). Diversity in the mass of central BHs stems primarily from diversity in the dark matter halo binding energy, as these quantities are strongly and positively correlated at fixed halo mass, such that $\sim L^*$ galaxies hosted by haloes that are more (less) tightly-bound develop central BHs that are more (less) massive than is typical for their halo mass. Variations in the halo gas fraction at fixed halo mass are reflected in both the soft X-ray luminosity and thermal Sunyaev-Zel’dovich flux, suggesting that the prediction of a strong coupling between the properties of galaxies and their halo gas fractions can be tested with measurements of these diagnostics for galaxies with diverse SFRs but similar halo masses.

I then examine the connection between the properties of the CGM and the quenching and morphological evolution of central galaxies in both the EAGLE and IllustrisTNG simulations. The simulations yield very different median CGM mass fractions, f_{CGM} , as a function of halo mass, M_{200} , with low-mass haloes being significantly more gas-rich in IllustrisTNG than in EAGLE. Nonetheless, in both cases scatter in f_{CGM} at fixed M_{200} is strongly correlated with the specific star formation rate and the kinematic morphology of central galaxies. This feedback elevates the CGM cooling time, preventing gas from accreting onto the galaxy to fuel star formation, and thus establishing a preference for quenched, spheroidal galaxies to be hosted by haloes with low f_{CGM} for their mass. In both simulations, f_{CGM} correlates negatively with the host halo's intrinsic concentration, and hence with its binding energy and formation redshift, primarily because early halo formation fosters the rapid early growth of the central black hole (BH). This leads to a lower f_{CGM} at fixed M_{200} in EAGLE because the BH reaches high accretion rates sooner, whilst in IllustrisTNG it occurs because the central BH reaches the mass threshold at which AGN feedback is assumed to switch from thermal to kinetic injection earlier. Despite these differences, there is consensus from these state-of-the-art simulations that the expulsion of efficiently-cooling gas from the CGM is a crucial step in the quenching and morphological evolution of central galaxies.

The above results suggest a connection between the assembly histories of dark matter haloes and the properties of their galaxy-CGM ecosystems. I clearly demonstrate this connection by performing a controlled and systematic experiment in which I adjust the assembly history of a single EAGLE halo hosting a moderately star-forming Milky Way-like galaxy by “genetically modifying” its initial conditions, keeping all other variables fixed. Shifting the halo assembly history to earlier times increases the integrated feedback injected by the AGN, ejecting a greater fraction of the CGM baryons and leading to the quenching and morphological transformation of the central galaxy. These effects can only have originated from differences in the assembly history of the halo, providing compelling evidence for this novel picture of the self-regulation of galaxy formation.

Contents

Declaration	ii
Abstract	iii
List of Figures	vii
Publications	ix
Acknowledgements	x
1 Introduction	1
1.1 The Λ CDM paradigm	2
1.2 Galaxy formation in dark matter haloes	5
1.2.1 Assembly of stellar mass through gas accretion	5
1.2.2 Feedback as a regulatory mechanism	7
1.3 The circumgalactic medium	11
1.3.1 Observational study of the CGM	13
1.3.1.1 X-ray emission	14
1.3.1.2 Thermal Sunyaev-Zel'dovich effect	15
1.3.1.3 Absorption features in background quasar spectra	16
1.3.2 The CGM in cosmological simulations	17
1.4 This thesis	19
2 A universe in a box: cosmological simulations of galaxy formation	22
2.1 Initial conditions and background cosmology	24
2.2 Gravity solver	26
2.3 Hydrodynamics	27
2.4 Characterising haloes and galaxies	28
2.5 Subgrid physics	29
2.5.1 Radiative cooling, photoheating and reionisation	30
2.5.2 Star formation	30
2.5.3 Stellar feedback	32
2.5.4 Black hole seeding, growth and feedback	34
2.6 Coupling to plasma emission models	36
3 Black hole feedback and the gas content of dark matter haloes	39
3.1 Introduction	39
3.2 Methods	40

3.3	The origin of scatter in halo gas fractions	42
3.4	Testing via complementary observables	49
3.5	Summary and discussion	53
4	The role of circumgalactic gas expulsion in galaxy evolution	57
4.1	Introduction	57
4.2	Methods	58
4.2.1	Identifying and characterising haloes and galaxies	59
4.2.2	Cooling rates and timescales	60
4.2.3	Feedback energetics	61
4.3	The correlation of galaxy and BH properties with the CGM mass fraction	62
4.4	The influence of feedback on the cooling time of circumgalactic gas	66
4.4.1	The effect of feedback on the CGM cooling time	67
4.4.2	Quenching and morphological evolution in response to elevation of the CGM cooling time	71
4.5	The origin of the diversity in CGM mass fractions at fixed halo mass	73
4.6	Summary and Discussion	80
5	The galaxy-CGM connection in dark matter haloes with controlled assembly histories	87
5.1	Introduction	87
5.2	Methods	88
5.2.1	Construction of the initial conditions	88
5.2.2	The EAGLE model	90
5.2.3	Identifying haloes, galaxies and their progenitors	91
5.2.4	Feedback energetics and cooling timescales	92
5.3	Evaluation of modified haloes and their assembly histories	93
5.4	The influence of halo assembly history on the galaxy-CGM ecosystem	98
5.4.1	The influence of halo assembly history on the CGM mass fraction	98
5.4.2	The influence of halo assembly history on BH growth and AGN feedback	102
5.4.3	The influence of halo assembly history on galaxy properties	107
5.5	The impact of AGN feedback on the thermodynamic state of the CGM	112
5.6	Summary and Discussion	116
6	Summary and Conclusions	122
6.1	Future work	125
6.1.1	Prospects for observational testing	125
6.1.2	Further theoretical study	126
6.1.2.1	The connection between CGM expulsion and galaxy mor- phology	126
6.1.2.2	Relationship with large-scale structure and the cosmic web	127
6.1.2.3	Breaking the degeneracy in galaxy formation	128
A	Comparison of simulated X-ray luminosities and tSZ fluxes with ob- servations	130
Bibliography		133

List of Figures

1.1	A comparison of the observed galaxy stellar mass function with the results of several simulations. Reproduced from Naab & Ostriker (2017).	8
1.2	The $z \sim 0.1$ stellar mass - halo mass relation. Reproduced from Behroozi et al. (2019).	9
2.1	APEC emission spectra for H, O, Ne, Mg, Si and Fe at a temperature of $T = 10^7$ K.	38
3.1	Present-day CGM mass fractions, f_{CGM} , as a function of halo mass, M_{200} in EAGLE, coloured by the residuals of the relationship between the total stellar fraction of the halo and M_{200}	42
3.2	The EAGLE $f_{\text{CGM}} - M_{200}$ relation, coloured by residuals about the relationships of various properties as a function of halo mass: the accretion rate of the central BH (\dot{M}_{BH}), the mass of the central BH (M_{BH}), the intrinsic binding energy of the inner halo (E_{DMO}^{2500}), and the star formation rate of the central galaxy (\dot{M}_*).	44
3.3	Halo baryon fractions, f_b , as a function of halo mass, M_{200} , in EAGLE simulations with and without AGN feedback, coloured by the residuals of the $E_{\text{DMO}}^{2500} - M_{200}$ relation.	47
3.4	Diffuse soft X-ray luminosity and thermal Sunyaev-Zel'dovich effect flux as a function of M_{200} for present-day EAGLE haloes, coloured by f_{CGM} , M_{BH} and M_*	50
4.1	The present-day $f_{\text{CGM}} - M_{200}$ relation in EAGLE and IllustrisTNG, coloured by the residuals of the $\dot{M}_{\text{BH}} - M_{200}$ and $M_{\text{BH}} - M_{200}$ relations.	62
4.2	The $f_{\text{CGM}} - M_{200}$ relation in EAGLE and IllustrisTNG, coloured by the residuals of the relationships between both the specific star formation rate (sSFR) and stellar co-rotational kinetic energy fraction (κ_{co}) of the central galaxy, with M_{200}	65
4.3	Cumulative distribution functions of the radiative cooling times of fluid elements comprising the CGM of present-day $M_{200} \sim 10^{12.5} \text{ M}_\odot$ haloes in EAGLE and IllustrisTNG.	68
4.4	The $f_{\text{CGM}} - M_{200}$ relation in EAGLE and IllustrisTNG, coloured by the residuals of the relationship between the characteristic CGM radiative cooling time ($t_{\text{cool}}^{\text{CGM}}$) and M_{200}	69
4.5	The sSFR - M_{200} and $\kappa_{\text{co}} - M_{200}$ relations in EAGLE and IllustrisTNG, coloured by the residuals of the $t_{\text{cool}}^{\text{CGM}} - M_{200}$ relation.	71

4.6	The $f_{\text{CGM}} - M_{200}$ relation in EAGLE and IllustrisTNG, coloured by the residuals of two relations with M_{200} : that of $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$, which is a proxy for the halo concentration, and that of $E_{\text{FB}}/E_{\text{bind}}^b$, the ratio of the total integrated feedback energy to the intrinsic binding energy of the halo baryons.	74
4.7	The $E_{\text{FB}}/E_{\text{bind}}^b - M_{200}$ relation in EAGLE and IllustrisTNG, coloured by the residuals of the $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200} - M_{200}$ relation.	77
5.1	Images of the dark matter distribution in haloes with fiducial and GM initial conditions, at $z = 0$ and $z = 2$	94
5.2	Images of the large-scale structure of dark matter and boundary particles, generated with fiducial and GM initial conditions, at $z = 0$ and $z = 2$. . .	95
5.3	DONLY and RECAL halo mass accretion histories for the three families of simulations.	96
5.4	The evolution of f_{CGM} for the three families of simulations.	99
5.5	The $f_{\text{CGM}} - M_{200}$ relation in EAGLE, coloured by the residuals of the $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200} - M_{200}$ relation, with the median evolution of the three families of simulation through $f_{\text{CGM}} - M_{200}$ space overplotted.	101
5.6	The evolution of M_{BH} for the three families of simulations.	103
5.7	The evolution of $E_{\text{AGN}}/E_{\text{bind}}^b$ for the three families of simulations, where E_{AGN} is the integrated energy injected by AGN feedback.	105
5.8	The evolution of the central galaxy sSFR for the three families of simulations.	108
5.9	The evolution of the central galaxy κ_{co} for the three families of simulations.	109
5.10	Face-on and edge-on images of the stellar distribution in the central galaxies of haloes generated from fiducial and GM initial conditions, with contours showing the star-forming gas.	111
5.11	Present-day distributions of the cooling time, temperature and hydrogen number density in the CGM of haloes generated from fiducial and GM initial conditions.	113
A.1	Comparison of EAGLE’s CGM scaling relations with observational measurements.	131

Publications

During the course of the preparation of this thesis, the contents of Chapters 3 and 4 have been submitted and accepted for publication in a refereed journal:

- *The gas fractions of dark matter haloes hosting simulated $\sim L^*$ galaxies are governed by the feedback history of their black holes*
Davies, J. J., Crain, R. A., McCarthy, I. G., Oppenheimer, B. D., Schaye, J., Schaller, M. and McAlpine, S., 2019, MNRAS 485, 3783
- *The quenching and morphological evolution of central galaxies is facilitated by the feedback-driven expulsion of circumgalactic gas*
Davies, J. J., Crain, R. A., Oppenheimer, B. D. and Schaye, J., 2020, MNRAS 491, 4462

The contents of Chapter 5 correspond to a third study currently in an advanced state of preparation for submission to MNRAS.

Other work completed over the course of the preparation of this thesis appears in:

- *Feedback from supermassive black holes transforms centrals into passive galaxies by ejecting circumgalactic gas*
Oppenheimer, B. D., **Davies, J. J.**, Crain, R. A., Wijers, N. A., Schaye, J., Werk, J. K., Burchett, J. N., Trayford, J. W., Horton, R., 2020, MNRAS 491, 2939

Acknowledgements

First and foremost, I thank my supervisor Rob Crain. In addition to being an endless source of scientific wisdom, he has been tremendously supportive throughout the past three and a half years, offering friendly guidance and down-to-earth advice for which I am truly grateful. I also thank my second supervisor Ian McCarthy, who is always willing to drop in and offer some profound scientific insight whenever I need it. I am also indebted to my fantastic collaborators, namely Ben Oppenheimer, Joop Schaye, Andrew Pontzen, Matthieu Schaller, Stuart McAlpine, Natasha Wijers and Alex Richings; their scientific insights and advice have enriched the content of this thesis beyond measure. Particular thanks go to Ben, for always being keen to advertise my work, and for inviting me over to the USA. I also thank Richard Bower, who introduced me to the world of simulations and pointed me towards Rob when asked about a good person to work for. I am grateful to Joel Pfeffer, Adrien Thob, Dylan Nelson, Jiang-tao Li and Wenting Wang for providing useful software and datasets, and I thank Ákos Bogdán, Ralph Kraft, Peter Mitchell, Annalisa Pillepich, James Trayford and John ZuHone for useful scientific discussions.

The Astrophysics Research Institute at LJMU has been a brilliant place to work for the past few years. Everyone in the department has been very kind and helpful; my particular thanks in this regard go to Phil James, Steve Longmore, Marie Martig, Sue Percival, Ricardo Schiavon and Renske Smit. Thank you to Anna and Danielle, without whom the department would surely collapse, and to Ben Clark and Dan Harman for their technical support. The PhD student population has grown dramatically during my time here, and it has been an absolute pleasure to get to know each and every one of you. To my cohort - Dan, Egidijus, Hannah, Kirsty, Maisie, Seb and Silvia - working (and in some cases, living) with you all has been brilliant. Special thanks go to my PhD sensei Ted Mackereth, who has given me fantastic advice over the years and deserves all

the success he has found, to all the students in the HPC group, who are always a good laugh, to my dear friends Meghan Hughes and Sarah McDonald, who are always up for a chat over a beer, and to Tom Sedgwick, for being a truly good and steadfast friend that I can always count on. Thank you.

I must thank my original physics crew from my time at Durham - Charlotte, Chloe, Conor, Ginny, Jake, Luke, Robbie and Sam - without you all I don't think I'd have survived my undergraduate degree. Thanks to everyone in the Southern Spirit Singers; singing with you keeps me sane, and is one of the best things that I do. Auntie Marie, Grandma and Grandad - thank you for the love and support that you have given me for as long as I can remember.

Above all else, I thank my sister Sarah and my parents, to whom I owe everything. You gave me the resilience I needed to pull this off, and the support I needed to see it through. I would never have made it this far without you.

The funding for this PhD was provided by an STFC doctoral studentship. Both the EAGLE and IllustrisTNG teams are acknowledged for making their simulation data available. The project made use of high performance computing facilities at Liverpool John Moores University, partly funded by the Royal Society and LJMU's Faculty of Engineering and Technology, and the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure. All analyses and plots presented in this thesis were produced using `Python`, most notably its packages `numpy`, `scipy`, `matplotlib`, `h5py`, `astropy`, `mpi4py`, `py-sphviewer`, `read_eagle` and `commah`.

*When I heard the learn'd astronomer,
When the proofs, the figures, were ranged in columns before me,
When I was shown the charts and diagrams, to add, divide, and measure them,
When I sitting heard the astronomer where he lectured with much applause in the
lecture-room,
How soon unaccountable I became tired and sick,
Till rising and gliding out I wander'd off by myself,
In the mystical moist night-air, and from time to time,
Look'd up in perfect silence at the stars.*

Walt Whitman

For Jim and Peg

Chapter 1

Introduction

The Solar System that we call home is hosted by a star that is one of billions in a vast, disk-like structure seen stretching across the night sky: the Milky Way. Up until the early 20th century, the Milky Way was thought to encompass the entire Universe, with the observed “spiral nebulae” such as Andromeda being part of this one single system. Ideas to the contrary were beginning to emerge at the time of the Great Debate on the scale of the Universe between Harlow Shapley and Heber Curtis in 1920, where Curtis argued that spiral nebulae were in fact “island universes” distinct from the Milky Way. This idea was soon confirmed by Edwin Hubble’s identification of Cepheid variable stars in several nebulae, which act as “standard candles” from which the distance to the nebulae could be obtained (e.g. [Hubble, 1925, 1926a](#)). The distances to objects such as Andromeda were found to be of order 10^5 parsecs, unequivocally demonstrating that such objects are extragalactic and fundamentally transforming our view of the Universe. Hubble’s discovery led to his famous classification of these extragalactic nebulae - now simply considered “other galaxies” - into groups based on their appearance: spirals, barred spirals, ellipticals and irregular galaxies ([Hubble, 1926b](#)), which would later become known as the “Hubble tuning fork” diagram. With this discovery, a whole new field of astronomy was born.

It is difficult to overstate how dramatically our knowledge of the Universe has advanced in the nearly 100 years since this watershed moment. A combination of advances in theory, observational instrumentation and computational modelling have revealed that the Milky Way is one of billions of galaxies in an expanding Universe of baryonic matter, dark matter and dark energy. Understanding how galaxies form and evolve into the population we see today is one of the most fundamental problems in astrophysics, and moves human knowledge closer to answering its ultimate question: where did it all come from?

While galaxies appear to be “island universes” surrounded by great empty voids when seen in optical images, they do not evolve in isolation. Modern models predict that galaxies are in fact embedded in a “cosmic web” of dark matter and gas; this web can be clearly seen in the distribution of the galaxies which trace it in large-scale structure surveys such as the 2dF Galaxy Redshift Survey (2dFGRS), Sloan Digital Sky Survey (SDSS) and Galaxy and Mass Assembly (GAMA) survey. To build a self-consistent model for how galaxies form and evolve into the observed population, we must understand how they interact with this medium that surrounds them. The processes through which this interaction takes place and the properties of the material which mediates this interaction are both poorly constrained by theory and observations at the present time, and are the principal focus of this thesis. The “big questions” that I aim to address in this work are:

- How does the process of galaxy formation impact the environments of galaxies?
- How do the properties of a galaxy’s environment affect its evolution?
- How are the properties of galaxies linked to the underlying nature of the universe?

The first two of these questions might be considered two sides of the same coin, and this reflects the self-regulatory nature of galaxy formation. Understanding the nature and mechanics of this self-regulation, how it might be disrupted or altered, and how it is linked to the initial conditions of the Universe will lead us to a deeper understanding of galaxy formation in a cosmological context.

To provide a foundation for the work presented in this thesis, I devote the remainder of this Chapter to a general introduction to the topic. In Section 1.1 I will introduce the current paradigm for our understanding of cosmology, and explain the successes it has brought. In Section 1.2 I will briefly describe our current theoretical picture for galaxy formation, and explain the importance of feedback processes in regulating this formation. In Section 1.3 I will motivate the importance of understanding the nature of the gas around galaxies - the “circumgalactic medium” - in answering the above questions, and detail our current knowledge of this elusive component of the Universe. Finally, at the end of the Chapter I will give an outline for the content of the rest of the thesis.

1.1 The Λ CDM paradigm

The currently-preferred “standard model” for cosmology, or the framework in which galaxy formation takes place, is the Λ CDM model. In this paradigm, the Universe is

geometrically flat (i.e. it satisfies Euclidean geometry) and expanding, and is primarily made up of baryonic matter, cold dark matter (CDM) and vacuum energy (represented by the analogous constant Λ in Einstein's field equations) which causes the Universe to accelerate in its expansion at late times. In this section, I will briefly explain how this consensus was reached.

On large scales, the properties of the Universe are dominated by the force of gravity, for which the most complete theory we have is Einstein's general relativity. A significant challenge to this theory arose with Hubble's discovery that there are other galaxies besides the Milky Way in the universe, and that they are receding from the Milky Way at a rate proportional to their distance from us (the [Hubble \(1929\)](#) law, where the constant of proportionality is the Hubble constant H_0). This discovery demonstrated that the Universe is expanding, challenging Einstein's steady-state and finite solution for general relativity in a universe of matter. [Friedmann \(1922\)](#) and [Lemaître \(1927\)](#) had previously (and independently) proposed solutions to Einstein's field equations which allowed for an expanding Universe before this discrepancy had arisen, and these solutions led to the Friedmann-Lemaître-Robertson-Walker (FLRW) metric for a homogeneous, isotropic, non-static Universe. In this metric, Einstein's field equations are exactly solved by the Friedmann equations, a set of differential equations which describe the expansion of spacetime as a function of energy density and spatial curvature.

A natural consequence of an expanding universe is that it has an origin ([Lemaître, 1931](#)), leading to the notion of a “Big Bang”: the expansion of space from a singularity. This idea was corroborated by the discovery of the cosmic microwave background (CMB, [Penzias & Wilson, 1965](#)), emitted at the epoch of recombination where the Universe became transparent to radiation. The CMB is broadly isotropic, but exhibits small-scale anisotropies that reflect primordial density fluctuations in the very early Universe. While the detection of the CMB effectively confirmed the Big Bang theory, two significant problems remained unsolved:

- As a result of the finite speed of light through spacetime, distant regions of the Universe should exhibit strong differences in their densities because they are not causally connected. Despite this, the early Universe seen in the CMB appears broadly homogeneous and isotropic - this is the “horizon problem”.
- The universe is observed to be spatially flat, which requires a very precise fine-tuning of the curvature at very early times. Any minor curvature (positive or negative) would be dramatically amplified as the Universe expands, yet no such curvature is observed - this is the “flatness problem”.

The theory of inflation ([Guth, 1981](#); [Linde, 1982](#)) resolves these issues by invoking a very short period of exponential expansion. This solves the above problems by assuming that the pre-inflation Universe is sufficiently compact that all scales are causally connected, giving homogeneity to the post-inflation Universe, and that any pre-inflation curvature is flattened by the exponential expansion. A consequence of this process is that the quantum fluctuations present prior to inflation are greatly amplified by the expansion into the density fluctuations seen as temperature anisotropies in the CMB, and which seed the large-scale structure of the Universe.

The central idea in the Λ CDM model is that the structures observed in the Universe - from dwarf galaxies to galaxy clusters - grew via gravitational collapse from these initial overdensities, with dynamics dominated by a hypothesised form of matter that interacts only via gravity. This existence of this “dark” matter can be inferred from several phenomena which suggest the presence of non-luminous, “missing” mass: the velocities of galaxies in clusters ([Zwicky, 1933, 1937](#)), the flat rotation curves of galaxies ([Rubin & Ford, 1970](#); [Rubin et al., 1980](#)) and the strong gravitational lensing observed around galaxy clusters ([Walsh et al., 1979](#)). Various non-luminous baryonic candidates for this missing mass have been proposed in the past, such as black holes, neutron stars, brown dwarfs and faint white dwarfs (collectively known as massive compact halo objects, or MACHOS). Observations of the CMB, the theory of Big Bang Nucleosynthesis and the lack of gravitational microlensing detections in the Milky Way collectively rule out these candidates, and suggest instead that dark matter is non-baryonic in nature, existing in the form of hypothesised particles such as axions, sterile neutrinos or weakly-interacting massive particles (WIMPs).

Evidence for the existence of dark matter led to the first models of structure formation in the Universe, such as those of [Press & Schechter \(1974\)](#), [White & Rees \(1978\)](#), and [Fall & Efstathiou \(1980\)](#), which led to the hypothesis that dark matter must be “cold” (mass $> 1 \text{ keV}$) rather than “hot” (mass $< 50 \text{ eV}$) ([Blumenthal et al., 1982](#); [Peebles, 1982](#); [White et al., 1983](#)) to produce the observed clustering of galaxies. In these models, dark matter undergoes gravitational collapse into primordial overdensities prior to recombination, while baryons and radiation remain coupled together. Following recombination, the resulting neutral gas falls into the gravitational potentials created by the dark matter, eventually cooling and condensing into the first stars and protogalaxies. These small systems do not evolve in isolation, but merge together to form progressively larger structures; galaxy formation is therefore “hierarchical”.

The above models assume that the Universe is dominated entirely by matter (CDM and baryons). By the 1990s, observations of large-scale structure (e.g. [Efstathiou et al.,](#)

1990) and the discovery of temperature anisotropies in the CMB by the Cosmic Microwave Background Explorer (Smoot et al., 1992; Wright et al., 1992) indicated that this is not the case, and that the energy density of the Universe is in fact dominated by vacuum energy, which contributes negative pressure to cosmological expansion and can be represented by a cosmological constant, Λ . Observations of type Ia supernovae (which act as standard candles for distance measurements due to their fixed absolute magnitude) by Riess et al. (1998) and Perlmutter et al. (1999) indicated that the Universe is accelerating in its expansion as a result of this negative pressure. These results established the hierarchical formation of structure with cold dark matter and dark energy, the Λ CDM model, as the leading theory in cosmology. The model exhibits a remarkable level of agreement with more modern measurements of the cosmic microwave background by the Wilkinson Microwave Anisotropy Probe (WMAP Spergel et al., 2007) and the Planck satellite (Planck Collaboration et al., 2018); fitting the parameters of the model to features in the CMB such as the power spectrum of the temperature fluctuations has constrained quantities such as the baryon, dark matter and dark energy content of the Universe with sub-percent precision.

1.2 Galaxy formation in dark matter haloes

The Λ CDM cosmology describes the expansion history of the Universe, and now gives a precise accounting of its contents: the total energy density is $\approx 70\%$ dark (vacuum) energy, $\approx 25\%$ dark matter, and $\approx 5\%$ baryonic matter, with negligible contributions from neutrinos and radiation. The field of galaxy formation is concerned with how the dark matter and baryons co-evolve from the epoch of recombination to form the wide variety of galaxies observed in the Universe today. In this Section I will briefly cover the most important processes that govern the formation of galaxies and the regulation of their evolution.

1.2.1 Assembly of stellar mass through gas accretion

The earliest galaxy formation models established that galaxies cannot be formed through gravitational processes alone, but require dissipative processes such as the radiative cooling of gas (Binney, 1977; Rees & Ostriker, 1977; Silk, 1977). In these models, primordial, diffuse, ionised gas clouds collapse under gravity and are heated by shocks to form quasi-hydrostatic coronae of hot material, which can then radiatively cool, condense and fragment to form stars in the centre. The large-scale distribution of galaxies is, however, set by purely gravitational clustering in a hierarchical fashion, such that small systems combine to make larger ones and have their small-scale structure erased as the

new system relaxes (Press & Schechter, 1974). White & Rees (1978) combined these ideas for systems containing 80% dark matter and 20% gas, showing that while the purely gravitational dark matter virialises, gas cooling through radiative processes such as thermal bremsstrahlung, recombination and line emission can fuel the formation of stars in the centre of the ‘halo’, producing an approximate match to the observed galaxy luminosity function. In subsequent years this picture was shown to account for other galaxy properties such as the kinematics of galaxies (Fall & Efstathiou, 1980), and was expanded upon by White & Frenk (1991) to predict the abundance and structure of dark matter haloes, the star formation rates at their centres, and the X-ray luminosities and metallicities of the included gas, as a function of mass and time. With this unification of gravitational and dissipative processes for systems of dark matter and gas, modern galaxy formation theory was born.

The modern perspective for how galaxies accrete gas from the intergalactic medium (IGM) for star formation considers two modes: the above “hot mode”, where gas is heated to high temperatures by gravitational shocks and then cools from hydrostatic equilibrium, and the “cold mode”, where gas reaches the galaxy through filaments of cold gas. The cold mode is important for accretion in all systems and at all redshifts, while the hot mode becomes more important in high-mass systems and at late times (Birnboim & Dekel, 2003; Kereš et al., 2005; van de Voort et al., 2011). Each of these modes provides galaxies with their interstellar medium (ISM), which constitutes their fuel supply for producing stars. When clouds of gas in the ISM reach a threshold mass such that their internal pressure cannot stabilise them against gravity, they collapse and form stars in Jeans (1902) instabilities. Parts of such clouds may also fragment and form stars if they locally become Jeans-unstable. The rate at which this process occurs (the star formation rate, SFR) is observed to be related to the gas surface density in the ISM, giving rise to the empirical Kennicutt-Schmidt relation (Kennicutt, 1998). On far larger scales, the cosmic star formation rate density describes the star formation activity of the whole galaxy population, and this is observed to rise with decreasing redshift until $z \sim 2$, after which it begins to fall (e.g. Madau et al., 1998; Bouwens et al., 2015).

Through this process, stars of a variety of masses are formed, the distribution of which is described by a stellar initial mass function (Salpeter, 1955; Chabrier, 2003). A given stellar population will lose a significant fraction of its initial mass over its lifetime, ejecting material into the surrounding gas and enriching it with metals (elements heavier than helium). Massive stars ($M_\star > 8 M_\odot$) evolve rapidly and lose mass in both stellar winds and core-collapse supernovae, which occur at the end of their lifetimes and enrich the surrounding gas with α -elements. Lower-mass stars are much longer-lived and drive strong stellar winds when they reach the asymptotic giant branch (AGB) stage of their evolution. Stars of this mass do not undergo core-collapse supernovae, but the

stellar cores remaining after the AGB stage can explode as Type Ia supernovae following accretion of material from a companion star, and this process provides significant iron enrichment to the ISM.

Galaxies do not form stars indefinitely, and can be “quenched” during their evolution. Galaxies that are actively forming stars appear blue, because they continuously form short-lived and massive blue stars, while quenched galaxies appear “red and dead” as their massive stars have long-since evolved to the supernova phase. This difference manifests itself as a bimodality in the galaxy population: surveys reveal that most galaxies are members of the “blue cloud” of spiral-like, star forming galaxies, or of the “red sequence” of elliptical, quenched galaxies on a plot of $u - r$ colour as a function of stellar mass (Baldry et al., 2004, 2012; Schawinski et al., 2014), with a small minority undergoing a transition between the two in the “green valley”. Galaxies quench when they run out of fuel for star formation, and this can occur for a number of reasons. The fuel supply could have been entirely converted into stars, or could have been stripped away by ram pressure as the galaxy passes through the gaseous halo of a more massive galaxy, group of galaxies, or cluster. The most relevant quenching mechanism for regulating the evolution of galaxies, however, is feedback, which can eject gas from the ISM into the galaxy’s surroundings.

1.2.2 Feedback as a regulatory mechanism

There are several characteristics of the observed galaxy population that must be reproduced by a successful galaxy formation model. Two of these observed characteristics are arguably the most fundamental:

- **The distribution of stellar masses** - this is quantified by the observed galaxy stellar mass function (GSMF, shown in Figure 1.1), which quantifies the comoving number density of galaxies at a given stellar mass.
- **The efficiency of galaxy formation** - this is quantified by the inferred stellar mass - halo mass (SMHM) relation, shown in Figure 1.2. It is not sufficient to simply form the correct number of galaxies of a given stellar mass; they must live in the right haloes.

The shape of the GSMF reflects the hierarchical, “bottom-up” nature of galaxy formation: the universe contains very many low-mass galaxies, and very few high-mass galaxies, with a characteristic “knee” at the Milky Way (L^*) mass scale. The GSMF also reveals information about the efficiency of galaxy formation; the observed relation

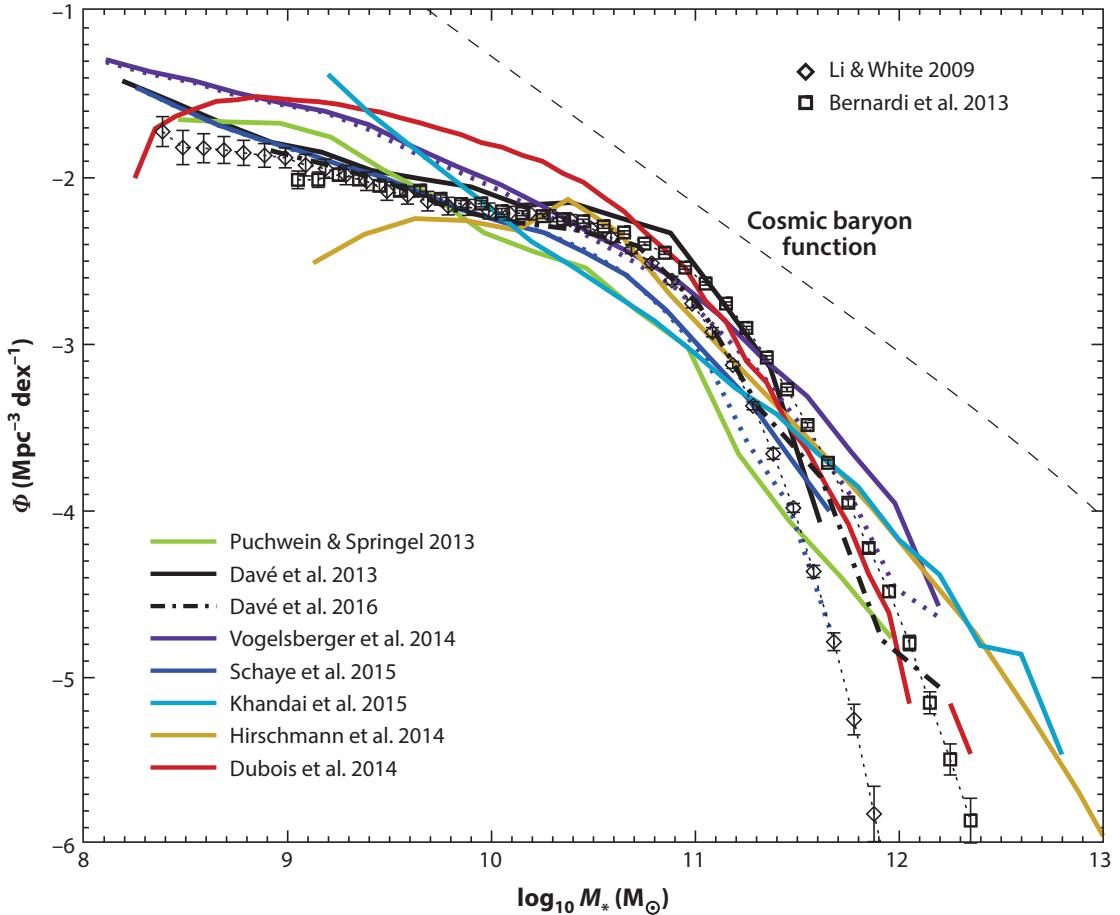


FIGURE 1.1: The observed galaxy stellar mass functions of Li & White (2009) and Bernardi et al. (2013), with the results of several cosmological, hydrodynamical simulations overlaid. The dashed line shows a hypothetical GSMF assuming the cosmic baryon fraction (i.e. perfectly efficient galaxy formation). Figure reproduced from Naab & Ostriker (2017).

lies below the prediction from a purely collisionless case assuming the stellar mass density is equal to the cosmic baryon budget (i.e. perfectly efficient galaxy formation). This deficit is the smallest for $\sim L^*$ galaxies at the knee of the GSMF, which dominate the present-day stellar mass density of the Universe, and is far greater at sub- and super- L^* masses. The inefficiency of galaxy formation is also clearly shown in the SMHM relation; the assembly of stellar mass is most efficient in haloes of mass $\sim 10^{12} M_\odot$, and is far less efficient at lower and higher masses. These observations suggest that the process of galaxy formation requires some source of feedback energy to reduce its efficiency in low- and high-mass haloes.

The necessity of feedback as a mechanism for regulating galaxy growth was recognised in some of the earliest galaxy formation models. White & Rees (1978) noted that some kind of feedback energy would need to be injected to prevent runaway cooling of gas into very dense regions, and early simulation work confirmed this “overcooling problem”:

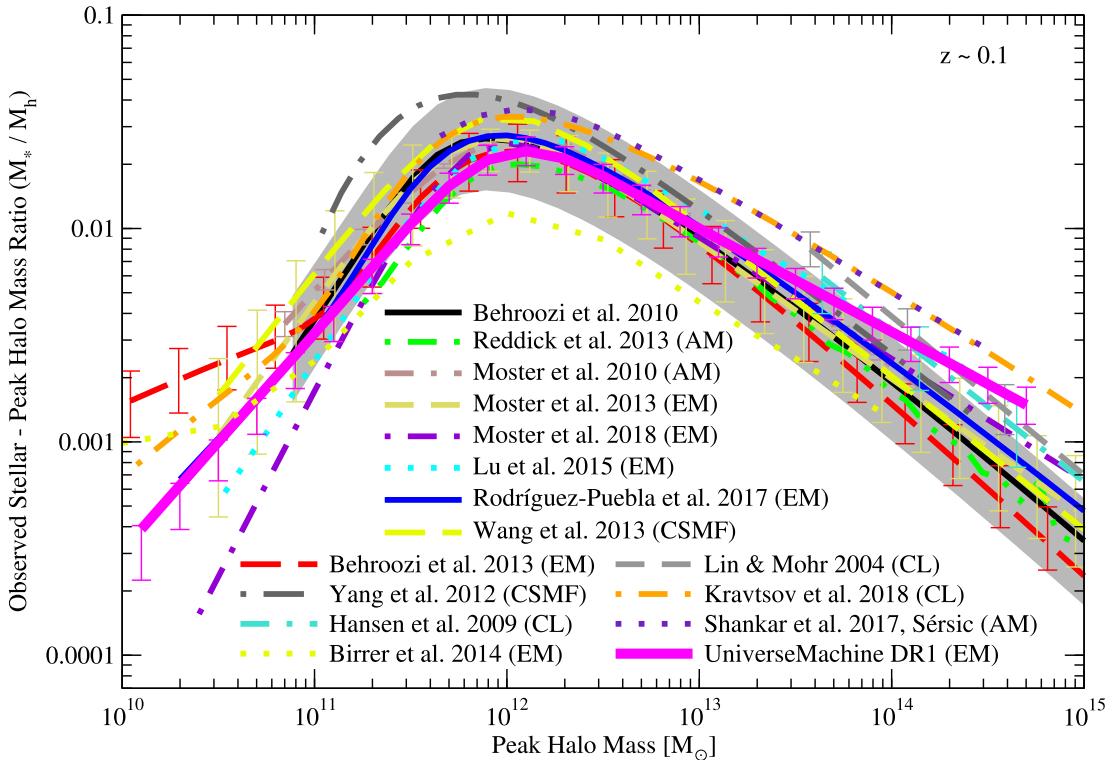


FIGURE 1.2: The present-day stellar mass - halo mass relation. Figure reproduced from Behroozi et al. (2019), who compiled several results obtained using abundance matching (AM), empirical modelling (EM), conditional stellar mass function (CSMF) modelling, and from cluster X-ray mass measurements (CL).

the stellar discs at the centres of haloes were too massive and too compact compared to observations (Katz & Gunn, 1991; Navarro & Benz, 1991; White & Frenk, 1991; Navarro & White, 1994; Balogh et al., 2001), and this problem becomes worse as the resolution of the simulation increases. The inclusion of feedback associated with star formation (stellar winds and supernovae) was soon shown to partly alleviate this overcooling; as the star formation rate increases, so does the energy injection rate and the ejection of fuel for star formation, thus enabling the self-regulation of the system. This mechanism is motivated by observations of outflows moving at $> 100 \text{ km s}^{-1}$ away from star-forming galaxies (Veilleux et al., 2005; Steidel et al., 2010; Martin et al., 2012; Rubin et al., 2014; Nielsen et al., 2020), with inferred mass loadings (mass outflow rate relative to the star formation rate) greater than unity (Bland-Hawthorn et al., 2007; Schroetter et al., 2015). This source of feedback is now a vital component of both semi-analytical models of galaxy formation (e.g. Cole et al., 2000) and cosmological simulations (e.g. Springel & Hernquist, 2003; Oppenheimer et al., 2010; Schaye et al., 2015; Pillepich et al., 2018a) for preventing the overproduction of galaxies of Milky Way-mass and below (hence fitting the low end of the GSMF), for fitting the SMHM relation below its peak (Mitchell et al., 2019), and for producing spiral galaxies with the realistic morphologies and sizes (e.g.

Governato et al., 2007; Guedes et al., 2011; Hopkins, 2013; Marinacci et al., 2014; Agertz & Kravtsov, 2016).

While feedback from star formation is able to reconcile the low-mass end of the GSMF with observational results, it is unable to regulate the growth of more massive galaxies (e.g. Crain et al., 2009). In the hierarchical, bottom-up picture, one would expect the most massive systems to form most of their stellar content at late times, yet observations indicate that most of their mass was assembled prior to $z = 1$. This is the phenomenon of “downsizing”, where the most massive systems form their mass at earlier times, contrary to expectations from a hierarchical model (Cowie et al., 1996; Neistein et al., 2006). For a model to reproduce this, massive galaxies must be quenched at early times, and early semi-analytical models indicated that feedback from star formation alone cannot achieve this; ejected material will simply re-cool in the deeper potentials of such systems (Benson et al., 2003). A simulation incorporating only stellar feedback will therefore not reproduce the sharp break seen in the GSMF at high stellar mass - another source of feedback is required.

Supermassive black holes (SMBHs) are now thought to be ubiquitous at the centres of galaxies (e.g. Genzel et al., 1997; Kormendy & Ho, 2013), and exhibit a conspicuously consistent mass ratio with the central stellar bulge (Magorrian et al., 1998; Häring & Rix, 2004) given the remarkable disparity in physical size (~ 9 orders of magnitude), suggesting a possible co-evolution of the SMBH and the host galaxy (though see Jahnke & Macciò, 2011, for an alternative explanation). The rest mass energy liberated as black holes (BHs) grow through accretion is capable of launching outflows (e.g. Silk & Rees, 1998a), which are observed at both high redshift and in the local Universe (e.g. Rupke & Veilleux, 2011; Maiolino et al., 2012; Harrison et al., 2014; Ciccone et al., 2015, 2016). This energy injection channel is termed active galactic nucleus (AGN) feedback, and it becomes important for regulating galaxy growth when feedback from star formation becomes inefficient and gas builds up in the halo centre, fuelling the rapid growth of the BH. Modern models predict that this process is triggered once the halo reaches a critical halo mass (Bower et al., 2006; Croton et al., 2006) - and therefore critical virial temperature - at which star formation-driven outflows are no longer buoyant (Bower et al., 2017; McAlpine et al., 2018). Simulations of AGN feedback indicate that it has a significant influence on the structure and star formation activity of massive galaxies (e.g. Springel et al., 2005; Hopkins et al., 2005; Sijacki et al., 2007; Booth & Schaye, 2009; Johansson et al., 2009; Dubois et al., 2013), leading to a quenching of star formation at early times and the process of downsizing. The inclusion of AGN feedback in semi-analytical models allowed them to match the GSMF at the high-mass end for the first time (Croton et al., 2006; Bower et al., 2006), and it continues to be a vital part of the models employed in cosmological simulations for this purpose.

Implementation of schemes for feedback from both star formation and AGN in simulations is challenging, as it is difficult to prevent numerical losses when injecting energy, and the resulting effect on the galaxy is strongly dependent on the chosen implementation (a limitation that was identified in early implementations of supernova feedback, e.g. [Navarro & White, 1993](#); [Okamoto et al., 2005](#)). I will return to this issue in Chapter 2, and outline how it is overcome in modern, state-of-the-art simulations.

1.3 The circumgalactic medium

While our understanding of the processes that govern galaxy formation has dramatically advanced in recent decades thanks to advances in observational techniques and sophisticated modelling, several key problems remain. One of the most vital descriptors of a galaxy is its ability to form stars, and two questions remain unanswered in this regard: how do galaxies maintain their star formation rates over long periods when the depletion time of their ISM is much shorter than the time for which they have been star-forming ([Tumlinson, Peeples & Werk, 2017](#)), and how is star formation quenched permanently? The former question implies some form of external inflowing fuel source which replenishes the ISM, and this has implications for the latter - if a galaxy is quenched by depletion of the ISM through either consumption, stripping or ejection, why then does this external fuel source not subsequently reignite star formation?

Another key problem lies in the inefficiency of star formation in galaxies. The baryon content of the Universe predicted by Big Bang Nucleosynthesis is constrained by observations of the CMB with sub-percent precision ([Planck Collaboration et al., 2018](#)), and yet galaxy surveys indicate that a mere 5% of this “budget” exists in the form of stars and stellar remnants (see e.g. [Balogh et al., 2001](#); [Li & White, 2009](#); [Baldry et al., 2012](#)). Even at their most efficient (L^* galaxies), only $\sim 20\%$ of the baryon content of dark matter haloes is typically converted into stars by the present day ([Behroozi et al., 2010](#)). The remainder of these baryons are not readily seen in observations, leading to the notion of a “missing baryons problem”. Another issue tied into this is the metal deficiency of star-forming galaxies, since only 20 – 25% of the metal mass believed to be produced by their stars can be accounted for in their stars and ISM ([Peeples et al., 2014](#)); this could be termed the “missing metals problem”.

As I briefly touched on at the start of this Chapter, the solutions to these problems require an understanding of the regulation of galaxy evolution through inflows, star formation, and feedback-driven outflows. These processes take place through an interface between galaxies and the intergalactic medium (IGM), termed the “circumgalactic medium” (CGM), which constitutes the gas in the outskirts of the galaxy and further

out in the dark matter halo. These gaseous reservoirs within dark matter haloes have long been a prediction of analytical models (e.g. [White & Rees, 1978](#); [White & Frenk, 1991](#); [Fukugita & Peebles, 2006](#)) and were invoked to explain the presence of absorption systems around the Milky Way as early as the 1950s ([Spitzer, 1956](#)) and around other galaxies in the 1980s ([Bergeron & Stasińska, 1986](#)). The CGM represents not only a putative fuel supply for replenishing the ISM of galaxies, but also a potential reservoir for the metals produced by star formation and a location for the recycling of outflowing material back into inflows (e.g. [Oppenheimer & Davé, 2008](#)). An understanding of the origin, composition, properties and evolution of the CGM is therefore vital to any theory of galaxy formation.

The most basic characterisation of the CGM is its mass fraction, f_{CGM} ; the mass content in gas relative to the mass of the underlying dark matter halo. One might expect that the majority of the galactic missing baryons lie in the CGM in the form of diffuse gas (e.g. [Crain et al., 2007](#)), such that f_{CGM} is approximately equal to the cosmic average, $\Omega_b/\Omega_0 \simeq 0.15$. Detailed observational measurements of both the stellar and (hot) gas phases exist for nearby galaxy groups (with halo masses of $M_{500} \gtrsim 10^{13} M_\odot$, where M_{500} is the mass within a sphere enclosing a mean density of 500 times the critical density) and clusters, and they indicate that the most massive bound systems ($M_{500} \sim 10^{15} M_\odot$) are indeed effectively ‘baryonically closed’ (e.g. [Allen et al., 2002](#); [Lin et al., 2004](#); [Gonzalez et al., 2013](#)). However, less massive galaxy groups exhibit significantly lower f_{CGM} (e.g. [Vikhlinin et al., 2006](#); [Pratt et al., 2009](#); [Sun et al., 2009](#); [Lin et al., 2012](#); [Lovisari et al., 2015](#)), and estimates of the mass of circumgalactic gas around Milky Way-mass galaxies based on absorption features in quasar sightlines or X-ray emission fall significantly short of the cosmic fraction (e.g. [Bregman, 2007](#); [Shull et al., 2012](#); [Werk et al., 2013](#)).

There are several possible explanations for the observed deficiency of baryons in lower mass haloes. The baryons could in fact reside in the CGM, but remain undetected by observations; given the difficulties inherent in detecting circumgalactic gas in haloes below the mass scale of galaxy groups (see Section 1.3.1) this is not an altogether remote possibility, though future advances in the sensitivity of space-borne X-ray telescopes promise to confirm or rule out this scenario. Alternatively, the baryons could have been accreted onto the halo and subsequently ejected beyond the virial radius by feedback, or been prevented from accreting onto the halo in the first place through a heating and pressurisation of the existing CGM.

The primary focus of this thesis is on the second of these possibilities, and what the resulting effects are on the formation of galaxies. This choice of scenario can be motivated from early analytical models and more modern simulation results: the rest-mass energy

required to grow central BHs (e.g. [Soltan, 1982](#)) typically exceeds the binding energy of their host galaxies by large factors, and may even exceed the binding energy of all baryons bound to their host dark matter haloes (e.g. [Silk & Rees, 1998b](#); [Booth & Schaye, 2010, 2011](#); [Oppenheimer, 2018](#)). The feedback associated with BH growth is therefore likely more than capable of ejecting a significant fraction of the CGM baryons.

In addition to its effects on galaxies themselves, feedback from accreting BHs is also invoked as a means of inducing the observed deviations from self-similarity in the radial profiles of the thermodynamic properties of circumgalactic and intragroup gas (e.g. [Si-jacki et al., 2007](#); [McCarthy et al., 2011](#); [Stott et al., 2012](#); [Planck et al., 2014](#); [Barai et al., 2016](#)), and it has become clear that there is an intimate connection between the regulation and quenching of star formation in massive galaxies, and the properties of the gas associated with their dark matter haloes (e.g. [Bower et al., 2008](#); [Stott et al., 2012](#); [Bower et al., 2017](#); [McAlpine et al., 2018](#); [McDonald et al., 2018](#)). A successful model of galaxy formation and evolution must therefore reproduce simultaneously the evolution of the stellar and gaseous matter bound to dark matter haloes, and the ability of BH-driven feedback to eject material from the CGM and transform its physical properties is likely vital to this process.

1.3.1 Observational study of the CGM

The principal difficulty in characterising the properties of the CGM, how it is affected by feedback processes, and how it shapes the process of galaxy evolution lies in how challenging it is to observe. Hydrodynamical simulations predict, and observations reveal, that diffuse gas extends beyond the ISM to the virial radius and beyond and is thermodynamically multiphase, consisting of hot, volume filling, diffuse gas ($T > 10^6$ K) potentially traced by X-ray emission and the thermal Sunyaev-Zel'dovich effect, warm-hot gas ($T \sim 10^{5-6}$ K) observable through quasar absorption spectroscopy in the UV band with high ions such as CIV, OVI and NeVII, cool gas clouds ($T \sim 10^{4-5}$ K) traced by low ions such as CII/III, SiII/III and MgII, clumps of cold neutral gas ($T < 10^4$ K) detectable in 21cm HI emission and molecular gas which has been detected in outflows through sub-mm observations with ALMA.

Assessing the mass content of the CGM contributed by each of these gas phases, along with characterising their physical properties and distributions within circumgalactic haloes, is the primary goal of current observational CGM studies. Building an holistic picture of the CGM around Milky Way-like galaxies from these observations is very difficult, because each of the above techniques presents its own unique challenges and difficulties in interpretation. Here I will discuss the techniques which are most commonly

used and the most relevant to this thesis, in terms of their limitations, existing results from each, and how advances in instrumentation will increase their power to constrain future galaxy formation models.

1.3.1.1 X-ray emission

Gas cooling via thermal bremsstrahlung and metal line cooling gives rise to X-ray emission in galactic haloes. In galaxy groups and clusters with high virial temperatures and near-closed baryon fractions, this emission is eminently detectable, and the total luminosity and characteristic temperature obtained are commonly used, typically under the assumption that the intragroup/intracluster medium is in hydrostatic equilibrium, to obtain the gas and halo masses of the system. At the virial temperatures of haloes hosting L^* galaxies ($T_{\text{vir}} \sim 10^6$ K), soft X-ray emission from the halo should, in the absence of feedback, be detectable to currently available X-ray telescopes (e.g. [Benson et al., 2000](#)). As noted previously, however, the content and structure of the CGM in galaxy-scale haloes are likely to be significantly affected by feedback, resulting in low densities that yield soft X-ray fluxes that are in general too faint for detection with current instrumentation.

The X-ray observatories *XMM-Newton* and *Chandra* lack the sensitivity required for detailed study of the CGM with reasonable integration times, and at present only a handful of convincing detections of extra-planar emission from very massive galaxies have been made with these telescopes (e.g. [Anderson & Bregman, 2011](#); [Dai et al., 2012](#); [Bogdan et al., 2013b, 2017](#); [Li et al., 2016, 2017](#); [Lakhchaura et al., 2019](#)). In addition to very long integration times, these observations require the complex removal of emission from point sources, X-ray binaries, cataclysmic variables, the “hot bubble” of the Milky Way halo and distant AGN. These detections are also typically limited to the inner ~ 50 kpc of the halo, where the strong dependence of X-ray emissivity on density ($L_X \propto n^2$) and metallicity mean that the emission is likely to be dominated by ejecta from supernovae and not representative of the diffuse halo ([Crain et al., 2010, 2013](#); [Li et al., 2014](#)). Examination of the hot component of the CGM of $\sim L^*$ galaxies is a leading motivation for forthcoming and proposed X-ray observatories such as *Athena* ([Barret et al., 2016](#)) and particularly *Lynx* ([Ozel, 2018](#)), which should possess the sensitivity to make such detections commonplace and facilitate detailed study of the hot-phase CGM.

The low surface brightness limitations inherent in targeted observations motivate the use of stacking low-resolution X-ray survey data about the co-ordinates of optically-selected galaxies. Detection of the hot CGM of L^* galaxies does, however, remain beyond the reach of this technique; by stacking *ROSAT* All-Sky Survey maps about the co-ordinates

of “locally brightest galaxies” (LBGs) selected from Data Release 7 of the Sloan Digital Sky Survey (SDSS-DR7 [Abazajian et al., 2009](#)), [Anderson et al. \(2015\)](#) and [Wang et al. \(2016\)](#) were only able to obtain convincing detections for supra- L^* galaxies. The low resolution of *ROSAT* also precludes the removal of X-ray bright point sources from the galactic disc, so any analysis with this technique must be restricted to the outer halo. However, the entire X-ray sky will soon be mapped at $30\times$ greater sensitivity and higher spatial resolution than *ROSAT* by the *eROSITA* instrument on the recently launched Spectrum-Roentgen-Gamma mission ([Merloni et al., 2012](#)), and these observations may soon afford a means of examining the hot CGM of $\sim L^*$ galaxies in a statistical sense.

1.3.1.2 Thermal Sunyaev-Zel'dovich effect

The thermal Sunyaev-Zel'dovich (tSZ) effect is the inverse Compton scattering of CMB photons by high energy electrons. When observing the CMB through a hot halo, the effect is observed as a frequency-dependent change in the CMB temperature that quantifies the integrated electron pressure along the line of sight. When combined with a model for the electron pressure profile, typically obtained through a combination of cluster X-ray observations and hydrodynamical simulations (e.g. [Arnaud et al., 2010](#); [Battaglia et al., 2012](#)), the effect enables the estimation of the mass and thermodynamic properties of the hot gas. In contrast with X-ray emission, the strength of the tSZ effect is linear with density and independent of metallicity, meaning that it will not be strongly weighted towards feedback-driven outflows and will give a less biased view of the hot-phase CGM.

High signal-to-noise tSZ studies of individual objects are beyond the capabilities of current instrumentation, so stacking methods must be employed to characterise the average properties of statistical samples. This technique has been used to detect intracluster gas in superclusters ([Tanimura et al., 2019b](#)) and to probe the contribution of gaseous filaments between galaxies to the Universe’s missing baryons ([Tanimura et al., 2019a](#); [de Graaff et al., 2019](#)), but most relevant to the study of the CGM is the work of the [Planck Collaboration et al. \(2013\)](#) and [Greco et al. \(2015\)](#). In a method analogous to the X-ray studies of [Anderson et al. \(2015\)](#), they stacked tSZ maps about the co-ordinates of LBGs from SDSS-DR7 and obtained convincing detections for the CGM of galaxies down to a stellar mass of $M_\star \sim 10^{11.1-11.3}$. Studying the CGM of L^* galaxies (which lie below this stellar mass) with this technique is rendered impossible by the Planck satellite’s $\simeq 10$ arcmin beam, which corresponds to scales significantly larger than the virial radius of nearby $\sim L^*$ galaxies, and as such this approach awaits the next generation of ground-based high-resolution CMB experiments such as CMB-S4 ([Abazajian et al., 2016](#)) and the Simons Observatory ([Ade et al., 2019](#)).

1.3.1.3 Absorption features in background quasar spectra

Since the study of circumgalactic gas around Milky Way-like galaxies is currently beyond the capabilities of the above techniques, our present picture of these systems is based primarily on the observation and interpretation of absorption systems seen in the light of distant quasars. Circumgalactic gas which lies along the line of sight to a bright background quasar produces absorption features in the observed quasar spectrum according to its chemical composition, and the strength of these lines (quantified by their equivalent width) can be used to compute the integrated column density of the corresponding ion in the system. This type of study is commonly undertaken in the UV and optical regimes with the Cosmic Origins Spectrograph aboard the Hubble Space Telescope, tracing the warm and cool phases of the CGM, though the technique is also used with X-ray spectra from *Chandra* or *XMM-Newton* to study high ions tracing the hot CGM.

Studies using this technique have produced a wide variety of constraints on the properties of the Universe's missing baryons. The hot halo of the Milky Way can be traced in sightlines to any distant X-ray luminous quasar, and studies of OVII and OVIII absorption have yielded constraints for its baryon content and structure ([Anderson & Bregman, 2010](#); [Miller & Bregman, 2013, 2015b](#)), and evidence for its rotation ([Hodges-Kluck et al., 2016](#)). On far larger scales, X-ray absorption lines have also been used to make tentative detections of baryons in the warm-hot intergalactic medium (WHIM) ([Nicastro et al., 2018](#); [Nicastro, 2018](#)).

Finding sightlines through the CGM of other galaxies is more challenging, and quasar sources are typically selected such that they probe a specific type of foreground system at a small impact factor, such as L^* galaxies (in e.g. COS-Halos, [Tumlinson et al. 2013](#), MAGIICAT, [Nielsen et al. 2013](#), and the Keck Baryonic Structure Survey, [Turner et al., 2014](#)), dwarf galaxies (COS-Dwarfs [Bordoloi et al., 2014](#)), galaxies with well-characterised ISM properties (COS-GASS [Borthakur et al., 2015](#)) and galaxies hosting AGN (COS-AGN, [Berg et al., 2018](#)). These studies have identified large circumgalactic reservoirs of neutral hydrogen (HI) around L^* galaxies ([Tumlinson et al., 2013](#)) and led to a characterisation of the surface density profile of such reservoirs and the contribution they could make to the galaxy's missing baryons ([Prochaska et al., 2017](#)). UV absorption surveys have shown that the CGM is a significant reservoir of metals tracing various gas phases (e.g. [Stocke et al., 2013](#); [Werk et al., 2013](#); [Lehner et al., 2014](#); [Turner et al., 2015](#)), leading to a characterisation of its physical/ionisation conditions ([Werk et al., 2014, 2016](#)) and kinematics ([Nielsen et al., 2017](#)), and yielding a census of the metals contained within the ISM and CGM ([Peeples et al., 2014](#)). Detected absorption systems suggest an intimate connection between the properties of the CGM and the central

galaxy: significant OVI absorption is found around actively star-forming galaxies, but is absent around passive galaxies of a similar stellar mass (Tumlinson et al., 2011), and the gas content of the ISM correlates with the strength of Lyman α absorption in the CGM (Borthakur et al., 2015). In the future, multi-object UV spectrographs on observatories such as the conceptual LUVOIR telescope will possess a higher resolution than Hubble’s COS and will benefit from larger telescope effective areas, facilitating the study of far larger samples of absorbers with this technique.

While detections of the CGM around L^* galaxies are eminently more feasible with this technique than with the others already discussed, translating the observed column densities and covering fractions of various ion species into the physical properties of the CGM presents serious challenges. One cannot ‘image’ individual extragalactic systems (though some galaxies can be probed with multiple background sources, see e.g. Bechtold et al., 1994; Dinshaw et al., 1995; Hennawi et al., 2006; Crighton et al., 2010; Lopez et al., 2018), meaning that radial trends must be inferred from samples of absorbers with diverse impact factors (e.g. Stocke et al., 2013; Tumlinson et al., 2013; Turner et al., 2014; Borthakur et al., 2015; Burchett et al., 2016; Bielby et al., 2019). The conversion from observables to physical conditions requires many assumptions, particularly in relation to the elemental abundances of, and ionisation conditions local to, the absorbing gas. Many of the ions most readily observed in the CGM are influenced by both collisional and radiative processes (e.g. Wiersma et al., 2009a) and can exhibit significant departures from ionisation equilibrium (e.g. Gnat & Sternberg, 2007; Oppenheimer & Schaye, 2013a,b; Segers et al., 2017; Oppenheimer et al., 2018). Interpretation of these observations is therefore challenging, and relies on sophisticated models, which I will discuss in the next section.

1.3.2 The CGM in cosmological simulations

The strong, non-linear coupling between the processes of star formation, heavy element synthesis, heating, radiative cooling and gas dynamics that are involved in the physical state of the CGM necessitates the use of hydrodynamical simulations when interpreting observations using any of the previously discussed methods. In simulations, the temperatures, densities and chemical enrichment of circumgalactic gas can be followed throughout cosmic time, enabling the X-ray emission, tSZ signal and absorption features predicted by the simulation model to be readily extracted and compared with observations. The tSZ signal can be readily extracted from the simulated electron pressure, while predictions of absorption and emission may be obtained by coupling the properties of the simulated CGM to photoionization and/or emission models such as CLOUDY

([Ferland et al., 1998](#)) in order to obtain the ionization balance and emission properties of gas in certain phases, which can then be used to produce realistic mock observations.

The ability to trace the origin, history and dynamics of the circumgalactic gas in these mock observations can produce valuable insights that would be impossible to obtain from observations alone. For example, [Crain et al. \(2010\)](#) showed, by computing X-ray luminosities of particles in the GIMIC simulations through coupling them with the Astrophysical Plasma Emission Code (APEC, [Smith et al., 2001](#)), that the X-ray luminosity of the CGM around star-forming galaxies is overwhelmingly dominated by material recently ejected from the ISM (through stellar feedback), and hence that observations of such systems will be strongly biased to these outflows. Similarly, producing mock spectra from simulated quasar sightlines can be used to determine the physical origin of different absorption species observed around galaxies; low ions such as MgII and SiIV have been shown to primarily trace gas close to the galaxy that is being accreted onto the ISM, while higher ions such as OVI and NeVIII trace lower density material further out in the halo that was previously ejected in an outflow ([Ford et al., 2013, 2014](#)).

Simulations can also be used to help interpret observed connections between galaxy and CGM properties. Both the EAGLE ([Schaye et al., 2015; Crain et al., 2015](#)) and IllustrisTNG ([Pillepich et al., 2018a; Nelson et al., 2018a; Springel et al., 2018](#)) simulation suites reproduce the positive correlation between the specific star formation rates of galaxies and the column density in OVI observed by [Tumlinson et al. \(2011\)](#), and the EAGLE simulations revealed that this correlation is predominantly caused by the observed passive galaxies existing in more massive haloes than the active sample, and that the higher virial temperature causes the column density of OVI to drop because it is collisionally ionized to OVII and OVIII ([Oppenheimer et al. 2016a](#), though ejection of OVI by AGN feedback in IllustrisTNG may play a role, see [Nelson et al., 2018a](#)).

A recently-identified limitation in the use of simulations to interpret observations of small-scale clouds and structure in the CGM is that their predictions regarding the properties of the CGM on these scales appear highly sensitive to their resolution ([van de Voort et al., 2019; Peeples et al., 2019](#)), though the average density and content of the CGM is relatively insensitive to this issue. The most salient issue for the purposes of this thesis is that the properties of the CGM (and indeed those of the IGM and the intragroup/intracluster medium) are also impacted markedly by the feedback processes that govern and regulate galaxy growth, which are the least well understood elements of galaxy formation theory. Even in state-of-the-art simulations, these processes are partially unresolved and must be treated with ‘subgrid’ routines, and choices relating to their numerical implementation can significantly influence the resulting properties of

the CGM (e.g. [van de Voort & Schaye, 2012](#); [Hummels et al., 2013](#); [Ford et al., 2016](#); [Rahmati et al., 2016](#); [Sembolini et al., 2016](#)).

In general, this sensitivity is greater than is the case for the stellar properties of the galaxies, with the latter often used as the benchmark against which the parameters of subgrid routines (particularly those describing feedback mechanisms) are calibrated. Simulations that yield similar galaxies need not therefore yield similar circumgalactic or intragroup gas distributions (see e.g. [McCarthy et al., 2017](#)), and at present the degree of consensus between state-of-the-art models in this regard is unclear; this might be termed the “degeneracy” in galaxy formation theory. Detailed observations of the CGM are therefore an urgently-needed constraint for future generations of numerical models, if this degeneracy is to be broken and a self-consistent model for galaxy formation, which reproduces the observed properties of both the stellar and gaseous components of galaxy-CGM systems, is to be found.

1.4 This thesis

As I have motivated throughout this chapter, a complete model of galaxy formation must explain how star formation is regulated through inflows and feedback, and how this self-regulation can be interrupted, leading to a quenching of the galaxy. To understand these processes, one must understand the cycling of baryons through the CGM and answer the three questions I posed at the start of this thesis:

1. How does the process of galaxy formation impact the environments of galaxies?
2. How do the properties of a galaxy’s environment affect its evolution?
3. How are the properties of galaxies linked to the underlying nature of the universe?

Finding conclusive answers to these questions through purely observational means is currently near-impossible; to summarise Section 1.3.1 earlier in this chapter, X-ray and tSZ methods lack both the sensitivity and resolution for detailed studies in all but the most massive systems, and absorption line studies are limited by their inherently 1D nature and the extensive assumptions required in their interpretation. For these reasons, I endeavour to address these questions in this thesis using cosmological, hydrodynamical simulations, which I introduce in **Chapter 2**.

Question 1 can be more succinctly phrased as “how does feedback affect the CGM”, and this is where I begin my analysis. The sensitivity of the properties of the CGM in cosmological simulations to the chosen prescription for feedback processes suggests that

feedback has a significant impact on the CGM in the real Universe, though one might argue that given this sensitivity, simulations are simply “getting out what they put in”. The approach I take is to use the EAGLE simulations, which implement energetically feasible prescriptions for feedback which are tailored only to produce a match to the observed stellar properties of galaxies, leaving the nature of the CGM a prediction of the simulation model. While not guaranteed to be a good representation of what happens in the real Universe, the effects of feedback on the CGM in EAGLE are well-motivated in that they result in the production of realistic galaxies, and the simulations can therefore be used to guide our intuition towards answers for the above questions.

In addressing Question 1, I also aim to answer “where are the missing baryons in galactic haloes?”. As I motivated in Section 1.3, feedback from growing SMBHs is likely able to eject baryons from the CGM, thus fundamentally transforming it. In **Chapter 3**, I show using the EAGLE simulations that the circumgalactic gas fractions of haloes hosting $\sim L^*$ galaxies are indeed regulated by the total integrated feedback injected by their central black holes, with gas-poor haloes hosting more massive SMBHs than gas-rich haloes at a fixed halo mass. I show that at fixed mass, gas-rich haloes tend to host more actively star-forming galaxies than their gas-poor counterparts, implicating this mechanism in quenching (Question 2), and show that the cosmological origin of this diversity lies in the binding energy (and therefore assembly time) of the underlying dark matter halo (Question 3). Finally, I demonstrate that these correlations are borne out in the halo X-ray luminosity and tSZ “flux”, which are proxies for the CGM gas fraction, paving the way for the testing of this picture with forthcoming instrumentation.

Chapter 4 is primarily focused on how this ejection of material from the CGM affects the evolution of the central galaxy. I show that gas-poor haloes are not only more likely to host quenched galaxies, but that they also tend to host galaxies with dispersion-dominated kinematics, implicating the CGM in a morphological transformation of the central. I show that AGN feedback acts to elevate the cooling time of the CGM, either by removing the most rapidly-cooling material or by reconfiguring it at a lower density, and that this inhibits the replenishment of the ISM, leading to quenching and facilitating morphological transformation. I demonstrate that this effect is stronger in haloes where more feedback has been injected relative to the intrinsic binding energy of the halo baryons (the system is pushed away from self-regulation), and that this occurs in the earliest forming and most concentrated haloes. I therefore link the quenching of galaxies to a cosmological origin via the effects of AGN feedback on the CGM, providing answers for my three key questions. The analysis in this chapter will be performed for both the EAGLE and IllustrisTNG simulation suites in parallel, to elucidate how the significant differences in feedback scheme affect the above picture. Remarkably, the overall process broadly remains the same, engendering confidence in the generality of my conclusions,

though there are differences in detail for which I offer observationally distinguishable predictions. Given that both EAGLE and IllustrisTNG produce broadly realistic galaxies, these differences highlight the “degeneracy” in galaxy formation models mentioned previously in this chapter, and that observations of the CGM are best placed to break it.

The results of Chapters 3 and 4 reveal a connection between the assembly histories of dark matter haloes and the properties of their galaxy-CGM ecosystems. Inferring this connection from correlations in large samples does, however, rely upon comparing *different* haloes, limiting my ability to control for other potential driving factors such as environment. In **Chapter 5**, I resolve this issue and clearly demonstrate the connections revealed in this thesis by performing a controlled and systematic experiment in which I adjust the assembly history of a single Milky Way-mass halo in an EAGLE zoom simulation by “genetically modifying” its initial conditions, keeping all other variables fixed. Shifting the halo assembly history to earlier times increases the integrated feedback injected by the AGN, ejecting a greater fraction of the CGM baryons and leading to the quenching and morphological transformation of the central galaxy. These effects can only have originated from differences in the assembly history of the halo, providing clear evidence supporting my conclusions.

Finally, in **Chapter 6**, I present a summary of the work in this thesis, outline some of the new questions it has raised, and suggest how future work might provide us with deeper insight into the galaxy-CGM connection.

Chapter 2

A universe in a box: cosmological simulations of galaxy formation

As motivated in the previous chapter, the difficulties involved in observationally characterising the origin and properties of the CGM demand that one turn to cosmological, hydrodynamical simulations of galaxy formation. In such simulations, there is no “missing baryons problem” - the evolution of the baryon content of a simulated universe can be followed throughout cosmic time, providing an unambiguous (though not necessarily authoritative) assessment of the fractions of the baryonic mass locked into stars, or distributed in the ISM, CGM or WHIM. Our confidence in the realism of hydrodynamical simulations stems primarily from their ability to reproduce the observed characteristics of the galaxy population, which has historically been poor. Over the past 5-10 years, however, the agreement between simulations and observations has dramatically improved. Current state-of-the-art simulations now produce realistic galaxy populations through the use of well-motivated physical prescriptions for sub-resolution non-linear processes, and they are now widely used to guide our understanding of the formation and evolution of galaxies and their CGM.

Prior to this revolution in the success of hydrodynamical simulations, the field of theoretical galaxy formation relied on the use of “semi-analytic” models, in which analytical prescriptions for baryonic processes such as gas accretion, radiative cooling, star formation, chemical evolution and feedback from star formation and growing black holes are used to ‘embed’ galaxies within large-scale cosmological simulations of dark matter. Such simulations are relatively straightforward to run due to the dissipationless nature of dark matter; one need only simulate the gravitational interactions of particles, with the aid of an efficient gravity solver to reduce the computational cost of N -body computation. The most notable large-scale simulations to be used for semi-analytical

studies are the Millennium simulation (Springel, 2005) and its successors, which at the time were the largest and most computationally expensive of their kind. From these simulations, the formation histories of dark matter haloes can be extracted as “merger trees”¹, which can then be coupled to semi-analytical models such as GALFORM (Cole et al., 2000; Lacey et al., 2016) and L-GALAXIES (Henriques et al., 2020) to produce realistic galaxy populations.

The primary advantage of the semi-analytical approach is that it is (in principle) computationally inexpensive relative to hydrodynamical simulations, allowing the creation of mock galaxy catalogues and the interpretation of observations in the context of a dark matter framework. The advantage of full hydrodynamics, however, is that the co-evolution of baryons and dark matter is included, and accretion and feedback processes between galaxies and the diffuse intergalactic medium can be followed, processes that are fundamental to the evolution of galaxies and the CGM and cannot be self-consistently modelled with semi-analytics. Recent improvements in the availability of high-performance computing resources, along with significant improvements in the modelling of feedback from supernovae and AGN, have recently allowed hydrodynamical simulations to attain a similar level of realism to semi-analytical models within cosmologically representative volumes (see e.g. Guo et al., 2016; Mitchell et al., 2018), and they are now at the forefront of our theoretical modelling of the Universe.

The addition of baryons to dark matter-only simulations requires prescriptions for complex multi-scale physics beyond N -body gravity, dramatically increasing their computational expense. This expense is significantly increased for larger simulation volumes, and even more so when the resolution of the simulation is increased, as smaller-scale baryonic physics must still be followed in detail. This naturally leads to two regimes for hydrodynamical simulations, and the choice between them typically depends on the desired use-case for the simulation.

The first regime is the large-scale periodic cosmological volume, in which the formation of large populations of galaxies can be followed, from large samples of lower-mass galaxies to small numbers of rarer objects such as galaxy groups and clusters. These simulations are typically used for examining correlations within large statistical samples, and to generate lightcones for comparison with galaxy surveys. They typically have volumes of $\sim 25 - 100$ comoving Mpc and resolution, expressed in terms of the initial masses of the simulated baryonic particles, of $\sim 10^6 \text{ M}_\odot$. Examples of such simulations are OWLs (Schaye et al., 2010), Illustris (Vogelsberger et al., 2014; Genel et al., 2014), HorizonAGN

¹Such trees can also be computed analytically from extended Press-Schechter theory (e.g. Parkinson et al., 2008).

([Dubois et al., 2014](#)), EAGLE ([Schaye et al., 2015; Crain et al., 2015](#)), MassiveBlack-II ([Khandai et al., 2015](#)), MUFASA ([Davé et al., 2016](#)), IllustrisTNG ([Pillepich et al., 2018a; Nelson et al., 2018a; Springel et al., 2018](#)) and SIMBA ([Davé et al., 2019](#)). Large-scale simulations can also be tailored specifically for the study of large-scale structure and cosmology, and for obtaining larger samples of groups and clusters, by significantly increasing the simulation volume and reducing the resolution accordingly. Examples of this technique include BAHAMAS ([McCarthy et al., 2017](#)) and the TNG300 simulation (part of the IllustrisTNG suite).

The other regime is the so-called “zoom” simulation technique, in which small regions are extracted from a large parent simulation and re-simulated at much higher resolution. These simulations are typically used when a specific type of system must be studied in great detail; examples include galaxy clusters and their environments in C-EAGLE/Hydrangea ([Barnes et al., 2017; Bahé et al., 2017](#)), FABLE ([Henden et al., 2018](#)), and the THREE HUNDRED project ([Cui et al., 2018](#)), the Local Group in APOSTLE ([Sawala et al., 2016](#)), Milky Way-like galaxies in Eris ([Guedes et al., 2011](#)), FIRE ([Hopkins et al., 2014](#)), FIRE-2 ([Hopkins et al., 2018](#)) and Auriga ([Grand et al., 2017](#)), and larger regions of the Universe in GIMIC ([Crain et al., 2009](#)). Zoom simulations allow for far higher resolution than is possible in large periodic volumes, and also permit the detailed study of rare objects such as rich galaxy clusters. The next generation of hydrodynamical simulations will attempt to bridge the gap between the resolution of zooms and the statistics of larger volumes, and simulations such as TNG50 (from the IllustrisTNG suite) have already made advances in this regard.

This thesis includes work in both of these regimes. The work in Chapter 3 was carried out with the EAGLE simulation suite, and in Chapter 4 I will utilise both EAGLE and IllustrisTNG (hereafter TNG) to assess the consensus on my conclusions between two of the foremost models that are presently available. In Chapter 5, I will return to using only the EAGLE model, but with zoom simulations whose initial conditions have been “genetically modified” to enable a tightly-controlled, systematic experiment. To serve as a technical reference for the rest of this thesis, this Chapter will detail the components of the current state-of-the-art in hydrodynamical simulations of galaxy formation, in the context of both the EAGLE and TNG simulation suites.

2.1 Initial conditions and background cosmology

In order to perform a cosmological simulation, the initial conditions of the dark matter and gas content of the simulation must be specified. Large scale simulations are run with periodic initial conditions, such that the mass distribution at one face of the cubic volume

is contiguous with the distribution at the opposite face. In this way, the gravitational and hydrodynamical forces acting on the material at the edge of the box are correct. Zoom simulations simulate a zoomed, high resolution region in concert with a resimulation of its parent periodic volume at lower resolution and without baryonic physics, in order to preserve the correct large-scale forces acting on the zoom region.

The first ingredient in the initial conditions is the background cosmology, which will define the content, primordial distribution and expansion of the simulated universe. Both EAGLE and TNG adopt the cosmological parameters advanced by the [Planck Collaboration et al. \(2014\)](#), for EAGLE, [Planck Collaboration et al. 2016](#) for TNG). For EAGLE (TNG) these are $\Omega_0 = 0.307(0.310)$, $\Omega_b = 0.04825(0.0486)$, $\Omega_\Lambda = 0.693(0.691)$, $\sigma_8 = 0.8288(0.8159)$, $n_s = 0.9611(0.9667)$, and $h = 0.6777(0.6774)$. A transfer function and power spectrum can then be generated from these parameters using software for modelling anisotropies in the CMB (e.g. CAMB [Lewis et al., 2000](#)), which are used in combination with a Gaussian white noise field such as *Panphasia* ([Jenkins, 2013](#)) to create a random density field with the correct linear power spectrum for the input cosmology. This field is then used to displace an initially unperturbed distribution of mass tracer particles, and the growing mode of gravitational instability is used to set their initial velocities.

Once these initial conditions have been generated, each particle is ‘split’ into a pair: a dark matter particle and a gas particle in EAGLE, or a dark matter particle and a gas cell in TNG. The co-ordinates of these particles are offset in opposite directions from the original tracer particles in a random orientation, with the offset governed by the mass ratio of the dark matter and gas elements. In concordance with the background cosmology, the mass ratio of the dark matter and gas elements is $1 : \Omega_b / (\Omega_0 - \Omega_b)$. These initial conditions now represent the state of a cosmological volume at $z = 127$, and the simulations are evolved from this point following the model described in the following sections.

In Chapter 5, I will make use of the genetic modification technique ([Roth et al., 2016](#); [Rey & Pontzen, 2018](#)), which makes alterations to the initial conditions of zoom simulations in such a way that they are still consistent with Λ CDM, and do not represent an extremely unlikely draw from the Gaussian white noise field. The method preserves the large-scale structure in which the zoom region evolves, and creates a controlled experiment in which the assembly history of the resimulated galaxy can be fine-tuned.

To evolve the simulation from its initial conditions, three sets of processes occur for each simulation timestep. Firstly, the gravitational interactions between the simulated particles are computed and the resulting accelerations are applied to them. This process alone is sufficient for a dark matter-only simulation; for gas particles, the hydrodynamical

forces acting on them must also be computed. Finally, the “subgrid” prescription for processes acting below the resolution of the simulation must be implemented. In the following sections I will give a brief overview of these processes.

2.2 Gravity solver

To compute the gravitational acceleration of a particle due to a simulated mass distribution, the gravity solver must find the potential by solving Poisson’s equation. The most direct method for this is to perform an N -body calculation, summing the interactions of one particle with every other particle. The disadvantage of this method is its $\mathcal{O}(N^2)$ computational scaling, making it prohibitively expensive for modern simulations following $\sim 10^9$ particles or more.

The cost of N -body simulations can be significantly reduced through the use of an hierarchical “octree”. As a result of the inverse-square scaling of gravitational forces, the potential at a given point is dominated by the matter closest to it, with the contribution of more distant material being minor. Short-range interactions must therefore be captured in detail, while the precision of interactions over a greater distance can be sacrificed, and approximated as a single interaction. The octree method (e.g. [Barnes & Hut, 1986](#)) therefore recursively splits the simulation volume into smaller cubes until each cube contains a chosen maximum number of particles. This spatial grouping allows the accurate calculation of short-range forces using the smallest cubes, while particles at long-range can be grouped into larger nodes. This greatly reduces the number of interactions which must be computed, reducing the complexity to $\mathcal{O}(N \log N)$.

Another means to reduce the computational cost of simulating gravity is the particle-mesh method, in which particles are assigned to nearby points on a continuous mesh, producing a 3D mass distribution. The potential is then simple to obtain: performing a Fast Fourier Transform to transfer the mesh into the spatial frequency domain converts Poisson’s equation to a simple linear equation for the potential, which can then be inverse transformed back to real space. The disadvantage of this method is its sensitivity to the resolution of the mesh; to capture smaller-scale interactions, a finer mesh must be used, increasing the memory footprint of the procedure. One solution to this is “adaptive mesh refinement”, where a finer mesh is used where higher resolution is needed, however most modern simulation codes, in particular the EAGLE solver GADGET3([Springel, 2005](#)) and the TNG solver AREPO ([Springel, 2010](#)), use a hybrid of the octree and particle-mesh methods (TreePM), where shorter-range interactions are computed using an octree, while those at longer range are computed from a mesh.

When simulating the interaction of point-like particles due to gravity, the gravitational force diverges as the distance between particles, r , goes to zero. To prevent spurious gravitational forces, the $1/r$ gravitational potential is replaced with $1/\sqrt{r^2 + \epsilon^2}$, where the parameter ϵ “softens” the gravitational interaction on small scales. At present, there is no strong consensus on the most applicable softening scale (e.g. [Ludlow et al., 2019](#)), and hence different simulations adopt different values: in EAGLE, the softening length is $\epsilon_{\text{com}} = 2.66$ comoving kpc (ckpc), limited to a maximum proper length of $\epsilon_{\text{prop}} = 0.7$ proper kpc (pkpc). In TNG, the softening length of dark matter and stellar particles is $\epsilon_{\text{com}} = 1.48$ ckpc, limited to a maximum proper length of $\epsilon_{\text{prop}} = 0.74$ pkpc, for gas cells it is 2.5 times the effective cell radius, and for black hole particles it scales as $\epsilon_{\text{BH}} = \epsilon_{\text{DM}}(m_{\text{BH}}/m_{\text{DM}})^{1/3}$.

2.3 Hydrodynamics

When simulating the evolution of baryons, the behaviour of a continuous distribution of gas must be modelled in such a way that it follows the laws of fluid dynamics (i.e. satisfies the Euler equations) and thermodynamics. There are two principal methods for achieving this. Eulerian schemes discretise the fluid at a fixed volume, simulating the fluid as a regular mesh of fluid elements between which mass can flow. Lagrangian schemes instead discretise the fluid by mass; fluid elements are represented by particles about which their density is distributed. Both methods have significant disadvantages: Eulerian methods typically suffer from poorer conservation of thermodynamic quantities and are not invariant under Galilean transforms, while Lagrangian schemes conserve mass by construction but suppress the formation of shocks and instabilities in the fluid. Both EAGLE and TNG employ quasi-Lagrangian schemes which are designed to overcome the typical shortcomings of both methods.

The EAGLE solver GADGET3 is an example of a smoothed particle hydrodynamics (SPH) scheme. In SPH, fluid elements are represented by particles about which their mass is distributed according to a kernel function, defined such that the distribution has a finite volume. This kernel is described by a smoothing length, h , usually defined such that the kernel contains a fixed number of neighbours. In this way, the smoothing adapts to the local particle number density. EAGLE utilises the Wendland C^2 kernel ([Dehnen & Aly, 2012](#)) with 58 nearest-neighbours. Following this smoothing of the fluid quantities over particles, the properties of the fluid at any point in space are given by the kernel-weighted sum over all nearby particles.

The primary issue with standard SPH is that particles must be assigned an artificial viscosity in order to capture shocks (EAGLE implements this following [Cullen & Dehnen](#),

2010), and this, in combination with the pressure gradient at contact discontinuities, causes unphysical surface tension which inhibits the development of instabilities and the mixing of gas phases (e.g. Agertz et al., 2007). To overcome this, EAGLE uses the Hopkins (2013) formulation of SPH, which defines the thermodynamic state of particles in terms of pressure and entropy and ensures that pressure varies smoothly across discontinuities. SPH is non-diffusive, therefore to model the mixing of different phases of gas and the transfer of heat between particles, EAGLE implements the thermal conduction switch of Price (2010). To ensure that very sudden increases in the particle energy are properly captured in the simulation, EAGLE limits the length of the timestep in such cases following Durier & Dalla Vecchia (2012). These methods, referred to as “ANARCHY SPH” greatly mitigate the issues inherent in traditional SPH schemes.

The TNG solver AREPO attempts to find a compromise between SPH and Eulerian schemes that inherits the advantages of both. AREPO models the fluid as an unstructured, moving mesh defined by a Voronoi tesselation. Fluid elements are represented by “cells” that track several conserved quantities: mass, momentum, energy and magnetic field, and the evolution of these quantities is obtained by solving the Riemann problem at cell boundaries. The inclusion of magnetic fields makes AREPO a magnetohydrodynamical (MHD) code. The code is “quasi-Lagrangian” because unlike in Eulerian schemes, the generating points of the mesh can move as the simulation evolves, following the local fluid velocity. The code therefore inherits the adaptive resolution and Galilean invariance of SPH, while benefitting from the use of a mesh, eliminating the need for artificial viscosity and capturing shocks and instabilities with high accuracy.

2.4 Characterising haloes and galaxies

Identifying haloes, subhaloes and galaxies in the simulation requires the use of algorithms that identify structures within the simulation. While the simulations are running, the friends-of-friends (FoF) algorithm (Davis et al., 1985) is applied to the dark matter particle distribution, with a linking length of 0.2 times the mean interparticle separation. This algorithm links structure together and identifies haloes based on purely spatial information, making it computationally inexpensive to run on-the-fly, and is used to seed black hole particles in haloes once they reach a threshold mass (see Section 2.5.4). Gas, stars and BHs are associated with the FoF group, if any, of their nearest dark matter particle.

Gravitationally-bound substructures within haloes are subsequently identified in post-processing using the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009), which searches for self-bound regions within FoF haloes which are bounded by saddle

points in the density distribution, now considering all particle types rather than solely the dark matter. The structure, or “subhalo” containing the particle with the lowest potential in the halo is designated the central subhalo, and others are considered satellite subhaloes. The location of this particle is considered the centre of the halo, and we characterise quantities such as halo mass and virial radius via the spherical overdensity mass (M_{200} , Lacey & Cole, 1994) about this location. This thesis is concerned almost exclusively with central galaxies and their haloes.

2.5 Subgrid physics

Hydrodynamical simulations of galaxy formation inevitably require the modelling of processes occurring on scales far smaller than can be resolved in the simulation (i.e. smaller than the mean interparticle separation or the gravitational softening scale). These processes govern some of the most vital aspects of galaxy formation: the radiative cooling of gas, star formation and its associated feedback, the formation of supermassive black holes, and the feedback associated with black hole growth. It is not possible to implement these processes in the simulation directly, therefore the standard approach is to use simplified “subgrid” prescriptions for their effect on the resolved elements in the simulation, which are based on analytical predictions or empirical laws obtained from observations. The chosen subgrid approach can significantly influence the outcome of the simulation, and as such the implementation of subgrid physics is the most uncertain aspect of modern models.

The subgrid approach shares similarities with semi-analytical modelling, and as such is subject to similar caveats: it is not possible for a subgrid model in which feedback energy is injected at a chosen efficiency to predict the properties of galaxies from first principles. If, for example, a galaxy is self-regulating its formation with feedback from star formation, the outflow rate will eventually balance the inflow rate, but the star formation rate for which this balance is achieved (and thus the stellar mass formed by the present day) is set by the chosen feedback efficiency. As motivated by Schaye et al. (2015) in the design of the EAGLE subgrid model, the current approach is to *calibrate* the parameters of the subgrid model in order to reproduce key properties of the observed galaxy population. After such a calibration, these observables are not predictions of the simulation, however the other emergent properties of the simulated universe *are*, and can be used to guide our intuition in the field of galaxy formation. This has particular relevance for the work in this thesis; the EAGLE subgrid model was not calibrated on the gaseous properties of galaxies, and the properties of the CGM are therefore predictions of the model.

EAGLE and TNG share similarities in some aspects of their subgrid physics, but also employ very different strategies in others. In this section I will briefly outline the methods used in both simulations, and how they were calibrated.

2.5.1 Radiative cooling, photoheating and reionisation

To follow the radiative cooling and heating of gas, both EAGLE and TNG employ the publicly available CLOUDY model (version 07.02, last described by Ferland et al., 1998), which includes the important atomic processes in the typical temperature range of gas in the simulations ($T \sim 10^4\text{--}10^8$ K). Following Wiersma et al. (2009a), the contributions to cooling from 11 elements (H, He, C, N, O, Ne, Mg, Si, S, Ca and Fe), can be tabulated as a function of hydrogen number density, n_{H} , temperature, T , redshift, z and in the case of the metal-free cooling contribution, the helium fraction $n_{\text{He}}/n_{\text{H}}$. These tabulated rates assume that the gas is in ionisation equilibrium, optically thin and dust free, and that it is photoionised by the cosmic microwave background and a time-varying UV/X-ray background from galaxies and quasars.

In the EAGLE simulations, the radiative cooling rates (per unit volume) of particles are then computed by summing the element-by-element contributions,

$$\Lambda = \Lambda_{\text{H,He}} + \sum_{i>\text{He}} \Lambda_{i,\odot} \frac{n_e/n_{\text{H}}}{(n_e/n_{\text{H}})_{\odot}} \frac{n_i/n_{\text{H}}}{(n_i/n_{\text{H}})_{\odot}}, \quad (2.1)$$

where $\Lambda_{\text{H,He}}$ is the metal-free contribution, $\Lambda_{i,\odot}$ is the contribution of element i for the solar abundances assumed in CLOUDY, n_e/n_{H} is the particle electron abundance, and n_i/n_{H} is the particle abundance in element i . The TNG simulations do not consider element-by-element contributions to radiative cooling, instead assuming solar abundance ratios and tabulating the net cooling rate for all 11 elements as a function of metallicity.

Despite both simulations adopting a cooling implementation based on that of Wiersma et al. (2009a), there are differences in their cooling rates, owing primarily to the adoption of different UV/X-ray background radiation models (EAGLE uses that of Haardt & Madau 2001 while TNG uses Faucher-Giguère et al., 2009) and, in TNG, the assumption of solar abundance ratios, the adoption of an HI self-shielding correction for high-density gas, and the suppression of the cooling rate in gas close to accreting BHs.

2.5.2 Star formation

In the real Universe, stars form as the result of Jeans instabilities in molecular clouds. This process cannot be directly implemented in the current generation of cosmological

simulations, because they lack the resolution to model the cold molecular gas phase, and because the physics of molecular clouds requires the inclusion of prohibitively expensive numerical methods such as radiative transfer and non-equilibrium chemistry. The approach taken by both EAGLE and TNG to overcome this issue is to impose an effective equation of state for the interstellar medium (ISM) and to implement star formation stochastically according to an empirical law.

The EAGLE approach is to impose a density-dependent temperature floor, $T_{\text{eos}}(\rho)$, which corresponds to an equation of state $P_{\text{eos}} \propto \rho^{4/3}$, normalised to $T_{\text{eos}} = 8000$ K at $n_{\text{H}} = 0.1 \text{ cm}^{-1}$, the approximate temperature and density of the warm ISM. In this prescription, the temperature of the ISM is not physical, but reflects the pressurisation of the ISM. For gas in this regime, star formation proceeds according to the observed Kennicutt-Schmidt star formation law ([Kennicutt, 1998](#)),

$$\dot{\Sigma}_{\star} = A \left(\frac{\Sigma_{\text{g}}}{1 \text{ M}_{\odot} \text{ pc}^{-2}} \right)^n, \quad (2.2)$$

where $\dot{\Sigma}_{\star}$ and Σ_{g} are the surface densities of stars and gas and A and n are free parameters. Following [Schaye & Dalla Vecchia \(2008\)](#), this can be written as a pressure law,

$$\dot{m}_{\star} = m_{\text{g}} A (1 \text{ M}_{\odot} \text{ pc}^{-2})^{-n} \left(\frac{\gamma}{G} f_{\text{g}} P \right)^{(n-1)/2}, \quad (2.3)$$

where m_{g} is the mass of the gas particle, $\gamma = 5/3$ is the heat capacity ratio, G is Newton's gravitational constant, f_{g} is the gas mass fraction (assumed to be unity) and P is the pressure. With the constants A and n determined from observations, star formation can then be implemented stochastically; at each simulation timestep Δt , the probability that a gas particle is transformed into a star particle is $\min(\dot{m}_{\star} \Delta t / m_{\text{g}}, 1)$. Star formation occurs in cold, dense gas, so EAGLE implements a density threshold above which stars can form,

$$n_{\text{H}}^{\star} = 0.1 \text{ cm}^{-3} \left(\frac{Z}{0.002} \right)^{-0.64}, \quad (2.4)$$

where Z is the metallicity of the gas. This metallicity dependence accounts for the fact that molecular gas can form at lower densities if the metallicity and dust content are higher. The TNG simulations implement a similar prescription to this in order to convert gas cells into star particles, but use a 3D Schmidt law rather than a 2D Kennicutt-Schmidt law, utilise a two-phase equation of state and do not include a metallicity dependence on their density threshold of $n_{\text{H}}^{\star} = 0.1 \text{ cm}^{-3}$, following [Springel & Hernquist \(2003\)](#).

Once star particles form in both simulations, they are treated as simple stellar populations with a [Chabrier \(2003\)](#) initial mass function and inherit the element abundances

of their parent gas particles/cells. Stellar evolution and the chemical enrichment of the ISM are implemented following [Wiersma et al. \(2009b\)](#); at each timestep, the fraction of the initial mass reaching the end of the main sequence is computed, the production of each of the 11 elements tracked in the simulations is obtained from pre-computed yield tables, and the lost mass and metals are distributed among the surrounding gas particles/cells.

2.5.3 Stellar feedback

As discussed earlier in this chapter, the recent improvement in the agreement between hydrodynamical simulations and observations stems from improvements in the subgrid treatment of feedback from both star formation and AGN. Neither EAGLE nor TNG possess the resolution required to model the development of outflows from feedback processes self-consistently, hence a subgrid treatment must be used to simulate the effect of energy injection from these sources on resolved scales. The properties of galaxies and their CGM are highly sensitive to the implementation and efficiency of feedback, and adjustment of the chosen prescription is the primary means by which EAGLE and TNG are calibrated.

The principal difficulty in implementing feedback lies in minimizing numerical losses and ensuring that the injected energy is efficient enough to drive a wind. When injecting the energy from Type II supernovae thermally, it is typically distributed over too much mass (at least one resolution element, $\sim 10^6 M_\odot$) in comparison to the “real” ejecta (usually $\lesssim 1 M_\odot$), and the resulting increase in temperature is very small, giving a very short cooling time for the particle. The injected energy is therefore radiated away before it can be converted to kinetic energy and drive a wind. This overcooling problem can be resolved by ensuring that the cooling time (i.e. the ratio of the internal energy of a particle to its cooling rate) is substantially longer than the sound crossing time, either by increasing the mass resolution of the simulation or ensuring that the temperature increment for the heated particles is large.

To resolve this issue, the EAGLE simulations implement the stochastic thermal feedback method of [Dalla Vecchia & Schaye \(2012\)](#), where a parameter f_{SF} ² is used to set the probability that an SPH neighbour of a newly formed star particle is heated by a fixed temperature increment of $\Delta T = 10^{7.5}$ K after a delay of 30 Myr (corresponding to the lifetime of the least massive progenitors of Type II supernovae). This large temperature jump ensures that a pressure gradient is created which drives outflows without the need to set an initial wind mass loading or velocity; the efficiency of the feedback can then

²This parameter is equivalent to f_{th} in the EAGLE reference articles.

be calibrated by adjusting f_{SF} . There remains a maximum density for which feedback is efficient in this prescription, an issue which can be mitigated by using a higher ΔT , however doing so will make individual feedback events less frequent, leading to poorer sampling of the feedback cycle. The chosen value therefore represents a compromise between these two regimes.

In this prescription, a value of $f_{\text{SF}} = 1$ corresponds to an injected energy of $1.74 \times 10^{49} \text{ erg M}_{\odot}^{-1}$ of stellar mass formed, corresponding to 10^{51} ergs per supernova in a Chabrier IMF, assuming stars of mass $6 - 100 \text{ M}_{\odot}$ explode as Type II supernovae. To account for variations in the efficiency of supernova feedback induced by local gas properties, f_{SF} is a function of both the metallicity Z and density $n_{\text{H,birth}}$ of the parent gas particle at the time it converted into a star particle,

$$f_{\text{SF}} = f_{\text{SF,min}} + \frac{f_{\text{SF,max}} - f_{\text{SF,min}}}{1 + \left(\frac{Z}{0.1Z_{\odot}}\right)^{n_Z} \left(\frac{n_{\text{H,birth}}}{0.67 \text{ cm}^{-3}}\right)^{-n_n}}, \quad (2.5)$$

where $Z_{\odot} = 0.0127$ is the solar metallicity and $n_Z = n_n = 2/\ln 10$. The metallicity dependence accounts for the stronger radiative losses expected at high metallicity, and the density dependence prevents excessive losses at high density. The asymptotic values $f_{\text{SF,min}} = 0.3$ and $f_{\text{SF,max}} = 3$ are the principal calibrated parameters which set the efficiency of the feedback; they were chosen such that the simulation provides a good match to the galaxy stellar mass function (GSMF) and to the sizes of galaxies as a function of stellar mass (see [Crain et al., 2015](#)).

The TNG model adopts a different approach to the driving of winds from stellar feedback. Rather than injecting the energy thermally, a kinetic wind scheme is used in which wind particles are stochastically and isotropically launched from star-forming gas. These wind particles have an assigned initial velocity which scales with the local one-dimensional dark matter velocity dispersion and has a redshift-dependent minimum value which makes the feedback more efficient in low-mass galaxies and at low redshift. To determine the mass loading of the wind, the energy must be specified, which is subject to an efficiency parameter that is a function of the metallicity of the stellar population's natal gas, $f_{\text{SF}}(Z)$ ³. Here $f_{\text{SF}} = 1$ corresponds to an expectation value of the injected energy $1.08 \times 10^{49} \text{ erg M}_{\odot}^{-1}$, which is lower than is the case for EAGLE since here the progenitors of core-collapse SNe are assumed to be those with mass $8 - 100 \text{ M}_{\odot}$. Ninety per cent of this energy is injected kinetically, with the remaining ten per cent injected into the wind particles via a thermal dump. To prevent overcooling, these particles are temporarily decoupled from the hydrodynamics scheme, enabling them to escape the galaxy without interacting with the ISM. The thermal energy injected prevents spurious star formation

³This parameter is equivalent to the dimensionless prefactors in the expression for e_w in the TNG reference articles.

on recoupling with the hydrodynamics. Further details on the parameterisation and implementation of this scheme are given in [Pillepich et al. \(2018b\)](#).

2.5.4 Black hole seeding, growth and feedback

Just as feedback from star formation is essential in obtaining a good match to the low-mass end of the GSMF, feedback from growing supermassive black holes (SMBHs) at the centres of galaxies is essential for preventing the overproduction of very massive galaxies. The formation of SMBHs and the detailed physics of their accretion disks, growth and associated feedback occur on scales well below the resolution of cosmological simulations and must be treated with subgrid prescriptions.

Both EAGLE and TNG follow [Springel et al. \(2005\)](#) and seed BH “sink” particles in haloes identified on-the-fly with FOF that attain a certain mass ($10^{10} M_{\odot} h^{-1}$ in EAGLE, $7.38 \times 10^{10} M_{\odot}$ in TNG) and do not already contain a BH. The particles are seeded with an initial mass of $10^5 M_{\odot} h^{-1}$ in EAGLE and $1.18 \times 10^6 M_{\odot}$ in TNG, and grow according to an Eddington-limited Bondi-Hoyle rate for spherically symmetric accretion, which depends on the BH mass, and the local gas temperature and density. The accretion rate is given by the minimum of the Eddington rate,

$$\dot{m}_{\text{Edd}} = \frac{4\pi G m_{\text{BH}} m_p}{\epsilon_r \sigma_T c}, \quad (2.6)$$

and

$$\dot{m}_{\text{acc}} = \begin{cases} \dot{m}_{\text{Bondi}} \times \min(C_{\text{visc}}^{-1} (c_s/V_\phi)^3, 1) & \text{for EAGLE} \\ \dot{m}_{\text{Bondi}} & \text{for TNG} \end{cases} \quad (2.7)$$

where \dot{m}_{Bondi} is the [Bondi & Hoyle \(1944\)](#) rate,

$$\dot{m}_{\text{Bondi}} = \begin{cases} \frac{4\pi G^2 m_{\text{BH}}^2 \rho}{(c_s^2 + v^2)^{3/2}} & \text{for EAGLE} \\ \frac{4\pi G^2 m_{\text{BH}}^2 \rho}{c_s^3} & \text{for TNG.} \end{cases} \quad (2.8)$$

Here m_{BH} is the BH mass, m_p is the proton mass, ϵ_r is the radiative efficiency of the BH accretion disk (0.1 in EAGLE, 0.2 in TNG), σ_T is the Thomson cross-section, c is the speed of light, ρ and c_s are density and sound speed of the gas around the BH and v is the relative velocity of the BH to its surrounding gas. The extra term for EAGLE

in Equation 2.7 incorporates the rotation speed of the gas around the BH, V_ϕ and a free parameter related to the viscosity of the accretion disk, $C_{\text{visc}} = 2\pi$, following Rosas-Guevara et al. (2015). The mass growth rate of the black hole in both simulations, accounting for the radiative efficiency of the accretion, is given by

$$\dot{m}_{\text{BH}} = (1 - \epsilon_r)\dot{m}_{\text{acc}}. \quad (2.9)$$

The injection of energy as a result of BH growth (AGN feedback) proceeds at a rate given by $f_{\text{AGN}}\dot{m}_{\text{acc}}c^2$ in both simulations, with f_{AGN} encoding the efficiency of the feedback. The mechanism through which this energy is injected is, however, very different for each simulation. In analogy with feedback from star formation, EAGLE injects AGN feedback energy thermally and stochastically, with a temperature increment of $\Delta T_{\text{AGN}} = 10^{8.5}$ K. BHs in the simulation carry a reservoir of feedback energy, E_{BH} , to which $\Delta E_{\text{BH}} = f_{\text{AGN}}\dot{m}_{\text{acc}}c^2\Delta t$ of energy is added at each timestep Δt . When enough energy has been accumulated to heat at least 1 particle by ΔT_{AGN} , the BH is allowed to stochastically heat each of its SPH neighbours by that temperature increment with a probability

$$P = \frac{E_{\text{BH}}}{\Delta\epsilon_{\text{AGN}}N_{\text{neighbours}}\langle m_g \rangle}, \quad (2.10)$$

where $\Delta\epsilon_{\text{AGN}}$ is the energy change per unit mass corresponding to ΔT_{AGN} , $N_{\text{neighbours}}$ is the number of SPH neighbours and $\langle m_g \rangle$ is their mean mass. The efficiency of this process is regulated by $f_{\text{AGN}} = \epsilon_r\epsilon_f$, where $\epsilon_f = 0.15$ is the calibrated parameter in the EAGLE model. Over a very wide range of non-zero values, this parameter only affects the BH mass (see e.g. Booth & Schaye, 2010), with the resulting outflow rate and reduction of star formation being insensitive to this choice; self regulation causes the accretion rate to adjust until inflows balance outflows. As such, this parameter is calibrated only on the BH mass - stellar mass relation. The most important factor in terms of the effects of the feedback is the choice of ΔT_{AGN} ; as with the EAGLE stellar feedback prescription, a balance must be found between the prevention of overcooling and a sufficient sampling of the feedback. A higher value is adopted for AGN feedback because the typical densities around SMBHs are much higher than is typical of the ISM.

In TNG, the energy injection rate is the same as for EAGLE, but here the feedback is injected in one of two modes. Feedback associated with high accretion rates (relative to the Eddington rate) is injected via a thermal dump, heating gas cells neighbouring the BH with an efficiency $f_{\text{AGN,thm}} = \epsilon_{f,\text{thm}}\epsilon_r$, where $\epsilon_f = 0.1$ is the calibrated parameter and $\epsilon_r = 0.2$. At low accretion rates, energy is injected kinetically in a direction that is chosen randomly for each injection event, with an efficiency $f_{\text{AGN,kin}}$ that scales with the local gas density up to a maximum of 0.2. In contrast to stellar feedback, gas cells which

receive energy from kinetic AGN feedback do not decouple from the hydrodynamics scheme. Here the injection velocity is governed by the mass of gas within the injection region and, in analogy to the stochastic heating used by EAGLE, a minimum injection energy is accumulated between individual injection events. This minimum injection energy is a function of the gas mass within the injection region, the one-dimensional dark matter velocity dispersion and a free parameter governing the “burstiness” of the feedback. Note that such a threshold is not implemented for the thermal AGN mode, where the energy is injected continuously. To prevent kinetic feedback from becoming a runaway process, the coupling efficiency $f_{\text{AGN,kin}}$ is reduced when the surrounding gas is at very low densities. The threshold separating the two injection modes is specified in terms of the Eddington ratio and scales as a function of the BH mass,

$$\chi = \min[0.1, \chi_0(m_{\text{BH}}/10^8 M_{\odot})^2], \quad (2.11)$$

where $\chi_0 = 0.002$. As discussed by [Weinberger et al. \(2017\)](#), this approach in principle allows for any BH, regardless of its mass, to deliver feedback in the thermal mode if the accretion rate is sufficiently high. However, this becomes rare once the BH reaches the pivot mass, effectively making the choice of $10^8 M_{\odot}$ a calibrated parameter that governs when AGN feedback switches from thermal to kinetic injection.

2.6 Coupling to plasma emission models

Once hydrodynamical simulations have been run, they can be used to create mock observables for comparison with observational data. One such observable which has particular relevance to this thesis is the X-ray luminosity of gas in dark matter haloes, which is a proxy for the gas mass. In this section I will briefly outline my method for computing the X-ray luminosities of fluid elements (gas particles/cells) in EAGLE and TNG.

The calculation of X-ray emission from gas in the simulations involves the integration of emission spectra generated by a plasma emission code. To obtain these spectra, I interpolate a pre-computed cooling table generated using the Astrophysical Plasma Emission Code (APEC, [Smith et al., 2001](#)), which assumes that the gas is an optically thin plasma in collisional ionisation equilibrium. The cooling table gives emission spectra at temperatures ranging from 10^4 to 10^9 K for the 11 chemical elements (H, He, C, N, O, Ne, Mg, Si, S, Ca, Fe) which are tracked in EAGLE and TNG. To begin computing the radiative cooling rate of a fluid element, the table is linearly interpolated to provide spectra corresponding to the temperature of the fluid element. Figure 2.1 shows example APEC spectra for six chemical elements at a temperature of 10^7 K, with the solar

abundance ratios of [Anders & Grevesse \(1989\)](#), which the APEC table assumes. These spectra highlight the importance of metal line cooling to emission in the soft, 0.5-2.0 keV X-ray passband (grey dashed lines), which matches that of *ROSAT* and will be used in Chapter 3.

The selected spectra for each chemical element are then integrated over the appropriate passband to obtain a set of cooling functions, λ . These cooling functions represent the radiative cooling rate per unit volume divided by $n_e n_H$, the product of the number densities in electrons and hydrogen, and hence have units of erg cm³ s⁻¹. I rescale these cooling functions to the individual abundances in the fluid element (using the smoothed abundances in EAGLE); the computed luminosities therefore do not assume solar abundance ratios. The total cooling function Λ_j for a fluid element j is then given by the sum of the cooling functions. Once the total cooling function for a fluid element has been computed, the X-ray luminosity is given by

$$L_{X,j} = n_{e,j} n_{H,j} V_j \Lambda_j \quad (2.12)$$

$$= \frac{X_{e,j}}{(X_{e,j} + X_{i,j})^2} \frac{\rho_j}{\mu_j m_H} \frac{m_{\text{gas},j}}{\mu_j m_H} \Lambda_j, \quad (2.13)$$

where n_i is the number density in ions, $X_e = n_e/n_H$, $X_i = n_i/n_H$, μ is the mean molecular weight, and ρ , m and V are the density, mass and volume respectively.

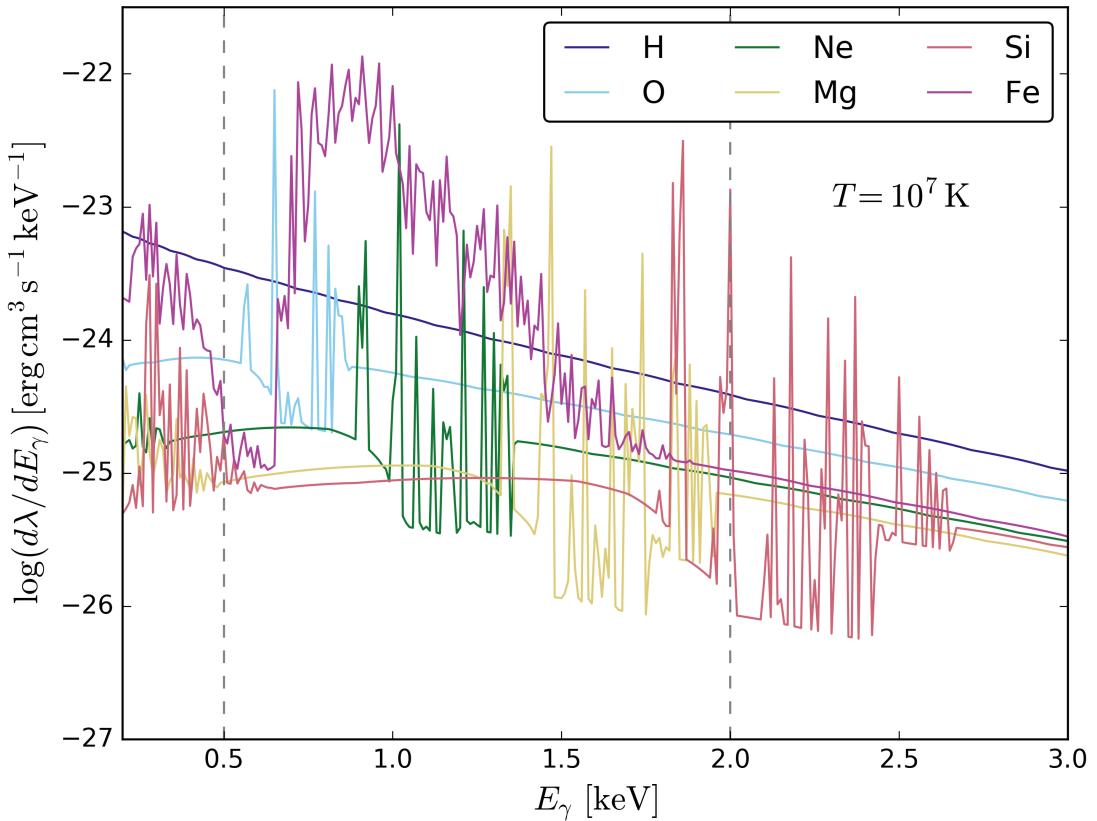


FIGURE 2.1: APEC emission spectra for H, O, Ne, Mg, Si and Fe at a temperature of $T = 10^7 \text{ K}$. Grey dashed lines indicate the *ROSAT* soft X-ray band.

Chapter 3

Black hole feedback and the gas content of dark matter haloes

3.1 Introduction

The primary goals of this thesis are to understand how the properties of circumgalactic haloes are affected by the process of galaxy evolution occurring at their centres, and to then understand the role the CGM plays in regulating the evolution of galaxies. In this Chapter I will focus on the first of these goals.

The most fundamental property of the CGM is simply “how much of it there is”, which is described by the circumgalactic gas fraction: the mass in gas within the halo virial radius relative to the mass of the halo. As I discussed in 1.3, this fraction is observed to fall well below the cosmic average in all but the most massive (cluster-scale) haloes, with only $\sim 5\%$ of the cosmic budget existing in the form of stars, leading to the notion of a “missing baryons problem”. Feedback processes are the likely culprit for this baryon deficit, yet the effects of feedback on the CGM remain poorly-constrained by observations and their physical efficiencies cannot (yet) be predicted from first principles, meaning that *ab initio* prediction of the relationship between the gas fraction and total mass of haloes is not yet feasible. Given that the gas fractions of haloes below the mass scale of galaxy groups remain essentially unconstrained, one must turn to cosmological simulations to examine the effects of feedback on the content of the CGM.

In this Chapter I examine the influence of galaxy properties on the present-day gas fractions of haloes in the EAGLE simulations of galaxy formation. EAGLE adopts the pragmatic approach of calibrating feedback efficiencies to ensure the reproduction of key properties of the galaxy population, such as their stellar and central BH masses and the

sizes of disc galaxies¹, and has been shown to reproduce a diverse range of observable properties of the galaxy population (e.g. Furlong et al., 2015, 2017; Trayford et al., 2015, 2017; Segers et al., 2016; Crain et al., 2017) and intergalactic gas, as probed by X-ray emission (e.g. Schaye et al., 2015) and absorption features in quasar sightlines (e.g. Oppenheimer et al., 2016a; Rahmati et al., 2015, 2016; Turner et al., 2017). The suite is therefore well suited to the study of the co-evolution of galaxies and their gaseous environments.

In Section 3.2 I briefly describe the simulations and my techniques for identifying and characterising galaxies and their haloes. In Section 3.3 I present the scaling relation between the halo gas fraction and halo mass, and examine the origin of scatter about it, whilst in Section 3.4 I investigate means by which these predictions of the simulations can be confronted with observational measurements. I summarise and discuss my findings in Section 3.5. In Appendix A I briefly compare the X-ray and thermal Sunyaev-Zel'dovich fluxes of my simulated haloes with current observational measurements. Throughout, I adopt the convention of prefixing units of length with ‘c’ and ‘p’ to denote, respectively, comoving and proper scales, e.g. cMpc for comoving megaparsecs.

3.2 Methods

Throughout this Chapter I will use the EAGLE simulations, which were introduced in detail throughout Chapter 2, to which I refer the reader for details of the simulations and the employed subgrid physics. I would however like to reiterate that the efficiency of stellar feedback was calibrated to reproduce the present-day stellar masses of galaxies and the sizes of galaxy discs, whilst the efficiency of AGN feedback was calibrated to reproduce the present-day scaling relation between the stellar mass of galaxies and that of their central BH. The gaseous properties of galaxies and their haloes were **not** considered during the calibration and may be considered predictions of the simulations.

I analyse four simulations from the EAGLE suite, focusing primarily on the simulation with the largest volume and greatest particle number, Ref-L100N1504, which evolves with the EAGLE Reference model a periodic cube of side $L = 100$ cMpc, populated with $N = 1504^3$ collisionless dark matter particles with mass $9.70 \times 10^6 M_\odot$ and an (initially) equal number of baryonic particles with mass $1.81 \times 10^6 M_\odot$. In order to compute the intrinsic binding energy of haloes in this simulation, i.e. that which emerges in the absence of the dissipative physics of galaxy formation, I also analyse a simulation starting from identical initial conditions but considering only collisionless gravitational dynamics, DMONLY-L100N1504. I briefly examine NOAGN-L050N0752, a simulation

¹Section 2.5 motivates this approach.

following a smaller $L = 50$ cMpc cubic volume at the same resolution, using a variation of the Reference model in which AGN feedback is disabled. To ensure that comparisons with this simulation are made on an equal footing I use Ref-L050N0752, a simulation of the same $L = 50$ cMpc volume using the EAGLE Reference model. In all cases a Plummer-equivalent gravitational softening length of $\epsilon_{\text{com}} = 2.66$ ckpc was used, limited to a maximum proper length of $\epsilon_{\text{prop}} = 0.7$ pkpc.

Haloes are identified by applying the friends-of-friends algorithm to the dark matter particle distribution, with a linking length of 0.2 times the mean interparticle separation. Gas, stars and BHs are associated with the FoF group, if any, of their nearest dark matter particle. Bound substructures are subsequently identified within haloes using the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009). I consider in my analysis present-day haloes with $M_{200} > 10^{11.5} M_\odot$, with each halo thus being resolved by at least $\sim 10^5$ particles. The typical present-day stellar mass of central galaxies hosted by haloes with $M_{200} \simeq 10^{11.5} M_\odot$ is $M_\star \simeq 10^{9.5} M_\odot$; as shown by Schaye et al. (2015), present-day galaxies in EAGLE with at least this mass exhibit a passive fraction that broadly agrees with observational measurements.

I compute the spherical overdensity mass (M_{200} , Lacey & Cole, 1994) of each halo about its most-bound particle, such that the mean density enclosed within a sphere of radius r_{200} is 200 times the critical density, ρ_c . More generally, halo properties are computed by aggregating the properties of all particles of the relevant type that reside within an appropriate aperture. I compute the inner halo binding energy by summing the binding energies of all particles within r_{2500} that comprise each halo's counterpart in the DMONLY-L100N1504 simulation. In common with a number of other studies based on the analysis of EAGLE (e.g. Schaller et al., 2015a; Matthee & Schaye, 2019), I use this ‘intrinsic’ binding energy, E_{DMO}^{2500} , in order to eliminate the influence of dissipative baryonic processes. The dissipation of baryons in, and their ejection from, the progenitors of haloes throughout their formation and assembly can markedly influence their structure, potentially masking or exaggerating the influence of the intrinsic properties of the haloes. I pair haloes with their counterparts using the bijective particle matching algorithm described by Schaller et al. (2015a), which successfully pairs 3411 of the 3543 haloes (96 percent) satisfying $M_{200} > 10^{11.5} M_\odot$ in Ref-L100N1504. Unpaired haloes are discarded, thus ensuring the same sample of haloes is used throughout.

Following Schaye et al. (2015), I compute the properties of central galaxies by aggregating the properties of the relevant particles that reside within 30 pkpc of the halo centre. I equate the BH mass of galaxies, M_{BH} , to the mass of their most-massive BH particle, which is almost exclusively coincident with the halo centre.

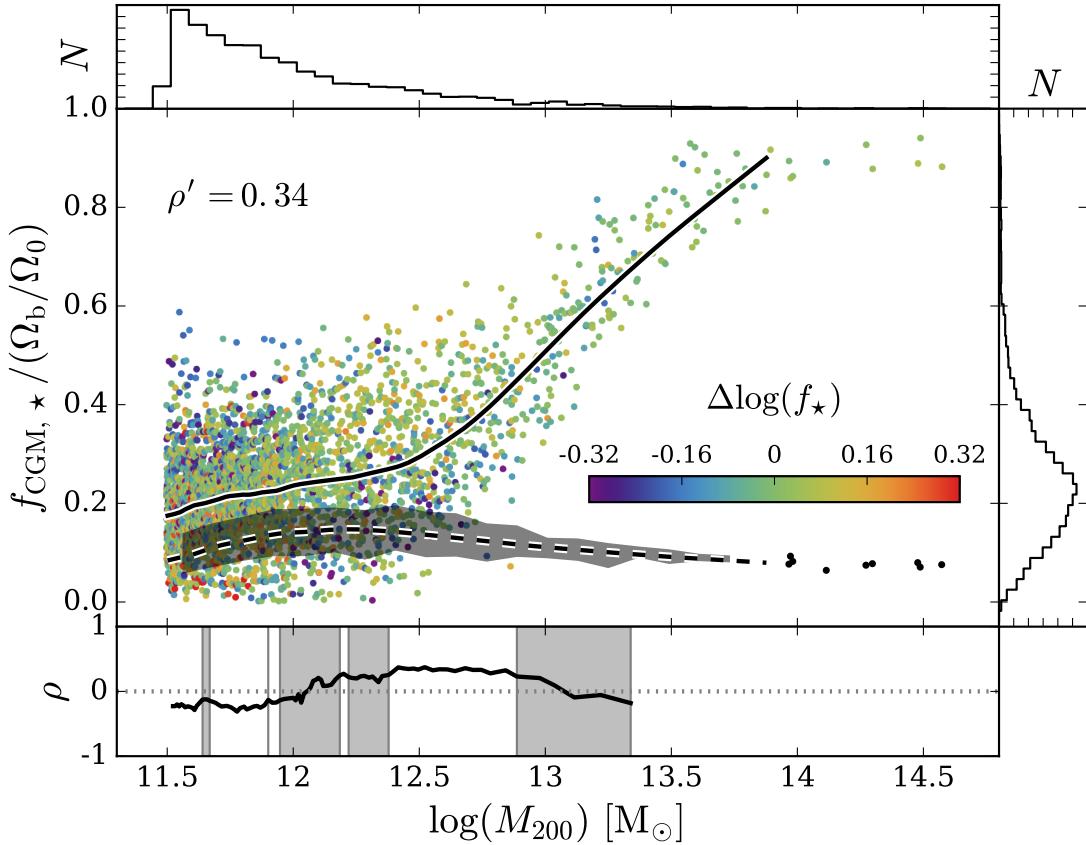


FIGURE 3.1: Present-day halo gas fractions, f_{CGM} , as a function of halo mass, M_{200} . Histograms of M_{200} and f_{CGM} are shown above and to the right of the main panel, respectively. The solid curve denotes the running median, whilst the dashed curve denotes the running median of the total stellar fraction of the halo, f_{\star} . The shaded region shows the 10th – 90th percentile scatter of f_{\star} . Symbols are coloured by the residuals of the relationship between f_{\star} and M_{200} , $\Delta \log_{10} f_{\star}$. The lower panel shows the running value of the Spearman rank coefficient, ρ , for the $\Delta f_{\text{CGM}} - \Delta \log_{10} f_{\star}$ relation. Shading denotes where the recovered correlation is not significant ($p > 0.01$).

3.3 The origin of scatter in halo gas fractions

In Fig. 3.1 I show, as a fraction of their mass M_{200} , the circumgalactic gas fraction, $f_{\text{CGM}} \equiv M_{\text{gas}}(r < r_{200})/M_{200}$, normalised by the cosmic baryon fraction, of present-day haloes in the Ref-L100N1504 simulation². This definition includes the contribution of interstellar gas, but this is in general a small fraction of the halo gas mass. The solid black line shows the running median of the gas fraction, $\tilde{f}_{\text{CGM}}(M_{200})$, computed via the locally-weighted scatterplot smoothing method (LOWESS, e.g. Cleveland, 1979). There is considerable scatter in f_{CGM} in relatively low-mass haloes, which declines for $M_{200} \gtrsim 10^{13} \text{ M}_\odot$.

²The baryon and stellar fractions of EAGLE haloes were presented and discussed by Schaller et al. (2015a).

For reference, the running median of the total stellar mass fraction of the halo, $\tilde{f}_*(M_{200})$ is also shown as a dashed black curve, where $f_* \equiv M_*(r < r_{200})/M_{200}$. The shaded region about this curve denotes the 10th–90th percentile scatter of f_* . The LOWESS curves are plotted within the interval for which there are at least 10 measurements at both higher and lower M_{200} ; in poorly-sampled high-mass bins, halo stellar fractions are plotted as individual black dots. Histograms of M_{200} and f_{CGM} are shown above and to the right, respectively, of the main panel. Gas fractions transition from $\simeq 0.3\Omega_b/\Omega_0$ below $M_{200} \simeq 10^{12.5} M_\odot$, rising steadily towards $\simeq 0.9\Omega_b/\Omega_0$ at $M_{200} \simeq 10^{14} M_\odot$, beyond which the trend flattens. This interval therefore represents a transition regime between EAGLE’s relatively gas-poor, low-mass haloes and their gas-rich, high-mass counterparts. Present-day $\sim L^*$ galaxies, with stellar mass similar to that of the Milky Way ($M_* \simeq 6 \times 10^{10} M_\odot$, e.g. [Bland-Hawthorn & Gerhard, 2016](#)) are thought to be hosted by haloes with mass $M_{200} \simeq 10^{12.5} M_\odot$ (e.g. [Moster et al., 2013](#))

Symbols are coloured by the residuals of the relationship between the stellar mass fraction and the halo mass, i.e for the i^{th} halo, $\Delta \log_{10} f_{*,i} = \log_{10} f_{*,i} - \log_{10} \tilde{f}_*(M_{200,i})$. Haloes denoted by red (blue) points therefore have a greater (lower) stellar mass fraction than is typical for their halo mass. Inspection of the symbol colours indicates that Δf_{CGM} and $\Delta \log_{10} f_*$ are not strongly correlated at any mass scale. I quantify the strength of the correlation with the Spearman rank correlation coefficient, ρ . Since the correlations can in principle exhibit a strong dependence on halo mass, I compute ‘running’ correlation coefficients from halo-mass ordered sub-samples. For bins whose median halo mass exhibits $M_{200} < 10^{12} M_\odot$, I use samples of 300 haloes with starting ranks separated by 50 haloes (i.e. 1-300, 51-350, 101-400 etc), otherwise I obtain superior sampling of the high-mass range with bins of 100 haloes separated by 25 haloes³. This diagnostic is shown in the bottom panel of Fig. 3.1. In this and subsequent figures I shade regions where the p -value exceeds 0.01, corresponding to $< 2.3\sigma$ confidence, to highlight where the recovered correlation is not significant. The correlation coefficient for the $\Delta f_{\text{CGM}} - \Delta \log_{10} f_*$ is relatively low ($|\rho| \lesssim 0.3$) at all halo masses, and is recovered with low confidence for much of the halo mass range. The coefficient for the 106 haloes in a 0.1 dex window about $M_{200} = 10^{12.5} M_\odot$ can however be recovered with significance, and its value, which for reference I denote ρ' , is 0.34. This constitutes a weak-to-moderate correlation for a narrow range in M_{200} , but more broadly it is evident that the diversity of the gas fractions exhibited by present-day $\sim L^*$ haloes does not emerge primarily as a consequence of some haloes converting more of their gas into stars than others. I note however that, as one approaches the ‘closed box’ regime of massive haloes with near-unity baryon fractions, one should expect a correlation between the scatter in $f_{\text{CGM}}(M_{200})$ and $f_*(M_{200})$ to emerge ([Farahi et al., 2018](#)).

³The resulting coefficients are not strongly sensitive to these choices.

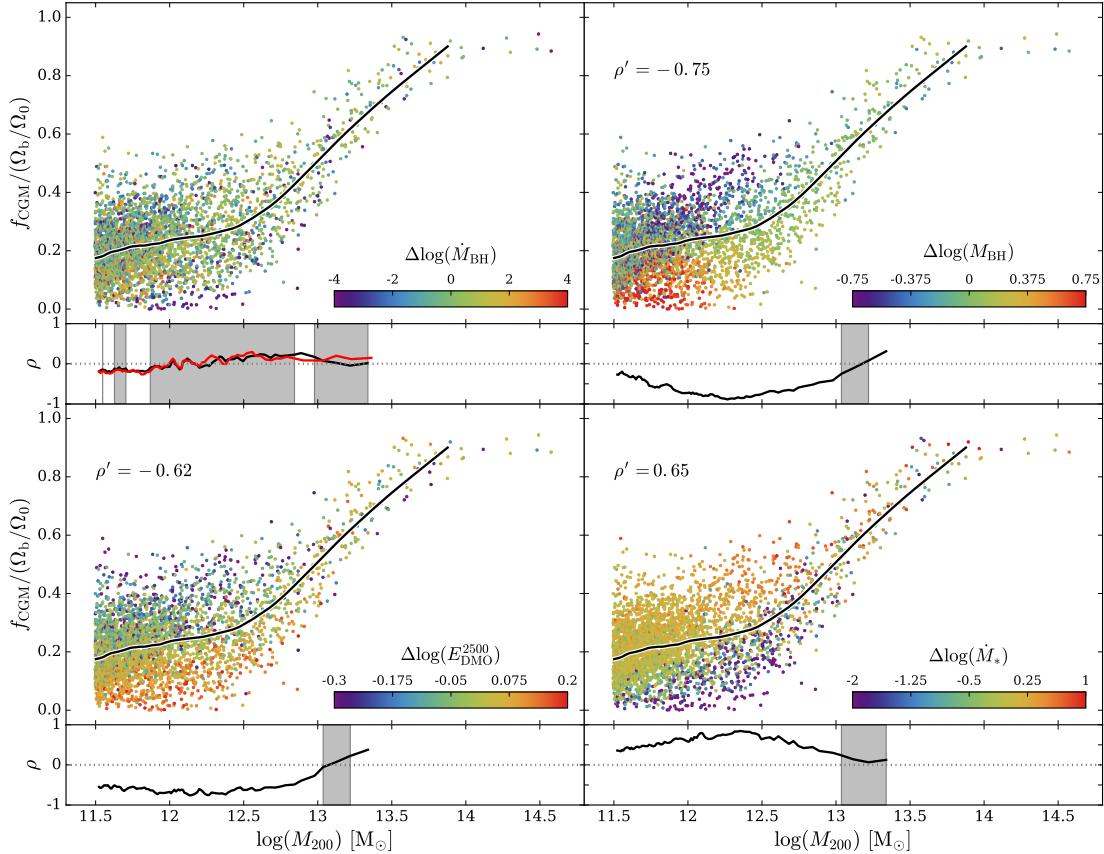


FIGURE 3.2: Present-day halo gas fractions, f_{CGM} , as a function of halo mass, M_{200} , with the solid curve denoting the running median. In each of the main panels, the symbols are coloured by residuals about the relationships of various properties as a function of halo mass: the accretion rate of the central BH (\dot{M}_{BH} , top left), the mass of the central BH (M_{BH} , top right), the intrinsic binding energy of the inner halo (E_{DMO}^{2500} , bottom-left), and the star formation rate of the central galaxy (M_{\star} , bottom-right). Beneath the main panels I show the running value of the Spearman rank correlation coefficient for the relationships between these residuals and Δf_{CGM} , the residuals about the $f_{\text{CGM}} - M_{200}$ relation. The red curve in the upper-left plot corresponds to the running ρ recovered when smoothing the BH accretion rate over a 100 Myr window. Shading denotes regions for which the correlation is recovered at low significance ($p > 0.01$). In the cases where correlations are significant, I quote the Spearman coefficient, ρ' , of the correlation computed for haloes within a 0.1dex window about $\log_{10} M_{200} [\text{M}_\odot] = 12.5$.

The rest-mass energy liberated throughout the growth of central BHs is comparable to the gravitational binding energy of the halo (Silk & Rees, 1998a, see also Oppenheimer 2018), and is thus expected to foster the expulsion of gas from galaxies and their haloes. It is reasonable to surmise then that AGN feedback should have a significant influence on the gas fraction of haloes (e.g. Puchwein et al., 2008; Bower et al., 2008; McCarthy et al., 2010, 2011; Bocquet et al., 2016; Pillepich et al., 2018a). In Fig. 3.2 I again show the present-day $f_{\text{CGM}} - M_{200}$ relation of Ref-L100N1504, with the symbols in the top row coloured by the residuals about the median relation between the characteristics of central BHs and halo mass. In the top-left panel they are coloured by the residuals about the running median of the BH accretion rate as a function of halo mass, $\Delta \log_{10} \dot{M}_{\text{BH}}$. As in

the case of the $\Delta \log_{10} f_\star$ correlation shown in Fig. 3.1, the Spearman rank correlation coefficient is relatively low ($|\rho| < 0.3$) at all halo masses, and is recovered with low confidence for much of the sampled mass range, including at $M_{200} = 10^{12.5} M_\odot$, hence I am unable to quote a significant value of ρ' for this diagnostic. Scatter in halo gas fractions is therefore not strongly correlated with scatter in the BH accretion rate at any M_{200} . McAlpine et al. (2017) recently showed that the accretion rate of BHs in EAGLE can vary by orders of magnitude on very short timescales ($\lesssim 10^5$ yr), so I have repeated this test after time-averaging the BH accretion rate over the preceding 100 Myr. I find similar results with this definition of \dot{M}_{BH} , as is evident from its running value of ρ , shown as a red curve.

Prior analyses of the EAGLE simulations (Bower et al., 2017; McAlpine et al., 2018) have revealed that the development of a hot ($T \gtrsim 10^6$ K), quasi-hydrostatic CGM in haloes with mass $M_{200} \gtrsim 10^{12} M_\odot$ inhibits the buoyant transport away from the galaxy of gas ejected from the interstellar medium (ISM) in stellar feedback-driven outflows. The resulting build-up of gas triggers non-linear growth of the BH, which accretes rapidly until the feedback associated with its growth becomes the dominant means of regulating the inflow of gas onto the galaxy. McCarthy et al. (2011) argue that the expulsion of gas from the progenitors of group- and cluster-scale haloes, which accompanies this onset of BH feedback, occurs primarily at early cosmic epochs ($1 \lesssim z \lesssim 3$) when their central BHs accreted most of their mass. I therefore colour the symbols of the top-right panel of Fig. 3.2 by the residuals about the running median of the BH mass as a function of halo mass, $\Delta \log_{10} M_{\text{BH}}$. In this case, the colouring reveals a striking negative correlation between the gas fraction (at fixed halo mass) and the BH mass, such that haloes whose central galaxies host atypically-massive central BHs exhibit systematically low gas fractions, and vice versa. The visual impression is corroborated by the Spearman rank correlation coefficient, which is significant and negative for all $10^{11.5} < M_{200} \lesssim 10^{13} M_\odot$. The coefficient for haloes with $M_{200} \simeq 10^{12.5} M_\odot$ is very strong, $\rho' = -0.75$. These results indicate that the halo gas fractions of $\sim L^*$ galaxies are regulated primarily by the evolutionary stage of their central BHs.

I turn next to the origin of the diversity of BH masses at fixed M_{200} . This question was explored by Booth & Schaye (2010, 2011) using the OWLS simulations (Schaye et al., 2010), who concluded that BH mass is governed primarily by the binding energy of the inner halo. I therefore colour the symbols of the bottom-left panel of Fig. 3.2 by the residuals about the running median of the binding energy of the halo as a function of halo mass. As motivated in Section 3.2, I use the intrinsic binding energy, E_{DMO}^{2500} , i.e. that recovered from each halo's counterpart in the DMONLY simulation measured within r_{2500} . This eliminates the effects of dissipative physical processes, which could potentially mask or exaggerate any correlations induced by the intrinsic binding energy of the halo.

The colouring reveals a striking negative correlation between the scatter in f_{CGM} and that of the binding energy of the inner halo⁴. The Spearman rank correlation coefficient is significant and negative for all $10^{11.5} < M_{200} \lesssim 10^{13} \text{ M}_\odot$, and exhibits a broad, strong minimum (characterised by $\rho < -0.5$ recovered at $p < 0.01$) for $M_{200} \lesssim 10^{12.5} \text{ M}_\odot$. The coefficient for haloes with $M_{200} \simeq 10^{12.5} \text{ M}_\odot$ is strong, with $\rho' = -0.62$. At fixed mass, more tightly-bound haloes require more energy to unbind gas from the inner halo, so their central BHs must grow to be more massive, reaching a higher peak luminosity, and thus they ultimately eject a greater fraction of the halo gas beyond r_{200} ⁵. More tightly-bound haloes at fixed mass also tend to be those with a higher concentration and an earlier formation time (e.g. Navarro et al., 2004); indeed Booth & Schaye (2010, 2011) found that halo concentration correlates with M_{BH} at fixed M_{200} . I have therefore examined the relationships between the scatter in gas fractions at fixed halo mass and the scatter in the Navarro et al. (1997, ‘NFW’) concentration and the halo assembly lookback time (computed as per Qu et al., 2017) of each halo’s counterpart in the DMONLY-L100N1504 simulation, and recover negative correlations that are again significant, albeit slightly weaker than is the case for E_{DMO}^{2500} .

The binding energy of the halo in the DMONLY simulation is effectively encoded within the phase-space configuration of the initial conditions, and residuals of f_{CGM} correlate with similar strength, but over a wider range in halo mass, to those of E_{DMO}^{2500} than with those of M_{BH} . Scatter in E_{DMO}^{2500} at fixed M_{200} might reasonably then be considered as the fundamental cosmological origin of the scatter in f_{CGM} at fixed M_{200} . However, the influence of the binding energy is physically ‘transmitted’ to the gas fraction by ejective feedback. The necessity of this conduit can be demonstrated using the NOAGN-L050N0752 EAGLE simulation, in which AGN feedback is disabled. Examination of the relationship between Δf_{CGM} and $\Delta \log_{10} E_{\text{DMO}}^{2500}$ in this simulation reveals a significantly weaker correlation than in the Reference simulation, which is driven largely by the lowest-mass haloes in the sample. The origin of the correlation here differs with respect to the Reference simulation; haloes with greater binding energy (at fixed mass) exhibit lower gas fractions in the NOAGN simulation because the positive correlation of the $\Delta f_\star - \Delta \log_{10} E_{\text{DMO}}^{2500}$ relation is much stronger than in the Reference simulation. The halo gas fraction is therefore depleted because of the condensation of gas into stars rather than its ejection by feedback.

⁴Using E_{DMO}^{500} or E_{DMO}^{200} instead yields similar results.

⁵This effect underpins the importance of using E_{DMO}^{2500} rather than E_{Ref}^{2500} . In general the dissipation of gas into stars results in $E_{\text{Ref}}^{2500} > E_{\text{DMO}}^{2500}$, but crucially the fractional difference in the two measures at fixed halo mass anti-correlates strongly with E_{DMO}^{2500} , because the formation of more massive BHs at the centres of haloes with greater E_{DMO}^{2500} enables them to eject a greater fraction of their gas. The use of E_{Ref}^{2500} therefore partially masks the underlying correlation between the inner binding energy and f_{CGM} .

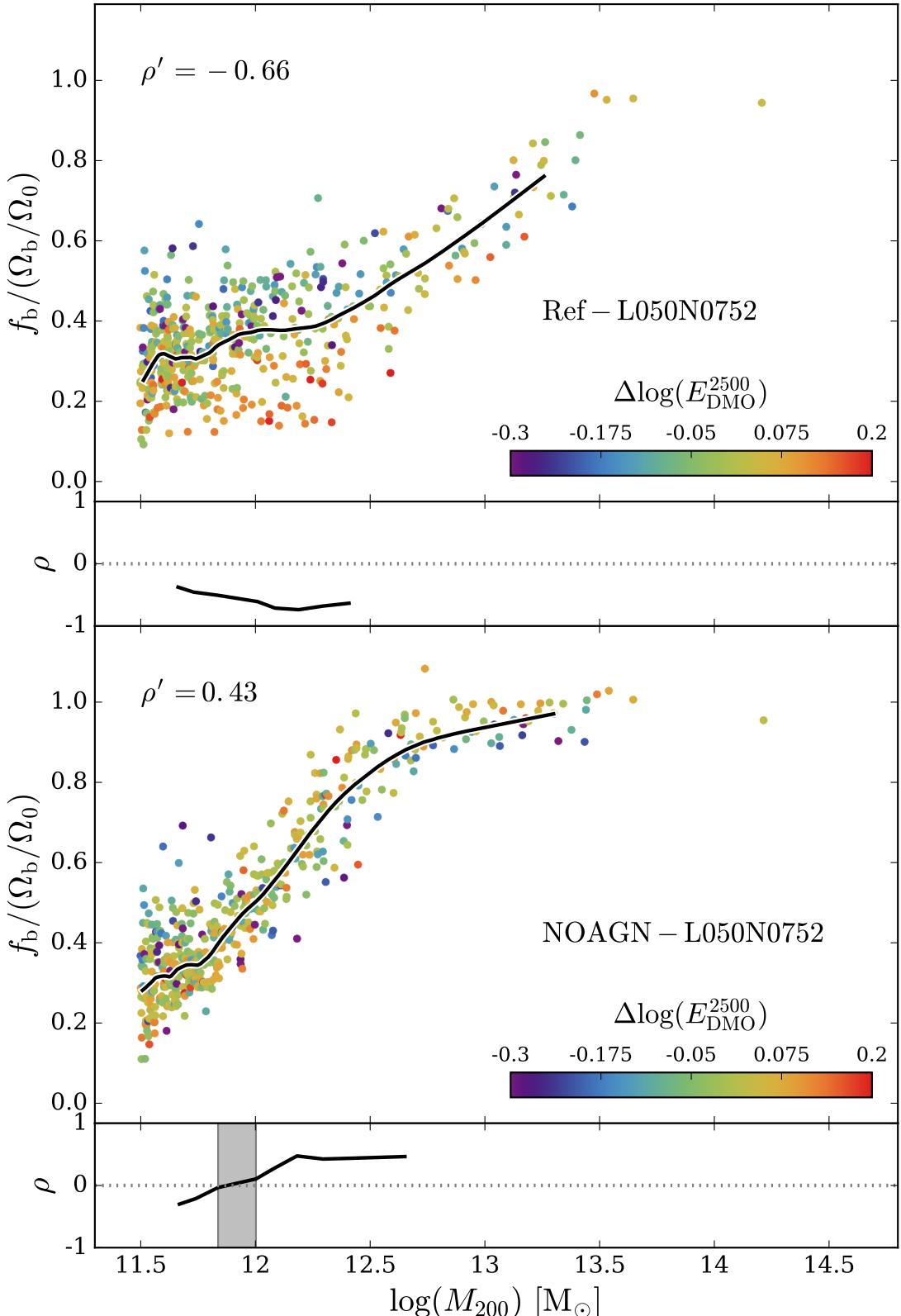


FIGURE 3.3: Present-day halo baryon fractions, f_b , as a function of halo mass, M_{200} , in the Ref-L050N0752 (*top*) and NOAGN-L050N0752 (*bottom*) simulations. In each case the solid curve denotes the running median, and the symbols are coloured by the residuals about the running median of the $E_{\text{DMO}}^{2500} - M_{200}$ relation, $\Delta E_{\text{DMO}}^{2500}$. Sub-panels show the running value of the Spearman rank correlation coefficient, ρ , for the relationships between these residuals and Δf_b , the residuals about the $f_b - M_{200}$ relation. Shading denotes regions for which the correlation is recovered at low significance ($p > 0.01$).

The transmission of the influence of halo binding energy via AGN feedback can be more concisely demonstrated by examination of the halo baryon fraction, $f_b \equiv [M_{\text{gas}}(r < r_{200}) + M_*(r < r_{200})]/M_{200}$. In Fig. 3.3 I show the present-day $f_b - M_{200}$ relations for Ref-L050N0752 (top) and NOAGN-L050N0752 (bottom), and colour the symbols by $\Delta \log_{10} E_{\text{DMO}}^{2500}$. Using the halo baryon fraction rather than the halo gas fraction takes account of the aforementioned additional condensation of gas into stars in the NOAGN simulation. In the absence of AGN feedback, haloes with $M_{200} \gtrsim 10^{12} M_\odot$ retain a significantly greater fraction of their baryons, as stellar feedback is unable to expel gas from massive haloes. In Ref-L050N0752, the $\Delta f_b - \Delta \log_{10} E_{\text{DMO}}^{2500}$ correlation for haloes with $M_{200} = 10^{11.5} - 10^{13} M_\odot$ is strong, significant and negative for $M_{200} > 10^{11.8} M_\odot$, with $\rho' = -0.66^6$. In NOAGN-L050N0752, the correlation is again mildly negative for $M_{200} \lesssim 10^{12} M_\odot$, but becomes moderately positive for more massive haloes, with $\rho' = 0.43$. Therefore, in the absence of AGN feedback, the intrinsic binding energy of the inner region of haloes has a markedly different influence on the baryon fractions for $M_{200} \gtrsim 10^{12} M_\odot$. This notwithstanding, I reiterate that residuals about the $f_{\text{CGM}} - M_{200}$ relation in the Reference model correlate strongly with those about the $E_{\text{DMO}}^{2500} - M_{200}$ relation over a broader range in halo mass than with those about the $M_{\text{BH}} - M_{200}$ relation, indicating that the binding energy may also influence halo gas fractions via other mechanisms, such as formation time.

It is interesting to note that the correlations between Δf_{CGM} and each of $\Delta \log_{10} M_{\text{BH}}$ and $\Delta \log_{10} E_{\text{DMO}}^{2500}$ change sign for $M_{200} \gtrsim 10^{13} M_\odot$, and become positive. The reversal of these correlations is likely a consequence of the declining efficiency of gas expulsion by AGN feedback in the most-massive haloes; as feedback becomes unable to eject gas from the assembling halo, a greater central binding energy only serves to inhibit gas expulsion.

[Matthee & Schaye \(2019\)](#) recently demonstrated that the star formation histories of EAGLE galaxies hosted by haloes with earlier formation times (at fixed halo mass) are systematically shifted to earlier cosmic epochs. A consequence of this effect is that these galaxies exhibit systematically lower present-day star formation rates (SFRs, \dot{M}_*). Since haloes with early formation times generally exhibit greater central binding energies, one might expect that scatter about the $f_{\text{CGM}} - M_{200}$ relation will correlate with the SFR. Returning briefly to Fig. 3.2, I show in the bottom-right panel the $f_{\text{CGM}} - M_{200}$ relation with symbols coloured by the residuals about the running median of the SFR as a function of halo mass, $\Delta \log_{10} \dot{M}_*$, which indeed reveals a striking positive correlation between the residuals about the running medians of the $f_{\text{CGM}} - M_{200}$ and $\dot{M}_* - M_{200}$ relations. The running coefficient of the Spearman rank correlation is positive for all

⁶I compute ρ' for $L = 50$ cMpc simulations within the larger window of 0.5 dex around $M_{200} = 10^{12.5} M_\odot$, to maintain adequate sampling.

halo masses, and for $M_{200} \lesssim 10^{13} M_\odot$ it resembles the inverse of that of the $\Delta f_{\text{CGM}} - \Delta \log_{10} M_{\text{BH}}$ correlation. The coefficient for haloes with $M_{200} \simeq 10^{12.5} M_\odot$ is $\rho' = 0.65$, highlighting the strength of this correlation. More gas-rich haloes at fixed mass, besides exhibiting relatively undermassive BHs, therefore also exhibit an elevated SFR.

A close connection between the SFR of galaxies and their interstellar gas content is well established by observations (e.g. [Kennicutt, 1998](#)), a finding whose reproduction and interpretation continues to attract considerable analytic and numerical effort (e.g. [Thompson et al., 2005](#); [Krumholz & Tan, 2007](#); [Semenov et al., 2016](#); [Orr et al., 2018](#)). To my knowledge, however, a correlation between the gas fractions of haloes (of similar mass) and the SFR of their central galaxies has not been demonstrated previously. The correlation of both the SFR and the BH mass with the halo gas fraction at fixed halo mass is helpful from the perspective of scrutinizing the predictions advanced here, as the SFR can be inferred from a diverse range of photometric and spectroscopic diagnostics; I turn to this scrutiny in the next section.

The correlation is also of intrinsic interest. Surveys of galaxies and associated absorption systems have revealed a positive correlation between the SFR of galaxies and the column density of the absorbers (e.g. [Chen et al., 2010](#); [Lan et al., 2014](#); [Rubin et al., 2018](#)), and a popular interpretation of this correlation is that the absorption column densities are enhanced by outflows driven by stellar feedback. The analysis presented here suggests that the correlation may not be causal, but rather a consequence of the negative correlation of both the halo gas fraction, and the SFR, with the halo binding energy at fixed halo mass. The first correlation is a consequence of more tightly-bound haloes requiring more massive BHs to unbind gas from their inner regions, whilst the second is due to more tightly-bound haloes collapsing earlier, shifting the growth of their central galaxy and BH (and the associated expulsion of their halo gas) to earlier times. Alternatively, the star formation activity of the central galaxy could be fundamentally tied to the expulsion of halo gas; I investigate this hypothesis in detail in Chapter 4.

3.4 Testing via complementary observables

The influence of the central BH mass on the scatter of halo gas fractions at fixed halo mass is an unambiguous prediction of the EAGLE simulations. However, it is not one that is trivial to confront with observations, since one requires measurements of f_{CGM} and M_{200} , and dynamical measurements of BH masses, for a large sample of galaxies. I therefore briefly explore in this section whether it might be possible to test the predictions of Section 3.3 with complementary observational diagnostics provided by extant or forthcoming facilities.

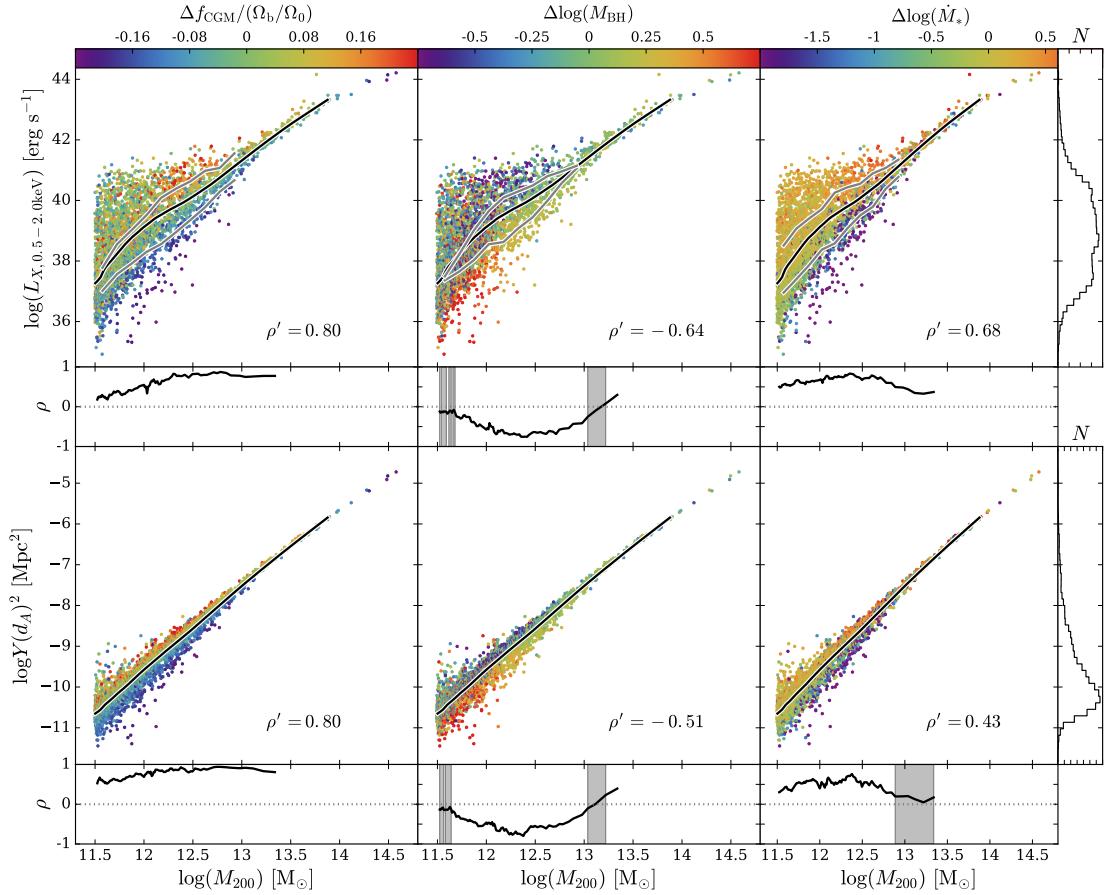


FIGURE 3.4: Present-day diffuse soft X-ray luminosity (L_X , top) and thermal Sunyaev-Zel'dovich effect flux (Yd_A^2 , bottom), as a function of halo mass, M_{200} . Histograms of L_X and Yd_A^2 are shown to the right of the main panel. The solid curve denotes the running median of each quantity. In each column the symbols are coloured by the residuals about the running median of f_{CGM} (left), M_{BH} (centre) and \dot{M}_* (right) as a function of M_{200} . Beneath the main panels I show the running value of the Spearman rank correlation coefficient for the relationships between these residuals and Δf_{CGM} , the residuals about the $f_{\text{CGM}} - M_{200}$ relation. Shading denotes regions for which the correlation is recovered at low significance ($p > 0.01$), whilst ρ' denotes the value of the Spearman rank correlation coefficient computed for haloes with mass $M_{200} \simeq 10^{12.5} M_\odot$. The grey curves on the upper row show the median of the $L_X - M_{200}$ relations of the halo sub-samples comprising the upper and lower quartiles of the diagnostic used for the symbol colouring.

I require diagnostics for which scatter about their running median as a function of halo mass correlates with scatter about the $f_{\text{CGM}} - M_{200}$ relation. I first consider the diffuse, soft ($0.5 \lesssim E_X \lesssim 2.0 \text{ keV}$) X-ray luminosity of the hot ($T \gtrsim 10^6 \text{ K}$), collisionally-ionized component of halo gas. Characterisation of the properties of the gas in haloes less massive than those of galaxy groups ($M_{500} \sim 10^{13} M_\odot$, $kT \sim 1 \text{ keV}$) remains challenging, with the extended hot CGM of only a handful of galaxies having been convincingly detected and characterised beyond the optical envelope of the galaxy (e.g. [Dai et al., 2012](#); [Bogdan et al., 2013b, 2017](#); [Li et al., 2016, 2017](#)), but forthcoming and proposed X-ray observatories such as *Athena* and *Lynx* promise to make such detections more

commonplace. Moreover, stacking low spatial resolution *ROSAT* All-Sky Survey X-ray maps about the coordinates of local, optically-selected galaxies has proven an effective means of characterising the relationship between galaxies and their gas content (Anderson et al., 2015; Wang et al., 2016). I therefore compute the soft X-ray luminosity of each halo by coupling the physical properties of its constituent gas particles (i.e. those within r_{200}) to the Astrophysical Plasma Emission Code (APEC, Smith et al., 2001), using the techniques described in Section 2.6.

Stacking about the coordinates of optically-selected galaxies has also recently been used to characterise the hot gas content of haloes via measurement of the thermal Sunyaev-Zel'dovich (tSZ) effect, the inverse Compton scattering of cosmic microwave background (CMB) photons by energetic electrons within the hot, ionized CGM (e.g. Planck Collaboration et al., 2013; Greco et al., 2015). The tSZ ‘flux’ can be defined as the Compton- y parameter integrated over the solid angle of the halo and is thus proportional to the total energy of the hot gas:

$$Y(< r_{200})d_A(z)^2 = \frac{\sigma_T}{m_e c^2} \int_0^{r_{200}} 4\pi P_e(r, z) r^2 dr, \quad (3.1)$$

where d_A is the angular diameter distance to the halo, σ_T is the Thomson cross-section, m_e the electron rest mass, and $P_e = n_e k_B T_e$ is the electron pressure with k_B being the Boltzmann constant. The flux therefore scales with the density of the hot gas, rather than the square of its density, as is the case for the collisional mechanisms that dominate the X-ray emissivity of diffuse plasmas. I compute Y by summing the contributions of the gas particles associated with each halo, as per McCarthy et al. (2017). Star-forming gas particles (i.e. those comprising the ISM) are assumed to be neutral and do not contribute to the flux.

The rows of Fig. 3.4 show the present-day $L_X - M_{200}$ (top) and $Yd_A^2 - M_{200}$ (bottom) relations⁷. Histograms of L_X and Yd_A^2 are shown to the right of the main panel. In each panel the solid black line denotes the LOWESS running median. In the left-hand column symbols are coloured by their residuals with respect to the running median of the $f_{\text{CGM}} - M_{200}$ relation. In both cases the colouring highlights that elevated values of the observable correspond to elevated values of f_{CGM} . The running values of the Spearman rank correlation coefficient demonstrate that scatter about both proxies correlates strongly, significantly and positively with that about f_{CGM} for all M_{200} . The X-ray luminosity is somewhat noisier than the tSZ flux, which is unsurprising since it is also

⁷Since I am concerned with the response of L_X and Yd_A^2 to deviations from the median scaling relations relating the properties of galaxies and their gaseous haloes to halo mass, precise correspondence between the properties of the simulated haloes and observational measurements is not crucial. Reasonable correspondence is however necessary to engender confidence in the realism of the simulations; I therefore compare the simulated scaling relations of L_X and Yd_A^2 as a function of halo mass with observational measurements in Appendix A.

sensitive to the metallicity of the halo (see e.g. Crain et al., 2013), and is more sensitive than the tSZ flux to the structure of the CGM.

The X-ray luminosity remains an attractive observable however, owing in particular to the dynamic range it displays: for haloes with $M_{200} \simeq 10^{12.5} M_\odot$ the 10th–90th percentile range spans 1.54 decades in L_X . Measurements of the X-ray luminosity therefore afford the opportunity to highlight readily the diversity of halo properties at fixed M_{200} . I demonstrate this quantitatively by showing, as grey curves, the median L_X of the subsets of haloes representing the upper and lower quartiles of the $f_{\text{CGM}} - M_{200}$ relation, in 10 bins of halo mass in the interval $10^{11.5} < M_{200} < 10^{13} M_\odot$. The separation of the subsets peaks at $M_{200} = 10^{12.2} M_\odot$, with the gas-rich subset of haloes exhibiting a median L_X that is 1.5 dex greater than that of the gas-poor subset.

To use L_X and Yd_A^2 to test the influence of BH mass on halo gas fractions, they must respond to scatter about the $M_{\text{BH}} - M_{200}$ relation in a similar fashion to f_{CGM} . In the centre column of Fig. 3.4, the symbols are coloured by $\Delta \log_{10} M_{\text{BH}}$. Similar to the top-right panel of Fig. 3.2, residuals about the relations can be seen to correlate negatively with $\Delta \log_{10} M_{\text{BH}}$. The running values of the Spearman rank correlation coefficient, shown below the main panels, behave similarly and become more strongly negative from relatively low masses to a peak at $M_{200} \simeq 10^{12.4} M_\odot$. Therefore, the haloes of galaxies with central BHs that are more (less) massive than is typical for their mass are systematically ‘faint’ (bright) in both L_X and Yd_A^2 . I again highlight the dynamic range of the X-ray luminosity and show as grey curves in the upper-centre panel the median L_X of the subsets representing the upper and lower quartiles of the $M_{\text{BH}} - M_{200}$ relation. The subsets are again most strongly separated at $M_{200} = 10^{12.2} M_\odot$, with the subset of haloes with under-massive BHs exhibiting a median $L_X = 1.2 \times 10^{40} \text{ ergs}^{-1}$, which is 1.4 dex greater than that of the over-massive BH subset, $L_X = 4.5 \times 10^{38} \text{ ergs}^{-1}$.

At present dynamical measurements of the masses of BHs are available for only $\sim 10^2$ galaxies (e.g. Kormendy & Ho, 2013; McConnell & Ma, 2013), presenting an obstacle to observational scrutiny of the influence of BH mass on halo gas fractions. However, as noted in Section 3.3, the correlations of Δf_{CGM} with ΔM_{BH} and $\Delta \dot{M}_*$ are of similar strength (but opposite sign). Identification of this correlation in observations would therefore corroborate EAGLE’s predictions concerning the origin of scatter in the $f_{\text{CGM}} - M_{200}$ scaling relation, hence it is important to establish how the proxies for f_{CGM} respond to scatter in SFR at fixed M_{200} .

In the right-hand column of Fig. 3.4 the symbols are coloured by $\Delta \log_{10} \dot{M}_*$. Encouragingly, I find that the residuals about the $\dot{M}_* - M_{200}$ relation correlate positively with the residuals about both the $L_X - M_{200}$ and $Yd_A^2 - M_{200}$ relations, with the correlations being strong and significant for $M_{200} \lesssim 10^{12.7} M_\odot$. Therefore, the haloes of galaxies

that exhibit a SFR that is high (low) for their halo mass are systematically ‘bright’ (faint) in both proxies. The grey curves of the upper-right panel show the median L_X of the subsets representing the upper and lower quartiles of the $\dot{M}_\star - M_{200}$ relation. The subsets are again most strongly separated at $M_{200} = 10^{12.2} M_\odot$, for which the upper quartile ($\dot{M}_\star > 1.8 M_\odot \text{ yr}^{-1}$) exhibit a median $L_X = 1.8 \times 10^{40} \text{ ergs}^{-1}$, which is 1.7 dex (a factor of $\simeq 50$) greater than that of the lower quartile ($\dot{M}_\star < 0.2 M_\odot \text{ yr}^{-1}$), $L_X = 3.8 \times 10^{38} \text{ ergs}^{-1}$.

3.5 Summary and discussion

I have examined the origin of scatter in the relationship between gas fraction, f_{CGM} , and mass, M_{200} , of the haloes with mass similar to those that host present-day $\sim L^*$ central galaxies ($10^{11.5} < M_{200} < 10^{13} M_\odot$) in the EAGLE simulations. I quantify the scatter by computing the difference between each halo’s gas fraction and the running median of the $f_{\text{CGM}} - M_{200}$ relation, $\tilde{f}_{\text{CGM}}(M_{200})$, constructed using the LOWESS locally-weighted scatterplot smoothing method.

My results are drawn primarily from the largest EAGLE simulation Ref-L100N1504, and its counterpart considering only collisionless gravitational dynamics, DMONLY-L100N1504. The parameters of the subgrid models governing feedback in the EAGLE Reference model were calibrated to ensure reproduction of key present-day properties of the galaxies, but the gaseous properties of galaxies and their haloes were not considered during the calibration and can be considered predictions of the simulations. I have also briefly studied the NOAGN-L050N0752 simulation in which AGN feedback is disabled, and its Reference model counterpart Ref-L050N0752.

My findings can be summarized as follows:

1. Scatter about the $f_{\text{CGM}} - M_{200}$ relation is not strongly correlated with residuals of the relationship between the stellar mass fraction of haloes and M_{200} . Low (high) halo gas fractions are therefore not generally a consequence of haloes having converted more (less) of their gas into stars throughout their assembly (Fig. 3.1).
2. Similarly, the scatter is neither strongly nor significantly correlated with the residuals of the relationship between the gas accretion rate of central BHs (whether measured instantaneously or time-averaged over 100 Myr) and M_{200} . Low (high) halo gas fractions are therefore not a consequence of relatively strong (weak) ongoing AGN feedback (Fig. 3.2, top-left).

3. The scatter correlates strongly, significantly and negatively with the residuals of the relationship between the present-day mass of central BHs and M_{200} . At $M_{200} = 10^{12.5} M_\odot$ the Spearman rank correlation coefficient is $\rho' = -0.75$. At fixed M_{200} therefore, galaxies that host more-massive central BHs reside within relatively gas-poor haloes, and vice versa. This suggests that the main cause of scatter in f_{CGM} at fixed M_{200} is differences in the mass of halo gas expelled by AGN feedback throughout the assembly of the halo (Fig. 3.2, top-right).
4. A corollary of (ii) is the implication that the scatter about the $f_{\text{CGM}} - M_{200}$ relation might be driven by a more fundamental process that fosters scatter in the $M_{\text{BH}} - M_{200}$ relation. [Booth & Schaye \(2010, 2011\)](#) previously highlighted with cosmological simulations that this scatter is driven by differences in the binding energy of haloes at fixed mass. I find that scatter about the $f_{\text{CGM}} - M_{200}$ relation indeed correlates strongly, significantly and negatively with the residuals of the $E_{\text{DMO}}^{2500} - M_{200}$ relation, where E_{DMO}^{2500} is the intrinsic binding energy of the halo, i.e. that which emerges in the absence of the dissipative physics of galaxy formation, measured within r_{2500} . This correlation is strong and significant over a broad range in M_{200} , and at $M_{200} = 10^{12.5} M_\odot$ the correlation coefficient is $\rho' = -0.62$. (Fig. 3.2, bottom-left).
5. Although reasonably interpreted as the fundamental origin of the scatter in f_{CGM} , the influence of the intrinsic binding energy of haloes is communicated via AGN-driven gas expulsion for $M_{200} \gtrsim 10^{12} M_\odot$. This is succinctly demonstrated by examination of the residuals about the $f_b - M_{200}$ and $E_{\text{DMO}}^{2500} - M_{200}$ relations (where f_b is the halo baryon fraction) in simulations with and without AGN feedback. In the Reference model simulation these residuals are strongly, significantly and negatively correlated for all M_{200} , whilst in the NOAGN model the correlation is weaker for relatively low-mass haloes ($M_{200} \lesssim 10^{12} M_\odot$), and becomes positive for more massive haloes (Fig. 3.3).
6. The scatter in $f_{\text{CGM}}(M_{200})$ correlates strongly, significantly and positively with the residuals of the relationship between the present-day SFR of central galaxies and M_{200} . The correlation is similar to that with ΔM_{BH} , but with opposite sign. At $M_{200} = 10^{12.5} M_\odot$ the correlation coefficient is $\rho' = 0.65$. Haloes with high (low) gas fractions for their mass therefore typically host central galaxies with high (low) SFRs. (Fig. 3.2, bottom-right).
7. I consider the diffuse soft X-ray luminosity of the hot component of halo gas (L_X) and the ‘flux’ of the thermal Sunyaev-Zel’dovich effect (Yd_A^2) as proxies for f_{CGM} . Scatter about the relation of these observables with M_{200} correlates positively with residuals of the $f_{\text{CGM}} - M_{200}$ relation, such that variations in the halo gas fraction

at fixed halo mass are echoed by the two observables. Residuals about the running median of both proxies also correlate negatively with scatter about the median BH mass at fixed M_{200} . At $M_{200} = 10^{12.5} M_\odot$ the associated correlation coefficients for L_X and Yd_A^2 are $\rho' = -0.62$ and $\rho' = -0.51$, respectively. This highlights that they respond to variations in BH mass in a similar fashion to f_{CGM} (Fig. 3.4, left and centre columns).

8. Scatter about the $L_X - M_{200}$ and $Yd_A^2 - M_{200}$ relations also correlates strongly and positively with scatter about the median SFR at fixed M_{200} . At $M_{200} = 10^{12.5} M_\odot$ the associated correlation coefficients for L_X and Yd_A^2 are $\rho' = 0.68$ and $\rho' = 0.43$, respectively. These correlations afford a route to observational scrutiny of the predictions of the simulations advanced here, without the need to acquire dynamical BH mass measurements for a large sample of galaxies. The simulations indicate that, for galaxies hosted by haloes of $M_{200} \simeq 10^{12.2} M_\odot$, the median X-ray luminosity of those with $\dot{M}_* > 1.8 M_\odot \text{ yr}^{-1}$ (the upper quartile in SFR) is a factor $\simeq 50$ higher than for those comprising the lower quartile ($\dot{M}_* < 0.2 M_\odot \text{ yr}^{-1}$, Fig. 3.4, right-hand column).

The discovery of scaling relations connecting central BHs (with optical accretion discs of scale $\sim 10^{-2} \text{ pc}$) with the properties of galaxies (on scales $\sim 10^3 \text{ pc}$, e.g. [Magorrian et al., 1998](#); [Kormendy & Ho, 2013](#)) has focussed intense interest on the possibility of an intimate physical connection between the two. The release of rest mass energy from the accretion of gas onto BHs has long been advocated as a means to regulate cooling flows onto massive galaxies (e.g. [Silk & Rees, 1998a](#)) at the centres of groups and clusters (on scales $\sim 10^5 - 10^6 \text{ pc}$, e.g. [Binney & Tabor, 1995](#)). Cosmological simulations still lack the physics and resolution required to capture the full complexity of the coupling between these phenomena across such a broad dynamic range, but my findings nonetheless indicate that central BHs can also have a significant influence on the structure and content of the CGM. It is an unambiguous prediction of the EAGLE simulations that scatter in the central BH mass, at fixed halo mass, markedly influences the gas fractions of the haloes that host present-day $\sim L^*$ central galaxies.

My findings also suggest that it is possible to corroborate or falsify EAGLE's predictions for the origin of scatter about the $f_{\text{CGM}} - M_{200}$ relation, using extant or forthcoming observations. I posit that the locally-brightest galaxy (LBG) sample from the New York University Value Added Galaxy Catalogue (Blanton et al. 2005, VAGC), based on the seventh data release of the Sloan Digital Sky Survey (SDSS/DR7, Abazajian et al. 2009), is well-suited to this purpose. It has been used previously as the basis for stacked measurements of both diffuse X-ray luminosity ([Anderson et al., 2015](#)) from the *ROSAT* All-Sky Survey and the tSZ flux ([Planck Collaboration et al., 2013](#); [Greco et al.,](#)

2015) from *Planck* maps. The acquisition of well-characterised rotation curves for this sample would further enable the identification of sub-samples with similar dynamical masses. At present, X-ray luminosity remains the preferred proxy for f_{CGM} , as the ~ 10 arcmin beam of *Planck*'s 100 GHz maps corresponds to a scale comparable to or larger than r_{200} for the majority of the LBG sample. However, characterisation of the tSZ flux within lower-mass haloes may soon be possible with the advent of next-generation, high-resolution, ground-based CMB experiments.

A further, complementary means of assessing variations in f_{CGM} is to search for variations in the column densities of absorption systems, revealed by the intersection of bright quasars with the CGM of a sample of nearby galaxies. Such an approach would be similar in spirit, but different in detail, to the COS-AGN survey (Berg et al., 2018). Extending the COS-Halos survey (Tumlinson et al., 2013), COS-AGN enabled a controlled comparison of the absorption systems associated with galaxies with and without AGN. Consistent with my finding that the present-day BH accretion rate has little impact on halo gas fractions at fixed M_{200} , Berg et al. (2018) found no significant differences between these samples when examining the equivalent width distributions of absorption systems tracing the inner CGM. To probe the scenario advanced here, it is necessary instead to compare COS-Halos with a sample of galaxies with similar halo masses but diverse central BH masses (or a suitable proxy). Such a survey would also enable scrutiny of EAGLE's prediction that gas-rich (gas-poor) haloes exhibit higher (lower) SFRs than is typical for their halo mass, as a consequence of the negative correlation of scatter about both the $f_{\text{CGM}} - M_{200}$ relation (due to the feedback history of the central BH) and the $\dot{M}_* - M_{200}$ relation (due to the shift of the SF history to earlier times) with scatter about the $E_{\text{DMO}}^{2500} - M_{200}$ relation. In a companion study to the work included in this Chapter (which was published as Davies et al., 2019), Oppenheimer et al. (2020) showed that such a study would be feasible, using the column densities and covering fractions of circumgalactic CIV, a readily-accessible absorption species in the local Universe.

Chapter 4

The role of circumgalactic gas expulsion in galaxy evolution

4.1 Introduction

The findings of Chapter 3 suggest a close coupling between BH-driven feedback and the CGM in the regulation (and quenching) of galaxy growth by star formation. In a companion study following the work in Chapter 3, Oppenheimer et al. (2020, hereafter O20) used high-cadence ‘snapshot’ outputs to show that the CGM mass fraction declines in response to expulsive outflows driven by episodes of BH-driven feedback, and that galaxies whose central BH injects, over its lifetime, an energy that is a greater fraction of the binding energy of its halo baryons, tend to exhibit lower gas fractions and redder colours. They further showed that the covering fraction of CIV and OVI absorption systems can be used as an effective observational proxy for the circumgalactic gas fraction. In a recent paper, Mitchell et al. (2019) presented outflow rates from galaxies and their haloes in the EAGLE simulations, showing that more gas leaves the halo than leaves the galaxy, indicating that circumgalactic gas is entrained in, and expelled by, galactic outflows.

In this Chapter I build on these studies by examining in detail how BH-driven feedback influences the CGM, and why this subsequently impacts the star formation activity of galaxies. I further examine whether the influence of the BH-CGM connection extends beyond star formation activity and might also be reflected in related properties such as galaxy morphology. In an effort to generalise my findings I present results throughout based on analyses of simulations from the EAGLE and IllustrisTNG (hereafter TNG) projects, both of which have released their particle data to the community (see McAlpine et al. 2016 and Nelson et al. 2019, respectively). These models broadly reproduce a

diverse range of properties of the observed galaxy population, in the local Universe and at earlier cosmic epochs, but they differ significantly in many respects, notably in terms of their hydrodynamics solvers and their subgrid routines for the injection of feedback energy from star formation and from the accretion of gas onto BHs. Comparison of the outcomes of these suites therefore represents a meaningful test of the degree to which there is consensus between state-of-the-art simulations in this challenging regime.

This Chapter will be structured as follows. In Section 4.2 I briefly describe the simulations, my techniques for identifying and characterising galaxies and their haloes, and the calculation of CGM cooling rates. In Section 4.3 I examine the correlation between the CGM mass fraction of present-day haloes and the properties of their central BHs, and between the CGM mass fraction and the both the specific star formation rate (sSFR) and the kinematic morphology of their central galaxies. In Section 4.4 I examine the influence of feedback on the cooling time of circumgalactic gas, and the consequent effect on galaxy properties. In Section 4.5 I explore the origin of differences in the efficiency of feedback in haloes of fixed present-day mass. I summarise my findings in Section 4.6. Throughout, I adopt the convention of prefixing units of length with ‘c’ and ‘p’ to denote, respectively, comoving and proper scales, e.g. cMpc for comoving megaparsecs.

4.2 Methods

My analyses are based on the EAGLE Ref-L100N1504 and TNG-100 cosmological hydrodynamical simulations of the formation and evolution of the galaxy population in a Λ-Cold Dark Matter (CDM) cosmogony. The simulations follow periodic comoving cubic volumes of similar side length ($\simeq 100$ cMpc), with comparable resolution in terms of both the mass of baryonic fluid elements ($\sim 10^6 M_\odot$) and the gravitational softening scale (~ 1 pkpc). They both therefore adequately resolve present-day galaxies of mass $M_\star \gtrsim 10^{9.5} M_\odot$ ($\sim 0.1L^*$), whilst following a sufficiently large sample to allow examination of trends at fixed galaxy or halo mass. Hereon, for brevity I simply refer to these simulations as the ‘EAGLE’ and ‘TNG’ simulations. The simulations and their respective subgrid physics models were described in Chapter 2, to which I refer the reader for any details. Of particular relevance for this Chapter are the subgrid prescriptions for stellar and AGN feedback, and the differences in their efficiencies and injection mechanisms. These are detailed in Sections 2.5.3 and 2.5.4. In this Section I will detail further methods that are relevant to this particular study, including the process of selecting galaxies and their haloes, and the computation of cooling rates, cooling timescales, and feedback energetics.

4.2.1 Identifying and characterising haloes and galaxies

As in Chapter 3, haloes and galaxies in both simulation suites are identified via a two-step process, beginning with the application of the friends-of-friends (FoF) algorithm to the dark matter particle distribution, with a linking length of 0.2 times the mean interparticle separation. Gas, stars and BHs are associated with the FoF group, if any, of their nearest dark matter particle. Bound substructures within haloes are subsequently identified using the SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009), and I characterise halo mass via the spherical overdensity mass (M_{200} , Lacey & Cole, 1994) about the coordinates of each halo’s most-bound particle. More generally, halo properties are computed by aggregating the properties of all particles of the relevant type that reside within an appropriate aperture. Following Schaye et al. (2015), I compute the properties of central galaxies in both simulations by aggregating the properties of the relevant particles that reside within 30 pkpc of the halo centre. I equate the BH mass of galaxies, M_{BH} , to the mass of their most-massive BH particle, which is almost exclusively coincident with the halo centre.

Throughout, I consider present-day haloes with $M_{200} > 10^{11.5} M_{\odot}$, such that haloes are sampled in both simulations by at least $\sim 10^5$ particles. The central galaxies hosted by the least massive haloes I examine have a typical mass of $M_{\star} \gtrsim 10^{10} M_{\odot}$, ensuring that they are sampled by at least $\sim 10^4$ stellar particles. As noted above, I match haloes in the Ref-L100N1504 and TNG100 simulations with their counterparts formed in the associated collisionless simulations, in order to compute the intrinsic properties of the haloes in the absence of the physics of galaxy formation. In both cases bijective matching algorithms are used, as discussed by Schaller et al. (2015b) and Nelson et al. (2015) for EAGLE and TNG, respectively. In Ref-L100N1504, this recovers matches for 3411 of the 3543 haloes satisfying my selection criterion, whilst in TNG100 5457 of the 5460 haloes are matched. I discard unpaired haloes from my analyses, irrespective of whether quantities drawn from the collisionless realisations are used, to ensure that a consistent sample from each simulation is used for all analyses.

For both simulations, I consider fluid elements (i.e. SPH particles in EAGLE and Voronoi cells in TNG) with a non-zero SFR to be the ISM, and non-star-forming fluid elements within r_{200} of the galaxy centre to be the CGM. Since EAGLE adopts a metallicity-dependent star formation threshold, I have explicitly checked the influence of adopting instead the fixed density threshold used by TNG ($n_{\text{H}} = 0.1 \text{ cm}^{-3}$), and I find no significant differences in any of the results presented hereafter.

4.2.2 Cooling rates and timescales

I use the radiative cooling time of circumgalactic gas as a diagnostic quantity in Sections 4.3 and 4.4. I compute cooling times both for individual fluid elements and integrated over all circumgalactic gas associated with haloes. The former I compute based on their internal thermal energy, u , and their bolometric luminosity, L_{bol} , via $t_{\text{cool}} = u/L_{\text{bol}}$. The bolometric luminosity is computed as $L_{\text{bol}} = n_{\text{H}}^2 \Lambda V$, where n_{H} is the fluid element's hydrogen number density, Λ is the (volumetric) cooling rate corresponding to its density, temperature and element abundances, in addition to the incident flux from the metagalactic UV/X-ray and cosmic microwave background radiation fields, and $V = m_{\text{g}}/\rho$ where ρ is the mass density of the fluid element¹. In analogy with observational estimates of coronal cooling times, I equate integrated CGM cooling times to the ratio of the total internal thermal energy of the CGM and its total bolometric luminosity:

$$t_{\text{cool}}^{\text{CGM}} = \frac{\sum_i u_i}{\sum_i L_{\text{bol},i}}, \quad (4.1)$$

where the sum runs over all fluid elements, i , comprising the CGM of a given halo.

Volumetric net radiative cooling rates are specified in the publicly-available TNG snapshots, but were not stored in EAGLE snapshots. I therefore recompute them for EAGLE using the [Wiersma et al. \(2009a\)](#) tabulated rates for each of the 11 tracked elements (H, He, C, N, O, Ne, Mg, Si, S, Ca and Fe), which were computed using CLOUDY version 07.02 ([Ferland et al., 1998](#)). The rates are tabulated as a function of hydrogen number density, n_{H} , temperature, T , and redshift, z . I interpolate these tables in $\log_{10} n_{\text{H}}$, $\log_{10} T$, z , and, in the case of the metal-free cooling contribution, the helium fraction $n_{\text{He}}/n_{\text{H}}$. I then compute contributions to the net cooling rate per unit volume element-by-element,

$$\Lambda = \Lambda_{\text{H,He}} + \sum_{i>\text{He}} \Lambda_{i,\odot} \frac{n_e/n_{\text{H}}}{(n_e/n_{\text{H}})_{\odot}} \frac{n_i/n_{\text{H}}}{(n_i/n_{\text{H}})_{\odot}}, \quad (4.2)$$

where $\Lambda_{\text{H,He}}$ is the metal-free contribution, $\Lambda_{i,\odot}$ is the contribution of element i for the solar abundances assumed in CLOUDY, n_e/n_{H} is the particle electron abundance, and n_i/n_{H} is the particle abundance in element i .

Despite both simulations adopting a cooling implementation based on that of [Wiersma et al. \(2009a\)](#), there are differences in their cooling rates, owing primarily to the adoption of different UV/X-ray background radiation models and, in TNG, the assumption of solar abundance ratios when computing the cooling rate, the adoption of an HI self-shielding

¹As discussed by [Schaye et al. \(2015\)](#), their Appendix A1), the use of a pressure-entropy SPH scheme (as in EAGLE) introduces a ‘weighted density’, $\bar{\rho}$, used in the conversion between thermodynamical quantities. For consistency with the rates used in the simulation, I use the physical density, ρ , rather than the weighted density, when computing the radiative cooling rate.

correction for high-density gas, and the suppression of the cooling rate in gas close to accreting BHs.

4.2.3 Feedback energetics

I use the integrated energy injected by feedback from star formation (E_{SF}) and black hole growth (E_{AGN}) as a diagnostic quantity in Section 4.5. Following the nomenclature introduced in Section 2.5.3, the total energy injected by stellar particle i in EAGLE is:

$$E_{\text{SF},i} = 1.74 \times 10^{49} \text{ erg} \left(\frac{m_{\star,\text{init},i}}{1 \text{ M}_{\odot}} \right) f_{\text{SF},i}(n_{\text{H},i}, Z_i), \quad (4.3)$$

and in IllustrisTNG it is:

$$E_{\text{SF},i} = 1.08 \times 10^{49} \text{ erg} \left(\frac{m_{\star,\text{init},i}}{1 \text{ M}_{\odot}} \right) f_{\text{SF},i}(Z_i), \quad (4.4)$$

where $m_{\star,\text{init},i}$ is the initial mass of the particle and the differing prefactors result from differences in the assumed mass range for the progenitors of core-collapse SNe. The total stellar feedback energy injected into a galaxy and its progenitors is therefore the sum of $E_{\text{SF},i}$ over its constituent stellar particles. Comparison of the characteristic values of E_{SF} in both simulations highlights that the difference in specific feedback energies is compensated by differences in the values of f_{SF} . This calculation includes the contribution of feedback injected by stars formed ex-situ to the main progenitor; I choose to include this contribution since it directly influences the CGM of the descendant galaxy.

Since EAGLE implements AGN feedback in a single mode, I compute the total feedback energy injected by the central BH of a galaxy via:

$$E_{\text{AGN}} = \frac{f_{\text{AGN}}}{1 - \epsilon_r} M_{\text{BHC}} c^2, \quad (4.5)$$

following the nomenclature of Section 2.5.4. This translates to 1.67% of the rest mass energy of the BH being coupled to the CGM. I note that this definition is an approximation, since it includes the contribution to the BH mass from seeds but, as shown by Booth & Schaye (2009), contribution of seed mass BHs to the cosmic BH mass density is small. IllustrisTNG implements AGN feedback in two modes, however the total energy injected through the thermal mode ($E_{\text{AGN,thm}}$) and the kinetic mode ($E_{\text{AGN,kin}}$) are each recorded by the snapshots, and do not need to be computed in post-processing. In analogy to the calculation of the stellar feedback energies, the value of E_{AGN} for both simulations explicitly includes the contribution of progenitor BHs that merged with the central BH at $z > 0$.

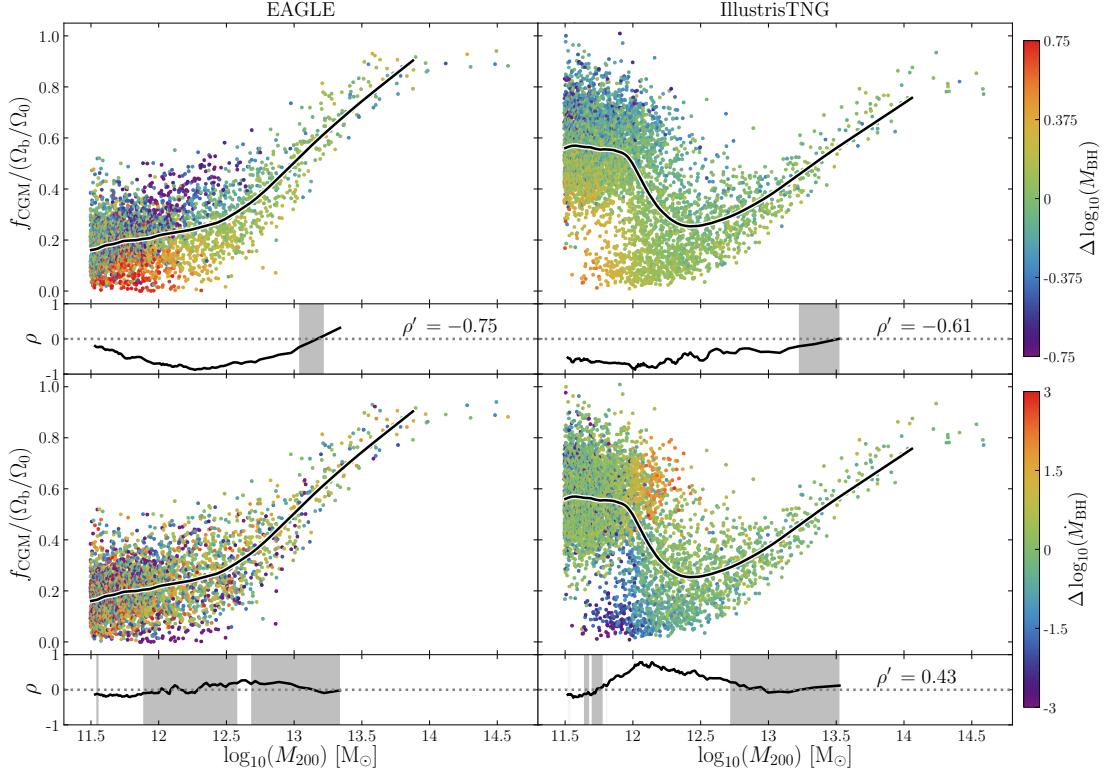


FIGURE 4.1: Present-day CGM mass fractions, $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, of haloes in the EAGLE Ref-L100N1504 (left column) and the TNG-100 (right column) simulations as a function of their mass, M_{200} . Fractions are normalised to the cosmic average baryon fraction, Ω_b/Ω_0 . Black curves denote running medians, $\bar{f}_{\text{CGM}}(M_{200})$. Symbols are coloured by the residuals about the running median, with respect to M_{200} , of (\log_{10} of) the mass of most-massive BH of the halo’s central galaxy (M_{BH} ; upper row) and of (\log_{10} of) its instantaneous present-day accretion rate (\dot{M}_{BH} ; lower row). Below each panel, I show running values of the Spearman rank correlation coefficient, ρ , of the Δf_{CGM} versus $\Delta \log_{10} M_{\text{BH}}$ and Δf_{CGM} versus $\Delta \log_{10} \dot{M}_{\text{BH}}$ relations, and shade regions where the correlation has low significance ($p > 0.01$). Where significant, I quote the correlation coefficients, ρ' , for haloes within a 0.1 dex window about $M_{200} = 10^{12.5} M_{\odot}$.

4.3 The correlation of galaxy and BH properties with the CGM mass fraction

I begin by examining, for both simulations, the relationship between the CGM mass fraction and halo mass, and the dependence of scatter about this relation on the present-day properties of the most-massive BH of the central galaxy. Fig. 4.1 shows, for both EAGLE Ref-L100N1504 (left column) and TNG-100 (right column), the circumgalactic gas mass fractions, f_{CGM} , of present-day haloes, normalised by the cosmic baryon fraction, as a function of halo mass, M_{200} . As noted in Section 4.2.1, and in contrast to the work in Chapter 3, I exclude the ISM from my definition of the CGM, such that $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, where M_{CGM} is the mass of all gas within r_{200} of the halo centre that is not star forming. The solid black line denotes the running median of the CGM

mass fraction, $\tilde{f}_{\text{CGM}}(M_{200})$, computed via the locally-weighted scatter plot smoothing method (LOWESS, e.g. [Cleveland, 1979](#)) and plotted within the interval for which there are at least 10 measurements at both higher and lower M_{200} . The points and median curves are identical in the upper and lower rows; I return to the differences between the rows shortly. Since the ISM generally constitutes only a small fraction of the halo gas mass, the $\tilde{f}_{\text{CGM}}(M_{200})$ curve in the panels of the left-hand column of Fig. 4.1 closely resemble those in Chapter 3 and the gas fraction plots presented by [Schaller et al. \(2015b\)](#). The CGM gas mass fractions of central galaxies in TNG, as a function of stellar mass, were presented by [Nelson et al. \(2018b, their Fig. 20\)](#).

Inspection of the two columns enables a comparison of the present-day CGM gas fractions that emerge in the two simulations. For haloes $M_{200} \gtrsim 10^{12.5} M_\odot$ the behaviour is qualitatively similar in both simulations, insofar that $\tilde{f}_{\text{CGM}}(M_{200})$ rises monotonically with increasing mass, though the fractions rise more quickly in EAGLE and asymptote towards a higher fraction: $\simeq 0.9\Omega_b/\Omega_0$ (the value expected in the absence of efficient feedback, e.g. [Crain et al., 2007](#)) for $M_{200} \gtrsim 10^{13.7} M_\odot$. However, the CGM fractions of less massive haloes differ markedly between the simulations. In EAGLE $\tilde{f}_{\text{CGM}}(M_{200})$ is a monotonic function for all M_{200} , such that the least-massive haloes I examine ($M_{200} = 10^{11.5} M_\odot$) typically exhibit low CGM mass fractions, $\tilde{f}_{\text{CGM}} < 0.2$. By contrast, haloes of $M_{200} < 10^{12} M_\odot$ in TNG have $\tilde{f}_{\text{CGM}} \simeq 0.55$, the CGM mass fraction declines abruptly to a minimum of $\tilde{f}_{\text{CGM}} \simeq 0.25$ at $M_{200} \simeq 10^{12.5} M_\odot$, before increasing again in massive haloes. There is also significantly greater diversity in f_{CGM} for low-mass haloes in TNG than in EAGLE: the interquartile range of f_{CGM} for haloes with $M_{200} \simeq 10^{12-12.5} M_\odot$ is 0.15 for EAGLE and 0.37 for TNG. The haloes that host sub- L_\star central galaxies are in general therefore significantly more gas-rich in TNG than in EAGLE. I note that neither scenario is ruled out by current observational measurements, and that both simulations exhibit cold gas (HI + H₂) fractions that are reasonably consistent with present constraints ([Crain et al., 2017; Stevens et al., 2019](#)).

In Chapter 3 I demonstrated that the residuals of the $\log_{10}(M_{\text{BH}}) - \log_{10}(M_{200})$ relation, $\Delta \log_{10} M_{\text{BH}}$, correlate strongly, negatively and significantly with the residuals of the $f_{\text{CGM}} - \log_{10}(M_{200})$ relation, Δf_{CGM} , in EAGLE, such that at fixed mass, haloes with a more-massive central BH therefore tend to exhibit systematically-lower CGM mass fractions. The sub-panels here confirm the impression given by inspection of the colouring of symbols in the upper row, namely that this correlation is exhibited by both simulations for $M_{200} \lesssim 10^{13} M_\odot$. The Spearman rank correlation coefficient between Δf_{CGM} and $\Delta \log_{10} M_{\text{BH}}$ for haloes within a 0.1 dex window centred on $M_{200} = 10^{12.5} M_\odot$, which I denote as ρ' , has a value of -0.75 for EAGLE and -0.61 for TNG, indicating a strong correlation for $\sim L^\star$ galaxies, which are thought to be hosted by haloes of approximately this mass (e.g. [Moster et al., 2013](#)).

I also demonstrated in the previous Chapter that there is no analogous correlation for the instantaneous present-day BH accretion rate, i.e. between $\Delta \log_{10} \dot{M}_{\text{BH}}$ and Δf_{CGM} , in EAGLE², a result that is reiterated by the lower-left panels of Fig. 4.1. However, inspection of the lower-right panels reveals that this is not the case for TNG. Here, I find a strong, positive correlation for haloes with mass in the range $M_{200} \simeq 10^{11.7-12.7} M_{\odot}$, which peaks at $M_{200} \sim 10^{12} M_{\odot}$, the halo mass at which the characteristic CGM mass fraction declines abruptly in TNG. The peak value of the Spearman rank correlation coefficient is particularly high, $\rho_{\text{max}} = 0.79$, and the value at $M_{200} = 10^{12.5} M_{\odot}$ is $\rho' = 0.43$. The marked difference of the characteristic CGM mass fractions as a function of halo mass, $\tilde{f}_{\text{CGM}}(M_{200})$ exhibited by the two simulations, and the dissimilarity of the correlation of scatter about it with respect to the present-day accretion rate of the central BH, signals significant differences in the means by which circumgalactic gas is expelled from haloes, and the epoch at which the expulsion takes place. I explore the origin of this dissimilarity further in Section 4.5.

I next turn to the connection between the CGM mass fraction of haloes and the properties of their central galaxies. Fig. 4.2 shows the same f_{CGM} versus M_{200} relation for EAGLE and TNG shown in Fig. 4.1, but here the symbols are coloured by residuals of the LOWESS median relationship between (\log_{10} of the) specific star formation rate (sSFR) and halo mass in the upper row, and between that of the co-rotational stellar kinetic energy fraction (κ_{co}) and halo mass in the panels of the lower row. To suppress noise in the sSFR, I average it over the preceding 300 Myr. I consider quenched galaxies to be those with $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$. The parameter κ_{co} denotes the fraction of a galaxy's stellar kinetic energy invested in co-rotation. Correa et al. (2017) showed that EAGLE galaxies with κ_{co} above (below) a value of 0.4 are typically star-forming discs (quenched ellipticals). I compute κ_{co} for galaxies in both EAGLE and TNG using the publicly-available routines of Thob et al. (2019), who also presented a detailed characterisation of the morphology and kinematics of EAGLE galaxies.

It is apparent from inspection of Fig. 4.2 that, despite the significant differences in $\tilde{f}_{\text{CGM}}(M_{200})$ for EAGLE and TNG, in both simulations gas-rich haloes preferentially host galaxies that are both more actively star forming, and exhibit greater rotational support. Inspection of the sub-panels confirms that $\Delta \log_{10}(\text{sSFR})$ correlates strongly, positively and significantly with Δf_{CGM} for $M_{200} \lesssim 10^{13} M_{\odot}$ in both simulations, with the correlation being strongest at $M_{200} \simeq 10^{12.3} M_{\odot}$ ($\rho_{\text{max}} = 0.85$) in EAGLE and at $M_{200} \simeq 10^{12.2} M_{\odot}$ ($\rho_{\text{max}} = 0.88$) in TNG. The correlation coefficient of the relation between Δf_{CGM} and $\Delta \log_{10}(\text{sSFR})$ for haloes within a 0.1 dex window centred on $M_{200} = 10^{12.5} M_{\odot}$ has a value of $\rho' = 0.66$ for EAGLE and 0.67 for TNG, indicating

²Since \dot{M}_{BH} can vary by orders of magnitude on short timescales (e.g. McAlpine et al., 2017), I repeated this analysis with the BH accretion rate time-averaged over 100 Myr, again finding no correlation.

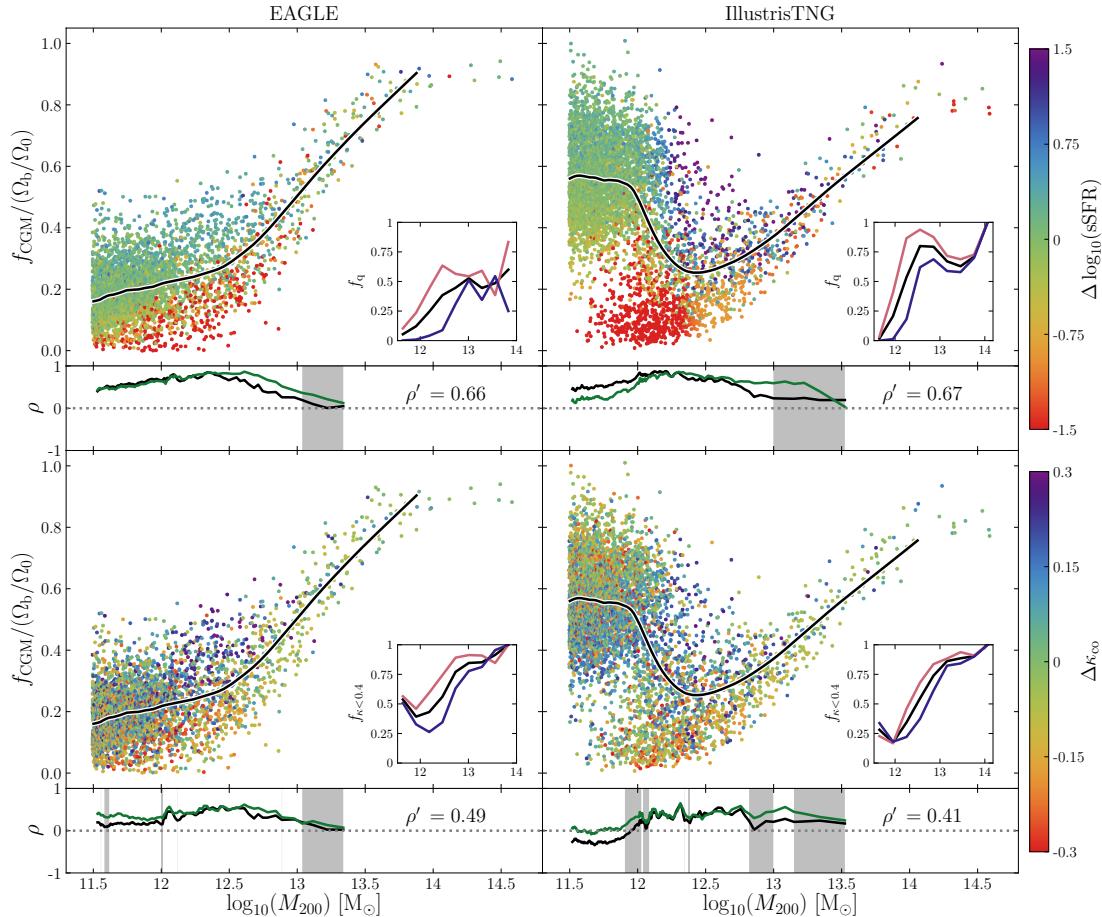


FIGURE 4.2: Present-day CGM mass fractions, $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, of haloes in EAGLE (left column) and TNG (right column) as a function of their mass, M_{200} . Fractions are normalised to the cosmic average baryon fraction, Ω_b/Ω_0 . Black curves denote running medians, $\hat{f}_{\text{CGM}}(M_{200})$. Symbols are coloured by the residuals about the running median, with respect to M_{200} , of the specific star formation rate (sSFR; upper row), and the fraction of stellar kinetic energy invested in co-rotation, (κ_{co} ; lower row). Below each panel, I show running values of the Spearman rank correlation coefficient, ρ , of the Δf_{CGM} versus $\Delta \log_{10} \text{sSFR}$ and Δf_{CGM} versus $\Delta \kappa_{\text{co}}$ relations, and shade regions where the correlation has low significance ($p > 0.01$). I quote the correlation coefficients, ρ' , for haloes within a 0.1 dex window about $M_{200} = 10^{12.5} M_\odot$. Green curves correspond to the Spearman rank correlation coefficients recovered if one instead measures f_{CGM} within $0.3r_{200}$. Inset panels show the quenched fraction (upper row) and the fraction with $\kappa_{\text{co}} < 0.4$ (lower row). Black curves correspond to all central galaxies, and blue and red curves show the fractions for the subsets of galaxies with CGM mass fractions greater than, or lower than, $\hat{f}_{\text{CGM}}(M_{200})$, respectively.

a particularly strong correlation for $\sim L^*$ galaxies. The Δf_{CGM} versus $\Delta \kappa_{\text{co}}$ relation is also strong and significant for $\sim L^*$ galaxies, albeit for a narrower range in M_{200} than is the case for the Δf_{CGM} versus $\Delta \log_{10}(\text{sSFR})$ relation and, consistent with the impression given by the symbol colouring, the correlation is weaker: I recover Spearman rank correlation coefficients at $M_{200} = 10^{12.5} \text{ M}_\odot$ of $\rho' = 0.49$ (EAGLE) and $\rho' = 0.41$ (TNG).

In order to obtain a sense of the connection between the CGM mass fraction on the one hand, and the sSFR and κ_{co} of the galaxies in an absolute sense on the other hand, the plots inset to the upper panels of Fig. 4.2 show the quenched (i.e. $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$) fraction as a function of M_{200} , whilst those in the lower panels show the fraction with an elliptical-like kinematic morphology, i.e. $\kappa_{\text{co}} < 0.4$. The curves are plotted over the same mass range for which there is a LOWESS measurement, sampled by 10 bins of equal size in $\Delta \log_{10} M_{200}$. Black curves show the fractions considering all central galaxies, whilst the blue and red curves show the fractions for the subset of galaxies with CGM mass fractions that are greater than or less than $\tilde{f}_{\text{CGM}}(M_{200})$, respectively (where f_{CGM} is measured within r_{200}). These plots show that for a given M_{200} , in both simulations central galaxies with low CGM mass fractions exhibit an elevated probability of being quenched and of being weakly rotation supported. The converse is also true: central galaxies with high CGM mass fractions exhibit an elevated probability of being actively star forming, and of being strongly rotation supported.

I stress that the existence of correlations between both Δf_{CGM} and ΔSSFR on the one hand, and Δf_{CGM} and $\Delta \kappa_{\text{co}}$ on the other, does not imply that both correlations necessarily emerge as a direct response to the same physical mechanism (e.g. AGN feedback). The connection between the evolution of galaxy colour and morphology in EAGLE galaxies was recently explored by (Correa et al., 2019), who reported only a weak connection. This suggests that the connection between the morphology of galaxies, their star formation rate, and their CGM mass fraction is more complex.

4.4 The influence of feedback on the cooling time of circumgalactic gas

Having demonstrated a connection between the properties of central galaxies and their CGM mass fractions in Section 4.3, I now turn to an examination of the effect of expulsive feedback on the properties of the CGM. I term such feedback ‘expulsive’ because O20 showed that periods of rapid BH growth are immediately followed by a decline in f_{CGM} using high-cadence ‘snapshot’ outputs from the EAGLE simulations, hence the strongly

negative $\Delta f_{\text{CGM}} - \Delta \log_{10} M_{\text{BH}}$ correlation originates from the ejection of CGM gas beyond r_{200} . However I note that feedback also heats and pressurises the remaining CGM, potentially inhibiting the further accretion of gas from the IGM, or the re-accretion of gas expelled by feedback (so-called ‘preventative feedback’).

I start by showing that present-day haloes (of fixed mass) with high (low) CGM fractions have relatively short (long) CGM cooling times (Section 4.4.1), indicating that the cooling time is elevated by the expulsion of circumgalactic gas. I then show that the properties of the central galaxies of haloes correlate significantly with the CGM cooling time (Section 4.4.2).

4.4.1 The effect of feedback on the CGM cooling time

In order to examine the influence of gas expulsion on the properties of the CGM, I isolate haloes within a 0.1 dex window about $M_{200} = 10^{12.5} M_{\odot}$, broadly the range for which the correlations shown in Fig. 4.2 are strongest. This yields 114 haloes for EAGLE and 111 for TNG. I rank the haloes according to their CGM mass fraction, f_{CGM} , and stack those in the upper and lower quartiles, respectively, to form CGM-rich and CGM-poor samples for each simulation. The CGM-rich stacks are thus comprised of haloes with $f_{\text{CGM}} > 0.36$ (EAGLE) and $f_{\text{CGM}} > 0.33$ (TNG), and the CGM-poor stacks are comprised of haloes with $f_{\text{CGM}} < 0.21$ (EAGLE) and $f_{\text{CGM}} < 0.17$ (TNG). Fig. 4.3 shows the cumulative mass distribution functions (CDFs) of the radiative cooling times, $\log_{10}(t_{\text{cool}})$, of fluid elements comprising the stacks, i.e. $M_{\text{CGM}}(< t_{\text{cool}})/M_{\text{CGM}}^{\text{tot}}$. Here $M_{\text{CGM}}^{\text{tot}}$ is the total mass of CGM fluid elements in each stack, such that each distribution asymptotes to unity. I normalise in this fashion to highlight differences in the relative distributions of cooling times in each stack, rather than differences between their CGM mass fractions. Blue and red curves correspond to the CGM-rich and CGM-poor stacks, respectively, for EAGLE (solid curves) and TNG (dotted curves). Vertical lines denote the median cooling time of each distribution.

The distributions are significantly different in EAGLE and TNG, with EAGLE haloes in this mass window exhibiting proportionately less circumgalactic gas with cooling times $t_{\text{cool}} \lesssim 0.1$ Gyr than is the case for their counterparts in TNG. This difference is likely a direct reflection of the different feedback implementations in the two simulations; I explore this in more detail in Section 4.5. This difference notwithstanding, an aspect common to both simulations is the paucity of gas with short cooling times in the CGM-poor haloes relative to their gas-rich counterparts. In both simulations, the CGM-poor haloes exhibit a relative paucity of efficiently-cooling gas with short-to-intermediate cooling times. This is the gas that would otherwise cool onto the ISM and replenish the

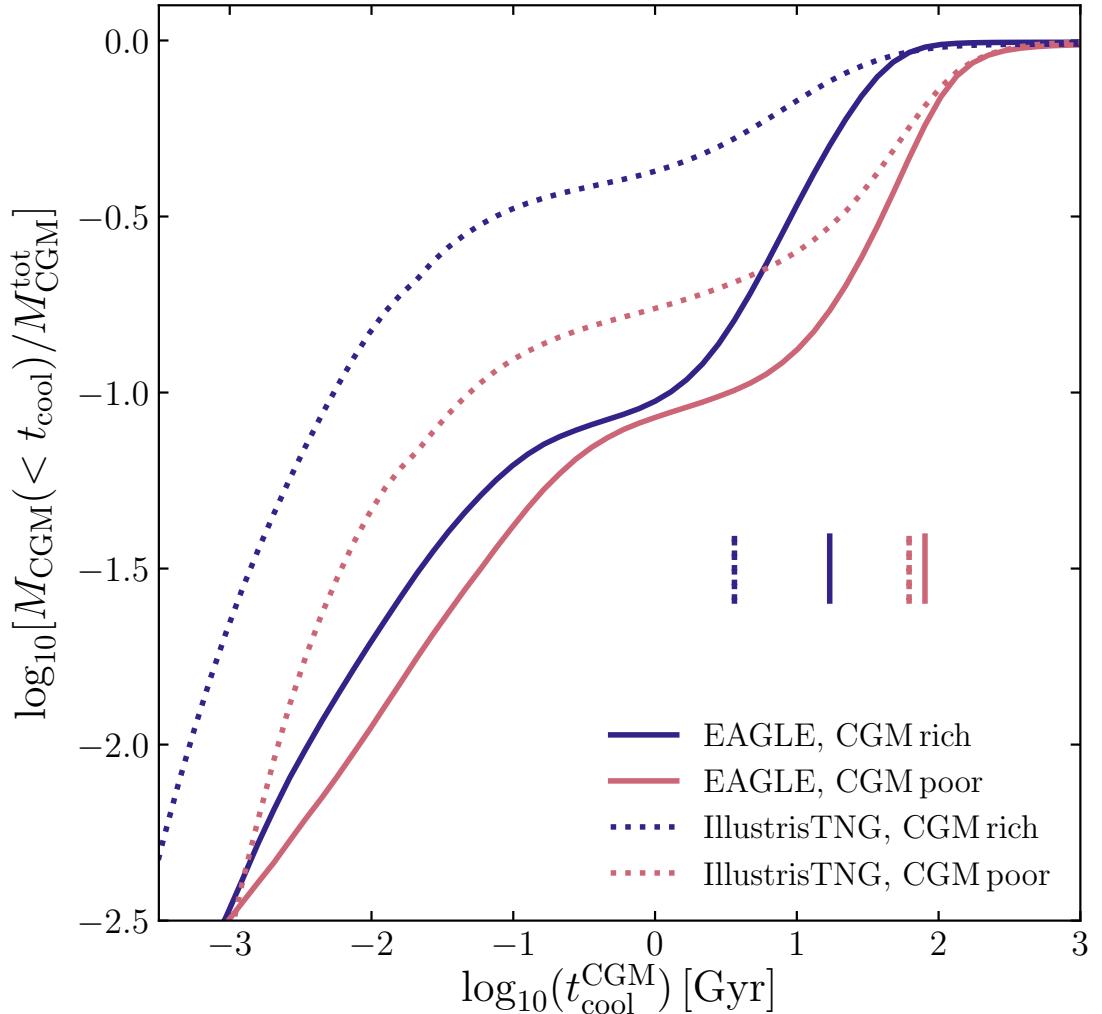


FIGURE 4.3: The cumulative distribution function of the radiative cooling times of fluid elements comprising the CGM of present-day haloes within a 0.1 dex window about $M_{200} = 10^{12.5} M_{\odot}$, in EAGLE (solid curves) and TNG (dotted curves). In each case, the haloes are ranked by their CGM mass fraction, f_{CGM} , and those comprising the upper and lower quartiles are stacked to form CGM-rich (blue curves) and CGM-poor (red curves) samples. Vertical lines denote the median cooling time of each stack. Despite the two simulations exhibiting significantly different CGM cooling time distributions for haloes of this mass, an aspect in common is the relative paucity of rapidly-cooling gas in the CGM-poor samples.

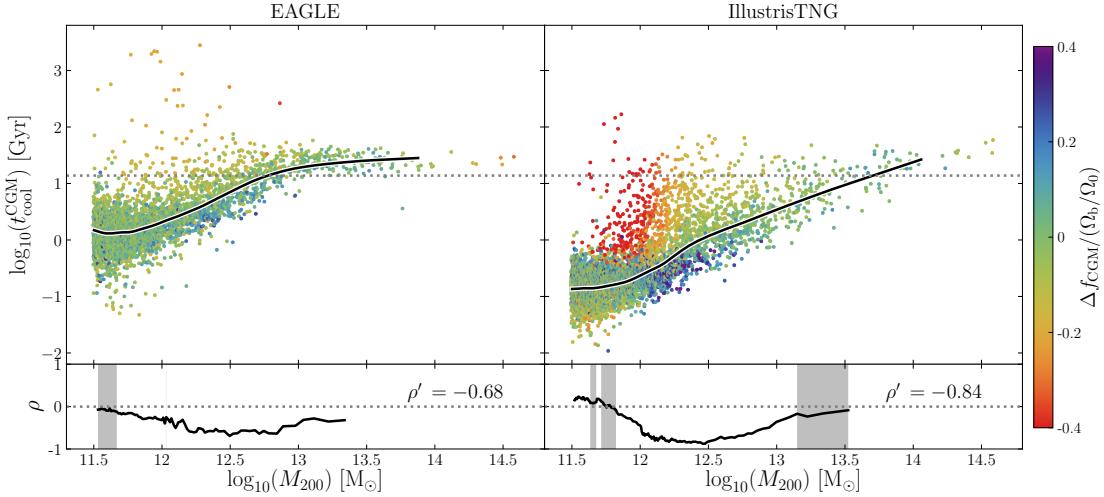


FIGURE 4.4: Present-day characteristic CGM radiative cooling time, $t_{\text{cool}}^{\text{CGM}}$, of haloes in the EAGLE (left) and TNG (right) simulations, as a function of halo mass, M_{200} . The dotted line shows the present-day Hubble time, t_{H} . Black curves denote running medians, $\tilde{t}_{\text{cool}}^{\text{CGM}}(M_{200})$. Symbols are coloured by residuals about the running median of the CGM mass fraction, $\tilde{f}_{\text{CGM}}(M_{200})$. The lower panels show the running Spearman rank correlation coefficient, ρ , of the $\Delta \log_{10} t_{\text{cool}}^{\text{CGM}}$ versus Δf_{CGM} relation. Grey shading denotes mass ranges where the correlation is not formally significant ($p > 0.01$). The quantity ρ' denotes the Spearman rank correlation coefficient for haloes within a 0.1 dex window about $M_{200} = 10^{12.5} M_{\odot}$. These panels highlight a strong and significant negative correlation over all masses sampled, such that haloes with low CGM mass fractions have systematically-longer CGM cooling times.

interstellar gas that is consumed by star formation or expelled by feedback. The paucity of efficiently-cooling gas is also highlighted by the significantly greater median t_{cool} of the CGM-poor haloes: in EAGLE, the median cooling time of the CGM-rich and CGM-poor stacks is, respectively, 22 Gyr and 80 Gyr. The corresponding values for TNG are 3.6 Gyr and 62 Gyr. The differing cooling times of the gas-rich and gas-poor samples stem almost entirely from their necessarily different characteristic CGM densities (since $\Lambda \propto n_{\text{H}}^2$).

I next seek to establish whether this behaviour is general, i.e. whether the CGM cooling time is elevated in response to the expulsion of circumgalactic gas in haloes of all masses probed by my sample. I therefore show in Fig. 4.4 the CGM cooling times (defined as per Eq. 4.1) as a function of M_{200} , and colour the symbols by residuals about the median CGM mass fraction, $\Delta f_{\text{CGM}}/(\Omega_{\text{b}}/\Omega_0)$. The $\tilde{t}_{\text{cool}}^{\text{CGM}}(M_{200})$ relation is qualitatively similar in both simulations; in both cases it is generally a monotonically-increasing function of M_{200} ³, but there are differences in detail that stem largely from the differences in the

³The CGM cooling time is a monotonic function of halo mass in TNG, despite the median CGM mass fraction exhibiting a minimum at $M_{200} \simeq 10^{12.5} M_{\odot}$. As shown by Nelson et al. (2018b), the CGM associated with TNG galaxies below this mass scale is dominated by cool gas ($T \ll 10^6 \text{ K}$), whilst that of more-massive haloes is dominated by hotter gas. Since the cool gas is rapidly cooling and the hot gas is quasi-hydrostatic, a monotonic relationship between CGM cooling time and halo mass still emerges.

$\tilde{f}_{\text{CGM}}(M_{200})$ relation. As presaged by the CDFs presented in Fig. 4.3, the characteristic CGM cooling time of present-day low-mass haloes is longer in EAGLE than in TNG: for the lowest-mass haloes in my sample, $M_{200} = 10^{11.5} M_\odot$, $\tilde{t}_{\text{cool}}^{\text{CGM}} \simeq 1$ Gyr in EAGLE and $\simeq 0.13$ Gyr in TNG, and at $M_{200} = 10^{12.5} M_\odot$ the difference is greater still, $\tilde{t}_{\text{cool}}^{\text{CGM}} \simeq 4$ Gyr in EAGLE and $\simeq 1.5$ Gyr in TNG. The CGM cooling time becomes similar to the Hubble time for haloes of $M_{200} \simeq 10^{13} M_\odot$ in EAGLE, whilst in TNG this threshold is reached at $M_{200} \simeq 10^{13.8} M_\odot$. As is clear from the symbol colouring, the particularly significant differences between the two simulations in low-mass haloes, whilst partly influenced by the structure and metallicity of the CGM, largely reflect differences in their CGM mass fractions. The latter are themselves a consequence of the different feedback implementations of the two simulations, which I return to in Section 4.5.

In both simulations, scatter about the $\tilde{t}_{\text{cool}}^{\text{CGM}}(M_{200})$ relation correlates strongly and negatively with the CGM gas fraction, f_{CGM} , over a wide range in halo mass. The $\Delta \log_{10} \tilde{t}_{\text{cool}}^{\text{CGM}}$ versus Δf_{CGM} relation is particularly strong over the halo mass range corresponding to the abrupt decline of f_{CGM} in TNG, and at $M_{200} \simeq 10^{12.5} M_\odot$ I recover $\rho' = -0.68$ for EAGLE and $\rho' = -0.84$ for TNG. The expulsion of a greater mass fraction of the CGM by feedback therefore unambiguously leads to an elevation of its cooling time in both simulations.

It is tempting to infer from comparison of the t_{cool} CDFs of the CGM-rich and CGM-poor populations shown in Fig. 4.3 that feedback processes *preferentially* eject circumgalactic gas with short cooling times. I note however that even in the case of CGM expulsion being agnostic to cooling time, the median cooling time of the remaining gas would increase in response to its reconfiguration at a lower density. An explicit demonstration that feedback preferentially expels rapidly-cooling gas would require the detailed tracking of fluid elements with high temporal resolution, which is beyond the scope of this study. Nonetheless, I posit that this is a plausible scenario, and note that it bears similarities to that advanced by McCarthy et al. (2011), who showed that the entropy excess of the IGrM associated with galaxy groups in the OWLS simulations (Schaye et al., 2010) is not primarily a consequence of heating of the observable IGrM, but rather the preferential expulsion of low-entropy intragroup gas (mostly from the progenitors of the present-day halo) by AGN feedback. The use of entropy as a diagnostic quantity is commonplace in the study of the IGrM and ICM, particularly by the X-ray astronomy community (e.g. Voit et al., 2003, 2005), but it is not so widely used by the galaxy formation community (though see e.g. Crain et al., 2010). For my purposes here it suffices to note that the cooling time and entropy of the CGM are very strongly and positively correlated: the Spearman rank correlation coefficient of the residuals about the $\tilde{t}_{\text{cool}}^{\text{CGM}}(M_{200})$ and $\tilde{S}(M_{200})$ relations at $M_{200} = 10^{12.5} M_\odot$ are $\rho' = 0.71$ in EAGLE and $\rho' = 0.91$ in TNG. Here, $S = T/n_e^{2/3}$, where n_e is the electron number density.

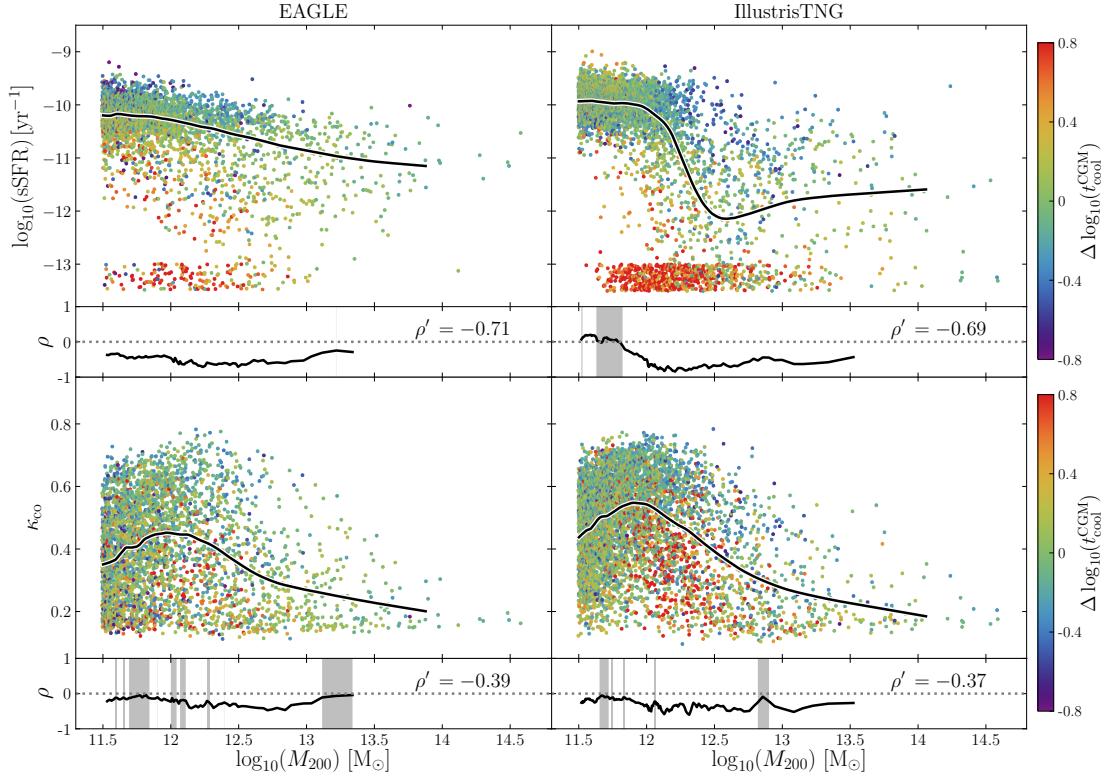


FIGURE 4.5: Present-day specific star formation rates (sSFR; upper row) and fractions of stellar kinetic energy invested in co-rotation (κ_{co} ; lower row), of the central galaxies of haloes in the EAGLE (left) and TNG (right) simulations, as a function of halo mass, M_{200} . Black curves denote running medians, $\text{sSFR}(M_{200})$ and $\tilde{\kappa}_{\text{co}}(M_{200})$. Symbols are coloured by residuals about the running median of the characteristic CGM radiative cooling time, $\tilde{t}_{\text{cool}}(M_{200})$. Below each panel I show the running values of the Spearman rank correlation coefficient, ρ , of the $\Delta \log_{10} \text{sSFR}$ versus $\Delta \log_{10} t_{\text{cool}}$ (upper row) and $\Delta \kappa_{\text{co}}$ versus $\Delta \log_{10} t_{\text{cool}}$ (lower row) relations, which are shaded where the correlation has low significance ($p > 0.01$). The quantity ρ' denotes the Spearman rank correlation coefficient for haloes within a 0.1 dex window about $M_{200} = 10^{12.5} \text{ M}_\odot$. All four panels exhibit negative correlations that are significant in particular halo mass regimes, with the $\Delta \log_{10} \text{sSFR}$ versus $\Delta \log_{10} t_{\text{cool}}$ correlation being particularly strong at intermediate mass.

This quantity is related to the specific thermodynamic entropy, s , via $s \propto \ln S$ and is therefore also conserved by adiabatic processes.

4.4.2 Quenching and morphological evolution in response to elevation of the CGM cooling time

The depletion of efficiently-cooling circumgalactic gas by feedback processes provides a potential explanation for the origin of the correlations shown in Fig. 4.2, which connect the properties of central galaxies to their CGM mass fraction. I therefore turn to an examination of the relations between the sSFR and kinematic morphology of galaxies, and the characteristic cooling time of their CGM. The upper row of Fig. 4.5 shows the

$\log_{10} \text{sSFR}(M_{200})$ relation of central galaxies for EAGLE (left) and TNG (right). For clarity, galaxies with $\log_{10} \text{sSFR} [\text{yr}^{-1}] < -13$ are randomly and uniformly assigned a value in the range $[-13.5, 13]$. The black curve denotes the running median of $\log_{10} \text{sSFR}$ as a function of M_{200} . Symbols are coloured by the residuals of the relationship between the (\log_{10} of the) CGM radiative cooling time, $\tilde{t}_{\text{cool}}^{\text{CGM}}(M_{200})$.

The central galaxies hosted by low-mass haloes ($M_{200} \lesssim 10^{12} M_{\odot}$) in both simulations exhibit $\log_{10} \text{sSFR} [\text{yr}^{-1}] \simeq -10$. In EAGLE, the characteristic sSFR of central galaxies hosted by more massive haloes declines gradually, reaching $\log_{10} \text{sSFR} [\text{yr}^{-1}] \simeq -11$ for $M_{200} \sim 10^{14} M_{\odot}$, whilst in TNG there is a steep and sudden decline to a minimum of $\log_{10} \text{sSFR} [\text{yr}^{-1}] \simeq -12$ at $M_{200} \sim 10^{12.5} M_{\odot}$, followed by a mild increase up to haloes of $M_{200} \simeq 10^{14} M_{\odot}$. Despite these significant differences, in both simulations there is a significant and negative $\Delta \log_{10} \text{sSFR}$ versus $\Delta \log_{10} t_{\text{cool}}^{\text{CGM}}$ relation of similar strength ($\rho' = -0.71$ in EAGLE, $\rho' = -0.69$ in TNG), such that low sSFRs are associated with long CGM cooling times. In EAGLE this correlation is strong and significant for all haloes examined, whilst in TNG the correlation appears abruptly at $M_{200} \simeq 10^{12} M_{\odot}$, coincident with the sharp decline in the sSFR. The cessation of star formation in central galaxies in concert with the expulsion of efficiently-cooling circumgalactic gas is therefore common to both simulations.

The lower row of Fig. 4.5 shows the $\kappa_{\text{co}}(M_{200})$ relation of central galaxies, with symbols again coloured by $\Delta \log_{10} t_{\text{cool}}^{\text{CGM}}$. The two simulations exhibit qualitatively similar trends, with the characteristic rotational support peaking in the central galaxies hosted by haloes with $M_{200} \simeq 10^{12} M_{\odot}$ (see also Clauwens et al., 2018), with peak median values of $\kappa_{\text{co}} \simeq 0.4$ in EAGLE and $\kappa_{\text{co}} \simeq 0.45$ in TNG. In both cases there is a significant and negative $\Delta \kappa_{\text{co}}$ versus $\Delta \log_{10} t_{\text{cool}}^{\text{CGM}}$ correlation, such that low rotation support in central galaxies is associated with long CGM cooling times. The expulsion of rapidly-cooling circumgalactic gas is therefore also implicated in the morphological evolution of the broader population of central galaxies in both simulations.

Consistent with the trends with Δf_{CGM} shown in Fig. 4.2, the sSFR is more strongly correlated with the CGM cooling time than κ_{co} is; I recover $\rho' = -0.39$ (EAGLE) and $\rho' = -0.37$ (TNG) for $\Delta \kappa_{\text{co}}$ versus $\Delta \log_{10} t_{\text{cool}}^{\text{CGM}}$. This finding is consistent with the conclusions of Correa et al. (2019), who identified only a weak connection between the evolution of colour and morphology in EAGLE galaxies. However, whilst the feedback-driven expulsion of circumgalactic gas is unlikely to influence the morphological evolution of central galaxies directly, a causal physical link between them is nevertheless plausible. It is well-established from numerical simulations that the presence of (cold) gas during mergers stabilises galaxy discs against transformation into spheroids, and enables the re-growth of disrupted discs (see e.g. Robertson et al., 2006; Hopkins et al., 2009; Font et al.,

2017). Since the expulsion of the efficiently-cooling component of the CGM suppresses the replenishment of cold interstellar gas in discs, this mechanism likely boosts the susceptibility of disc disruption via gravitational instability and mergers and inhibits the regrowth of a disc component in quenched galaxies, thus facilitating morphological evolution and yielding the positive correlation between Δf_{CGM} and $\Delta \kappa_{\text{co}}$.

4.5 The origin of the diversity in CGM mass fractions at fixed halo mass

I now turn to an examination of why there is significant diversity in the CGM mass fractions of present-day haloes at fixed mass in both simulations. As discussed in Section 4.3 and shown in Fig. 4.1, EAGLE and TNG exhibit similar relations between the scatter about $f_{\text{CGM}}(M_{200})$ and $M_{\text{BH}}(M_{200})$ at $z = 0$, but markedly different relations between the scatter about $f_{\text{CGM}}(M_{200})$ and $\dot{M}_{\text{BH}}(M_{200})$. Given that both simulations were calibrated to reproduce key stellar properties of the galaxy population (and also some properties of the intragroup/intracluster gas in the case of TNG), this is a significant outcome, because it illustrates that reproduction of the calibration diagnostics does not isolate a truly unique ‘solution’ to the implementation of feedback processes in galaxy formation models.

In Chapter 3 I showed that, in EAGLE, the scatter in f_{CGM} at fixed M_{200} correlates strongly and negatively with the mass of the halo’s central BH. Their interpretation was that scatter in the binding energy of haloes (at fixed M_{200}) drives scatter in the mass of the central BH (see also Booth & Schaye, 2010, 2011). Haloes with more tightly-bound centres therefore foster the growth of more massive central BHs⁴, injecting more feedback energy into the CGM and thus lowering their CGM mass fraction. In a follow-up study, O20 showed that scatter in f_{CGM} correlates with the ratio of the cumulative BH feedback energy injected throughout the formation history of the galaxy, E_{AGN} , to the binding energy of the baryons in its halo, $E_{\text{bind}}^{\text{b}}$. Moreover, they showed that this ratio is an effective means of separating red, quenched galaxies from blue, star-forming galaxies in EAGLE. Here I seek to test these conclusions more forensically, and establish a sense of their generality.

Fig. 4.6 shows the $f_{\text{CGM}}(M_{200})$ relation of present-day haloes, in both EAGLE (left) and TNG (right). In the upper row, symbols coloured by the residuals about the running median, with respect to M_{200} , of the quantity $V_{\text{DMO}}^{\text{max}}/V_{\text{DMO}}^{200}$, where $V_{\text{DMO}}^{\text{max}}$ is the maximum of the radial circular velocity profile, $V_c(r) = [GM(< r)/r]^{1/2}$, of the halo’s counterpart in

⁴The same interpretation applies to lower-mass haloes if I replace M_{BH} with M_{\star} (Matthee et al., 2017).

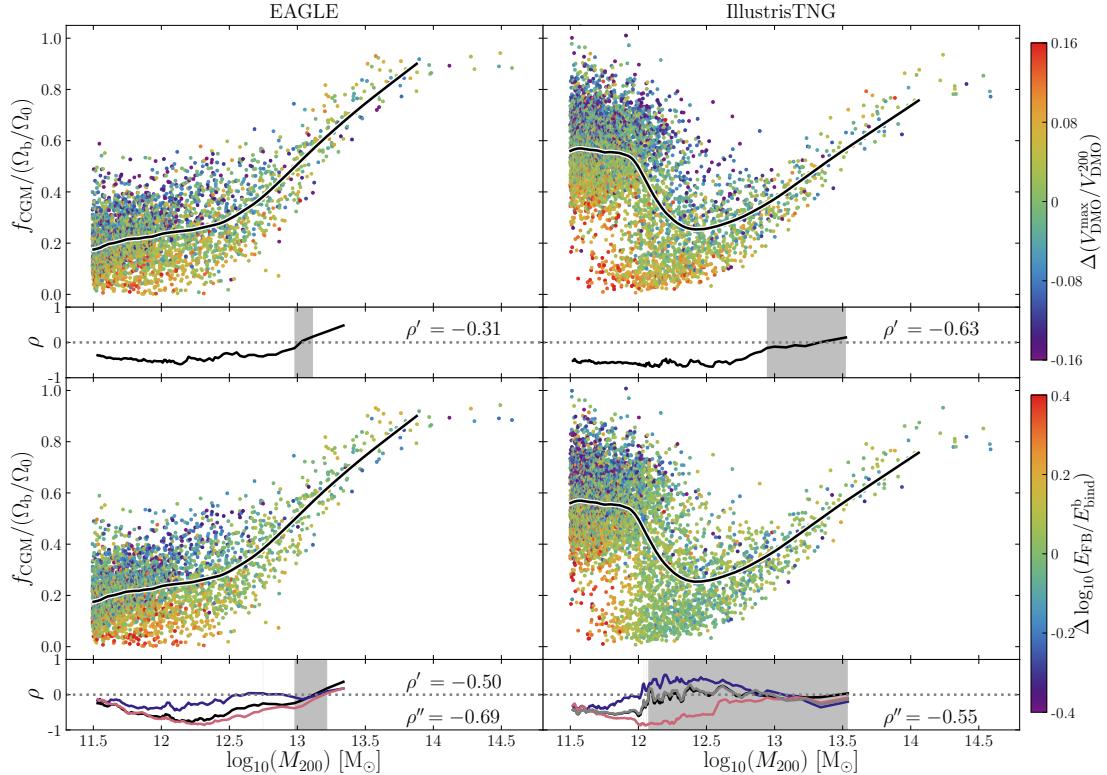


FIGURE 4.6: Present-day CGM mass fraction, $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, of haloes in the EAGLE (left) and TNG (right) simulations, as a function of halo mass, M_{200} . Fractions are normalised to the cosmic average baryon fraction, Ω_b/Ω_0 . Black lines correspond to running medians, $\tilde{f}_{\text{CGM}}(M_{200})$. In the upper row, symbols are coloured by residuals about the running median of the quantity $V_{\text{DMO}}^{\text{max}}/V_{\text{DMO}}^{200}$, which is a proxy for the concentration, formation time and binding energy of haloes of fixed mass (see text for details). In the lower row, they are coloured by residuals about the running median of the quantity $\log_{10} E_{\text{FB}}/E_{\text{bind}}^b$, where E_{FB} is the total feedback energy liberated by the galaxy and its progenitors, and E_{bind}^b is the binding energy of the halo's baryons. Sub-panels show with black curves the running values of the Spearman rank correlation coefficient of the relations between residuals about the plotted running medians, and the colour-coded quantity. These are shaded where the correlation has low significance ($p > 0.01$). For the lower row I also show the running Spearman rank correlation coefficient if one considers the individual contributions to E_{FB} from star formation and AGN, i.e. for EAGLE E_{SF} (blue) and E_{AGN} (red) and for TNG E_{SF} (blue), $E_{\text{AGN,thm}}$ (grey) and $E_{\text{AGN,kin}}$ (red). The quantity ρ'' is the equivalent of $\rho' = \rho(M_{200} = 10^{12.5} M_\odot)$ but considering only the main expulsive feedback mode in each simulation, i.e. AGN feedback in EAGLE and kinetic AGN feedback in TNG.

the respective DMONLY simulation⁵ (DMONLY-L100N1504 for EAGLE and TNG100-Dark for TNG), and V_{DMO}^{200} is the counterpart's virial circular velocity, $V_c(r = r_{200})$. The quantity $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$ is a simple and direct proxy for the intrinsic halo concentration, and hence correlates strongly and positively with the halo binding energy⁶ and formation time (e.g. [Navarro et al., 2004](#)). This test reveals that there is a negative correlation between this proxy for the concentration of haloes, and their CGM mass fraction. The correlation is significant over a wide range in halo mass ($M_{200} \lesssim 10^{12.8} M_\odot$) for both simulations, though the strength of the correlation is weaker in EAGLE than in TNG, with Spearman rank correlation coefficients of $\rho' = -0.31$ (EAGLE) and $\rho' = -0.63$ (TNG) at $M_{200} = 10^{12.5} M_\odot$. My earlier finding that the early collapse of haloes (of fixed present-day mass) results in the expulsion of a greater fraction of their baryons therefore applies not only to EAGLE, but also (and more strongly) to TNG.

In the lower row of Fig. 4.6, the symbols are coloured by the residuals about the running median of the cumulative energy injected by feedback relative to the CGM binding energy, $\log_{10}(E_{\text{FB}}/E_{\text{bind}}^b)$, where $E_{\text{FB}} = E_{\text{SF}} + E_{\text{AGN}}$. Recall that for TNG the latter term has contributions from the thermal and kinetic modes, which have differing subgrid efficiencies, $f_{\text{AGN,thm}}$ and $f_{\text{AGN,kin}}$. I equate E_{bind}^b to the intrinsic binding energy of the halo (i.e. that of the halo's counterpart in the matched collisionless simulation), normalised by the cosmic baryon fraction, $E_{\text{bind}}^b = (\Omega_b/\Omega_0)E_{\text{DMO}}^{200}$, where the superscript denotes that I consider the binding energy of the halo within r_{200} . I compute E_{DMO}^{200} by summing the binding energies of all particles within this radius, and thus self-consistently account for variations in halo structure at fixed mass.

Previous studies have shown that, in both simulations, it is AGN feedback that dominates the expulsion of gas from (massive) haloes ([Bower et al., 2017; Nelson et al., 2018b](#)), and this conclusion is not specific to EAGLE and TNG (see e.g. [Tremmel et al., 2017](#)). In sub-panels of the lower row, therefore, I also show the running Spearman rank correlation coefficient that one recovers if considering the individual contributions to E_{FB} from star formation and the growth of BHs, i.e. for EAGLE E_{SF} (blue) and E_{AGN} (red) and for TNG E_{SF} (blue), $E_{\text{AGN,thm}}$ (grey) and $E_{\text{AGN,kin}}$ (red). The quantity ρ'' is the equivalent of ρ' but considering only the main expulsive feedback mode in each simulation, i.e. AGN feedback in EAGLE and kinetic AGN feedback in TNG.

⁵I use ‘intrinsic’ measurements from the DMONLY simulation, because the expulsion of baryons from haloes in the simulations including baryon physics can induce systematic changes of their properties (e.g. the central binding energy or concentration) of a magnitude comparable to the intrinsic scatter. This can mask genuine underlying correlations between the properties of the haloes, and those of their central galaxies and the CGM.

⁶ $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$ correlates strongly and positively with the intrinsic binding energy of haloes used in Chapter 3, E_{DMO} , over the full range of halo masses I explore in both simulations ($\rho > 0.88$).

These panels reveal both similarities and differences between the simulations. At first glance, it appears that the origin of diversity in $f_{\text{CGM}}(M_{200})$ is different in the two simulations. As previously reported by O20, in EAGLE there is a strong, negative correlation between Δf_{CGM} and $\Delta \log_{10}(E_{\text{FB}}/E_{\text{bind}}^{\text{b}})$, over a wide range in halo mass ($M_{200} \lesssim 10^{13} M_{\odot}$), with a Spearman rank correlation coefficient of $\rho' = -0.50$ at $M_{200} = 10^{12.5} M_{\odot}$. I recover an even stronger correlation when considering only the contribution to E_{FB} from AGN feedback, with $\rho'' = -0.69$, indicating that the overall correlation is driven primarily by AGN feedback. In TNG, there is no significant correlation between Δf_{CGM} and $\Delta \log_{10}(E_{\text{FB}}/E_{\text{bind}}^{\text{b}})$ for $M_{200} \gtrsim 10^{12} M_{\odot}$. However, I do recover a strong, negative correlation between these quantities, over a wide halo mass range ($M_{200} \lesssim 10^{13} M_{\odot}$), if I consider only the contribution to E_{FB} from the kinetic mode of AGN feedback. In this case, the Spearman rank correlation coefficient is $\rho'' = -0.55$ at $M_{200} = 10^{12.5} M_{\odot}$. This marked difference between the overall trend and that for only kinetic AGN indicates that for TNG haloes with $M_{200} \gtrsim 10^{12} M_{\odot}$, f_{CGM} is governed almost exclusively by kinetic AGN feedback, despite this mode not dominating the overall feedback energy budget. In both simulations then, it appears that the diversity in $f_{\text{CGM}}(M_{200})$ is driven primarily by halo-to-halo differences in the energy ‘budget’ of feedback that effectively couples to the gas (i.e. AGN feedback in EAGLE and kinetic AGN feedback in TNG), relative to the binding energy of the halo baryons.

I examine the energetics of feedback in greater detail in Fig. 4.7, which shows $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ as a function of M_{200} for EAGLE (left) and TNG (right). Black curves show running medians. I also show the running median contributions from the individual energy injection mechanisms as secondary lines, i.e. for EAGLE SF feedback (blue) and AGN feedback (red), and for TNG SF feedback (blue), kinetic AGN feedback (red) and thermal AGN feedback (grey). I stress that E_{FB} , being a cumulative measure of energy injection throughout the formation and assembly of the galaxy, need not closely reflect the dominant energy injection mechanism at the present day.

The functional form of the overall relationship is broadly similar in both simulations, but there are differences. In EAGLE, galaxies hosted by haloes $M_{200} \lesssim 10^{12.5} M_{\odot}$ typically inject $E_{\text{FB}} \simeq 5E_{\text{bind}}^{\text{b}}$ over their lifetime. For haloes $M_{200} \lesssim 10^{12} M_{\odot}$, this energy is dominated by the contribution of stellar feedback injected throughout the formation and assembly of the central galaxy’s main progenitor. In more massive haloes, the ratio declines gradually and monotonically, such that the ratio approaches unity for haloes of $M_{200} \sim 10^{13.5} M_{\odot}$. For haloes $M_{200} \gtrsim 10^{12.7} M_{\odot}$, the energy injected over the lifetime of the galaxy by AGN feedback dominates marginally over that from SF feedback. I note that, since the growth of massive galaxies is dominated by mergers rather than in-situ star formation (e.g. Qu et al., 2017), the latter was primarily injected prior to the central galaxy becoming massive. As noted in Section 4.2.3, my calculation of E_{AGN}

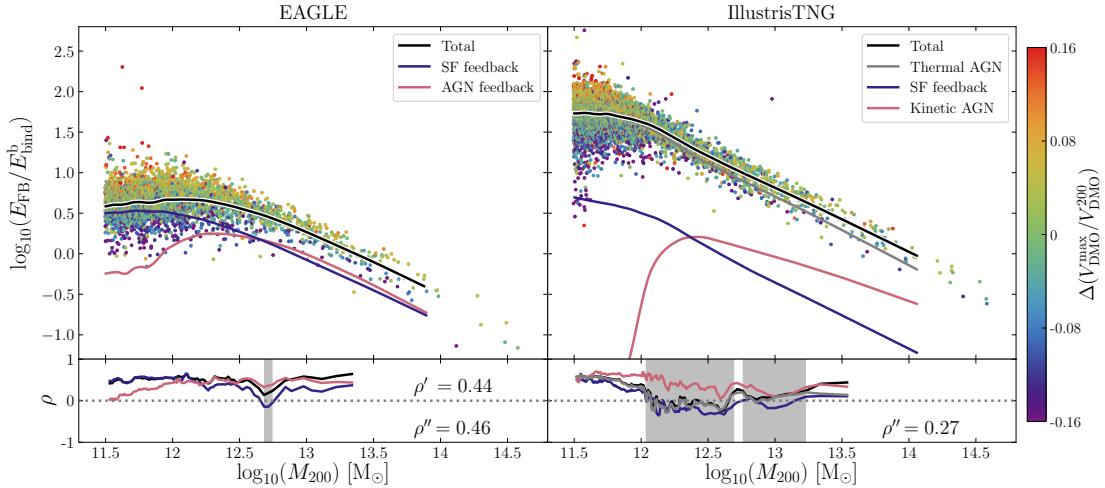


FIGURE 4.7: Present-day ratio of the total energy injected by feedback processes to the binding energy of halo baryons, $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$, as a function of M_{200} . Black lines correspond to the running median of this quantity considering all contributions to E_{FB} , blue lines correspond to the contribution from stellar feedback. Red lines correspond to the running median of AGN feedback in EAGLE and kinetic-mode AGN feedback in TNG, and grey lines correspond to thermal mode AGN feedback in TNG. Symbols are coloured by residuals about the running median of the quantity $V_{\text{DMO}}^{\text{max}}/V_{\text{DMO}}^{200}$ which is a proxy for the concentration, and inner binding energy, of the halo (see text for details). Sub-panels show with black curves the running values of the Spearman rank correlation coefficient of the relations between residuals about the plotted running medians, and that denoted by the colouring. These are shaded where the correlation has low significance ($p > 0.01$). I also show the running Spearman rank correlation coefficient if one considers the individual contributions to E_{FB} , i.e. for EAGLE E_{SF} (blue) and E_{AGN} (red) and for TNG E_{SF} (blue), $E_{\text{AGN,thm}}$ (grey) and $E_{\text{AGN,kin}}$ (red). The quantity ρ'' is the equivalent of $\rho' = \rho(M_{200} = 10^{12.5} M_{\odot})$ but considering only the main expulsive feedback mode in each simulation, i.e. AGN feedback in EAGLE and kinetic-mode AGN feedback in TNG.

for EAGLE haloes includes the contribution of seed mass BHs to the mass of the central BH; in haloes of mass $M_{200} \ll 10^{12} M_{\odot}$, these contributions may dominate (see e.g. Bower et al., 2017) and hence in this regime E_{AGN} should be considered an upper limit. The decline of $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ towards greater halo masses reflects the decreasing ‘ability’ of feedback mechanisms to unbind a large fraction of the baryons associated with group- and cluster-scale haloes. However, I remark that the regulation of the growth of the central galaxies hosted by these haloes does not require the majority of the IGrM/ICM to become unbound since, as is clear from Fig. 4.4, the majority of this gas has a cooling time significantly longer than the present-day Hubble time.

In TNG, galaxies hosted by haloes of $M_{200} \lesssim 10^{12} M_{\odot}$ typically inject $E_{\text{FB}} \simeq 50 E_{\text{bind}}^{\text{b}}$ over their lifetime, i.e. an order of magnitude more than for EAGLE, the majority of which is contributed by the thermal AGN mode. The ratio declines gradually and monotonically towards greater halo masses, reaching unity for the central galaxies hosted by haloes of $M_{200} \sim 10^{14} M_{\odot}$. For all haloes examined, the thermal AGN mode dominates

the injection of feedback energy over the lifetime of the galaxy. However, as shown by the significantly stronger correlation (at fixed mass) of the gas fraction with $E_{\text{AGN,kin}}$ than with E_{FB} (see Fig. 4.6), it is the kinetic AGN mode that governs the CGM gas fraction. Weinberger et al. (2018, see also Henden et al. 2018), notes that the thermal dump implementation of AGN feedback in TNG leads to the injected energy being distributed over a relatively large mass of gas, producing only a small heating increment. Such small increments often lead to numerical losses, as the heated gas radiates the injected energy on a timescale shorter than a sound crossing time across a resolution element (Dalla Vecchia & Schaye, 2012). It is therefore plausible that, despite the thermal AGN mode being the dominant channel by which energy is injected into haloes in TNG, numerical losses result in this mode having little impact on the evolution of the CGM. In contrast, the pulsed kinetic AGN mode imposes a minimum injection energy per feedback event to ensure that individual injection events are numerically, as well as physically, efficient. In this sense, this scheme is similar to the stochastic thermal heating method of Booth & Schaye (2009), used by the OWLS and EAGLE simulations to overcome numerical losses.

The use of a calibrated⁷ pivot mass in the expression that governs the transition of AGN feedback from thermal to kinetic mode in TNG (see Section 2.5.4 in Chapter 2) effectively imprints a mass scale at which expulsive feedback becomes efficient. This is clear from inspection of the red curve in the right-hand panel of Fig. 4.7, which shows a sharp transition in the energetics of kinetic-mode AGN feedback for present-day haloes in the mass range $M_{200} = 10^{12-12.5} \text{ M}_\odot$. This mass scale corresponds closely to that for which the CGM mass fraction reaches a minimum in TNG, and is likely the cause of the significantly greater diversity of f_{CGM} at this mass scale in TNG than in EAGLE. The contribution of kinetic-mode AGN feedback becomes greater than that of stellar feedback at $M_{200} \simeq 10^{12.3} \text{ M}_\odot$, and for more massive haloes it dominates strongly over stellar feedback, with $E_{\text{AGN,kin}} \simeq 4.4E_{\text{SF}}$ at $M_{200} = 10^{14} \text{ M}_\odot$. At this mass scale, $E_{\text{AGN,thm}} \simeq 2.5E_{\text{AGN,kin}}$, but I reiterate that the thermal injection took place prior to the central galaxy's BH (and hence the galaxy itself) becoming massive, and that thermal mode AGN feedback appears to be numerically inefficient in TNG. I remark that the sum of the energies injected by stellar and kinetic AGN in TNG is comparable to the total energy injected into EAGLE haloes, making it likely that the energy injected by these efficient mechanisms is a reasonable estimate of the energy required to regulate galaxy growth to the observed level.

⁷The AGN feedback model in EAGLE also includes a parameter, C_{visc} , which modulates the BH accretion rate. Although this parameter was calibrated (see Crain et al., 2015), it does not influence the (subgrid) AGN feedback efficiency and, as shown by Bower et al. (2017), its value has little bearing on when galaxies quench.

Symbols in Fig. 4.7 are coloured by the residuals about the running median, with respect to M_{200} , of $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$. In sub-panels, I show the running Spearman rank correlation coefficients for the relations between the residuals about the medians of $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ and $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$. As per Fig. 4.6, I show values for the total energy injected (black) and also those recovered for the individual feedback mechanisms, i.e. for EAGLE E_{SF} (blue) and E_{AGN} (red) and for TNG E_{SF} (blue), $E_{\text{AGN,thm}}$ (grey) and $E_{\text{AGN,kin}}$ (red). The quantity ρ'' is again the equivalent of ρ' but considering only AGN feedback in EAGLE and only kinetic AGN feedback in TNG.

The curves in the sub-panels highlight revealing differences between the simulations. In EAGLE, residuals about the median of $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ as a function of M_{200} correlate significantly with those about the median of $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$ for, effectively, haloes of all masses. The Spearman rank correlation coefficient at $M_{200} = 10^{12.5} \text{ M}_{\odot}$ is $\rho' = 0.44$. The trend is dominated by SF feedback at low halo masses, and by AGN feedback for $M_{200} \gtrsim 10^{12.3} \text{ M}_{\odot}$, such that in this case the Spearman rank correlation coefficient at $M_{200} = 10^{12.5} \text{ M}_{\odot}$ is $\rho'' = 0.46$. The behaviour is markedly different in TNG. There is a positive overall correlation for $M_{200} \lesssim 10^{12} \text{ M}_{\odot}$, but at higher masses the $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ ratio is effectively independent of $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$. However, if one again focuses only on the expulsive kinetic AGN feedback, a positive correlation similar to that seen in EAGLE is recovered, with Spearman rank correlation coefficient at $M_{200} = 10^{12.5} \text{ M}_{\odot}$ of $\rho'' = 0.27$. Massive haloes that are more tightly-bound than is typical for their mass therefore appear to foster the formation of central BHs that are slightly more massive than is typical, resulting in the injection of more energy from the efficient feedback mechanisms in both EAGLE and TNG. I stress that this fact alone does not guarantee that such haloes will foster a higher $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ ratio at fixed M_{200} , since it is *necessary* to inject more feedback energy in such haloes simply to offset their higher binding energy. However, my findings indicate that BH growth in tightly-bound haloes results in the ‘overshoot’ of E_{FB} relative to $E_{\text{bind}}^{\text{b}}$.

I speculate that the cause of this overshoot differs in the two simulations. Since EAGLE adopts a fixed subgrid efficiency for AGN feedback ($f_{\text{AGN}} = 0.015$), the energy injection rate is simply proportional to the BH accretion rate, i.e. $\dot{E}_{\text{AGN}} \propto \dot{M}_{\text{BH}}$, where the latter is the minimum of the Bondi-Hoyle ($\propto M_{\text{BH}}^2$) and Eddington ($\propto M_{\text{BH}}$) rates. Early growth of the BH therefore enables it to reach higher accretion rates, and hence higher AGN energy injection rates, sooner. The expulsion of circumgalactic gas in EAGLE therefore occurs at $z \sim 1 - 3$ when BH accretion rates peak, resulting in the absence of a strong correlation between CGM gas fractions and the BH accretion rate at $z = 0$ (Fig. 4.1, bottom-left). In TNG, early growth of the BH enables it to reach the calibrated ‘pivot’ mass scale, at which AGN feedback switches from the numerically inefficient thermal mode, to the efficient kinetic mode, sooner. CGM expulsion in TNG

is thus driven by high efficiency, low accretion rate kinetic-mode AGN feedback at later epochs, imprinting a strong, positive present-day Δf_{CGM} versus $\Delta \dot{M}_{\text{BH}}$ relation (Fig. 4.1, bottom-right). I note that not all of the ‘additional’ energy is likely to be used to expel gas from the halo, because the characteristic density of circumgalactic gas is greater at high redshift ($n_{\text{H}} \propto (1+z)^3$), thus influencing the cooling rate ($\Lambda \propto n_{\text{H}}^2 \propto (1+z)^6$), and hence the cooling time ($t_{\text{cool}} \propto (1+z)^{-3}$) of gas. Feedback energy injected at early times is therefore likely to be more strongly influenced by physical radiative losses. However, as is clear from Figs. 4.6 and 4.7, the early growth of BHs in tightly-bound haloes at fixed mass results in the expulsion of a greater CGM mass fraction in both simulations.

4.6 Summary and Discussion

I have investigated the connection between the properties of the circumgalactic medium (CGM) and the quenching and morphological evolution of galaxies in two state-of-the-art cosmological hydrodynamical simulations of the galaxy population. This study was motivated by the discovery in the EAGLE simulations of several strong correlations linking the properties of the CGM mass fraction, f_{CGM} , of dark matter haloes of fixed present-day halo mass with the properties of the central galaxy and its central BH, and the intrinsic properties of the halo itself, presented in Chapter 3 of this thesis and in Oppenheimer et al. (2020, O20). These correlations are indicative of an important physical role for the CGM in quenching galaxy growth and potentially also in mediating their morphological evolution.

My results are based on analyses of the EAGLE (Ref-L100N1504) and IllustrisTNG (TNG-100) simulations, both of which follow a periodic comoving cubic volume of side length ~ 100 cMpc, with gravitational force softening scales of ~ 1 pkpc and baryonic mass resolution $\sim 10^6 M_{\odot}$. They offer sufficiently large samples of well-resolved galaxy+CGM systems to allow the examination of correlations in properties at fixed halo mass. The ill-constrained parameters governing the efficiency of feedback mechanisms in both simulations were calibrated to ensure the reproduction of key present-day galaxy properties (the gas mass fractions of group-scale haloes were also considered during the calibration of TNG). Both simulations have been shown to reproduce a diverse range of galaxy properties, at the present-day and earlier times, that were not considered during the calibration.

The simulations are therefore similar in aims and scope, but they differ significantly in many aspects of their implementation. They adopt markedly different hydrodynamics solvers, and the subgrid treatments governing a number of unresolved physical processes

in the simulations, most notably feedback, are implemented in very different ways. Comparison of the relationships between the properties of galaxies, the CGM and the dark matter haloes that emerge from EAGLE and TNG therefore offers a meaningful and revealing test of the degree to which there is consensus between state-of-the-art simulations, in a regime for which their outcomes were not calibrated.

My findings can be summarised as follows:

1. The relation between the present-day CGM mass fraction of dark matter haloes, f_{CGM} , and their mass, M_{200} , differs significantly in the EAGLE and TNG simulations. Low-mass haloes ($M_{200} = 10^{11.5} M_{\odot}$) are typically gas-poor in EAGLE ($f_{\text{CGM}} < 0.2$), whilst they are relatively gas-rich in TNG ($f_{\text{CGM}} \simeq 0.55$). The CGM mass fraction is a monotonically-increasing function of halo mass in EAGLE, reaching $f_{\text{CGM}} \simeq 0.3$ at $M_{200} = 10^{12.5} M_{\odot}$ and then steepening to asymptote to $f_{\text{CGM}} \simeq 0.9$ for $M_{200} \gtrsim 10^{13.7} M_{\odot}$. In contrast, the relation in TNG initially declines with increasing halo mass, reaching a minimum of $f_{\text{CGM}} \simeq 0.25$ at $M_{200} \simeq 10^{12.5} M_{\odot}$, before reverting to a monotonically-increasing function of halo mass that reaches $f_{\text{CGM}} \simeq 0.8$ for $M_{200} \gtrsim 10^{14} M_{\odot}$ (Fig. 4.1).
2. There is significantly greater scatter about the present-day median f_{CGM} for relatively low-mass haloes in TNG than in EAGLE. This scatter is particularly strong in the host haloes of $\sim L^*$ galaxies in TNG, corresponding to the mass scale for which AGN feedback becomes injected primarily in the kinetic mode. For haloes of $M_{200} \simeq 10^{12-12.5} M_{\odot}$ the 10th – 90th percentile ranges are 0.15 (EAGLE) and 0.37 (TNG) (Fig. 4.1).
3. In both simulations, this scatter about the present-day median f_{CGM} correlates strongly, negatively and significantly with scatter in the mass of the halo's central BH, M_{BH} . Haloes of fixed mass whose central galaxy has a more massive BH than is typical therefore exhibit systematically lower CGM mass fractions. In EAGLE, scatter about the median f_{CGM} is uncorrelated with the present-day accretion rate of the central BH, and hence with AGN luminosity, but in TNG these quantities are predicted to be strongly and positively correlated, particularly so at the halo mass at which the characteristic CGM mass fraction declines abruptly ($M_{200} \simeq 10^{12} M_{\odot}$), such that BHs hosted by CGM-rich haloes are accreting rapidly at $z = 0$. This indicates dissimilarity in the means by which circumgalactic gas is expelled in the two simulations, and the epoch at which it occurs (Fig. 4.1).
4. In both simulations, scatter about the median f_{CGM} correlates strongly, positively and significantly with scatter in the central galaxy's specific star formation rate (sSFR), and with the fraction of its stellar kinetic energy invested in co-rotation,

κ_{co} . Galaxies with higher-than-typical CGM mass fractions have an elevated probability of being star-forming ($\text{sSFR} > 10^{-11} \text{ yr}^{-1}$) and having strong rotational support ($\kappa_{\text{co}} > 0.4$), whilst galaxies with lower-than-typical CGM mass fractions have an elevated probability of being quenched and having low rotational support. These correlations are indicative of a causal connection between the internal properties of central galaxies and the state of their CGM (Fig. 4.2).

5. The circumgalactic gas associated with central $\sim L^*$ galaxies has significantly different radiative cooling time distributions in EAGLE and TNG, with haloes in the latter exhibiting more gas with cooling times $t_{\text{cool}} \lesssim 0.1 \text{ Gyr}$. Nonetheless, in both cases comparison of present-day haloes with high and low CGM mass fractions highlights that the latter have elevated characteristic cooling timescales as a consequence of expulsion of efficiently-cooling gas (Fig. 4.3).
6. The relation between the characteristic cooling time of the CGM at the present day, $t_{\text{cool}}^{\text{CGM}}$, and halo mass, M_{200} , is qualitatively similar in the two simulations, but with differences in detail that stem largely from differences in their respective $f_{\text{CGM}}(M_{200})$ relations. In both cases $t_{\text{cool}}^{\text{CGM}}$ is a monotonically-increasing function of M_{200} , but for haloes of $M_{200} = 10^{11.5} M_{\odot}$, $t_{\text{cool}}^{\text{CGM}} \simeq 1 \text{ Gyr}$ in EAGLE and $\simeq 0.13 \text{ Gyr}$ in TNG, reflecting the higher f_{CGM} of low-mass haloes in the latter. The CGM cooling time becomes similar to the Hubble time in present-day haloes of $M_{200} \gtrsim 10^{13} M_{\odot}$ in EAGLE, and $M_{200} \gtrsim 10^{13.8} M_{\odot}$ in TNG. (Fig. 4.4).
7. Scatter about the median $t_{\text{cool}}^{\text{CGM}}(M_{200})$ correlates strongly and negatively with scatter about the median CGM gas fraction, $f_{\text{CGM}}(M_{200})$, in both simulations. Therefore, the elevation of the CGM cooling time in response to the expulsion of circumgalactic gas, shown in Fig. 4.3 for $\sim L^*$ galaxies, is a mechanism that applies to haloes of all masses explored here (Fig. 4.4).
8. In both simulations, scatter about the running medians (as a function of M_{200}) of both the sSFR and κ_{co} of central galaxies correlates negatively with scatter in $t_{\text{cool}}^{\text{CGM}}(M_{200})$, i.e., central galaxies in haloes with shorter cooling times tend to have higher sSFRs and greater rotational support. This suggests that the long-term evolution of both of these quantities is linked to the expulsion from the CGM of gas that would otherwise cool and replenish interstellar gas. The correlation is stronger for the sSFR than for κ_{co} , likely reflecting that the physical connection between CGM expulsion and morphological evolution is indirect. It is plausible that CGM expulsion facilitates morphological evolution by suppressing the replenishment of the ISM, making discs more susceptible to disruption by mergers and gravitational instability, and by inhibiting the regrowth of a disc component in quenched galaxies (Fig. 4.5).

9. In both simulations, scatter about the running median of $f_{\text{CGM}}(M_{200})$ correlates strongly and negatively with the ratio $V_{\text{DMO}}^{\max}/V_{\text{DMO}}^{200}$, for $M_{200} \lesssim 10^{12.8} M_{\odot}$. Here, V^{\max} is the maximum of the halo's circular velocity profile, V^{200} is the circular velocity at the virial radius, and the DMO subscript denotes that the measurement applies to the halo's counterpart identified in a simulation with identical initial conditions but considering only collisionless dynamics. This ratio is a proxy for the halo concentration and thus correlates strongly with the halo formation time (Fig. 4.6).
10. In EAGLE, scatter about the median $f_{\text{CGM}}(M_{200})$ correlates negatively with scatter about the median of the ratio $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}(M_{200})$. Here E_{FB} is the total energy injected into the halo by feedback from the central galaxy and its progenitors, and $E_{\text{bind}}^{\text{b}}$ is the binding energy of the halo's baryons. For haloes of $M_{200} \gtrsim 10^{12} M_{\odot}$ the overall relation is driven by energy injection from AGN feedback. In TNG, these quantities do not correlate, but this is a consequence of the feedback energy budget being dominated by thermal mode AGN feedback, which suffers from numerical overcooling in TNG. If one considers only the contribution to E_{FB} from the efficient kinetic AGN mode, a negative correlation is recovered, similar to that in EAGLE. In both simulations, diversity in f_{CGM} is therefore driven primarily by variations in the energy injected by efficient feedback processes, relative to the binding energy of the halo's baryons (Fig. 4.6).
11. The functional form of the relationship between $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}$ and M_{200} is broadly similar in the two simulations, but there are differences. In EAGLE, galaxies hosted by haloes $M_{200} \lesssim 10^{12.5} M_{\odot}$ typically inject $E_{\text{FB}} \simeq 5E_{\text{bind}}^{\text{b}}$, and for haloes $M_{200} \lesssim 10^{12.0} M_{\odot}$, the energy is dominated by feedback from star formation (SF). In more massive haloes, the ratio declines gradually and monotonically, approaching unity for haloes of $M_{200} \simeq 10^{13.5} M_{\odot}$. For haloes of $M_{200} \gtrsim 10^{13} M_{\odot}$, AGN feedback marginally contributes more to the cumulative energy budget than SF feedback. In TNG, haloes with $M_{200} \lesssim 10^{12} M_{\odot}$ typically inject $E_{\text{FB}} \gtrsim 50E_{\text{bind}}^{\text{b}}$, i.e. an order of magnitude more than for EAGLE, and for more massive haloes this declines monotonically, reaching $E_{\text{FB}} \simeq E_{\text{bind}}^{\text{b}}$ at $M_{200} = 10^{14} M_{\odot}$. For all M_{200} , thermal mode AGN feedback dominates the energy budget. Kinetic-mode AGN becomes important abruptly in haloes $M_{200} \simeq 10^{12.3} M_{\odot}$, and dominates strongly over SF feedback in massive haloes. Despite these differences, the sum of the energies injected as feedback that efficiently couples to the gas in TNG (SF feedback and kinetic-mode AGN) is comparable to the total energy injected into EAGLE haloes (Fig. 4.7).

12. Scatter about the running median of $E_{\text{FB}}/E_{\text{bind}}^{\text{b}}(M_{200})$ correlates positively with residuals about the running median of $V_{\text{DMO}}^{\text{max}}/V_{\text{DMO}}^{200}(M_{200})$ in EAGLE. In TNG, these quantities do not correlate for $M_{200} \gtrsim 10^{12} M_{\odot}$, however if one considers only the contribution to E_{FB} from the efficient kinetic AGN feedback mode, a positive correlation is also recovered. This indicates that central galaxies hosted by high-concentration haloes inject relatively more energy via efficient feedback, relative to the binding energy of their haloes, providing a plausible explanation for the negative correlation of the CGM mass fraction with halo concentration at fixed M_{200} in both simulations (Fig. 4.7).

I noted in Chapter 3 that a key prediction stemming from the EAGLE simulations is that the present-day CGM gas fraction of haloes is connected to the intrinsic properties of their haloes, such as their binding energy or concentration, an effect that is physically ‘transmitted’ by AGN feedback. Here I have shown that the same holds for TNG. I note that [Terrazas et al. \(2019\)](#) recently concluded that galaxies in TNG are quenched when the energy injected by their central BH in the kinetic mode exceeds the binding energy of gas within the effective radius. The correlations presented here, from both the EAGLE and TNG simulations, indicate that these intrinsic halo properties also influence readily-observable properties of present-day galaxies, such as their star formation rate and morphology. These halo properties are effectively encoded within the phase-space configuration of the initial conditions; if the growth of BHs and their influence on the CGM is sufficiently realistically captured by the current generation of state-of-the-art cosmological simulations, it appears that galaxies may be affected by halo assembly bias as a consequence of efficient AGN feedback.

In both EAGLE and TNG, the influence of halo properties on central galaxies is primarily a consequence of the expulsion of circumgalactic gas (or a reconfiguration of intragroup/intracluster gas in the inner halo). In both simulations, the expulsion of circumgalactic gas leads to the elevation of the characteristic CGM cooling time, and depletes haloes of gas that would otherwise replenish interstellar gas consumed by star formation or expelled by feedback processes. Efficient feedback also heats and pressurises the remaining CGM, possibly also contributing to the elevated cooling time by inhibiting the accretion of gas from the IGM, or the re-accretion of gas expelled by feedback, onto the CGM (so-called ‘preventative feedback’). The paucity of efficiently-cooling circumgalactic gas leads to the preferential quenching of central galaxies hosted by high-concentration haloes. On longer timescales, it also facilitates their evolution towards an early-type morphology. In both simulations, the CGM is modulated by AGN feedback at a similar mass scale for which galaxies become quenched: the corollary of the

results presented here is therefore that the feedback-driven expulsion of circumgalactic gas is predicted to be a crucial, but largely over-looked, step in these processes.

It is encouraging that the same trends are seen in both the EAGLE and TNG simulations, two state-of-the-art cosmological hydrodynamical simulations of galaxy formation with significantly different hydrodynamics solvers and subgrid implementations of unresolved physical processes, as this signals consensus in regard to this conclusion. However, there are two significant caveats. Firstly, it is important to recognise that the two suites share significant similarities; in particular, the fashion by which BHs are seeded, and then grow and merge, is similar in both cases, being based on the scheme introduced by [Springel et al. \(2005\)](#). BHs are thus seeded at similar stages of the formation and assembly of haloes in the two simulations.

Secondly, and perhaps more importantly, the physical origin of the correlation between scatter about the CGM mass fraction and the halo concentration (at fixed mass) is different in the two simulations. Although in both simulations scatter in f_{CGM} at fixed halo mass appears to be a consequence of halo-to-halo variations in the amount of energy injected via efficient feedback relative to the binding energy of the halo baryons, the cause of these variations differs. In EAGLE, which adopts a fixed AGN feedback efficiency, the expulsion of the CGM is simply a response to high BH accretion rates. In TNG, it is a response to the onset of kinetic AGN feedback. Therefore, the CGM mass fraction is depleted in early forming, high-concentration haloes in EAGLE because the central BH is able to reach high BH accretion rates sooner, and in TNG because the central BH reaches the calibrated pivot mass for the transition between thermal and kinetic feedback sooner. Typically, high BH accretion rates in EAGLE occur at earlier epochs than the BH pivot mass is reached in TNG, leading to marked difference in the present-day $\Delta f_{\text{CGM}} - \Delta \dot{M}_{\text{BH}}$ relations exhibited by the two simulations.

These differences lead to significant, and in principle testable, differences in scaling relations involving, for example, the relationship between the column density of CGM OVI absorbers and the specific star formation rate of central galaxies at fixed halo mass (see e.g. [Oppenheimer et al., 2016b](#); [Nelson et al., 2018b](#)), and the relationship between the present-day CGM mass fraction of haloes and the accretion rate of their central BHs, and hence the luminosity of their AGN (as shown in Fig. 4.1). In particular, while TNG predicts a strong anti-correlation between the CGM mass fraction and the AGN luminosity of haloes with $M_{200} \sim 10^{12} M_{\odot}$, EAGLE predicts no such relation. I anticipate that the question of which of these scenarios is the more realistic might also be meaningfully addressed with observations of diffuse circumgalactic gas enabled by future X-ray observatories such as *Athena* and *Lynx*. I emphasise, however, that despite these differences, both EAGLE and TNG predict that the ejection of circumgalactic

gas by AGN feedback is a crucial step in the quenching and morphological evolution of galaxies.

Chapter 5

The galaxy-CGM connection in dark matter haloes with controlled assembly histories

5.1 Introduction

The central result of Chapters 3 and 4 is that the assembly history of a dark matter halo (set only by cosmology) can be directly connected to the quenching of its central galaxy, and may also be implicated in a morphological transformation from late-type to early-type. Differences in the integrated energy from BH feedback (induced by variations in assembly history) regulate the baryon content of the CGM, and in doing so modulate the ability of CGM gas to cool onto the ISM, leading to quenching and facilitating morphological transformation.

These conclusions were reached through the identification of correlations at fixed halo mass in large periodic simulation volumes, which contain diverse populations of $\sim L^*$ galaxies. Using these simulations permits the comparison of haloes with markedly different assembly histories, and following my results, very different present-day CGM mass fractions and central galaxy properties. In examining these correlations, however, I am ultimately comparing *different* haloes, which limits my ability to conclusively link the assembly histories of dark matter haloes to the properties of galaxies and their CGM as other driving factors, such as the environment of the halo, may be important.

One might perform increasingly large-volume simulations with populations that can be sub-sampled ever more finely to control for these effects, however this would be very costly and inefficient. An alternative method, which yields a far ‘cleaner’ test of these

conclusions, is to modify the initial conditions of an individual halo, such that the assembly history is tightly-controlled, but the large-scale environment of the system remains fixed, thus eliminating any potentially confounding variables. This is the approach I take in this chapter, in which I utilise the “genetic modification” technique described by [Roth et al. \(2016\)](#), [Pontzen et al. \(2017\)](#) and [Rey & Pontzen \(2018\)](#) to modify the initial conditions of a zoom simulation of a single dark matter halo, thus creating a controlled experiment in which the effects of assembly history on the evolution of a single galaxy-CGM ecosystem can be directly tested.

This chapter will be structured as follows: in Section 5.2 I outline the rationale by which my candidate dark matter halo was selected, describe how the initial conditions were generated and genetically modified, give details on the simulation model, explain how the progenitors of the system are tracked through the simulation, and explain how certain diagnostic quantities were calculated. In Section 5.3 I evaluate the properties and assembly histories of the modified haloes, before examining the effects of a modified assembly history on the galaxy-CGM ecosystem in Section 5.4. Within the latter section, I examine the effects of assembly history on the CGM mass fraction (Section 5.4.1), on the growth of supermassive BHs and their associated AGN feedback (Section 5.4.2) and on the properties of the central galaxy (Section 5.4.3). Finally, in Section 5.5 I examine the effects of the ejection of baryons from the CGM on its thermodynamic state. I summarise my results in Section 5.6. Throughout, I adopt the convention of prefixing units of length with ‘c’ and ‘p’ to denote comoving and proper scales respectively, e.g. ckpc for comoving kiloparsecs.

5.2 Methods

5.2.1 Construction of the initial conditions

My analyses are based on a suite of simulations that follow the formation and evolution of the environment of an individual central galaxy, including its dark matter halo and circumgalactic medium, in its full cosmological context. This is most efficiently achieved via the adoption of ‘zoomed’ initial conditions (see e.g. [Katz & White, 1993](#); [Bertschinger, 2001](#)), whereby only the object of interest is followed at the high resolution and with hydrodynamics, whilst the remainder of the periodic volume is followed with purely collisionless dynamics and at reduced resolution.

To obtain a fiducial case of a present-day $\sim L^*$ galaxy with a broadly typical specific star formation rate and circumgalactic gas fraction, I identify candidate *galaxies* for resimulation from a parent volume evolved with a detailed galaxy formation model,

rather than following the more common practice of identifying candidate dark matter haloes from a simulation evolved with purely collisionless dynamics. I identify candidate galaxies from a periodic simulation of uniform resolution whose initial conditions were generated with the GENETIC software; use of this simulation rather than, for example, simulations from the EAGLE suite, simplifies the later process of applying modifications to multi-resolution zoom initial conditions. This ‘parent’ simulation adopts the cosmological parameters advocated by the [Planck Collaboration et al. \(2016\)](#), $h = 0.6727$, $\Omega_0 = 0.3156$, $\Omega_\Lambda = 0.6844$, $\sigma_8 = 0.831$ and $n_s = 0.9645$. It is $L = 50$ cMpc on a side, and is populated with $N = 512^3$ collisionless dark matter particles of mass $3.19 \times 10^7 M_\odot$ and an (initially) equal number of baryonic particles of mass $5.6 \times 10^6 M_\odot$. The cosmological parameters and particle masses are therefore sufficiently similar to those of the standard-resolution EAGLE simulations (i.e. those of the same resolution as its largest-volume run, Ref-L100N1504) such that simulating the volume with the EAGLE Reference model yields a galaxy population of similar realism to those simulations.

In Chapter 4, I noted that in both EAGLE and TNG, the influence of efficient AGN feedback on the mass fraction of the CGM, and by extension the properties of galaxies, is most apparent in haloes of present day mass $M_{200} \sim 10^{12.5} M_\odot$. I therefore sought central galaxies hosted by haloes of this mass, and identified as my resimulation target a present-day central galaxy of stellar mass $M_\star = 4.3 \times 10^{10} M_\odot$, hosted by a halo of mass $M_{200} = 3.4 \times 10^{12} M_\odot$ and virial radius $r_{200} = 318$ kpc, as my exemplar galaxy. The galaxy exhibits an extended stellar disc, has a stellar half-mass radius of $r_{\star,1/2} = 7.5$ kpc, and is quiescently star-forming ($sSFR = 10^{-10.2} \text{ yr}^{-1}$). The CGM mass fraction of the host halo, f_{CGM} , normalised by the cosmic baryon fraction, Ω_b/Ω_0 , is 0.31. Following the methods in Chapter 4, I define $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, where M_{CGM} is the total mass of all gas within the virial radius that is not star-forming. The centre of the nearest halo of at least equal mass is 3.3 Mpc from the centre of the target halo, a separation of more than 6 times the virial radius of the more massive halo. This is advantageous as it yields a compact and simply-connected Lagrangian region when the matter comprising the halo is traced back to the initial conditions, and thus improves the efficiency of the zoomed resimulation in terms of both compute time and memory footprint.

To construct the multi-resolution zoom initial conditions, I first identify all dark matter particles at $z = 0$ residing within a sphere of radius $r = 3r_{200}$, centred on the potential minimum of the halo, and trace them to their coordinates in the unperturbed (effectively $z = \infty$) particle distribution, which is achieved in practice by tracing the particles to their coordinates in the initial conditions at $z = 99$ and then subtracting the linear Zel’dovich displacement. The zoomed initial conditions of the halo, corresponding to the unmodified assembly history, are then constructed by masking the Lagrangian region defined by this particle selection in the unperturbed mass distribution; the enclosed

mass distribution is then resampled at higher resolution with a factor of 8 more particles (which represent both baryonic and dark matter), whilst the remainder of the volume is resampled with a factor of 8 fewer particles, which act as low-resolution boundary particles to provide the correct large-scale gravitational forces. The Zel'dovich displacements corresponding to the original phases and power spectrum (the latter now sampled to the higher and lower Nyquist frequencies of the high-resolution and boundary particles, respectively) are then reapplied.

I then apply the linear genetic modification technique of Roth et al. (2016) and Pontzen et al. (2017)¹ to these initial conditions (the ‘Fiducial’ case) to construct a pair of complementary initial conditions, designed to yield modified halo assembly histories whilst maintaining approximately the same present-day halo mass. I have modified the initial conditions to yield assembly histories shifted, with respect to the Fiducial case, to both an earlier time (‘GM-early’) and a later time (‘GM-late’). Specifically, the initial conditions were adjusted such that at $z = 99$, the matter that will eventually comprise the halo’s main progenitor at $z = 2$ has a mean overdensity differing from the Fiducial case by factors of 1.05 and 0.95 respectively. The adjustments simultaneously ensure that the mean overdensity (at $z = 99$) of the matter that ultimately constitutes the halo at $z = 0$ remains unchanged.

In common with the construction of the EAGLE initial conditions (see Appendix B4 of Schaye et al., 2015), the final step is to replace the high-resolution particles in each case with a pair of particles consisting of a gas particle and a dark matter particle, with a gas-to-dark matter mass ratio of $\Omega_b/(\Omega_0 - \Omega_b)$. The masses of the gas, dark matter and low-resolution mass tracer particles are therefore $m_g = 7.35 \times 10^5 M_\odot$, $m_{dm} = 3.98 \times 10^6 M_\odot$, and $m_{lr} = 3.02 \times 10^8 M_\odot$. The particle pairs are positioned such that their centre of mass corresponds to the position of the replaced particle, with the gas and dark matter particles moved in the $(1, 1, 1)$ and $(-1, -1, -1)$ coordinate directions, respectively.

5.2.2 The EAGLE model

To evolve the genetically-modified initial conditions in time, I utilise the EAGLE simulation code, which is described in detail in Chapter 2. The particle mass resolution of the initial conditions is higher than that of the simulations studied previously in this thesis, resulting in a reduction of numerical losses in both stellar and AGN feedback. The galaxy population evolved with the EAGLE Reference model falls below the observed galaxy stellar mass function (GSMF) at the mass scale of my selected system (Schaye et al., 2015), indicating that feedback is too efficient; an increase in resolution

¹Further information on the technique is also given in Rey & Pontzen (2018).

will only exacerbate this issue. I therefore utilise the recalibrated (RECAL) parameter values for EAGLE’s subgrid feedback prescriptions, described by Schaye et al. (2015, their Table 3), which provide a better match to the GSMF at higher resolution. For each set of GM initial conditions, I have also run counterpart simulations containing only collisionless gravitational dynamics (DMONLY), “non-radiative” simulations implementing hydrodynamics, but no star formation, BH formation, feedback or radiative cooling (NONRAD), and full-physics simulations where no black holes are seeded and no AGN feedback occurs (NOAGN).

The subgrid prescriptions for both star formation and feedback in the EAGLE model are stochastic in nature. The conversion of star-forming gas particles to star particles proceeds according to a probability set by the Kennicutt-Schmidt law, and the feedback energy liberated by processes associated with star formation or BH growth is used to set the probability that neighbouring gas particles are heated by a fixed temperature increment. These probabilities are governed by a preset ‘seed’ for star formation and feedback, defined before the simulation is run. To test the influence of this stochasticity on my simulations evolved with the RECAL model, I evolve the three sets of initial conditions with nine different seeds each (including the standard EAGLE seed), for a total of 27 simulations. Performing these simulations has the added benefit of improving the sampling of particle-based measurements through stacking, without the need to increase the resolution of the simulation, which would require recalibration of the subgrid model.

5.2.3 Identifying haloes, galaxies and their progenitors

As in Chapters 3 and 4, haloes are identified in the simulations through the application of the friends-of-friends (FoF) algorithm to the dark matter distribution, with a linking length of 0.2 times the mean interparticle separation. Gas, star and BH particles are then assigned to the FoF group (if any) of their nearest dark matter particle. The SUBFIND algorithm (Springel et al., 2001; Dolag et al., 2009) is then used to identify bound substructures within haloes. Throughout this chapter, the properties of haloes, such as the spherical overdensity mass (M_{200}) and CGM mass fraction (f_{CGM}), are computed within a radial aperture r_{200} about the halo’s most bound particle which encloses a mean density equal to 200 times the critical density, ρ_{crit} . Galaxy properties, such as the specific star formation rate, are computed by aggregating the properties of the relevant particles within 30pkpc of the halo centre, following Schaye et al. (2015). For the wider EAGLE population in the Ref-L100N1504 simulation, I construct a sample of present-day haloes of mass $M_{200} > 10^{11.5} \text{ M}_\odot$ matching that of Chapters 3 and 4.

To obtain the merger history of the genetically-modified galaxy-halo system, I first identify the 100 most bound dark matter particles² in the fiducial, unmodified system (using the standard EAGLE feedback seed) at $z = 0$. The main progenitor of the system in each prior snapshot is then defined as the subhalo (as defined by SUBFIND) containing the greatest number of these particles, yielding a fiducial merger history. At each snapshot output, I wish to compare the system with the same object in the other simulations which either have GM initial conditions, or employ different physical models (NOAGN/DMONLY), however applying this technique to each simulation individually can result in the tracing of different objects to the fiducial case. With the fiducial merger history in hand, I therefore cross-match the 100 most bound dark matter particles in the subhalo *in each snapshot* with particles in every other simulation, in order to identify the counterpart subhalo containing the greatest number of these particles. This method yields a reliable tracking of the same object across all simulations, though I highlight where the identification of the system is ambiguous throughout this chapter.

5.2.4 Feedback energetics and cooling timescales

I use the total energy injected by AGN feedback relative to the binding energy of the halo baryons as a diagnostic quantity in Section 5.4.2. The total energy injected over the lifetime of a BH of mass M_{BH} at time t is given by

$$E_{\text{AGN}}(t) = \frac{\epsilon_f \epsilon_r}{1 - \epsilon_r} (M_{\text{BH}}(t) - M_{\text{BH,seed}}) c^2, \quad (5.1)$$

where $M_{\text{BH,seed}} = 10^5 M_\odot/h$ is the seed mass of the black hole, c is the speed of light, $\epsilon_r = 0.1$ is the radiative efficiency assumed for the BH accretion disk, and $\epsilon_f = 0.15$ is a calibrated parameter regulating the coupling of feedback energy to the surrounding gas. Approximately 1.67% of the rest mass energy of gas accreted by the BH is therefore coupled to the CGM. I subtract off the contribution from the BH’s seed mass, since it has not been injected into the gas in the simulation. The “intrinsic” binding energy of the baryons at time t , $E_{\text{bind}}^b(t)$, is obtained by calculating the binding energy of the halo in an equivalent, collisionless, dark-matter only simulation³, and multiplying this by the cosmic baryon fraction: $E_{\text{bind}}^b(t) = (\Omega_b/\Omega_0) E_{\text{DMO}}^{200}(t)$.

I investigate the distribution of the radiative cooling times of circumgalactic gas particles in Section 5.5, which I compute based on their internal thermal energy, u , and their bolometric luminosity, L_{bol} , via $t_{\text{cool}} = u/L_{\text{bol}}$. The bolometric luminosity is given

²The obtained merger history is not strongly sensitive to this choice.

³I utilise ‘intrinsic’ binding energy measurements from the DONLY simulation, because the inclusion of baryonic physics can systematically alter the binding energy of the underlying dark matter structure to a degree comparable with the intrinsic scatter at a given M_{200} .

by $L_{\text{bol}} = n_{\text{H}}^2 \Lambda V$, where n_{H} is the hydrogen number density of the gas particle, Λ is its (volumetric) radiative cooling rate specified by its temperature, density, element abundances, and the incident flux from the CMB and metagalactic UV/X-ray radiation fields, and $V = m_{\text{g}}/\rho$ where m_{g} is the particle mass and ρ is its mass density.

Following the subgrid prescription for radiative cooling in EAGLE, I compute the volumetric net radiative cooling rates using the [Wiersma et al. \(2009a\)](#) tabulated rates, which were computed using CLOUDY version 07.02 ([Ferland et al., 1998](#)). The tables specify the cooling rate as a function of hydrogen number density, n_{H} , temperature, T , and redshift, z for each of the 11 tracked elements (H, He, C, N, O, Ne, Mg, Si, S, Ca and Fe) in EAGLE, and I interpolate them in $\log_{10} n_{\text{H}}$, $\log_{10} T$, z , and, in the case of the metal-free cooling contribution, the helium fraction $n_{\text{He}}/n_{\text{H}}$. The element-by-element contributions are then used to compute the net cooling rate for the particle:

$$\Lambda = \Lambda_{\text{H,He}} + \sum_{i>\text{He}} \Lambda_{i,\odot} \frac{n_e/n_{\text{H}}}{(n_e/n_{\text{H}})_{\odot}} \frac{n_i/n_{\text{H}}}{(n_i/n_{\text{H}})_{\odot}}, \quad (5.2)$$

where $\Lambda_{\text{H,He}}$ is the metal-free contribution, $\Lambda_{i,\odot}$ is the contribution of element i for the solar abundances assumed in CLOUDY, n_e/n_{H} is the particle electron abundance, and n_i/n_{H} is the particle abundance in element i .

5.3 Evaluation of modified haloes and their assembly histories

I begin with an evaluation of the assembly histories and present-day properties of the haloes yielded by my application of the genetic modification scheme to the initial conditions of the Fiducial case. Recall that my aim is to systematically shift the assembly of the halo to earlier or later times, but without inducing strong changes to the mass of the halo at $z = 0$. The cleanest assessment of the success of the technique comes from a comparison of the DMONLY haloes, in the absence of baryon physics.

Fig. 5.1 shows maps of the dark matter surface density of the GM-early (left column), Fiducial (centre column) and GM-late (right column) haloes in the DMONLY simulations, at the present day (upper row) and at $z = 2$ (lower row). The white circle on each panel denotes the virial radius, $r_{200}(t)$. The field of view in each case is 1.31 cMpc, which corresponds to $4r_{200}$ for the Fiducial case at $z = 0$. The images present a striking representation of the effect of the technique: the structure of the Fiducial and modified haloes at the present day is broadly similar, but at $z = 2$ major differences in the structure of the halo's main progenitor are evident, with the assembly of the GM-early

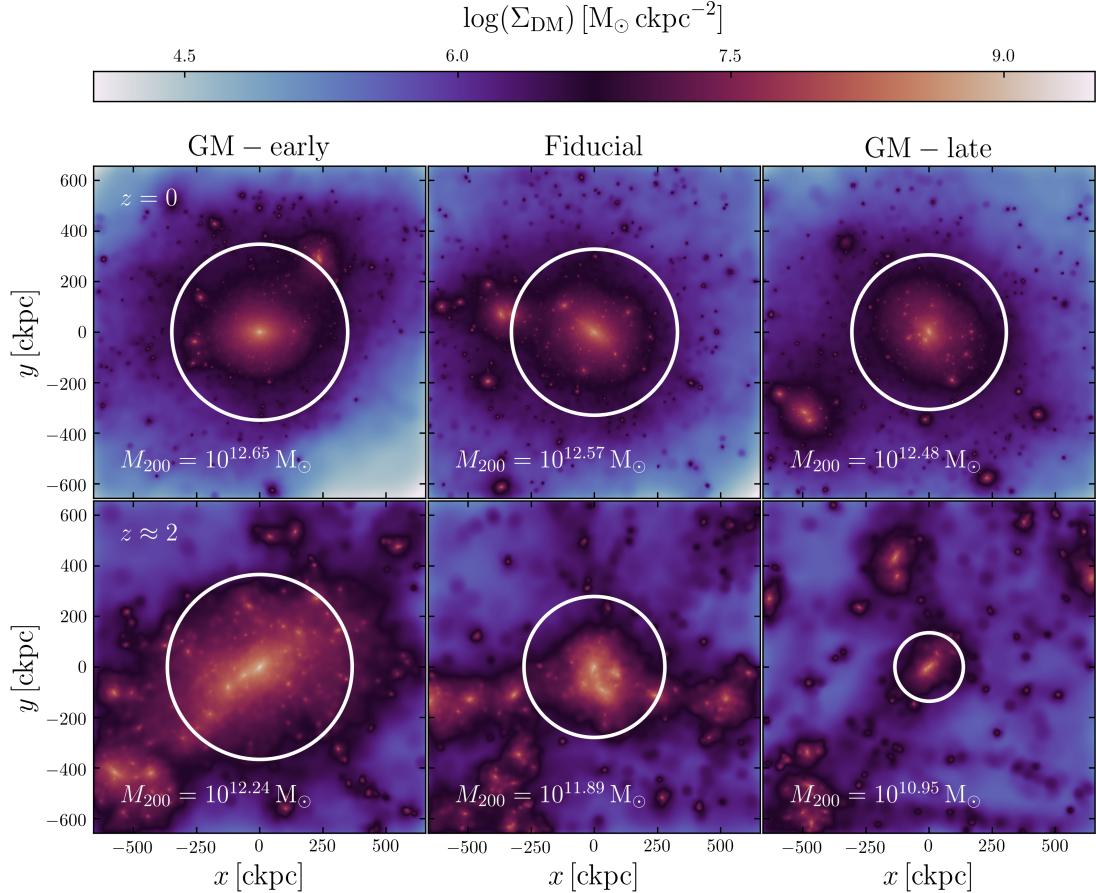


FIGURE 5.1: Surface density maps of the dark matter distribution for the GM-early (left column), Fiducial (centre column) and GM-late (right column) haloes in the DMONLY simulations, at $z = 0$ (upper row) and at $z \approx 2$ (lower row). The white circle on each panel denotes the virial radius, r_{200} . The field of view in each case is 1.31 cMpc, which corresponds to $4r_{200}$ for the Fiducial case at $z = 0$. The halo mass, M_{200} , is quoted on each panel.

(GM-late) case being significantly advanced (delayed) with respect to the Fiducial case. The present-day halo masses of the GM-early and GM-late cases deviate with respect to Fiducial by +0.08 dex and -0.09 dex. In contrast, at $z = 2$ the GM-early case has formed $\simeq 40\%$ of its final mass, in comparison to $\simeq 20\%$ for the Fiducial case and only $\simeq 3\%$ for the GM-late case.

For completeness, Fig. 5.2 repeats the dark matter surface density maps of Fig. 5.1, but with a much larger field of view of 25 cMpc and considering also the low-resolution boundary particles, to show the present-day large-scale structure of the halo in the three simulations. In these figures, the white circles denote a scale corresponding to $5r_{200}(t)$. It is clear that the genetic modifications enable controlled adjustment of the assembly history of the target halo without inducing changes to the large-scale structure, even at early times.

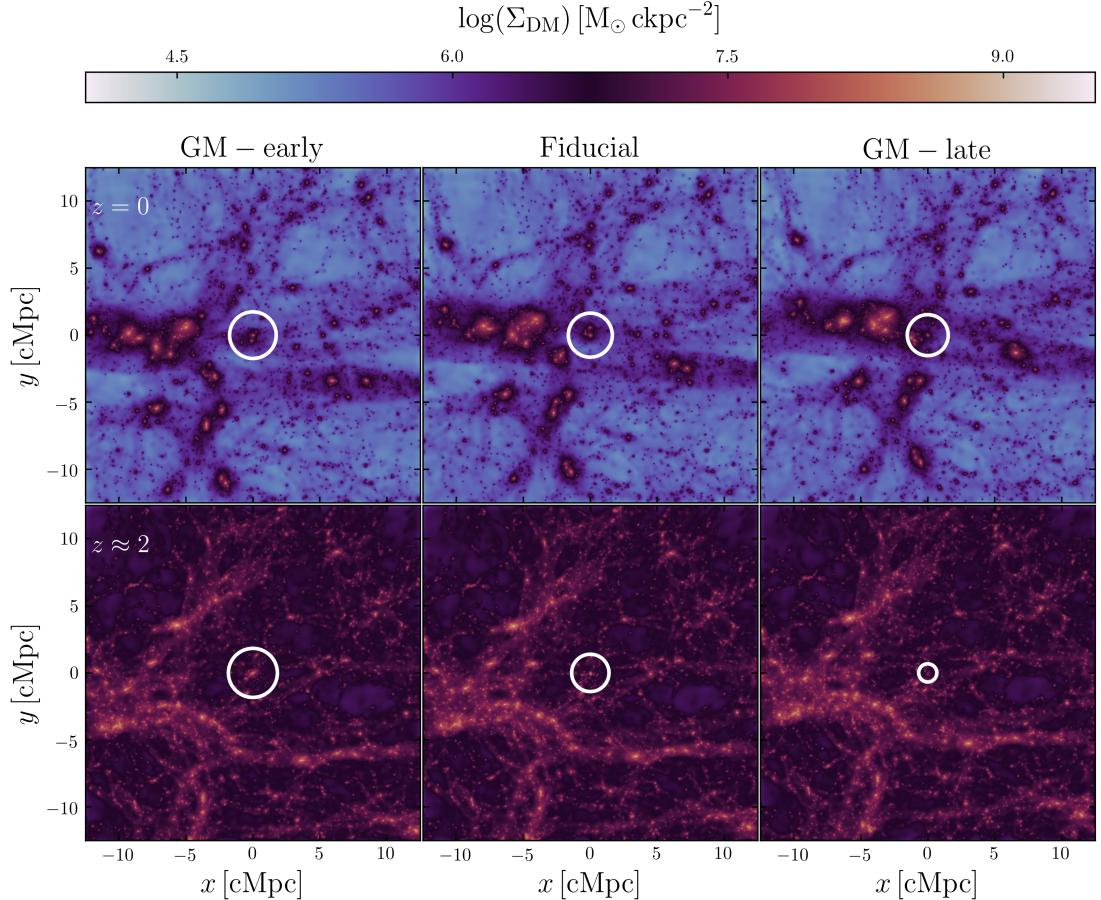


FIGURE 5.2: As Fig. 5.1, but with a 25 cMpc field of view. The surface density distribution of low-resolution boundary particles is included in addition to the dark matter distribution to reveal the large-scale structure in which the high-resolution zoom region is embedded. To establish a sense of scale, white circles are overlaid with radii equal to $5r_{200}$.

Fig. 5.3 shows mass accretion history of the haloes, defined as the halo mass of the main progenitor of the present-day halo, $M_{200}(t)$. The figure shows both the DMONLY (left) and RECAL (right) simulations. The tracks in each case are plotted with dotted lines until $M_{200}(t)/M_{200}(z = 0) = 0.05$, since in the early stages of halo assembly the identification of the main progenitor can be ambiguous. I focus first on the DMONLY panel since, as noted above, this case represents the ‘purest’ test of the genetic modifications, since halo’s assembly is not influenced by any differences in the baryonic physics potentially induced by the modifications. The sub-panel shows the relative difference of the GM-early and GM-late simulations with respect to the Fiducial case.

As is critical for my purposes, the genetic modification process induces very strong deviations from the Fiducial halo mass growth in the DMONLY modified realisations. At $z = 2$, the GM-early system has already assembled a halo mass of $M_{200} = 1.75 \times 10^{12} M_{\odot}$ (40% of its final mass), while the Fiducial halo has a mass of $M_{200} = 7.76 \times 10^{11}$

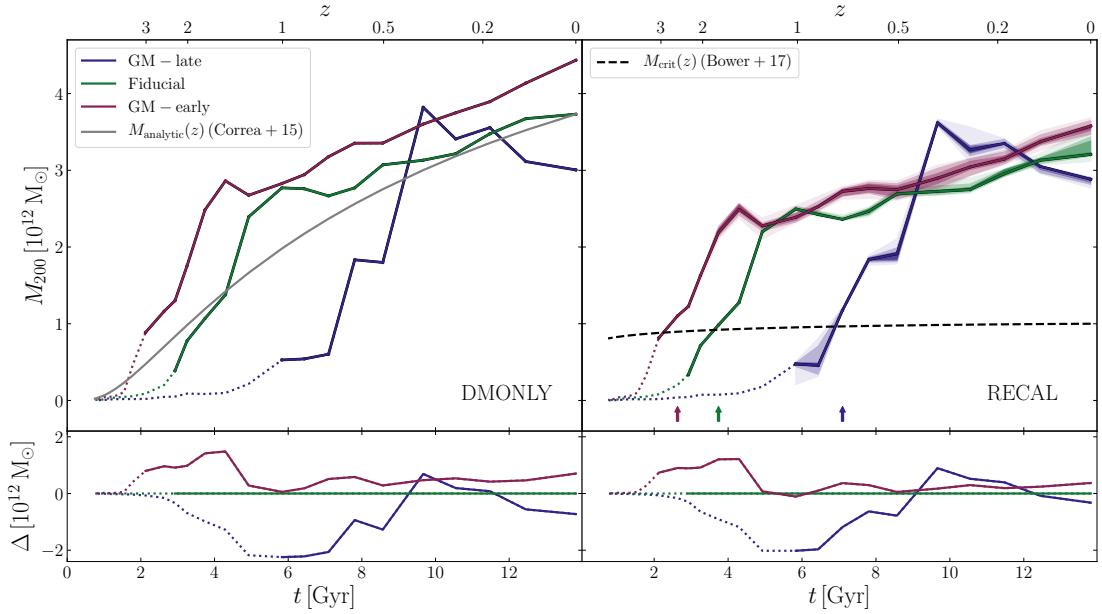


FIGURE 5.3: The halo mass accretion histories, $M_{200}(t)$ of the three GM cases in the DMONLY (left column) and RECAL (right column) simulations. Solid lines show the median M_{200} for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z = 0) < 0.05$. Sub-panels show the deviation, $\Delta(t)$, of the median value from that of the Fiducial case. The grey curve denotes an analytical prediction for the mass accretion history of the Fiducial halo, $M_{\text{analytic}}(z)$, derived from the methods of Correa et al. (2015). The black dashed line denotes the redshift-dependent critical halo mass, $M_{\text{crit}}(z)$, where the buoyant transport of outflows from EAGLE’s stellar feedback ceases to be efficient (Bower et al., 2017). The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

M_\odot (20% of its final mass) and the GM-late halo has assembled a halo mass of only $M_{200} = 8.92 \times 10^{10} M_\odot$ (3% of its final mass). By $z = 1$, the evolutionary tracks for the GM-early and Fiducial cases converge, and the halo grows in mass in a similar fashion for both thereafter. In the GM-late simulations, my halo grows in mass far more steadily, only attaining (and briefly exceeding) the mass of the other realisations after $z = 0.5$.

To assess whether the growth of the Fiducial halo is typical for its mass, I show in the left-hand panel of Fig. 5.3 the analytical prediction for the mass growth of the halo, $M_{\text{analytic}}(z)$, derived from the extended Press-Schechter formalism by Correa et al. (2015), for the appropriate background cosmology (that of the Planck Collaboration et al., 2016). The mass evolution of the Fiducial halo initially lies below this prediction, but then exceeds it at $z \approx 1.7$ and remains more massive than predicted until $z = 0$. I note therefore that the mass accretion history of the Fiducial halo is not entirely representative for haloes of its mass: it acquires half of its present-day mass by $z = 1.74$, approximately 2.1 Gyr earlier than the $z = 1.00$ predicted by Correa et al. (2015).

However as is clear from the GM-early and GM-late cases, significantly greater deviations from the typical are feasible.

I now turn to the halo assembly histories in the simulations with full baryonic physics, shown in the right panel of Fig. 5.3. Here, and in subsequent figures of this style, for the GM-early, Fiducial and GM-late families I show the evolution derived from all nine of the simulations run in each case (each with a different random seed). The solid lines do not represent the evolution of the system for any particular seed, but show the median value of the quantity of interest at each epoch. This median is used when plotting the relative differences in the sub-panel in all plots of this type. The distribution of values across the nine simulations in each of the GM-early, Fiducial and GM-late families is illustrated with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. I quantify the scatter induced by differing seeds with the interquartile range (IQR); in a rank-ordered sample of values taken from the nine simulations comprising each of the GM-early, Fiducial and GM-late families, the third and seventh values are good approximations for the 25th and 75th percentiles, and I quote the difference between these as the IQR throughout. I also show as a dotted line on Fig. 5.3 the redshift-dependent critical halo mass introduced by [Bower et al. \(2017\)](#), namely the mass at which buoyant transport of winds from EAGLE’s stellar feedback is expected to cease to be efficient, $M_{\text{crit}} = (\Omega_0(1+z)^3 + \Omega_\Lambda)^{-1/24} \times 10^{12} \text{ M}_\odot$. [Bower et al. \(2017\)](#) noted that the rapid growth of central BHs tends to begin when haloes reach this mass, and interpreted this as a signature of BH fuelling by cooling flows from the quasi-hydrostatic CGM that builds in response to the cessation of buoyant transport. The epoch at which this threshold is reached by the median of each of family is denoted by an arrow and has value $t = 2.63$ Gyr (GM-early), $t = 3.74$ Gyr (Fiducial) and $t = 7.10$ Gyr (GM-late). I discuss the consequences of the significant differences in these values in following sections.

As might be expected, the halo assembly histories in Recal are qualitatively similar to those of their DMONLY counterparts, but there are significant quantitative differences, most plainly the reduction in mass associated with the expulsion of a significant fraction of the circumgalactic medium. Here the present-day halo masses are $M_{200} = 3.58 \times 10^{12} \text{ M}_\odot$ (GM-early, IQR= 0.009 dex), $3.21 \times 10^{12} \text{ M}_\odot$ (Fiducial, IQR= 0.024 dex), and $2.88 \times 10^{12} \text{ M}_\odot$ (GM-late, IQR= 0.008 dex). The results of Chapters 3 and 4 indicate that this reduction should be stronger for earlier-assembling systems, which is the case here; as a result, the three families converge to more similar present-day masses than in the DMONLY case, with the GM-early and GM-late cases deviating with respect to Fiducial by 0.05dex and -0.05dex, respectively. As might be naively expected, there is little scatter between runs adopting different seeds.

The above results demonstrate the precision of the genetic modification scheme; I can tune the initial conditions such that the assembly history of an individual system is systematically shifted to significantly earlier or later times than the fiducial case, while only introducing a scatter of ± 0.05 dex in the final halo mass in a full-physics simulation. As an example of this precision, in the context of the wider galaxy population in the largest EAGLE simulation volume (Ref-L100N1504, utilised in Chapters 3 and 4), a shift of 0.1 dex about $M_{200} = 10^{12.5} M_\odot$ gives a difference in the median f_{CGM} of only 0.028. Any more significant differences found will therefore be the result of the modified assembly history of the halo. My set of 27 zoom simulations now represents a controlled experiment in which I can directly test the impact of halo assembly history on the evolution of a galaxy and its CGM.

5.4 The influence of halo assembly history on the galaxy-CGM ecosystem

In this section I turn to an examination of the influence of halo assembly history on the properties of the central galaxy, its central supermassive BH and its circumgalactic gas, using my suite of zoom simulations. I begin by exploring the evolution of the CGM mass fraction in Section 5.4.1, the evolution of the central SMBH mass in Section 5.4.2, and then turn to the evolution of the central galaxy’s specific star formation rate and morphology in Section 5.4.3.

5.4.1 The influence of halo assembly history on the CGM mass fraction

Fig. 5.4 shows the evolution of the CGM mass fraction, $f_{\text{CGM}}(t)$, normalised by the cosmic baryon fraction, Ω_b/Ω_0 , for my three families of simulations. Following the methods in Chapter 4, I define $f_{\text{CGM}} \equiv M_{\text{CGM}}/M_{200}$, where M_{CGM} is the total mass of all gas within the virial radius that is not star-forming. The GM-late halo has the highest CGM mass fraction at the present day, $f_{\text{CGM}}/(\Omega_b/\Omega_0) = 0.50$ (IQR= 0.06), followed by the Fiducial halo (0.31, IQR= 0.05), and the GM-early halo (0.15, IQR= 0.07).

In all three families, f_{CGM} is low at early epochs (when the identification of the main progenitor can be ambiguous), likely because the potential of the nascent structure is too low to accrete photoionised gas after the epoch of reionisation. As the halo grows, f_{CGM} quickly increases to a peak, $\max(f_{\text{CGM}}) = (0.62, 0.58, 0.76)$ for GM-early, Fiducial and GM-late, respectively, but then begins to decline towards its $z = 0$ value over several Gyr, with the decline broadly commencing when the halo mass reaches M_{crit} (denoted by the coloured arrows). Oppenheimer et al. (2020) showed that this is not coincidental:

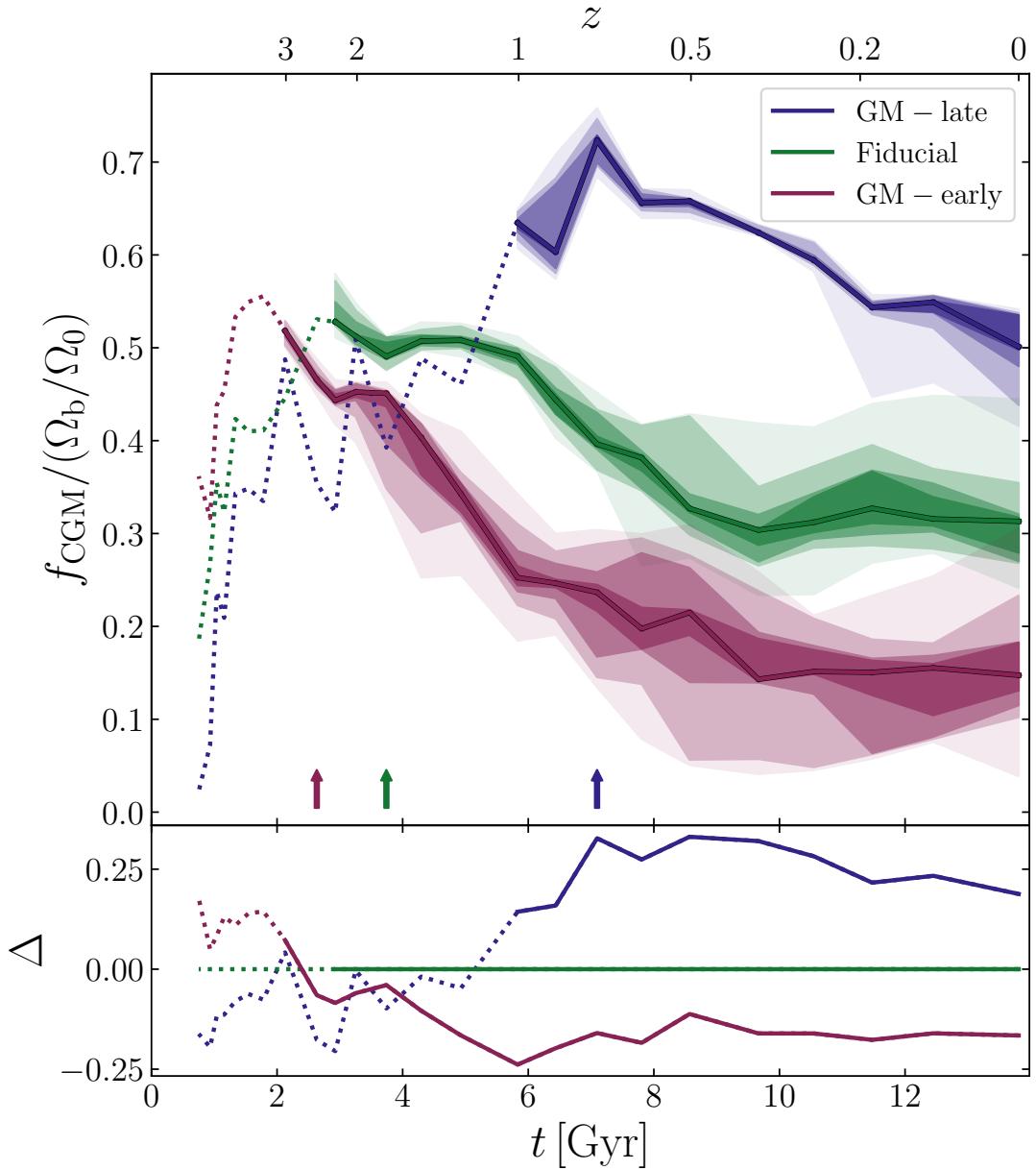


FIGURE 5.4: The evolution of the CGM mass fraction, $f_{\text{CGM}}(t) \equiv M_{\text{CGM}}(t)/M_{200}(t)$, normalised by the cosmic average baryon fraction, Ω_b/Ω_0 , for the three families of simulations. Solid lines show the median value for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z=0) < 0.05$. Sub-panels show the deviation, $\Delta(t)$, of the median value from that of the Fiducial case. The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

episodes of feedback driven by the rapid growth of the central BH in EAGLE, which Bower et al. (2017) found to be coincident with the halo reaching M_{crit} , are followed by explosive circumgalactic gas outflows. As noted in Section 5.3, the modified haloes reach M_{crit} significantly earlier (GM-early) and later (GM-late) than the Fiducial case, resulting in markedly different evolution of f_{CGM} in each case. By $z = 1$ the GM-early halo is already significantly depleted of circumgalactic gas, $f_{\text{CGM}}/(\Omega_b/\Omega_0) = 0.25$, IQR= 0.04, while the Fiducial halo (0.49, IQR= 0.02) is only slowly being depleted of baryons prior to a more rapid depletion at $z < 1$, and the GM-late halo (0.63, IQR= 0.01) does not begin to be depleted until $z \lesssim 0.7$. As is clear from the figure of the IQR values, the seed-to-seed scatter in f_{CGM} for any given family of simulations is more significant than is the case for the halo assembly history, insofar that the scatter is larger in comparison to the separation of the median values of the three families. However, the evolutionary tracks are clearly separated at all times with no overlap in their IQR. Critically, the strong correlation between the halo assembly history and the present-day value of f_{CGM} from large galaxy samples in EAGLE and TNG (seen in Chapters 3 and 4) is reproduced here in direct response to systematic adjustment of the assembly history of an individual halo.

Fig. 5.5 shows the present-day $f_{\text{CGM}} - M_{200}$ relation of haloes of $M_{200} > 10^{11.5} \text{ M}_\odot$ in the EAGLE Ref-L100N1504 simulation. The running median relation obtained with the LOWESS algorithm (Cleveland, 1979) is shown as a black line, whilst the equivalent relations for $z = 1$ and $z = 3$ are denoted by grey lines. Symbols are coloured by the residuals about the LOWESS running median, with respect to M_{200} , of (\log_{10} of) the intrinsic central binding energy of the halo, E_{DMO}^{2500} . The latter is an excellent proxy for the halo assembly time that is simple to compute without the need to examine merger trees. The overlaid tracks show the evolution of the (median) $f_{\text{CGM}} - M_{200}$ relation of the GM-early, Fiducial and GM-late simulations; values at $z = 3$, $z = 1$ and $z = 0$ are denoted by large triangle, circle and square symbols, respectively. At the present day, the CGM mass fraction of the Fiducial halo is similar to the running median value for the EAGLE population. The halo was not selected such that this is achieved by construction, but the selection of a reasonably quiescently star-forming present-day galaxy from a simulation of similar resolution evolved with the EAGLE Reference model (see Section 5.2) makes it unlikely that the Fiducial case would deviate far from the median relation in EAGLE.

The haloes with adjusted assembly histories yield present-day circumgalactic gas fractions that reside towards the extreme ranges of the scatter for EAGLE haloes of $M_{200} \approx 10^{12.5} \text{ M}_\odot$. If one compares with EAGLE haloes in a 0.2dex wide bin centred on this mass, the GM-late halo represents a $+2.0\sigma$ outlier in terms of f_{CGM} , while the GM-early halo is a -1.3σ outlier. Clearly, if these systems were to occur “organically” in

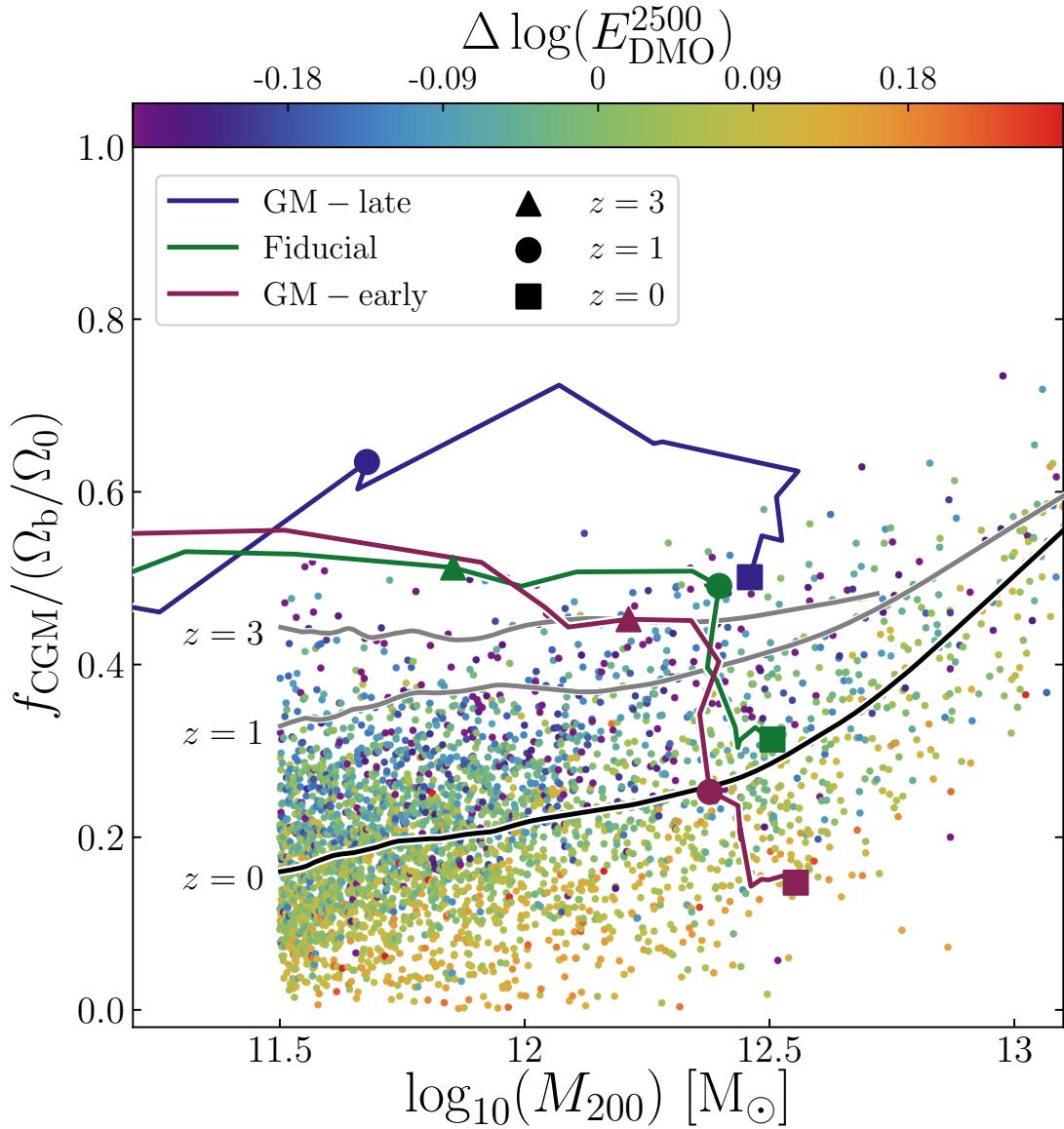


FIGURE 5.5: Present-day CGM mass fractions, $f_{\text{CGM}}(t) \equiv M_{\text{CGM}}(t)/M_{200}(t)$, of haloes in the EAGLE Ref-L100N1504 simulation, normalised by the cosmic average baryon fraction, Ω_b/Ω_0 , as a function of halo mass, M_{200} , with the solid curve denoting the running median. Symbols are coloured by residuals about the relationship between the intrinsic binding energy of the inner halo, E_{DMO}^{2500} , and M_{200} . The evolution of the median $f_{\text{CGM}}/(\Omega_b/\Omega_0)$ with the median M_{200} for the three families of simulations is overlaid, with their locations on the plot at $z = 0$, $z = 1$ and $z = 3$ marked with squares, circles and triangles respectively. The median $f_{\text{CGM}}/(\Omega_b/\Omega_0) - M_{200}$ relations at $z = 1$ and $z = 3$ in the Ref-L100N1504 simulation are shown with grey lines.

the simulation, they would be amongst the most CGM-poor (GM-early) and CGM-rich (GM-late) haloes of their mass. Moreover, as is clear from the symbol colouring, the GM-late (GM-early) cases yield f_{CGM} values similar to haloes in the EAGLE simulation with low (high) values of E_{DMO}^{2500} . I conclude from this controlled experiment that halo assembly history has a strong influence on the present-day circumgalactic mass fraction of haloes of fixed mass.

5.4.2 The influence of halo assembly history on BH growth and AGN feedback

The results in Chapter 4 identified that the correlations between f_{CGM} and the properties of galaxies (such as the sSFR and its morphology), and between f_{CGM} and proxies for the halo assembly history, are mediated by the expulsion of circumgalactic gas due to efficient AGN feedback. The latter therefore acts as an “astrophysical conduit” between the initial conditions and the present-day physical state of the galaxy. I showed that in both EAGLE and TNG, these effects are manifest in haloes that are sufficiently massive to host central BHs capable of efficiently delivering feedback energy to the galaxy+CGM ecosystem. In EAGLE this corresponds to haloes with $M_{200} > M_{\text{crit}}$, i.e. those for which the buoyant transport of outflows heated by stellar feedback ceases to be efficient, resulting in the establishment of a quasi-hydrostatic hot halo, whilst in TNG it corresponds to the haloes that host massive BHs ($M_{\text{BH}} \gtrsim 10^8 M_{\odot}$), since AGN feedback is typically delivered in the efficient kinetic mode for such haloes in that model.

I therefore turn next to an examination of the influence of halo assembly history on the growth of the central BH. Fig. 5.6 shows the evolution of the mass of the central BH⁴, $M_{\text{BH}}(t)$, for the three families of simulations. Comparison of the median curves reveals that the central BHs in the GM-early family reach a greater present-day mass, $\log(M_{\text{BH}} / M_{\odot}) = 8.08$, IQR= 0.30dex, than is the case for the Fiducial family, $\log(M_{\text{BH}} / M_{\odot}) = 7.90$, IQR= 0.08dex), having commenced their rapid growth phases earlier. By $z = 1$, the median M_{BH} of the GM-early family has reached 59 percent of its final value, while the median for the Fiducial case has reached only 20 percent of its final value. In marked contrast to these families, BHs in the GM-late family of simulations remain close to the seed mass until $z \simeq 1$, and do not grow rapidly until $z \simeq 0.3$. In consequence, they attain a significantly lower final mass, $\log(M_{\text{BH}} / M_{\odot}) = 7.28$, IQR= 0.20dex.

Note that the broad shaded regions in Fig. 5.6 indicate that the black hole mass can vary due to the stochastic nature of EAGLE’s feedback scheme; the evolution is broadly the

⁴I define the central BH as the most massive BH particle in the system.

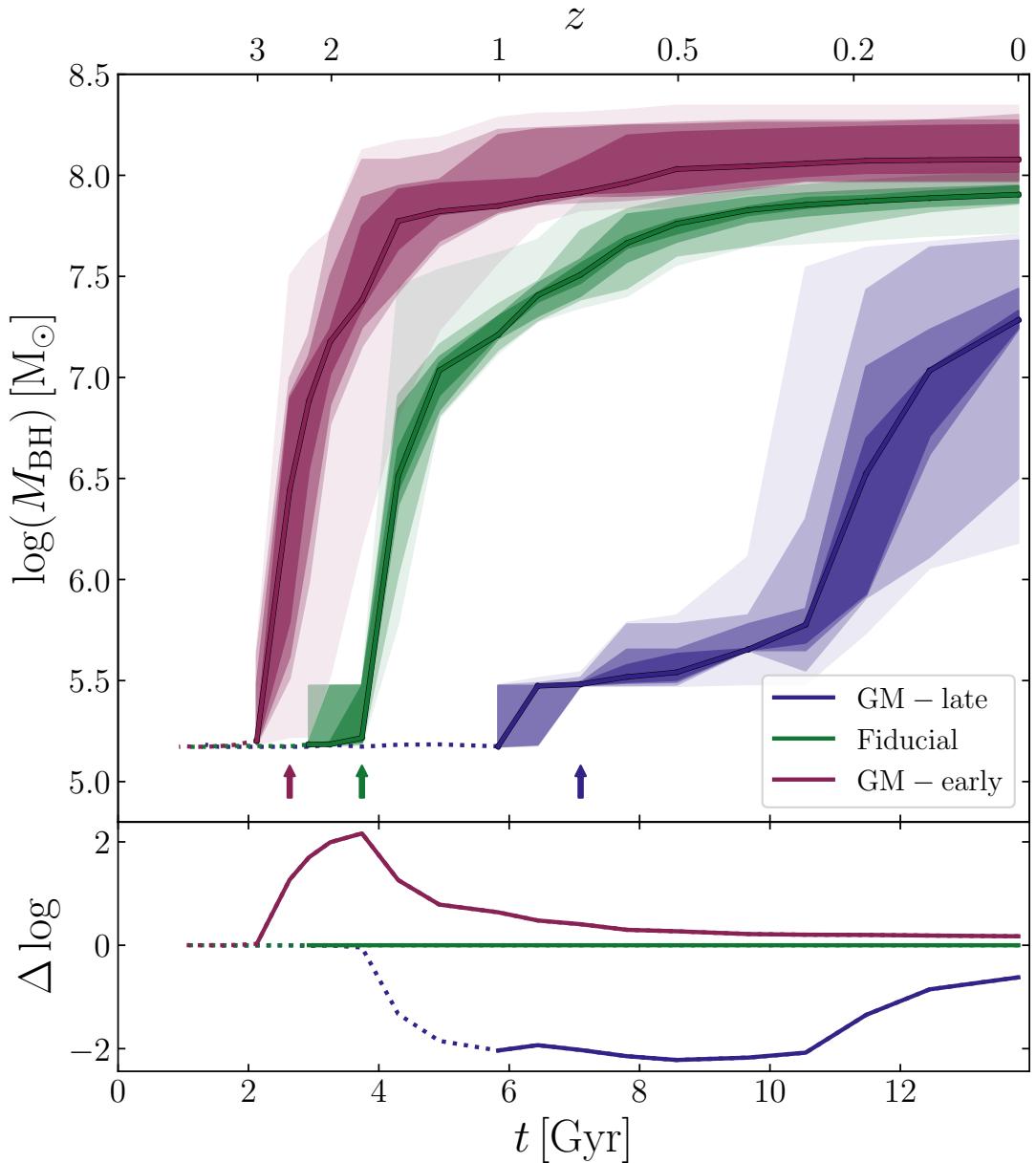


FIGURE 5.6: The evolution of the black hole mass, $M_{\text{BH}}(t)$, for the three families of simulations. Solid lines show the median value for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z = 0) < 0.05$. Sub-panels show the deviation, $\Delta(t)$, of the logarithm of the median black hole mass from that of the Fiducial case. The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

same for each GM case, however differences in the feedback seed cause a spread in the obtained BH mass as a function of time. Despite the scatter, there is clear separation between the three cases, with overlap between only the most extreme cases.

Comparison with Fig. 5.4 illustrates that, as per the findings of Oppenheimer et al. (2020), a strong decline in f_{CGM} generally follows shortly after periods of rapid growth of the central BH. Adjustment of the halo’s mass accretion history therefore clearly has a very strong effect on the growth history of the central BH, and I interpret the strong correlation between f_{CGM} and the halo assembly history to be a consequence of the influence of the latter on the BH growth history.

The early collapse of dark matter haloes (of a fixed present-day mass) leads to a higher concentration (e.g. Neto et al., 2007) and hence central binding energy, so one might expect from self-regulation arguments that the ratio of energy injected by feedback to the binding energy of the baryons asymptotes to the same value regardless of assembly time. However, in Chapter 4 I showed that the ratio of the injected feedback energy to the (intrinsic) binding energy of the halo correlates positively with assembly time in both EAGLE and TNG, such that the additional energy injected by the central galaxy of early-forming haloes “overshoots” the additional binding energy resulting from their higher concentration.

I therefore examine whether this ratio changes in response to adjustments to the halo assembly history. Fig. 5.7 shows the evolution of the $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ ratio, where E_{AGN} is the total energy injected by AGN feedback (defined per eq. 5.1). I focus only on the energy injected by AGN, since it is this mechanism that is principally responsible for gas expulsion in haloes where $M_{200} \gtrsim 10^{12.5} M_{\odot}$ in EAGLE. $E_{\text{bind}}^{\text{b}}$ is the intrinsic binding energy of the halo’s baryons, which was computed from the particle distribution of the halo in its counterpart DMONLY simulation, and thus self-consistently accounts for differences in the structure of the halo induced by their mass accretion histories (see Section 5.2.4).

Prior to the onset of the efficient growth of the BH, $E_{\text{AGN}} \ll E_{\text{bind}}^{\text{b}}$. Once $M_{200} \simeq M_{\text{crit}}$, the rapid growth of the BH results in a rapid increase of E_{AGN} , such that the ratio $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ stabilises at a value of order unity. In general, the ratio settles at a value greater than unity because radiative cooling inhibits the unbinding of circumgalactic gas. Since the halo reaches M_{crit} at very different times for the three families, the increase in $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ follows suit. By $z = 1$, the energy injected via AGN feedback has already exceeded the binding energy of the baryons in the GM-early case ($E_{\text{AGN}}/E_{\text{bind}}^{\text{b}} \simeq 2.5$, IQR= 3.23), and in the Fiducial case the binding energy is about to be exceeded (0.62, IQR= 0.18). At $z = 1$, the injected energy remains much less than the intrinsic binding energy in the GM-late case ($E_{\text{AGN}}/E_{\text{bind}}^{\text{b}} \sim 10^{-3}$, IQR= 0.113).

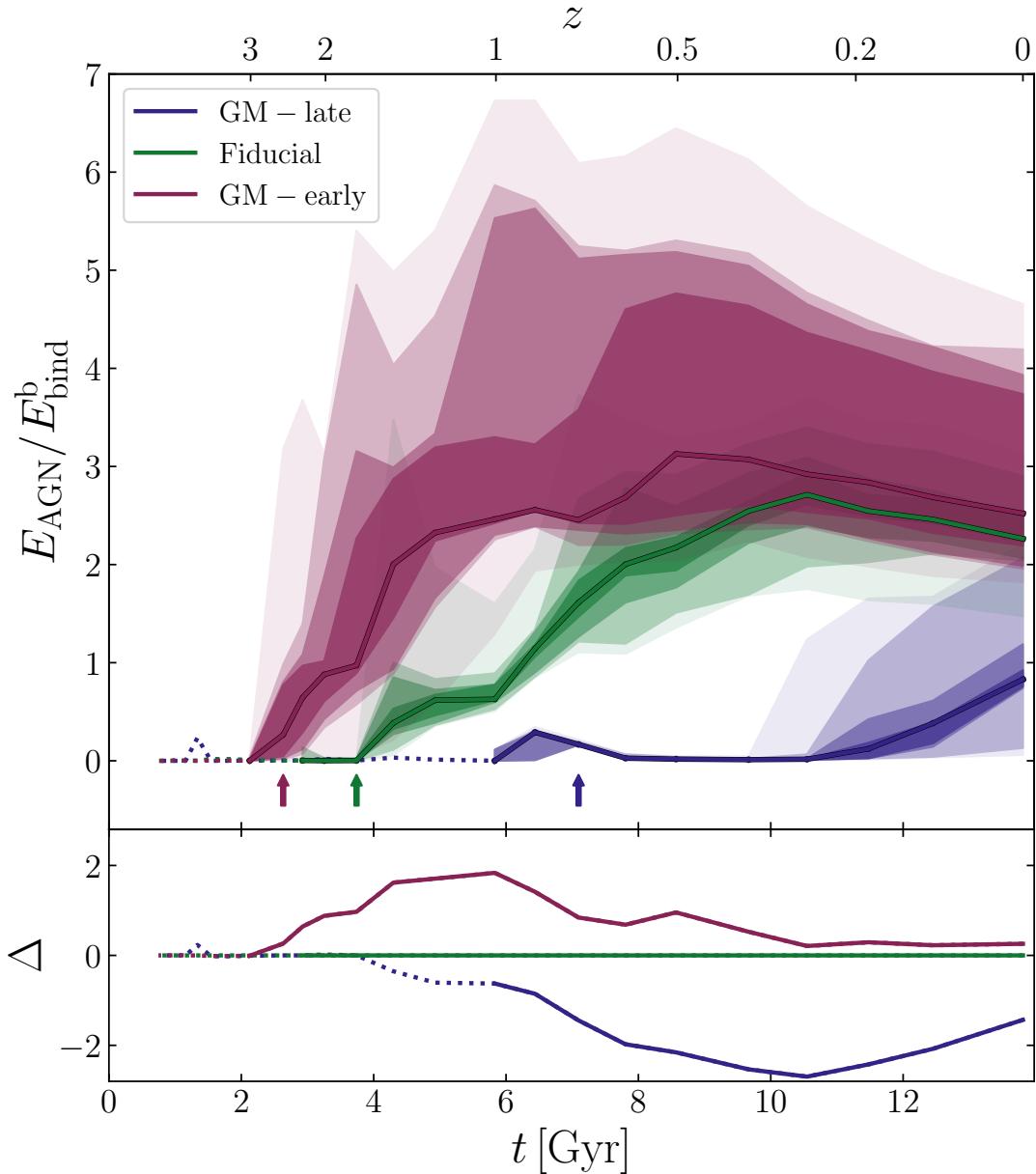


FIGURE 5.7: The evolution of the $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ ratio for the three families of simulations, where E_{AGN} is the total energy injected by AGN feedback and $E_{\text{bind}}^{\text{b}}$ is the intrinsic binding energy of the halo's baryons. Solid lines show the median value for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z=0) < 0.05$. Sub-panels show the deviation, $\Delta(t)$, of the median value from that of the Fiducial case. The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

The evolution of the $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ ratio provides an intuitive explanation for the CGM mass fractions at $z = 1$, shown in Figs. 5.4 and 5.5; the CGM of the halo in the GM-early family has already been depleted of a significant fraction of its mass, in the Fiducial family it is about to be depleted, and in the GM-late family it remains gas-rich. The final median ratio for the GM-early case is highest ($E_{\text{AGN}}/E_{\text{bind}}^{\text{b}} = 2.52$, IQR= 1.94), followed by the Fiducial case (2.26, IQR= 0.44), while the ratio for the GM-late case is far lower (0.83, IQR= 0.45). The scatter in the ratio is equivalent to the scatter in M_{BH} (since $E_{\text{bind}}^{\text{b}}$ is computed from a DMONLY run and is independent of the chosen seed) and is greatest for the GM-early case. While the GM-early and Fiducial cases did not overlap in the IQR of their final M_{BH} , there is significant overlap for $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$ because the final median values are more similar; while the median E_{AGN} in the GM-early case is a factor 1.49 times larger than for the Fiducial case, $E_{\text{bind}}^{\text{b}}$ is a factor 1.34 times greater, resulting in this small difference in ratio. It is however clear that the assembly history of the halo directly influences how much energy, beyond that required to unbind the halo’s baryons, is injected into the galaxy+CGM ecosystem by the central BH.

It is plausible that there is a close coupling of this ratio to the CGM mass fraction such that the injection of more energy relative to the binding energy of the baryons yields a lower f_{CGM} , and this is largely supported by the monotonic behaviour seen here. The timing of this energy injection is likely to affect the efficiency of baryon ejection; on one hand, energy injected at earlier times could be less efficient at driving a wind, since the mean density of the surrounding gas is higher ($n_{\text{H}} \propto (1+z)^3$), leading to greater radiative losses ($\Lambda \propto n_{\text{H}}^2 \propto (1+z)^6$). On the other hand, the ratio converges to similar values by the present day for the GM-early and Fiducial cases, despite the GM-early case injecting feedback energy earlier, indicating that this is not a major factor. Earlier injection instead appears to be more efficient, since the GM-early case has a lower final f_{CGM} than the Fiducial case (with little overlap in the scatter), but reaches a similar final $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$.

The results presented in this Section and in Section 5.4.2 provide complementary evidence supporting the conclusions of Chapters 3 and 4. In the galaxy populations of EAGLE and TNG, the binding energy of the underlying dark matter halo (a good proxy for assembly time) correlates positively with M_{BH} and $E_{\text{AGN}}/E_{\text{bind}}^{\text{b}}$, which both in turn correlate negatively with f_{CGM} , thus connecting the assembly histories of dark matter haloes with their circumgalactic mass fractions. Here I have shown that the modification of an individual halo’s mass accretion history enables the testing of these effects in a controlled fashion that minimises the influence of other driving influences such as the halo’s environment. These experiments show that BH growth and circumgalactic gas

expulsion respond to changes in halo assembly history in a fashion consistent with the correlations seen in the EAGLE and TNG galaxy populations.

5.4.3 The influence of halo assembly history on galaxy properties

I now explore whether adjustment of the halo’s mass accretion history influences the evolution of the central galaxy. Fig. 5.8 shows the evolution of the sSFR of the central galaxy of the halo in the three families. The sSFR is computed at each snapshot epoch, averaged over the preceding 300 Myr to suppress noise in the star formation rate, which can vary significantly on shorter timescales (see e.g. [McAlpine et al., 2017](#)). For clarity, I impose a minimum value of 10^{-13} yr^{-1} , and note that in any case it is not feasible to infer lower, non-zero values from observational measurements. I note that the evolution of the sSFRs of the EAGLE galaxy population was explored in detail by [Furlong et al. \(2015\)](#).

The sSFR decreases with advancing time in all three families. In the Fiducial and GM-late cases, the decline is relatively shallow, and in all simulations from these families the galaxy remains actively star forming at $z = 0$, with medians of $\text{sSFR} > 10^{-10.6} \text{ yr}^{-1}$ and $\text{sSFR} = 10^{-10.0} \text{ yr}^{-1}$, respectively. In contrast, the median sSFR for the GM-early family declines rapidly at $z \approx 2$, in concert with the rapid growth of its BH, and effectively ‘quenches’ (i.e. its sSFR drops below 10^{-11} yr^{-1}) at $z = 0.86$, and remains quenched at all subsequent times, with median $\text{sSFR} = 10^{-12.37} \text{ yr}^{-1}$ at $z = 0$. The scatter between runs of differing initial random seed is mild for the Fiducial and GM-late cases, with present-day IQR values of 0.20 dex and 0.16 dex, respectively. The scatter is much more significant for the GM-early case, with IQR at $z = 0$ of 1.54dex, driven primarily by sampling noise at low SFRs. In consequence, three of the nine simulations from this family yield galaxies that remain marginally star forming at $z = 0$. This scatter notwithstanding, it is clear that the median trends of the three families are unambiguously influenced by the halo assembly history. By adjusting this property for an individual halo, I am able to convert a galaxy with a reasonably normal present-day star formation rate into either a more vigorously star-forming galaxy, or a quenched one.

To demonstrate that AGN feedback is an essential ingredient in this chain of processes, I add dashed lines to Fig. 5.8, to show the median sSFR evolution for each family in simulations with AGN feedback disabled (i.e. where no BHs are seeded). There is little difference for the GM-late case, indicating that AGN feedback has not significantly affected the sSFR, and in the Fiducial case removing the AGN mildly elevates the sSFR for $z \lesssim 1.5$. Disabling AGN feedback in the GM-early case results in the galaxy never quenching, retaining a present-day sSFR of $10^{-8.37} \text{ yr}^{-1}$.

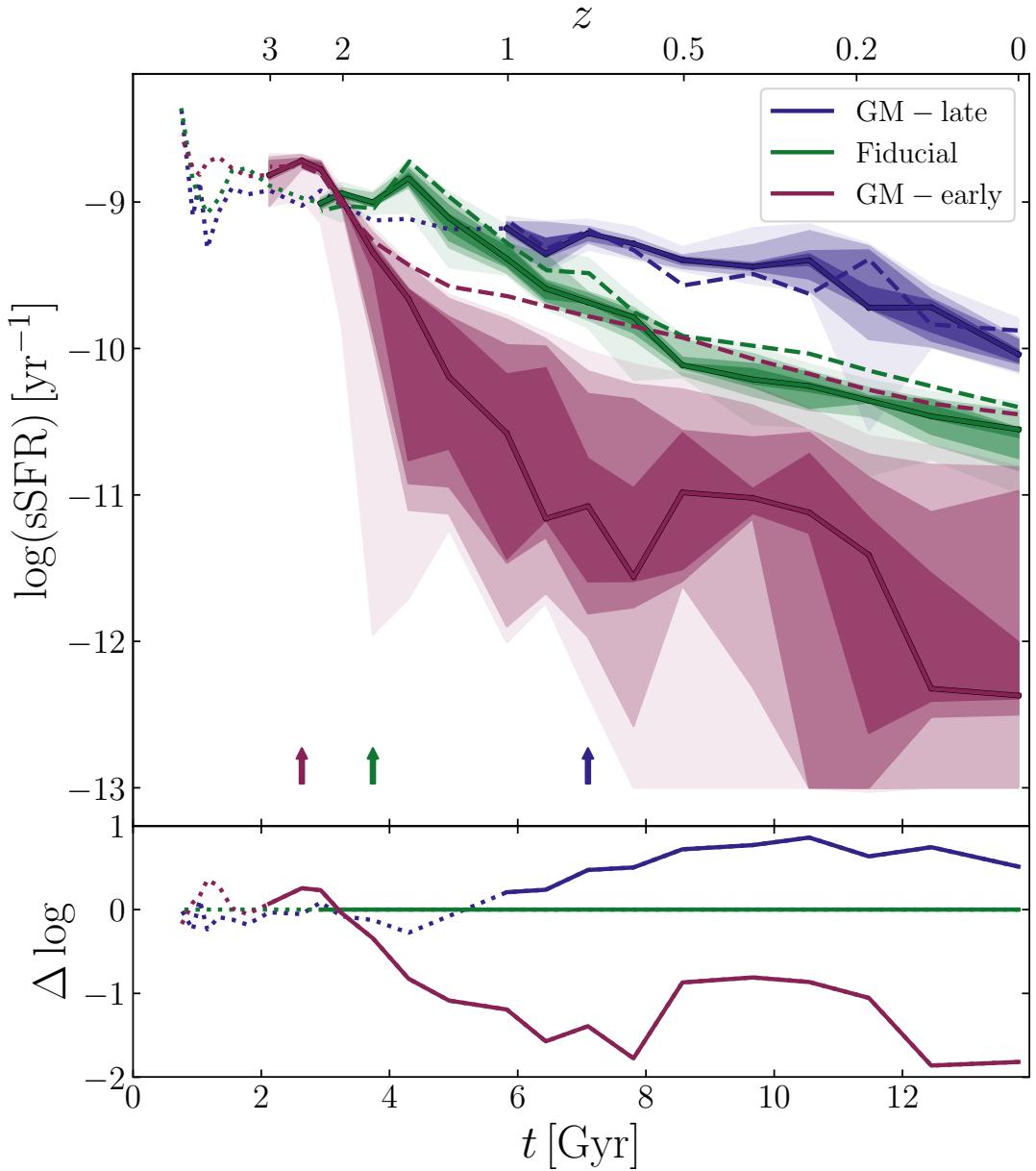


FIGURE 5.8: The evolution of the specific star formation rate (sSFR) of the central galaxy for the three families of simulations. Solid lines show the median value for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z = 0) < 0.05$. I impose a minimum value of 10^{-13} yr^{-1} . Dashed lines show the median sSFR in simulations where AGN feedback is disabled (i.e. where no BHs are seeded). Sub-panels show the deviation, $\Delta(t)$, of the logarithm of the median sSFR from that of the Fiducial case. The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

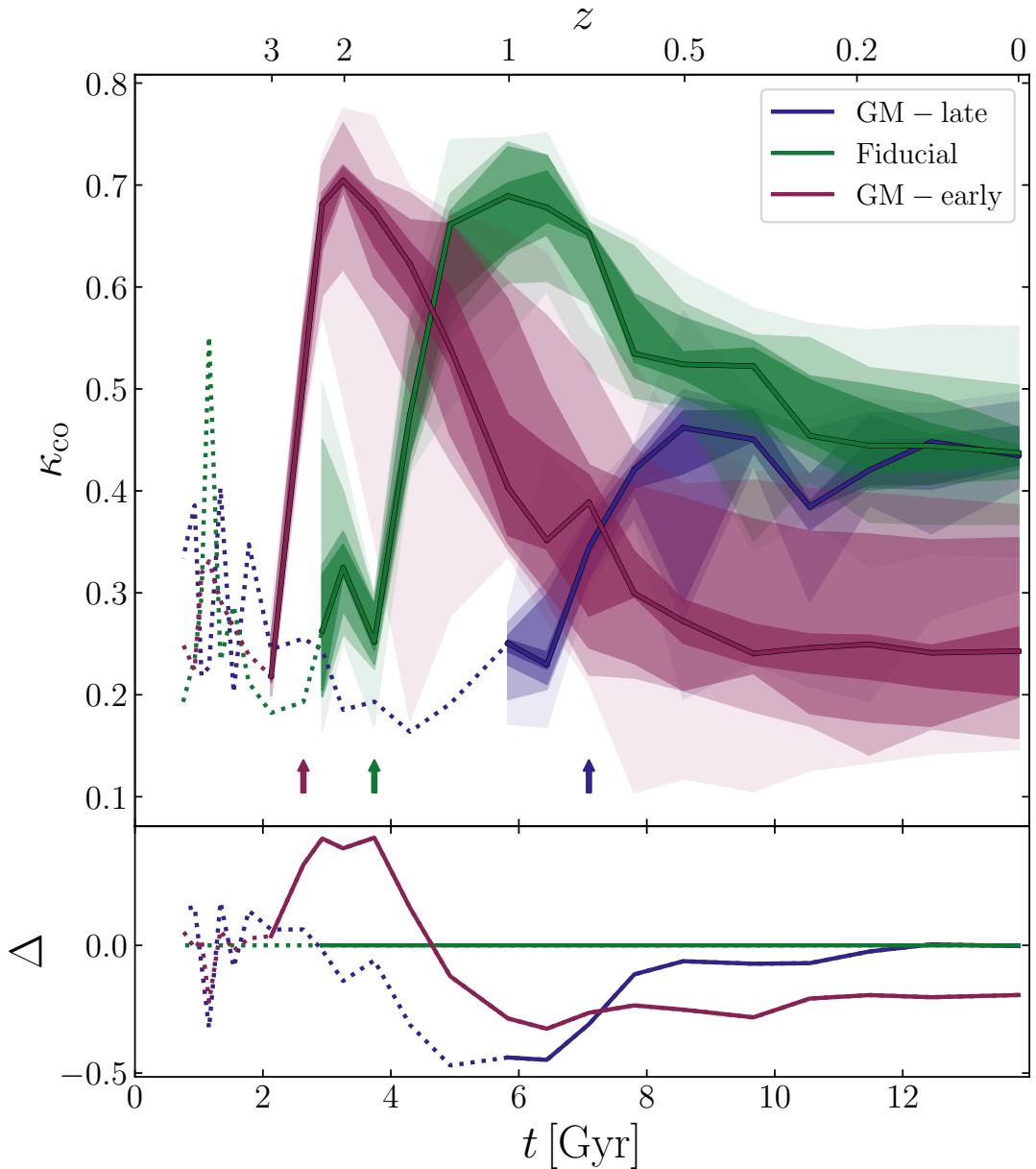


FIGURE 5.9: The evolution of the stellar co-rotational kinetic energy fraction, κ_{co} , of the central galaxy for the three families of simulations. Solid lines show the median value for nine initial random seeds for feedback and star formation; the distribution across the nine seeds is shown with progressively lighter shading between pairs of seeds that give increasingly divergent results from the median value. Evolutionary tracks are shown as dotted lines where $M_{200}(t)/M_{200}(z = 0) < 0.05$. Sub-panels show the deviation, $\Delta(t)$, of the median value from that of the Fiducial case. The first snapshot output where $M_{200}(t) > M_{\text{crit}}(z)$ is highlighted with a coloured arrow for each GM case.

The correlations identified in Chapter 4 suggested that the processes outlined thus far are not only implicated in the quenching of central galaxies, but that they may also facilitate a kinematical and/or morphological transformation. In EAGLE and TNG, there exists a positive correlation between proxies for halo assembly time (at fixed halo mass) and the degree of rotational support in the stellar kinematics of their central galaxies. I therefore examine whether the halo assembly history influences the (kinematically-quantified) morphology of the central galaxy.

Fig. 5.9 shows the evolution of the stellar co-rotational kinetic energy fraction, κ_{co} ⁵, of the main progenitor of the central galaxy, for the three families of simulations. κ_{co} is defined as the fraction of the kinetic energy of the stars in the galaxy that is invested in co-rotational motion, for which a threshold value of 0.4 has been shown to separate star-forming discs ($\kappa_{\text{co}} > 0.4$) from quenched ellipticals ($\kappa_{\text{co}} < 0.4$) in EAGLE (Correa et al., 2017). In all three cases, the kinematics of the central galaxy initially exhibits little co-rotational motion, before κ_{co} then rises to maximal values of $\kappa_{\text{co}} = 0.70$ (GM-early), $\kappa_{\text{co}} = 0.69$ (Fiducial) and $\kappa_{\text{co}} = 0.46$ (GM-late) at $z = 2$, $z = 1$ and $z = 0.5$ respectively, signalling the formation of a rotationally-supported disc. The GM-late galaxy then retains approximately this maximum κ_{co} until the present day, while κ_{co} declines for the Fiducial and GM-early galaxies after reaching a maximum. At the present day, the GM-late and Fiducial galaxies are spiral-like according to the threshold $\kappa_{\text{co}} = 0.4$, with similar median values of $\kappa_{\text{co}} = 0.43$ and $\kappa_{\text{co}} = 0.44$ respectively. By this same definition, the GM-early galaxy becomes elliptical, with a median $\kappa_{\text{co}} = 0.24$.

Each of the three families of simulations begin to exhibit strong co-rotational kinematics at different times; comparison of the evolution in κ_{co} with Fig. 5.3 illustrates that this is the result of differences in the halo mass assembly. Comparison with Fig. 5.4 then reveals that the erasing of this co-rotational motion in the GM-early and Fiducial galaxies occurs in concert with the depletion of baryons from the CGM. The scatter induced by using different initial seeds is mild for the GM-late and Fiducial cases, with present-day IQR values of 0.05 and 0.04 respectively, while it is far stronger for the GM-early case (IQR= 0.16). The $\kappa_{\text{co}} - f_{\text{CGM}}$ correlation found in Chapter 4 was weaker than other correlations such as sSFR – f_{CGM} , and the overlap in the scatter for the three cases here reflects this. Despite this, eight of the nine seeds yield a GM-early galaxy with $\kappa_{\text{co}} < 0.4$ at the present day, and the morphologies of galaxies can clearly be changed from spiral-like to elliptical through a controlled modification of the halo assembly history.

This connection between kinematics and assembly history is not a trivial result given my findings regarding quenching, since Correa et al. (2019) identify only a weak correlation

⁵I compute κ_{co} using the publicly-available routines of Thob et al. (2019).

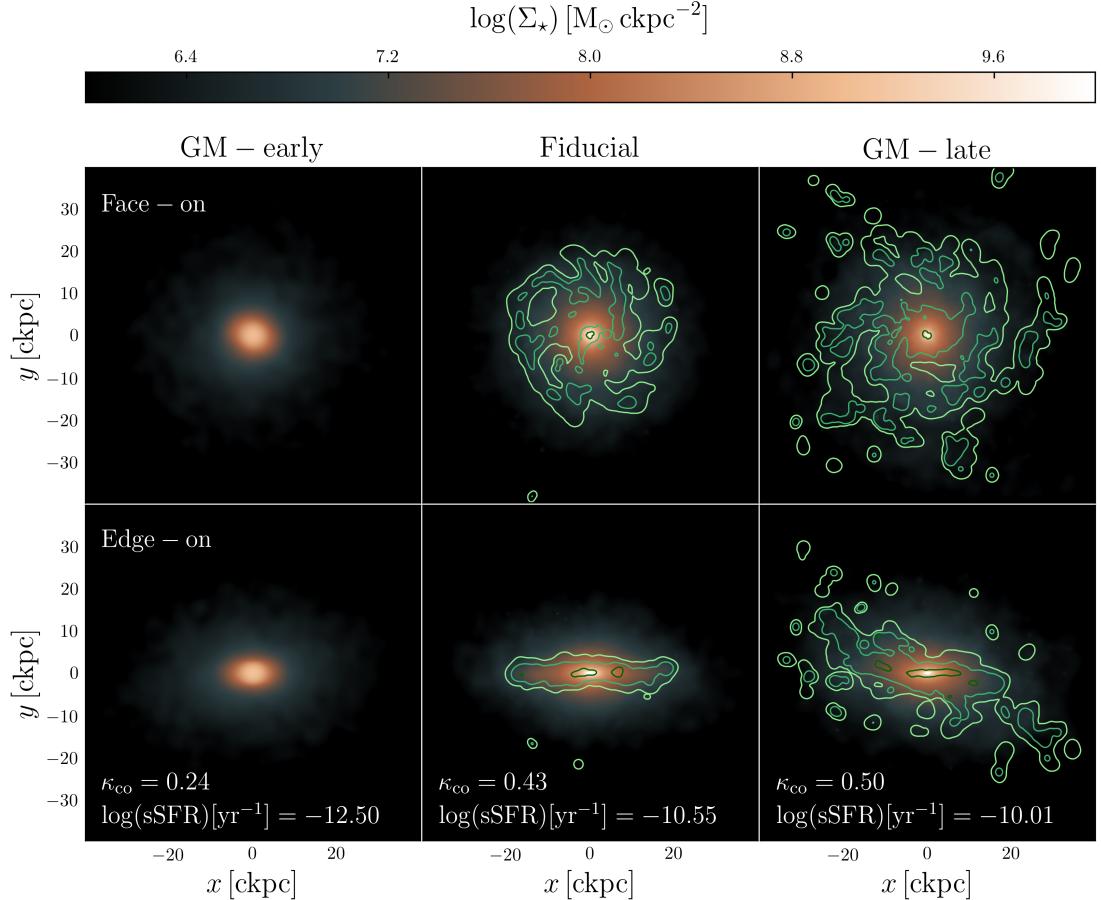


FIGURE 5.10: Face-on (upper row) and edge-on (lower row) present-day surface density maps of the stellar distribution for the central galaxies of GM-early (left column), Fiducial (centre column) and GM-late (right column) haloes produced with the standard EAGLE random seed for star formation and feedback. The overlaid contours show the star-forming gas distribution, enclosing hydrogen column densities of $N_{\mathrm{H}} = 10^{20}$, 10^{21} , and $10^{22} \mathrm{cm}^{-2}$ in progressively darker shades of green. The field of view (and depth) of the maps is 80 ckpc. The specific star formation rate (sSFR) and stellar co-rotational kinetic energy fraction (κ_{co}) are quoted for each galaxy in the lower row.

between galaxy colour and morphology in EAGLE. I speculated in Chapter 4 that the connection arises because the elevation of the CGM cooling time in response to circumgalactic gas expulsion inhibits the replenishment of cold gas in galaxy discs, which would otherwise stabilise them against transformation by mergers, tidal interactions and gravitational instabilities, and enable re-growth of the disc (see e.g. Robertson et al., 2006; Hopkins et al., 2009; Font et al., 2017). This reasoning is compatible with my results; the transformation in the Fiducial and GM-early cases is gradual and happens over several gigayears, contrary to the rapid transformation one would expect from a major merger; rotational motion in the disc is instead gradually transformed into random motion.

To illustrate the transformative effects of a modified assembly history on the properties of the central galaxy, I show in Fig. 5.10 face-on and edge-on images of a present-day

galaxy from each family. Each family comprises nine realisations, so I show here the galaxy in the simulation from each family that adopts the same initial seed as used by the EAGLE suite. The images are surface density maps of the stellar distribution with a field of view, and depth, of 80 ckpc. The overlaid green contours show the distribution of star-forming gas within the same volume; the three progressively darker contours enclose hydrogen column densities of $N_{\mathrm{H}} = 10^{20}$, 10^{21} , and 10^{22} cm^{-2} respectively. The Fiducial and GM-late haloes host actively star-forming spiral galaxies, with the GM-late galaxy exhibiting greater rotational support ($\kappa_{\mathrm{co}} = 0.50$) than the Fiducial galaxy ($\kappa_{\mathrm{co}} = 0.43$). The GM-early halo hosts a quenched galaxy with a slightly flattened elliptical morphology ($\kappa_{\mathrm{co}} = 0.24$). The Fiducial and GM-late galaxies host discs of star-forming gas, sustaining specific star formation rates of $10^{-10.0} \text{ yr}^{-1}$ and $10^{-10.6} \text{ yr}^{-1}$ respectively; the GM-late galaxy hosts significantly more star-forming gas (quantified by the ISM mass fraction, $f_{\mathrm{ISM}} \equiv M_{\mathrm{ISM}}/M_{\star} = 0.24$) than the Fiducial galaxy ($f_{\mathrm{ISM}} = 0.08$). In contrast, the GM-early galaxy is entirely devoid of star-forming gas, with an instantaneous sSFR of zero; the value quoted in Fig. 5.10 is integrated over the preceding 300 Myr for consistency with the results shown in Fig. 5.8.

5.5 The impact of AGN feedback on the thermodynamic state of the CGM

I return now to the thermodynamic state of the CGM in the genetically modified haloes. In Chapter 4 I showed that in both EAGLE and IllustrisTNG, haloes that are gas-poor exhibit longer halo-averaged cooling times (and characteristic entropies) than gas-rich haloes at fixed mass. The radiative cooling time distributions of gas particles in gas-poor haloes were also shown to be systematically greater than those of particles in gas-rich haloes of the same halo mass. The effect of this shift is that the CGM is less-readily able to cool and replenish the ISM when it is depleted by star formation or feedback processes, ultimately facilitating quenching and morphological transformation. While it is clear that the AGN-driven removal of baryons from the system is responsible for this shift in the halo cooling time, it is not clear precisely how the shift is realised, and how this influences the replenishment of the ISM.

I therefore now compare the present-day radiative cooling times (t_{cool}), temperatures (T) and densities (n_{H}) of particles in the CGM (i.e. within r_{200}) of the GM-early family of simulations, which have experienced strong AGN feedback, with those in the GM-late family, which have experienced comparatively little. Fig. 5.11 shows the distributions of these quantities for both the full-physics (RECAL) simulations and for simulations

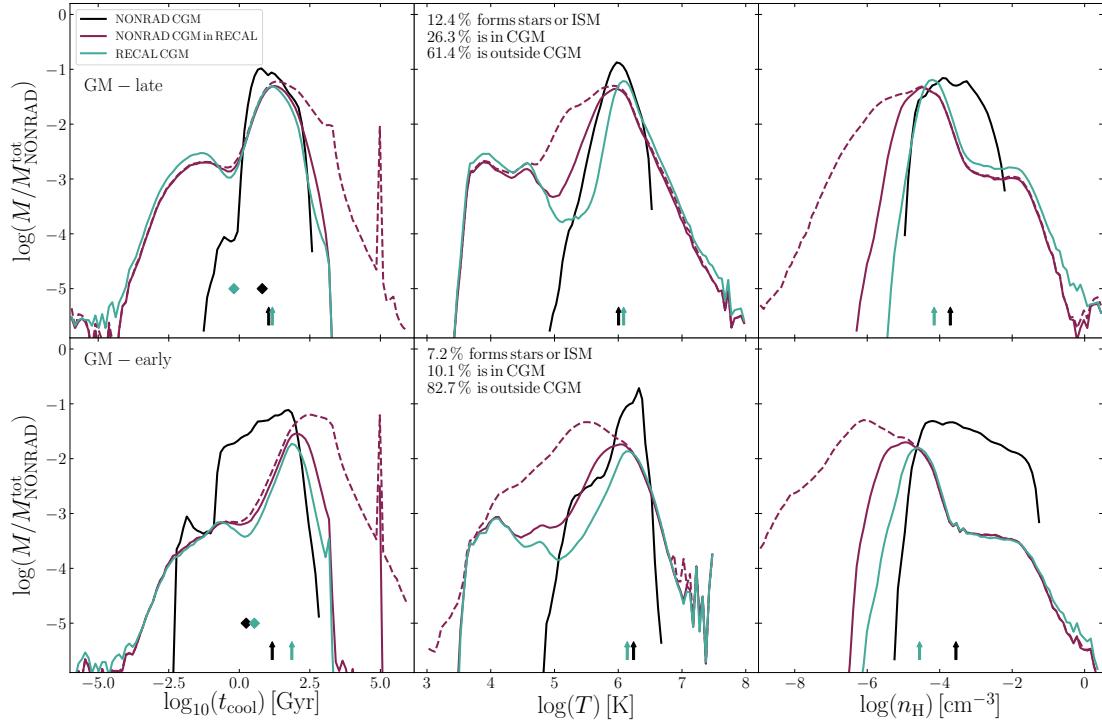


FIGURE 5.11: Present-day distributions of the cooling times (t_{cool} , left column), temperatures (T , centre column) and hydrogen number densities (n_{H} , right column) of three subsets of particles in simulations evolved from the GM-late (upper row) and GM-early (lower row) initial conditions. Distributions are shown for the particles in the CGM (i.e. they are not star-forming and are within the virial radius r_{200}) in non-radiative simulations (NONRAD CGM), particles in the CGM in full-physics simulations (RECAL CGM) and particles in full-physics simulations which were in the NONRAD CGM subset (NONRAD CGM in RECAL). I stack these particle subsets generated with all nine simulations in each family, and compute the distributions in 100 logarithmically-spaced bins as a mass fraction normalised to the total mass of the NONRAD stack. Particles at distances greater than $2r_{200}$ from the halo centre are excised from each distribution, though I show the unexcised distributions with dashed lines. Particles experiencing a net heating have their t_{cool} fixed to 10^5 Gyr, and star-forming particles are excised from “NONRAD CGM in RECAL”. Median values for each distribution are shown with arrows, and mean t_{cool} values, obtained by dividing the total internal energy of the stack by the total luminosity, are shown with diamonds in the left column.

adopting the same initial conditions but in which radiative cooling is disabled⁶ (NONRAD), and hence no stars form and no energy is injected by feedback processes. The distributions are generated by ‘stacking’ all nine simulations from each family, thus improving the particle sampling by a factor of nine with respect to an individual halo, and suppressing the influence of stochasticity from EAGLE’s sub-grid implementations of star formation and feedback from star formation and BH growth. They are computed in 100 logarithmically-spaced bins as a mass fraction normalised to the total mass of the NONRAD stack. The median values of each distribution are marked with arrows below the curves in all panels.

⁶In this case, the cooling times I compute (following Section 5.2.4) are based on the temperatures and densities of particles as though they were cooling.

The gas comprising the CGM in the NONRAD simulations effectively represents the baryonic mass that “should” be associated with the halo in the absence of feedback. Examination of the properties of this set of fluid elements in the RECAL simulations (shown as “NONRAD CGM in RECAL”) highlights the influence of feedback processes on the physical state of the CGM. These particles need not necessarily comprise the CGM in RECAL, and the difference between this set of fluid elements and those actually comprising the CGM in the RECAL simulation represents those baryons that have been ejected. Some of this material is ejected with sufficient momentum that it travels beyond the region sampled with gas particles by the zoomed simulation, after which it ceases to experience pressure forces and effectively expands adiabatically to artificially-low densities (and long cooling times). I excise any particles residing beyond $2r_{200}$ at $z = 0$ from the distributions to account for this, but show the unexcised distributions with dashed lines for completeness.

A subset of the particles in the NONRAD CGM form stars in RECAL, and therefore do not appear in the “NONRAD CGM in RECAL” distributions. Similarly, another subset comprises the present-day ISM in RECAL. Since EAGLE (intentionally) treats the multiphase ISM as a single-phase medium, the SPH temperature of star-forming particles is more a reflection of their pressure than their internal energy. I therefore excise both of these particle subsets from “NONRAD CGM in RECAL”. I set the cooling times of particles that are experiencing a net heating (and hence have a negative t_{cool}) to 10^5 Gyr to ensure that they are included in the distribution. In the centre panels of Fig. 5.11, I quote the mass fraction of the NONRAD CGM that, in RECAL, is found in the form of stars or the ISM, and also the fractions that are found within and without the CGM.

In the GM-late case, the median cooling times for the NONRAD and RECAL CGM are similar, and are comparable to the Hubble time, t_H , at 11.1 Gyr and 14.8 Gyr respectively. In RECAL, galaxies in the GM-late systems are actively star-forming and are being fuelled by cooling from the CGM; this can be seen in the tail to short cooling times in the RECAL distribution. The bulk of the distribution represents the “hot” CGM and has a similar shape to the non-radiative case, though at a lower normalisation because 61% of the NONRAD baryons have been ejected from the halo. The “NONRAD CGM in RECAL” distribution shows the fate of these baryons: they have long cooling times and predominantly reside beyond $2r_{200}$. The similarity of the RECAL and NONRAD distributions indicates that the expulsion of this gas has not strongly affected the radiative cooling of the CGM.

In the GM-early case, the halo has experienced much more efficient gas expulsion. Of the NONRAD CGM particles, 83% have been ejected in the RECAL simulations, and

thus exhibit long cooling times (> 100 Gyr) and reside far beyond the virial radius. The t_{cool} distribution of particles remaining in the CGM is shifted to much longer timescales: the median t_{cool} is $\simeq 80$ Gyr, in comparison to a median of 14.7 Gyr for the NONRAD simulations⁷. In the GM-early family, the halo has substantially less mass with a cooling time shorter than the Hubble time, inhibiting the cooling of circumgalactic gas onto the ISM and by extension inhibiting the resumption of star formation in the central galaxy. To more closely mimic the measurement of halo cooling that would be derived from observational measurements, I compute a mean value for the stacks by dividing the total internal energy of the stacked particles by their total luminosity, denoting these values with diamond symbols on Fig. 5.11. In the GM-late case, this value is lower in RECAL (650 Myr) than in NONRAD (6.5 Gyr) because the mean is strongly weighted to the rapidly cooling gas in the full-physics simulation, while in the GM-early case, it is higher (3.42 Gyr vs. 1.76 Gyr).

To understand the cause of this shift in the cooling time, I first examine the particle temperatures in the central panels of Fig. 5.11. The RECAL temperature distributions are broadly similar for the GM-late and GM-early families, but the different normalisations of their histograms reflect their different f_{CGM} values. Both families exhibit a dominant hot ($T \sim 10^6$ K) component whose distributions (and median temperatures) closely follow that recovered from their counterpart NONRAD simulations. Both also exhibit cool ($T \sim 10^{4-5}$ K) gas components that are absent from their non-radiative counterparts, and also contain gas that has been heated to significantly higher temperatures than seen in NONRAD. The latter result indicates that, contrary to the results for group-scale haloes presented by McCarthy et al. (2011), feedback processes in $\sim L^*$ galaxies heat a fraction of the CGM to higher temperatures than is produced by gravitational shock heating. I speculate that this may be a consequence of the adoption of a small amount of artificial entropy diffusion in EAGLE (see Appendix A3 of Schaye et al., 2015), but further investigation of this is beyond the scope of this study. The material ejected from the CGM exhibits temperatures below the median temperature of the gas comprising the CGM, owing to its adiabatic expansion into the vacuum beyond the baryon-sampled region of the zoom simulation. For the purposes of understanding the shift to longer t_{cool} in the GM-early case, the key result here is that the bulk of the CGM is not significantly heated by feedback, and retains the characteristic temperature that it would exhibit if simulated without radiative cooling.

The cooling rates of gas particles are particularly sensitive to their density, since $\Lambda \propto n_{\text{H}}^2$. Comparison of the RECAL and NONRAD distributions in the right-hand column of Fig. 5.11 reveals that the density structure of the CGM is significantly modified

⁷The small difference in the NONRAD cooling time between the GM-early and GM-late cases is due to their differing halo concentrations.

by radiative processes. In the GM-late case, the lowest-density material in the outer halo retains approximately the same density distribution with and without radiative processes, however the most dense material has been substantially reconfigured. A small fraction of this gas has cooled into stars or the ISM (12% of the budget), but the majority has been ejected beyond r_{200} and exists at low densities as a result of expansion into the vacuum beyond the baryon-sampled region of the zoom simulation, as shown by the sample matched between NONRAD and RECAL. These effects reduce the median density of the CGM from $1.96 \times 10^{-4} \text{ cm}^{-3}$ in NONRAD to $7.03 \times 10^{-5} \text{ cm}^{-3}$ in RECAL, a reduction by a factor of 2.8. In the GM-early case, this effect is much stronger; the ejection of a greater fraction of the CGM by feedback reconfigures the CGM at a lower density, shifting the entire distribution and moving the median from $2.80 \times 10^{-4} \text{ cm}^{-3}$ in NONRAD to $2.78 \times 10^{-5} \text{ cm}^{-3}$ in RECAL, a reduction by a factor of 10.1. This corresponds to a factor ~ 100 increase in the characteristic cooling time for the halo, though the actual median cooling time for the stack only increases by a factor of 5; I speculate that this is due to the enrichment of the CGM with metals in the radiative simulation, which facilitate more rapid cooling through line emission. Nevertheless, the removal of low-density gas from the system, coupled with a five-fold increase in the halo-wide cooling time, is sufficient to prevent efficient cooling onto the ISM and, in general, to quench the central galaxies in the GM-early family of simulations.

5.6 Summary and Discussion

I have investigated the impact of the assembly history of a dark matter halo on the properties of its central galaxy and circumgalactic medium (CGM) in a set of ‘zoom’ simulations of a single halo where the assembly history is systematically adjusted via the “genetic modification” (GM) technique (Roth et al., 2016; Pontzen et al., 2017, see also Rey & Pontzen 2018). This study was motivated by the identification of several correlations at fixed halo mass in the EAGLE and IllustrisTNG cosmological, hydrodynamical simulations of galaxy formation, which indicate that differences in assembly history drive the scatter in the CGM mass fractions of dark matter haloes, mediated by differences in the integrated feedback from active galactic nuclei (AGN) injected into the system. These correlations, presented in Chapters 3 and 4, also indicate that differences in the expulsion of CGM baryons by AGN feedback play a key role in determining the present-day properties of central galaxies, particularly in regards to their star formation activity and morphology. These correlations, identified in large samples of galaxy-halo systems in cosmologically-representative volumes, indicate that the quenching and morphological transformation of galaxies can be directly linked to the assembly history of their host dark matter haloes, which is determined only by cosmology. A limitation

of these results, however, is that they are ultimately obtained by comparing *different* haloes, which may reside in very different environments, impacting the evolution of the system.

To rigorously test the impact of differences in the halo assembly history on galaxies and their surrounding gas, I perform a set of zoomed resimulations of the *same* halo, but with modified initial conditions designed such that a halo of approximately the same mass is assembled by $z = 0$, but where the mean overdensity of the mass constituting the halo at $z = 2$ is higher or lower than the fiducial system. This technique shifts the halo assembly history to earlier or later times, while keeping the large-scale environment fixed and the initial conditions consistent with a Λ CDM cosmology, thus providing a controlled numerical experiment to test the above hypotheses. I use three sets of initial conditions: a fiducial case with no modification, and ‘GM-early’ and ‘GM-late’ cases where the assembly history is shifted to earlier or later times respectively.

The system used in this study was selected from a periodic simulation run with the Reference EAGLE model (Schaye et al., 2015; Crain et al., 2015) to be a present-day moderately star-forming ($\text{sSFR} = 10^{-10.2} \text{ yr}^{-1}$) central galaxy of stellar mass $M_{\star,30\text{kpc}} = 4.3 \times 10^{10} M_{\odot}$ hosted by a halo of mass $M_{200} = 3.4 \times 10^{12} M_{\odot}$, chosen to match the halo mass scale at which the correlations in Chapters 3 and 4 are the strongest. The GM resimulations were carried out with the ‘Recal’ EAGLE model (Schaye et al., 2015); I also utilise counterpart simulations containing only collisionless gravitational dynamics (DMONLY), simulations implementing hydrodynamics, but no star formation, BH formation, feedback or radiative cooling (NONRAD), and full-physics simulations where no black holes are seeded and no AGN feedback occurs (NOAGN). To quantify the effects of the stochasticity inherent in the EAGLE subgrid prescription for feedback, I evolve the simulation from the initial conditions with nine different ‘seeds’ for the feedback scheme. I therefore quote the emergent properties of the halo and galaxy at a given epoch as a median value for the nine seeds, and quantify the spread with the interquartile range (IQR).

My results can be summarised as follows:

1. In the absence of baryonic physics, the three GM cases assemble similar halo masses of $4.43 \times 10^{12} M_{\odot}$ (GM-early), $3.73 \times 10^{12} M_{\odot}$ (Fiducial) and $2.88 \times 10^{12} M_{\odot}$ (GM-late) by $z = 0$, but exhibit strong differences in their assembly histories induced by the GM initial conditions; as required, the GM-early system assembles the earliest, followed by the Fiducial system, then the GM-late system. At $z = 2$, where the state of the halo is specified by the GM technique, the GM-early system has already assembled 40% of its final mass, while the Fiducial and GM-late systems have

assembled 20% and 3% of their final masses respectively (Fig. 5.3, left column). The assembly history of the Fiducial system is mildly early-forming for its mass, in comparison with the analytical prescription of Correa et al. (2015).

2. When baryonic physics is included, the halo mass assembly histories are broadly similar to the DMONLY case, and exhibit the required shift in assembly time. They reach lower final masses than in DMONLY as a result of the ejection of baryons through feedback processes, though the choice of feedback seed causes very little scatter. At $z = 0$, the final halo masses are $3.58 \times 10^{12} M_{\odot}$ (GM-early, IQR= $0.07 \times 10^{12} M_{\odot}$), $3.21 \times 10^{12} M_{\odot}$ (Fiducial, IQR= $0.18 \times 10^{12} M_{\odot}$), and $2.88 \times 10^{12} M_{\odot}$ (GM-late, IQR= $0.05 \times 10^{12} M_{\odot}$) (Fig. 5.3, right column).
3. The CGM mass fraction (normalised by the cosmic baryon fraction Ω_b/Ω_0) declines for all three GM cases once the halo mass exceeds the critical mass, $M_{\text{crit}}(z)$ (as defined by Bower et al. (2017)), at which the central BH is expected to begin rapidly growing and injecting AGN feedback energy. This depletion occurs earlier for the GM-early case ($z \approx 2$) than for the Fiducial ($z \approx 1$) and GM-late ($z \approx 0.7$) cases. In concordance with the correlations found between f_{CGM} and proxies for the halo assembly time in Chapters 3 and 4, the GM-late halo is the most gas-rich at the present day ($f_{\text{CGM}}/(\Omega_b/\Omega_0) = 0.50$, IQR= 0.06), followed by the Fiducial halo (0.31, IQR= 0.05), then the GM-early halo (0.15, IQR= 0.07), demonstrating a strong influence of assembly time on the baryon content of the CGM in this controlled experiment (Fig. 5.4).
4. The GM realisations of a single halo are an excellent probe of the scatter in the present-day $f_{\text{CGM}} - M_{200}$ relation in the EAGLE Ref-L100N1504 simulation. The Fiducial system lies very close to the median f_{CGM} at $M_{200} \sim 10^{12.5}$, while the GM-early system lies 1.3σ below it and the GM-late system lies 2.0σ above it. The locations of each GM case within the wider population vary strongly with redshift; at $z = 1$, the GM-early system is already very gas-poor relative to the median, but the Fiducial system remains gas-rich as its CGM mass fraction is yet to decline (Fig. 5.5).
5. The onset of rapid BH growth is coincident with $M_{200}(z)$ reaching the $M_{\text{crit}}(z)$ threshold in all GM cases, and is also coincident with the expulsion of baryons and decline in f_{CGM} as predicted by Oppenheimer et al. (2020). Previous simulation results indicate that earlier-forming and more tightly-bound haloes foster the growth of more massive BHs (Booth & Schaye, 2010, 2011), and the GM-early system does indeed host a more massive central BH ($\log(M_{\text{BH}})[M_{\odot}] = 8.08$, IQR= 0.30 dex) at the present day than the Fiducial system (7.90, IQR= 0.08 dex), with the GM-late system forming the lowest-mass BH (7.28, IQR= 0.20 dex). The differences in halo

binding energy induced by the genetic modification therefore modulate the total injected AGN feedback, and subsequently, the baryon content of the CGM (Figure 5.6).

6. After M_{200} exceeds $M_{\text{crit}}(z)$, the total energy injected by the AGN, E_{AGN} , increases to become comparable with the ‘intrinsic’ binding energy of the halo baryons, E_{bind}^b , albeit over a longer time period for the GM-late case than the Fiducial and GM-early cases. The onset of baryon depletion in each system is coincident with $E_{\text{AGN}} \simeq E_{\text{bind}}^b$. The final ratio of these quantities is highest for the GM-early case ($E_{\text{AGN}}/E_{\text{bind}}^b = 2.52$, IQR= 1.94), followed by the Fiducial case (2.26, IQR= 0.44), and the GM-late case (0.83, IQR= 0.45), indicating that earlier halo assembly induces a greater shift away from self-regulation for the system (Fig. 5.7).
7. The specific star formation rate (sSFR) of the central galaxy declines in concert with the ejection of baryons from the halo in all three GM systems. The GM-late and Fiducial haloes remain actively star-forming at the present day (assuming a threshold sSFR of 10^{-11} yr^{-1} for quenching), with sSFR of $10^{-10.03} \text{ yr}^{-1}$ (IQR= 0.16 dex) and $10^{-10.55} \text{ yr}^{-1}$ (IQR= 0.20 dex) respectively. In contrast, the gas-poor GM-early system is quenched at the present day for six of the nine feedback seeds used, with a median sSFR of $10^{-12.37} \text{ yr}^{-1}$, (IQR= 1.54 dex). A systematic shift in the assembly history of a dark matter halo can therefore result in the production of a quenched system rather than an active one, and I demonstrate that AGN feedback is crucial for mediating this connection; in the absence of such feedback, all three GM cases remain actively star-forming at the present day (Fig. 5.8).
8. The expulsion of baryons from the GM-early and Fiducial systems is also accompanied by a marked decrease in the degree of rotational support in the stellar disc. This support is quantified by the stellar co-rotational kinetic energy fraction κ_{co} for which a threshold value of 0.4 separates star-forming discs ($\kappa_{\text{co}} > 0.4$) from quenched ellipticals ($\kappa_{\text{co}} < 0.4$) in EAGLE (Correa et al., 2017). In the GM-late case, the central galaxy undergoes a morphological transformation from spiral to elliptical, with a present-day κ_{co} of 0.24 (IQR= 0.16), while the Fiducial ($\kappa_{\text{co}} = 0.44$, IQR= 0.05) and GM-late ($\kappa_{\text{co}} = 0.43$, IQR= 0.04) central galaxies remain spiral-like according to this definition. A stronger ejection of baryons from the CGM induced by an earlier halo assembly history therefore appears to facilitate a morphological transformation of the central galaxy (Fig. 5.9, see also the images in Fig. 5.10).
9. The injection of AGN feedback energy into the CGM induces a strong shift in its cooling time. In the baryon-rich CGM of the GM-late system, where AGN feedback has had little effect on its contents, the distribution of the particle cooling times,

t_{cool} is similar to that of a non-radiative (NONRAD) simulation of the same halo, with similar median $t_{\text{cool}} = 11.1$ Gyr (NONRAD) and 14.8 Gyr (RECAL). In the baryon-poor GM-early systems, which have experienced stronger AGN feedback, the cooling time distribution is shifted to higher t_{cool} relative to the NONRAD case, elevating the median t_{cool} from 14.7 Gyr to 72.9 Gyr. The majority of the baryons ejected from this system exist beyond $2r_{200}$ and at very long cooling times (Fig. 5.11, left column). This elevation of the cooling time and overall depletion of the circumgalactic reservoir prevents the replenishment of the ISM, leading to the quenching of the central galaxy.

10. The elevation of the CGM cooling time is not the result of a substantial heating by AGN feedback. In both the GM-early and GM-late systems, while a small fraction of the CGM baryons are heated to higher temperatures than are present in the non-radiative case, and a subset of the CGM is able to cool to lower temperatures, the bulk of the halo gas follows the NONRAD temperature distribution (Fig. 5.11, central column).
11. The CGM t_{cool} distribution is shifted to longer times following AGN feedback as a result of changes to the density distribution. In the GM-late case, the bulk of the baryons have the same density distribution in the full-physics and NONRAD realisations, though the most dense component of the NONRAD CGM either cools to higher densities and forms stars, or is ejected from the halo, reducing the median density of the CGM by a factor of 2.8. In the GM-early case, strong AGN-driven baryon ejection reconfigures the CGM at a lower density, with the densest material being the most affected. The majority of the material which *should* be in the halo (since it remains within r_{200} in the NONRAD realisation) is ejected beyond the r_{200} and exists at low densities ($n_{\text{H}} \lesssim 10^{-6} \text{ cm}^{-3}$). This reconfiguration reduces the median density by a factor of 10.1 relative to the NONRAD case, and since the cooling function $\Lambda \propto n_{\text{H}}^2$, this significantly elevates the CGM cooling time, facilitating quenching and morphological transformation (Fig. 5.11, right column).

The results presented in this Chapter conclusively demonstrate that the assembly histories of present-day $\sim L^*$ haloes have a significant impact on the properties of their central galaxies. Here I can identify a clear sequence of events: haloes which form earlier than is typical for their mass foster the growth of more massive SMBHs, which inject more integrated feedback energy into their surrounding gas relative to the binding energy of the halo's baryons. This ejects a larger fraction of the baryons in the CGM, removing the densest gas and reconfiguring the CGM at a lower mean density, raising the cooling time and preventing replenishment of the ISM. This leads to the quenching and morphological transformation of the central galaxy. While the results of Chapters

3 and 4 showed this indirectly through identification of correlation in large statistical samples, here I can have greater confidence that the above sequence of events is driven by differences in the assembly history, because I consider the evolution of only one halo where all other variables are kept fixed.

I have only considered the EAGLE simulation model in this Chapter, however the results in Chapter 4 demonstrated that this picture is similarly valid in the TNG model. If this controlled experiment were replicated using identical initial conditions, but with the TNG model, the overall outcome would likely be the same, but with differences in detail stemming from the use of very different subgrid prescriptions for the feedback associated with star formation and BH growth. As shown in Chapter 4, the onset of efficient baryon ejection in EAGLE occurs when high BH accretion rates are reached, since the efficiency of AGN feedback is fixed; this happens at earlier times for earlier-assembling systems. In contrast, the onset of efficient ejection in TNG occurs when the BH accretion rate is low compared to the Eddington rate, and the AGN injects feedback energy in the high-efficiency ‘kinetic’ mode. This is typically the case once M_{BH} is greater than the “pivot mass” of $10^8 M_{\odot}$. While this threshold mass is likely reached at earlier times for SMBHs in earlier-assembling haloes (leading to the correlations presented in Chapter 4), the onset of ejective feedback need not happen at early ($z \sim 2$) times as it typically does for EAGLE, likely producing very different evolutionary tracks for quantities such as M_{BH} , f_{CGM} in TNG. Performing the same controlled experiment with the TNG model will likely highlight revealing differences induced by the choice of subgrid model, and this is a prime candidate for future work.

Chapter 6

Summary and Conclusions

At the start of this thesis, I posed three “big questions”:

- How does the process of galaxy formation impact the environments of galaxies?
- How do the properties of a galaxy’s environment affect its evolution?
- How are the properties of galaxies linked to the underlying nature of the universe?

In **Chapter 1**, I outlined the importance of these questions in the context of galaxy formation in a Λ CDM cosmology, and described our current understanding of the immediate environments of galaxies: the circumgalactic medium (CGM). I explained why the questions are currently difficult to address conclusively through purely observational means, and motivated the use of simulations to assist in the interpretation of existing results, and to make predictions for future observations.

Throughout this thesis, I have presented novel results which address these questions, obtained through analysis of two suites of cosmological, hydrodynamical simulations (introduced in **Chapter 2**) that represent the current state-of-the-art, and through performing controlled numerical galaxy formation experiments using the powerful “genetic modification” technique. In this chapter I will summarise these results, and connect them to describe a process in which the present-day properties of galaxies can be linked to the assembly histories of their host dark matter haloes, which are established purely by cosmology, and that the key to this connection is the interaction of galaxies with their circumgalactic medium.

To address the first of the above questions, **Chapter 3** commenced with an investigation into the relationship between AGN feedback and the gas content of the CGM in the EAGLE simulations. EAGLE represents an ideal suite for such a study since the parameters

governing the efficiency of feedback were calibrated solely on the stellar properties of galaxies, effectively leaving the properties of the CGM as a prediction of the simulation. I showed that the present-day circumgalactic gas mass fractions of dark matter haloes hosting Milky Way-mass galaxies in EAGLE are insensitive to the intensity of ongoing AGN feedback, but are strongly correlated with the integral of such feedback over the lifetime of the halo and its progenitors. The scatter in the gas fractions of simulated dark matter haloes therefore originates from the diversity in the growth histories of their central black holes, such that at fixed halo mass, gas-poor haloes host over-massive central black holes and vice versa. I showed that this diversity stems from the assembly history of the host dark matter halo, since earlier-assembling systems are more tightly-bound and typically foster the growth of more massive black holes. I also demonstrated that these correlations are borne out in observational proxies for the halo gas fraction, such as measurements of the soft X-ray luminosity and thermal Sunyaev Zeldovich effect, providing a means to test these conclusions with forthcoming observational facilities.

Another significant finding presented in Chapter 3 was that EAGLE haloes of fixed mass which are gas-rich preferentially host central galaxies with higher star-formation rates than average (and vice-versa), suggesting that the AGN-driven ejection of baryons from the halo subsequently impacts the evolution of the central galaxy. I explored the origin of this connection in **Chapter 4**, finding that AGN feedback acts to elevate the cooling time of the surrounding CGM through gas ejection, curtailing the supply of fresh gas to the ISM and leading to quenching. I showed that gas-poor haloes exhibit longer CGM cooling times than average, because they have experienced a greater integrated energy injection from explosive feedback relative to the binding energy of their baryons. These effects have implications for the galaxy beyond its star formation activity - I also show that gas-poor haloes preferentially host galaxies with dispersion-dominated kinematics, while gas-rich haloes of the same mass typically host rotation-supported galaxies.

The results of Chapters 3 and 4 connected the quenching and morphological transformation of galaxies to the assembly histories of their host dark matter haloes, through differences in the central black hole growth history and its effect on the ability of the CGM to cool. I established the generality of these conclusions in Chapter 4 by analysing the IllustrisTNG simulations in concert with EAGLE, finding that the above conclusions still apply. This consensus between two state-of-the-art simulations was not necessarily expected, given that they employ very different hydrodynamical solvers and subgrid prescriptions for feedback. These differences cause the simulations to exhibit very different median CGM mass fractions as a function of halo mass, with low-mass IllustrisTNG haloes being significantly more gas-rich than in EAGLE. The expulsion of CGM gas is a response to the central black hole reaching high accretion rates in EAGLE, whereas in IllustrisTNG it is initiated by the onset of a kinetic feedback mode. This leads to

significant differences between the simulations; the IllustrisTNG population exhibits a strongly positive correlation between the CGM mass fraction and the present-day black hole mass accretion rate which is absent in the EAGLE population. This difference is in principle distinguishable by observations, highlighting the importance of the CGM in constraining feedback models for future generations of hydrodynamical simulations.

The above results are based upon the identification of correlations at fixed halo mass in large statistical samples of galaxies, within simulations of cosmologically representative volumes. A significant disadvantage of this method is that one is ultimately comparing *different* haloes, in which other factors (such as the large-scale environment) might drive, or at least influence, the identified correlations. In **Chapter 5** I performed a more tightly-controlled test of the connection between halo assembly, feedback, the CGM and galaxy formation, by systematically adjusting the assembly history of a single Milky Way-mass halo in an EAGLE ‘zoom’ simulation. This adjustment was achieved through “genetic modification” of the initial conditions of the simulation such that the halo assembly history is shifted to earlier or later times, but the large-scale environment and $z = 0$ halo mass is kept constant. I show that shifting the halo assembly to earlier times leads to the growth of a more massive black hole, and that the greater integrated energy injection from AGN feedback associated with this ejects a greater fraction of the CGM mass. I demonstrate that this ejection reconfigures the CGM at a lower density, elevating its cooling time and facilitating the quenching and morphological transformation of the central galaxy. These effects *must* be the result of a modified assembly history, since all other variables are controlled in this experiment.

This sequence of events, outlined over Chapters 3, 4, and 5, addresses the three “big questions” collectively, forming an intuitive picture of the role that feedback, the CGM and the halo assembly history play in the regulation of galaxy formation:

- **Galaxy formation affects the galaxy environment.** As $\sim L^*$ galaxies form, AGN feedback associated with the growth of the central black hole ejects baryons from the CGM, reconfiguring it at a lower density and elevating its cooling time.
- **The properties of the CGM affect the evolution of the central galaxy.** Replenishment of the ISM is inhibited in gas-poor haloes with long cooling times, leading to the quenching of the central galaxy and facilitating a morphological transformation from spiral-like to elliptical. This process appears to be *necessary* for long-term quenching in EAGLE and IllustrisTNG; simply ejecting star-forming gas from the galaxy is not sufficient, and the efficiently-cooling component of the CGM must also be ejected.

- **The above processes can be directly connected to a cosmological origin.**

The scatter in the present-day black hole mass, and hence the CGM mass fraction and central galaxy properties, is governed by the assembly history of the halo. This connection is demonstrated by both correlations at fixed mass in cosmological simulations, and confirmed by the systematic experiment in Chapter 5.

6.1 Future work

6.1.1 Prospects for observational testing

The correlations between the properties of galaxies (star formation rate, colour, morphology, black hole mass) and the CGM mass fraction of the host halo, predicted in this thesis, are ripe for observational testing. Conclusive observational evidence of these correlations at fixed mass would lend strong credence to the results presented here, however for the reasons discussed in Chapter 2, such evidence will be challenging to obtain.

In Chapter 3, I showed in Fig. 3.4 that both the soft ($0.5 \lesssim Ex \lesssim 2.0$ keV), diffuse X-ray luminosity and the ‘flux’ from the thermal Sunyaev Zel’dovich (tSZ) effect are effective proxies for the CGM mass fraction, and correlate strongly with both the black hole mass (negatively) and star formation rate (positively) at fixed halo mass in EAGLE. As discussed in Section 1.3.1.1, however, detection of diffuse X-ray emission in the CGM of Milky Way-like galaxies remains beyond the capabilities of current observational facilities, however future telescopes such as *Athena* (Barret et al., 2016) and *Lynx* (Özel, 2018) promise to make such detections routine. Similarly, as discussed in Section 1.3.1.2, studying the CGM of Milky Way-like galaxies through the tSZ effect is precluded by the Planck satellite’s $\simeq 10$ arcmin beam, which corresponds to scales significantly larger than the virial radius of nearby $\sim L^*$ galaxies. This approach therefore awaits the next generation of ground-based high-resolution CMB experiments such as CMB-S4 (Abazajian et al., 2016) and the Simons Observatory (Ade et al., 2019).

Other observational routes are, however, more immediately feasible. At X-ray wavelengths, the *eROSITA* instrument on the recently launched Spectrum-Roentgen-Gamma mission (Merloni et al., 2012) will soon map the entire sky at $30\times$ greater sensitivity than the *ROSAT* all-sky survey, averaging a 2 ksec integration time upon its final release, which will comprise 4 years of observations. It will offer $15''$ spatial resolution, allowing X-ray profiles of nearby ($z \approx 0.01$) L^* haloes to be resolved. I am involved with a follow-up study to the results presented herein (Oppenheimer et al., submitted to *ApJL*), in which we demonstrate using EAGLE and IllustrisTNG that the stacking of *eROSITA* X-ray maps about the co-ordinates of blue, star forming galaxies and about

red, quenched galaxies at fixed stellar/halo mass should reveal differences in the diffuse X-ray luminosity originating from differences in the CGM mass fraction. Stacking about the co-ordinates of galaxies with under-massive or over-massive central black holes (using, for example, the sample of dynamically-measured masses described in [Terrazas et al. \(2017\)](#)) should also reveal differences, giving credence to the results in this thesis.

Another more immediately accessible proxy for the CGM mass fraction is the column densities (or covering fractions) of CIV absorption systems in the UV spectra of background quasars. As outlined by [Oppenheimer et al. \(2020\)](#), this ion is an attractive tracer because the 1548, 1551 Å doublet can be readily observed by the Cosmic Origins Spectrograph (COS, aboard the Hubble Space Telescope) in the local Universe ($z \lesssim 0.01$) where dynamically-measured black hole masses are available. At $z = 0$, CIV traces metals at temperatures of $10^4 - 10^5$ K and densities of $10^{-5} - 10^{-3}$ cm $^{-3}$ (e.g. [Rahmati et al., 2016](#)); it is photoionised by the UV background at $T < 10^5$ K at low densities and collisionally ionised at $T \sim 10^5$ K at higher densities, tracing a well-defined region of temperature-density phase space in the CGM. In a companion study to the work presented in this thesis, [Oppenheimer et al. \(2020\)](#) showed that the covering fractions and column densities of CIV in the CGM of Milky Way-like EAGLE galaxies significantly decrease (along with the CGM mass fraction) following episodes of rapid AGN feedback, and that red, quenched systems with overmassive black holes exhibit lower CIV covering fractions overall. Observational confirmation of this correlation should be eminently feasible with COS, and this would provide strong evidence in favour of the ideas presented in this thesis.

6.1.2 Further theoretical study

The results I have presented in this thesis predict that the effects of feedback on the CGM play a crucial, yet hitherto overlooked role in the quenching and morphological transformation of galaxies. The origin of the observed bimodality in the colours and morphologies of galaxies is one of the most important questions in the field of galaxy formation, and the framework for the creation of this bimodality presented in this thesis has therefore opened up many avenues for further research. I will discuss some of these in this section.

6.1.2.1 The connection between CGM expulsion and galaxy morphology

The results in Chapter 4 provide a prediction (from two state-of-the-art simulation models) that the assembly histories and CGM mass fractions of haloes are connected with the morphology and kinematics of their central galaxies, and the results in Chapter 5

unambiguously demonstrated that modifying a halo’s assembly history strongly changes the emergent central galaxy morphology. While the connection is clear, this thesis has not given conclusive evidence for *why* ejecting material from the CGM should transform the morphology of the central galaxy. It is intuitive to understand that the quenching of star formation is driven by the cessation of cooling from the CGM onto the ISM, however it is not clear why this process should alter the existing kinematics of the stellar disc.

In Chapter 4 I speculated that a lack of accretion onto the ISM in gas-poor haloes results in less cold gas in the disc, which would otherwise stabilise the disc against mergers and instabilities, and that a lack of replenishment would inhibit the re-growth of a disc. Testing this hypothesis and establishing a physical connection between the morphologies of galaxies and the content of their CGM will require further simulation work. The genetic modification technique will likely be a valuable tool in this endeavour, as it can be used to adjust both the overall assembly history and the significance of individual mergers (as shown by Pontzen et al., 2017), allowing detailed examination of the role of mergers in this process.

6.1.2.2 Relationship with large-scale structure and the cosmic web

Recent analyses of the galaxy population observed by SDSS and the Horizon-AGN and SIMBA simulations with the DISPERSE algorithm (Sousbie, 2011) have shown that galaxies in haloes of a given mass that are “better connected” to the cosmic web (i.e. by a greater number of filaments) tend to be more massive, less star-forming and less rotationally supported (Kraljic et al., 2020). This appears counterintuitive; one might expect that haloes connected to the cosmic web by a greater number of filaments would have access to a greater supply of gas for star formation, and host active galaxies at their centres.

Guided by the results in this thesis, I postulate that haloes which are better connected are those that assemble the earliest, and hence form the most massive black holes, giving rise to the previously discussed effects on the CGM. Future simulation work could demonstrate that gas-poor haloes are in fact better connected to the cosmic web than their gas-rich counterparts, and that AGN feedback is key to offsetting the increased gas inflow provided by this connectivity. This work would clearly link the star formation activity of a galaxy with its place in the large-scale structure of the Universe through the effects of AGN on the CGM.

6.1.2.3 Breaking the degeneracy in galaxy formation

At several points in this thesis, I have referred to a ‘degeneracy’ in galaxy formation theory. At present, several models produce realistic galaxy populations which provide a good match to many of the characteristics of the observed galaxy population, such as the galaxy stellar mass function, the sizes of galaxies and the galaxy colour bimodality. These models achieve this success despite great differences in their implementation of unresolved ‘subgrid’ physical processes. The most notable differences concern the implementation of feedback processes, about which there is no strong consensus between simulations. The efficiencies of feedback processes are typically calibrated to reproduce the above characteristics of the stellar components of galaxies, leaving the state of the gaseous universe as a prediction of the simulation; the properties of the CGM are therefore best placed to break this degeneracy.

The ability of the CGM to distinguish between subgrid feedback models has been highlighted in this thesis. In Chapter 4, I showed that the critical difference between my findings in EAGLE and IllustrisTNG lies in the initiation of efficient ejective feedback. In EAGLE, CGM ejection depends only on the central black hole accretion rate, and is therefore likely to occur at early times ($z \sim 2$) when accretion rates are high. In contrast, ejection in IllustrisTNG is initiated by the onset of kinetic AGN feedback, which in turn depends on the *mass* of the black hole. This difference, along with EAGLE’s greater efficiency at removing gas from low-mass haloes with feedback from star formation, leads to significant differences in the relationship between the CGM mass fraction and halo mass in lower mass haloes ($M_{200} \lesssim 10^{12.5} M_\odot$). The CGM mass fractions of haloes in this mass range are currently observationally unconstrained, however future detections with more sensitive instrumentation will conclusively rule out one of the two scenarios which, I reiterate, both produce realistic galaxies. In anticipation of these future constraints, the next logical step is to examine the redshift evolution of this relationship in both simulations to identify the epoch at which they diverge, and investigate what the consequences are for the properties of the CGM and galaxy.

Given the sensitivity of the CGM to the adopted subgrid model, it will be vital to investigate how sensitive the conclusions presented in this thesis are to the details of the chosen prescription for black hole seeding, growth and feedback. While the coupling of feedback energy from black holes to the CGM differs between EAGLE and IllustrisTNG, the method for seeding black holes and calculating their accretion rates is similar in both simulations: black holes are seeded at the centres of haloes once a halo mass threshold is reached, and proceed to grow according to a Bondi-Hoyle-like prescription for spherically symmetric accretion. It would therefore be of great interest to extend my analysis of

AGN and the properties of the CGM to a simulation model in which these processes are implemented in a different fashion which is potentially more realistic.

Prime candidates for this include the prescription implemented in the SIMBA (Davé et al., 2019) simulations, itself a modified version of the Anglés-Alcázar et al. (2017) model, where an additional stellar mass threshold is implemented for the seeding of black holes, and the accretion of cold ($T < 10^5$ K) gas proceeds according to the torque-limited accretion model of Hopkins & Quataert (2011). This accretion model links the infall rate onto the black hole to the properties of the inner galactic disc, and in contrast to Bondi-Hoyle accretion, does not require the black hole to regulate its own growth, naturally producing the black hole mass - stellar mass relation without calibration. Another option is the prescription employed in the ROMULUS (Tremmel et al., 2017) simulations, where the seeding of black holes is determined by the properties of the star-forming gas, and where the angular momentum of the gas surrounding the black hole is included in the calculation of the accretion rate. Another approach is to significantly increase the resolution of the simulation in the vicinity of the black hole in order to resolve the Bondi-Hoyle radius. This can be achieved through the use of a moving-mesh code such as AREPO, where the quasi-Lagrangian mesh can be dynamically ‘refined’ to a much higher mass resolution in a small region (Curtis & Sijacki, 2015).

To examine the sensitivity of my results to the choice of model, I envisage a tightly-controlled and systematic experiment in which a suite of identical simulations are run with a base model such as EAGLE, but in which various implementations of black hole seeding, growth and feedback are used. Such an experiment could benefit from the use of genetically-modified initial conditions, which could be employed to systematically adjust the impact of this feedback on the CGM, as in Chapter 5. This would provide insight into how the chosen subgrid implementation of black hole growth and feedback directly impacts the evolution of the host system over cosmic time and produce observable diagnostics for each case which can be tested, providing new constraints for future simulation models.

Appendix A

Comparison of simulated X-ray luminosities and tSZ fluxes with observations

In this short Appendix I briefly compare the observable properties of haloes in the EAGLE simulation with observational measurements. In the panels of Fig. A.1, I show the mean (solid lines) and median (dashed lines) values of scaling relations as a function of halo mass, with the 10th – 90th percentile scatter shown via shading. In bins sampled by fewer than 10 galaxies, haloes are plotted individually. I remind the reader that, when comparing the simulations with measurements derived from stacked observational data, it is more appropriate to compare with the mean of the simulated sample rather than the median.

The upper panel of Fig. A.1 focusses on the $L_X - M_{200}$ relation. The observational measurements are comprised of *XMM-Newton* and *Chandra* direct detections of diffuse emission from the environments of massive, isolated spiral galaxies, compiled and homogeneously re-analysed by Li et al. (2017). Specifically, these are observations of NGC 1961 (Bogdan et al., 2013a; Anderson et al., 2016) and NGC 6753 (Bogdan et al., 2013a), labelled here as ‘‘Massive Spirals’’, the CGM-MASS sample of Li et al. (2017), a measurement of the Milky Way’s X-ray luminosity from Snowden et al. (1997) with uncertainty estimates from Miller & Bregman (2015a), and a sample of inclined disc galaxies observed with *Chandra* and presented by Li et al. (2014). Since direct detections of diffuse X-ray emission from galaxies are generally limited to $r < 0.1r_{200}$, the luminosities of EAGLE haloes in this panel were computed using a 3-dimensional spherical aperture of this radius. The range of luminosities displayed by EAGLE galaxies is generally compatible with the scatter of the observational measurements, but I

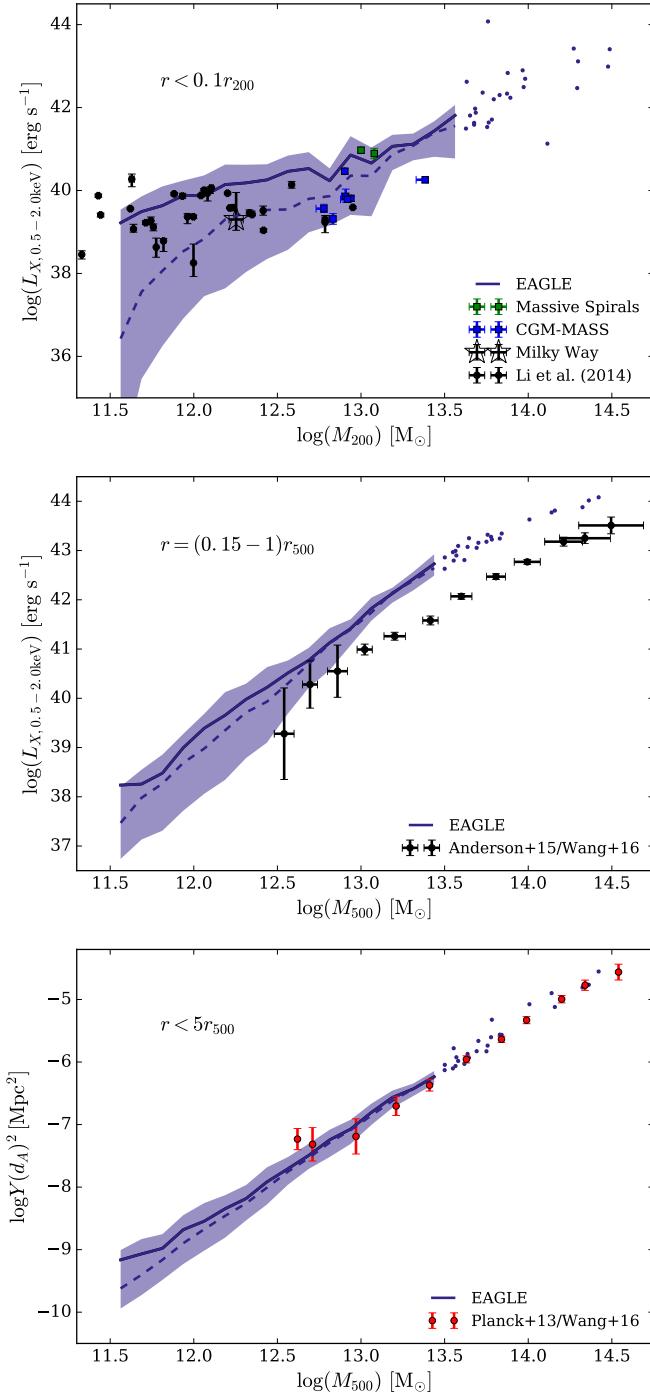


FIGURE A.1: Comparison of EAGLE’s CGM scaling relations with observational measurements. *Top panel*: the $L_X - M_{200}$ relation, where L_X is computed within a spherical aperture of $r = 0.1r_{200}$. Solid (dashed) lines indicating the binned mean (median) values, and shading indicates the 10th – 90th percentile scatter. Galaxies in bins sampled by fewer than 10 galaxies are shown individually. Observational measurements are taken from the homogeneously-reanalysed compilation of Li et al. (2017). *Centre panel*: the $L_X - M_{500}$ relation, where L_X is computed within a spherical annulus of $0.15 < r/r_{200} < 1$. Observational measurements are based on X-ray and weak lensing shear maps stacked about the coordinates of massive local galaxies, as per Wang et al. (2016). *Bottom panel*: the $Y d_A^2 - M_{500}$ relation, where Y is computed in a spherical aperture of $r = 5r_{500}$. As above, observational measurements are based on the stacking analysis of Wang et al. (2016), here using Planck CMB maps.

caution that direct detection observations are likely strongly biased towards the most X-ray-luminous galaxies at fixed mass.

In the centre panel, I compare the X-ray luminosity of EAGLE galaxies to the $L_{\text{X}} - M_{500}$ scaling relation presented by [Wang et al. \(2016\)](#), who used weak lensing halo mass estimates to recalibrate the relation originally presented by [Anderson et al. \(2015\)](#), derived by stacking X-ray maps from the *ROSAT* All-Sky Survey about the coordinates of optically-selected massive galaxies in the local Universe. The low spatial resolution of the maps precludes the excision of bright X-ray point sources, therefore I compare to the “CGM” luminosities of [Anderson et al. \(2015\)](#), computed using an aperture of $0.15 < r/r_{200} < 1$. This comparison reveals that the extended gas haloes of very massive galaxies, and galaxy groups, in EAGLE are too X-ray-luminous. This shortcoming was previously highlighted by [Schaye et al. \(2015\)](#), who showed that the issue can be mitigated via the use of a higher temperature increment in EAGLE’s stochastic heating implementation of AGN feedback. I elected to present results here from the Ref-L100N1504 simulation rather than the AGNdT9-L050N0752, which uses this higher heating temperature, since the latter affords a factor of 8 poorer sampling of the galaxy population. Moreover, I focus here primarily on $\sim L^*$ galaxies, for which the correspondence with the observed X-ray luminosity is reasonable for the Reference model.

Finally, I show in the bottom panel the stacked tSZ signal for the same sample used by [Anderson et al. \(2015\)](#) and [Wang et al. \(2016\)](#), as presented by [Planck Collaboration et al. \(2013\)](#). Following [McCarthy et al. \(2017\)](#), I convert the observed values of $Y(< r_{500})$ back into the measured flux $Y(< 5r_{500})$, and compare with tSZ flux of EAGLE galaxies computed within the same spherical aperture. The simulations reproduce the observed flux on these relatively large scales very well; a more stringent test of the simulation based on the tSZ flux awaits the availability of panoramic CMB maps with high spatial resolution.

I reiterate that the gaseous properties of galaxies and their haloes were not considered during the calibration of the parameters governing energetic feedback in EAGLE, and they may therefore be considered as predictions of the simulations. Whilst it is possible to identify differences in detail, the observable properties of the extended gaseous environments of $\sim L^*$ galaxies in EAGLE are in sufficiently good agreement with observational measurements to engender confidence in this aspect of the simulations.

Bibliography

- Abazajian K. N., et al., 2009, ApJS, 182, 543
- Abazajian K. N., et al., 2016, arXiv e-prints, p. arXiv:1610.02743
- Ade P., et al., 2019, J. Cosmology Astropart. Phys., 2019, 056
- Agertz O., Kravtsov A. V., 2016, The Astrophysical Journal, 824, 79
- Agertz O., et al., 2007, Monthly Notices of the Royal Astronomical Society, 380, 963
- Allen S. W., Schmidt R. W., Fabian A. C., 2002, Monthly Notices of the Royal Astronomical Society, 334, L11
- Anders E., Grevesse N., 1989, Geochimica Cosmochimica Acta, 53, 197
- Anderson M. E., Bregman J. N., 2010, The Astrophysical Journal, 714, 320
- Anderson M. E., Bregman J. N., 2011, The Astrophysical Journal, 737, 22
- Anderson M. E., Gaspari M., White S. D. M., Wang W., Dai X., 2015, Monthly Notices of the Royal Astronomical Society, 449, 3806
- Anderson M. E., Churazov E., Bregman J. N., 2016, Monthly Notices of the Royal Astronomical Society, 455, 227
- Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017, Monthly Notices of the Royal Astronomical Society, 472, L109
- Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, Astronomy & Astrophysics, 517, A92
- Bahé Y. M., et al., 2017, Monthly Notices of the Royal Astronomical Society, 470, 4186
- Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, The Astrophysical Journal, 600, 681
- Baldry I. K., et al., 2012, Monthly Notices of the Royal Astronomical Society, 421, 621

- Balogh M. L., Pearce F. R., Bower R. G., Kay S. T., 2001, Monthly Notices of the Royal Astronomical Society, 326, 1228
- Barai P., Murante G., Borgani S., Gaspari M., Granato G. L., Monaco P., Ragone-Figueroa C., 2016, Monthly Notices of the Royal Astronomical Society, 461, 1548
- Barnes J., Hut P., 1986, Nature, 324, 446
- Barnes D. J., et al., 2017, Monthly Notices of the Royal Astronomical Society, 471, 1088
- Barret D., et al., 2016, in Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray. p. 99052F ([arXiv:1608.08105](https://arxiv.org/abs/1608.08105)), doi:10.1117/12.2232432
- Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, The Astrophysical Journal, 758, 75
- Bechtold J., Croots A. P. S., Duncan R. C., Fang Y., 1994, ApJ, 437, L83
- Behroozi P. S., Conroy C., Wechsler R. H., 2010, The Astrophysical Journal, 717, 379
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3143
- Benson A. J., Bower R. G., Frenk C. S., White S. D. M., 2000, Monthly Notices of the Royal Astronomical Society, 314, 557
- Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, The Astrophysical Journal, 599, 38
- Berg T. A. M., Ellison S. L., Tumlinson J., Oppenheimer B. D., Horton R., Bordoloi R., Schaye J., 2018, Monthly Notices of the Royal Astronomical Society, 478, 3890
- Bergeron J., Stasińska G., 1986, Astronomy & Astrophysics, 169, 1
- Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, Monthly Notices of the Royal Astronomical Society, 436, 697
- Bertschinger E., 2001, ApJS, 137, 1
- Bielby R. M., et al., 2019, Monthly Notices of the Royal Astronomical Society, 486, 21
- Binney J., 1977, The Astrophysical Journal, 215, 483
- Binney J., Tabor G., 1995, Monthly Notices of the Royal Astronomical Society, 276, 663
- Birnboim Y., Dekel A., 2003, Monthly Notices of the Royal Astronomical Society, 345, 349

- Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529
- Bland-Hawthorn J., Veilleux S., Cecil G., 2007, *Ap&SS*, 311, 87
- Blumenthal G. R., Pagels H., Primack J. R., 1982, *Nature*, 299, 37
- Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, *Monthly Notices of the Royal Astronomical Society*, 456, 2361
- Bogdán Á., et al., 2013a, *The Astrophysical Journal*, 772, 97
- Bogdán Á., Forman W. R., Kraft R. P., Jones C., 2013b, *The Astrophysical Journal*, 772, 98
- Bogdán Á., Bourdin H., Forman W. R., Kraft R. P., Vogelsberger M., Hernquist L., Springel V., 2017, *The Astrophysical Journal*, 850, 98
- Bondi H., Hoyle F., 1944, *Monthly Notices of the Royal Astronomical Society*, 104, 273
- Booth C. M., Schaye J., 2009, *Monthly Notices of the Royal Astronomical Society*, 398, 53
- Booth C. M., Schaye J., 2010, *Monthly Notices of the Royal Astronomical Society*, 405, L1
- Booth C. M., Schaye J., 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 1158
- Bordoloi R., et al., 2014, *The Astrophysical Journal*, 796, 136
- Borthakur S., et al., 2015, *The Astrophysical Journal*, 813, 46
- Bouwens R. J., et al., 2015, *The Astrophysical Journal*, 803, 34
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *Monthly Notices of the Royal Astronomical Society*, 370, 645
- Bower R. G., McCarthy I. G., Benson A. J., 2008, *Monthly Notices of the Royal Astronomical Society*, 390, 1399
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *Monthly Notices of the Royal Astronomical Society*, 465, 32
- Bregman J. N., 2007, *Annual Review of Astronomy and Astrophysics*, 45, 221
- Burchett J. N., et al., 2016, *The Astrophysical Journal*, 832, 124
- Chabrier G., 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763

- Chen H.-W., Wild V., Tinker J. L., Gauthier J.-R., Helsby J. E., Shectman S. A., Thompson I. B., 2010, *ApJ*, 724, L176
- Cicone C., et al., 2015, *Astronomy & Astrophysics*, 574, A14
- Cicone C., Maiolino R., Marconi A., 2016, *Astronomy & Astrophysics*, 588, A41
- Clauwens B., Schaye J., Franx M., Bower R. G., 2018, *Monthly Notices of the Royal Astronomical Society*, 478, 3994
- Cleveland W. S., 1979, *Journal of the American Statistical Association*, 74, 829
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *Monthly Notices of the Royal Astronomical Society*, 319, 168
- Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1514
- Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C. R., 2017, *Monthly Notices of the Royal Astronomical Society*, 472, L45
- Correa C. A., Schaye J., Trayford J. W., 2019, *Monthly Notices of the Royal Astronomical Society*, 484, 4401
- Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, *AJ*, 112, 839
- Crain R. A., Eke V. R., Frenk C. S., Jenkins A., McCarthy I. G., Navarro J. F., Pearce F. R., 2007, *Monthly Notices of the Royal Astronomical Society*, 377, 41
- Crain R. A., et al., 2009, *Monthly Notices of the Royal Astronomical Society*, 399, 1773
- Crain R. A., McCarthy I. G., Frenk C. S., Theuns T., Schaye J., 2010, *Monthly Notices of the Royal Astronomical Society*, 407, 1403
- Crain R. A., McCarthy I. G., Schaye J., Theuns T., Frenk C. S., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 3005
- Crain R. A., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1937
- Crain R. A., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 4204
- Crighton N. H. M., Morris S. L., Bechtold J., Crain R. A., Jannuzzi B. T., Shone A., Theuns T., 2010, *Monthly Notices of the Royal Astronomical Society*, 402, 1273
- Croton D. J., et al., 2006, *Monthly Notices of the Royal Astronomical Society*, 365, 11
- Cui W., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 2898

- Cullen L., Dehnen W., 2010, *Monthly Notices of the Royal Astronomical Society*, 408, 669
- Curtis M., Sijacki D., 2015, *Monthly Notices of the Royal Astronomical Society*, 454, 3445
- Dai X., Anderson M. E., Bregman J. N., Miller J. M., 2012, *The Astrophysical Journal*, 755, 107
- Dalla Vecchia C., Schaye J., 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 140
- Davé R., Thompson R., Hopkins P. F., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3265
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *Monthly Notices of the Royal Astronomical Society*, 486, 2827
- Davies J. J., Crain R. A., McCarthy I. G., Oppenheimer B. D., Schaye J., Schaller M., McAlpine S., 2019, *Monthly Notices of the Royal Astronomical Society*, 485, 3783
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *The Astrophysical Journal*, 292, 371
- Dehnen W., Aly H., 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 1068
- Dinshaw N., Foltz C. B., Impey C. D., Weymann R. J., Morris S. L., 1995, *Nature*, 373, 223
- Dolag K., Borgani S., Murante G., Springel V., 2009, *Monthly Notices of the Royal Astronomical Society*, 399, 497
- Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, *Monthly Notices of the Royal Astronomical Society*, 433, 3297
- Dubois Y., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 1453
- Durier F., Dalla Vecchia C., 2012, *Monthly Notices of the Royal Astronomical Society*, 419, 465
- Efstathiou G., Sutherland W. J., Maddox S. J., 1990, *Nature*, 348, 705
- Fall S. M., Efstathiou G., 1980, *Monthly Notices of the Royal Astronomical Society*, 193, 189
- Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, *Monthly Notices of the Royal Astronomical Society*, 478, 2618

- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *The Astrophysical Journal*, 703, 1416
- Ferland G. J., Korista K. T., Verner D. A., Ferguson J. W., Kingdon J. B., Verner E. M., 1998, *Publications of the Astronomical Society of the Pacific*, 110, 761
- Font A. S., McCarthy I. G., Le Brun A. M. C., Crain R. A., Kelvin L. S., 2017, *Publ. Astron. Soc. Australia*, 34, e050
- Ford A. B., Oppenheimer B. D., Davé R., Katz N., Kollmeier J. A., Weinberg D. H., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 89
- Ford A. B., Davé R., Oppenheimer B. D., Katz N., Kollmeier J. A., Thompson R., Weinberg D. H., 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 1260
- Ford A. B., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 1745
- Friedmann A., 1922, *Zeitschrift fur Physik*, 10, 377
- Fukugita M., Peebles P. J. E., 2006, *The Astrophysical Journal*, 639, 590
- Furlong M., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 4486
- Furlong M., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 465, 722
- Genel S., et al., 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 175
- Genzel R., Eckart A., Ott T., Eisenhauer F., 1997, *Monthly Notices of the Royal Astronomical Society*, 291, 219
- Gnat O., Sternberg A., 2007, *ApJS*, 168, 213
- Gonzalez A. H., Sivanandam S., Zabludoff A. I., Zaritsky D., 2013, *The Astrophysical Journal*, 778, 14
- Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, *Monthly Notices of the Royal Astronomical Society*, 374, 1479
- Grand R. J. J., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 467, 179
- Greco J. P., Hill J. C., Spergel D. N., Battaglia N., 2015, *The Astrophysical Journal*, 808, 151
- Guedes J., Callegari S., Madau P., Mayer L., 2011, *The Astrophysical Journal*, 742, 76
- Guo Q., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 3457

- Guth A. H., 1981, Phys. Rev. D, 23, 347
- Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays. p. 64 ([arXiv:astro-ph/0106018](https://arxiv.org/abs/astro-ph/0106018))
- Häring N., Rix H.-W., 2004, ApJ, 604, L89
- Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, Monthly Notices of the Royal Astronomical Society, 441, 3306
- Henden N. A., Puchwein E., Shen S., Sijacki D., 2018, Monthly Notices of the Royal Astronomical Society, 479, 5385
- Hennawi J. F., et al., 2006, The Astrophysical Journal, 651, 61
- Henriques B. M. B., Yates R. M., Fu J., Guo Q., Kauffmann G., Srisawat C., Thomas P. A., White S. D. M., 2020, Monthly Notices of the Royal Astronomical Society, 491, 5795
- Hodges-Kluck E. J., Miller M. J., Bregman J. N., 2016, The Astrophysical Journal, 822, 21
- Hopkins P. F., 2013, Monthly Notices of the Royal Astronomical Society, 428, 2840
- Hopkins P. F., Quataert E., 2011, Monthly Notices of the Royal Astronomical Society, 415, 1027
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005, The Astrophysical Journal, 630, 705
- Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, The Astrophysical Journal, 691, 1168
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, Monthly Notices of the Royal Astronomical Society, 445, 581
- Hopkins P. F., et al., 2018, Monthly Notices of the Royal Astronomical Society, 480, 800
- Hubble E. P., 1925, The Astrophysical Journal, 62, 409
- Hubble E. P., 1926a, The Astrophysical Journal, 63, 236
- Hubble E. P., 1926b, The Astrophysical Journal, 64, 321
- Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168
- Hummels C. B., Bryan G. L., Smith B. D., Turk M. J., 2013, Monthly Notices of the Royal Astronomical Society, 430, 1548

- Jahnke K., Macciò A. V., 2011, *The Astrophysical Journal*, 734, 92
- Jeans J. H., 1902, *Philosophical Transactions of the Royal Society of London Series A*, 199, 1
- Jenkins A., 2013, *Monthly Notices of the Royal Astronomical Society*, 434, 2094
- Johansson P. H., Burkert A., Naab T., 2009, *The Astrophysical Journal*, 707, L184
- Katz N., Gunn J. E., 1991, *The Astrophysical Journal*, 377, 365
- Katz N., White S. D. M., 1993, *The Astrophysical Journal*, 412, 455
- Kennicutt Jr. R. C., 1998, *ARA&A*, 36, 189
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *Monthly Notices of the Royal Astronomical Society*, 363, 2
- Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1349
- Kormendy J., Ho L. C., 2013, *Annual Review of Astronomy and Astrophysics*, 51, 511
- Kraljic K., et al., 2020, *Monthly Notices of the Royal Astronomical Society*, 491, 4294
- Krumholz M. R., Tan J. C., 2007, *The Astrophysical Journal*, 654, 304
- Lacey C., Cole S., 1994, *Monthly Notices of the Royal Astronomical Society*, 271, 676
- Lacey C. G., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3854
- Lakhchaura K., Truong N., Werner N., 2019, *Monthly Notices of the Royal Astronomical Society: Letters*, 488, L134L142
- Lan T.-W., Ménard B., Zhu G., 2014, *The Astrophysical Journal*, 795, 31
- Lehner N., O'Meara J. M., Fox A. J., Howk J. C., Prochaska J. X., Burns V., Armstrong A. A., 2014, *The Astrophysical Journal*, 788, 119
- Lemaître G., 1927, *Annales de la Société Scientifique de Bruxelles*, 47, 49
- Lemaître G., 1931, *Monthly Notices of the Royal Astronomical Society*, 91, 483
- Lewis A., Challinor A., Lasenby A., 2000, *The Astrophysical Journal*, 538, 473
- Li C., White S. D. M., 2009, *Monthly Notices of the Royal Astronomical Society*, 398, 2177

- Li J.-T., Crain R. A., Wang Q. D., 2014, Monthly Notices of the Royal Astronomical Society, 440, 859
- Li J.-T., Bregman J. N., Wang Q. D., Crain R. A., Anderson M. E., 2016, The Astrophysical Journal, 830, 134
- Li J.-T., Bregman J. N., Wang Q. D., Crain R. A., Anderson M. E., Zhang S., 2017, The Astrophysical Journal Supplement Series, 233, 20
- Lin Y.-T., Mohr J. J., Stanford S. A., 2004, The Astrophysical Journal, 610, 745
- Lin Y.-T., Stanford S. A., Eisenhardt P. R. M., Vikhlinin A., Maughan B. J., Kravtsov A., 2012, ApJ, 745, L3
- Linde A. D., 1982, Physics Letters B, 108, 389
- Lopez S., et al., 2018, Nature, 554, 493
- Lovisari L., Reiprich T. H., Schellenberger G., 2015, Astronomy & Astrophysics, 573, A118
- Ludlow A. D., Schaye J., Bower R., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3663
- Madau P., Pozzetti L., Dickinson M., 1998, The Astrophysical Journal, 498, 106
- Magorrian J., et al., 1998, AJ, 115, 2285
- Maiolino R., et al., 2012, Monthly Notices of the Royal Astronomical Society, 425, L66
- Marinacci F., Pakmor R., Springel V., 2014, Monthly Notices of the Royal Astronomical Society, 437, 1750
- Martin C. L., Shapley A. E., Coil A. L., Kornei K. A., Bundy K., Weiner B. J., Noeske K. G., Schiminovich D., 2012, The Astrophysical Journal, 760, 127
- Matthee J., Schaye J., 2019, Monthly Notices of the Royal Astronomical Society, 484, 915
- Matthee J., Schaye J., Crain R. A., Schaller M., Bower R., Theuns T., 2017, Monthly Notices of the Royal Astronomical Society, 465, 2381
- McAlpine S., et al., 2016, Astronomy and Computing, 15, 72
- McAlpine S., Bower R. G., Harrison C. M., Crain R. A., Schaller M., Schaye J., Theuns T., 2017, Monthly Notices of the Royal Astronomical Society, 468, 3395

- McAlpine S., Bower R. G., Rosario D. J., Crain R. A., Schaller M., Schaye J., Theuns T., 2018, preprint, p. arXiv:1805.08293 ([arXiv:1805.08293](https://arxiv.org/abs/1805.08293))
- McCarthy I. G., et al., 2010, Monthly Notices of the Royal Astronomical Society, 406, 822
- McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, Monthly Notices of the Royal Astronomical Society, 412, 1965
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, Monthly Notices of the Royal Astronomical Society, 465, 2936
- McConnell N. J., Ma C.-P., 2013, The Astrophysical Journal, 764, 184
- McDonald M., Gaspari M., McNamara B. R., Tremblay G. R., 2018, The Astrophysical Journal, 858, 45
- Merloni A., et al., 2012, arXiv e-prints, p. arXiv:1209.3114
- Miller M. J., Bregman J. N., 2013, The Astrophysical Journal, 770, 118
- Miller M. J., Bregman J. N., 2015a, The Astrophysical Journal, 800, 14
- Miller M. J., Bregman J. N., 2015b, The Astrophysical Journal, 800, 14
- Mitchell P. D., et al., 2018, Monthly Notices of the Royal Astronomical Society, 474, 492
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2019, arXiv e-prints,
- Moster B. P., Naab T., White S. D. M., 2013, Monthly Notices of the Royal Astronomical Society, 428, 3121
- Naab T., Ostriker J. P., 2017, ARA&A, 55, 59
- Navarro J. F., Benz W., 1991, The Astrophysical Journal, 380, 320
- Navarro J. F., White S. D. M., 1993, Monthly Notices of the Royal Astronomical Society, 265, 271
- Navarro J. F., White S. D. M., 1994, Monthly Notices of the Royal Astronomical Society, 267, 401
- Navarro J. F., Frenk C. S., White S. D. M., 1997, The Astrophysical Journal, 490, 493
- Navarro J. F., et al., 2004, Monthly Notices of the Royal Astronomical Society, 349, 1039

- Neistein E., van den Bosch F. C., Dekel A., 2006, Monthly Notices of the Royal Astronomical Society, 372, 933
- Nelson D., et al., 2015, Astronomy and Computing, 13, 12
- Nelson D., et al., 2018a, Monthly Notices of the Royal Astronomical Society, 475, 624
- Nelson D., et al., 2018b, Monthly Notices of the Royal Astronomical Society, 477, 450
- Nelson D., et al., 2019, Computational Astrophysics and Cosmology, 6, 2
- Neto A. F., et al., 2007, Monthly Notices of the Royal Astronomical Society, 381, 1450
- Nicastro F., 2018, arXiv e-prints, p. arXiv:1811.03498
- Nicastro F., et al., 2018, Nature, 558, 406
- Nielsen N. M., Churchill C. W., Kacprzak G. G., Murphy M. T., 2013, The Astrophysical Journal, 776, 114
- Nielsen N. M., Kacprzak G. G., Muzahid S., Churchill C. W., Murphy M. T., Charlton J. C., 2017, The Astrophysical Journal, 834, 148
- Nielsen N. M., Kacprzak G. G., Pointon S. K., Murphy M. T., Churchill C. W., Davé R., 2020, arXiv e-prints, p. arXiv:2002.08516
- Okamoto T., Eke V. R., Frenk C. S., Jenkins A., 2005, Monthly Notices of the Royal Astronomical Society, 363, 1299
- Oppenheimer B. D., 2018, Monthly Notices of the Royal Astronomical Society, 480, 2963
- Oppenheimer B. D., Davé R., 2008, Monthly Notices of the Royal Astronomical Society, 387, 577
- Oppenheimer B. D., Schaye J., 2013a, Monthly Notices of the Royal Astronomical Society, 434, 1043
- Oppenheimer B. D., Schaye J., 2013b, Monthly Notices of the Royal Astronomical Society, 434, 1063
- Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, Monthly Notices of the Royal Astronomical Society, 406, 2325
- Oppenheimer B. D., et al., 2016a, Monthly Notices of the Royal Astronomical Society, 460, 2157
- Oppenheimer B. D., et al., 2016b, Monthly Notices of the Royal Astronomical Society, 460, 2157

- Oppenheimer B. D., Segers M., Schaye J., Richings A. J., Crain R. A., 2018, Monthly Notices of the Royal Astronomical Society, 474, 4740
- Oppenheimer B. D., et al., 2020, Monthly Notices of the Royal Astronomical Society, 491, 2939
- Orr M. E., et al., 2018, Monthly Notices of the Royal Astronomical Society, 478, 3653
- Özel F., 2018, Nature Astronomy, 2, 608
- Parkinson H., Cole S., Helly J., 2008, Monthly Notices of the Royal Astronomical Society, 383, 557
- Peebles P. J. E., 1982, ApJ, 263, L1
- Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, The Astrophysical Journal, 786, 54
- Peeples M. S., et al., 2019, The Astrophysical Journal, 873, 129
- Penzias A. A., Wilson R. W., 1965, The Astrophysical Journal, 142, 419
- Perlmutter S., et al., 1999, The Astrophysical Journal, 517, 565
- Pillepich A., et al., 2018a, Monthly Notices of the Royal Astronomical Society, 473, 4077
- Pillepich A., et al., 2018b, Monthly Notices of the Royal Astronomical Society, 473, 4077
- Planck Collaboration et al., 2013, Astronomy & Astrophysics, 557, A52
- Planck Collaboration et al., 2014, Astronomy & Astrophysics, 571, A16
- Planck Collaboration et al., 2016, Astronomy & Astrophysics, 594, A13
- Planck Collaboration et al., 2018, preprint, p. arXiv:1807.06209 ([arXiv:1807.06209](https://arxiv.org/abs/1807.06209))
- Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, Monthly Notices of the Royal Astronomical Society, 438, 195
- Pontzen A., Tremmel M., Roth N., Peiris H. V., Saintonge A., Volonteri M., Quinn T., Governato F., 2017, Monthly Notices of the Royal Astronomical Society, 465, 547
- Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, Astronomy & Astrophysics, 498, 361
- Press W. H., Schechter P., 1974, The Astrophysical Journal, 187, 425
- Price D. J., 2010, Monthly Notices of the Royal Astronomical Society, 401, 1475

- Prochaska J. X., et al., 2017, *The Astrophysical Journal*, 837, 169
- Puchwein E., Sijacki D., Springel V., 2008, *ApJ*, 687, L53
- Qu Y., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 1659
- Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, *Monthly Notices of the Royal Astronomical Society*, 452, 2034
- Rahmati A., Schaye J., Crain R. A., Oppenheimer B. D., Schaller M., Theuns T., 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 310
- Rees M. J., Ostriker J. P., 1977, *Monthly Notices of the Royal Astronomical Society*, 179, 541
- Rey M. P., Pontzen A., 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 45
- Riess A. G., et al., 1998, *AJ*, 116, 1009
- Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, *The Astrophysical Journal*, 645, 986
- Rosas-Guevara Y. M., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 454, 1038
- Roth N., Pontzen A., Peiris H. V., 2016, *Monthly Notices of the Royal Astronomical Society*, 455, 974
- Rubin V. C., Ford W. Kent J., 1970, *The Astrophysical Journal*, 159, 379
- Rubin V. C., Ford W. K. J., Thonnard N., 1980, *The Astrophysical Journal*, 238, 471
- Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, *The Astrophysical Journal*, 794, 156
- Rubin K. H. R., Diamond-Stanic A. M., Coil A. L., Crighton N. H. M., Moustakas J., 2018, *The Astrophysical Journal*, 853, 95
- Rupke D. S. N., Veilleux S., 2011, *The Astrophysical Journal*, 729, L27
- Salpeter E. E., 1955, *The Astrophysical Journal*, 121, 161
- Sawala T., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 457, 1931
- Schaller M., et al., 2015a, *Monthly Notices of the Royal Astronomical Society*, 451, 1247
- Schaller M., et al., 2015b, *Monthly Notices of the Royal Astronomical Society*, 451, 1247

- Schawinski K., et al., 2014, Monthly Notices of the Royal Astronomical Society, 440, 889
- Schaye J., Dalla Vecchia C., 2008, Monthly Notices of the Royal Astronomical Society, 383, 1210
- Schaye J., et al., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1536
- Schaye J., et al., 2015, Monthly Notices of the Royal Astronomical Society, 446, 521
- Schroetter I., Bouché N., Péroux C., Murphy M. T., Contini T., Finley H., 2015, The Astrophysical Journal, 804, 83
- Segers M. C., Schaye J., Bower R. G., Crain R. A., Schaller M., Theuns T., 2016, Monthly Notices of the Royal Astronomical Society, 461, L102
- Segers M. C., Oppenheimer B. D., Schaye J., Richings A. J., 2017, Monthly Notices of the Royal Astronomical Society, 471, 1026
- Sembolini F., et al., 2016, Monthly Notices of the Royal Astronomical Society, 459, 2973
- Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2016, The Astrophysical Journal, 826, 200
- Shull J. M., Smith B. D., Danforth C. W., 2012, The Astrophysical Journal, 759, 23
- Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, Monthly Notices of the Royal Astronomical Society, 380, 877
- Silk J., 1977, The Astrophysical Journal, 211, 638
- Silk J., Rees M. J., 1998a, Astronomy & Astrophysics, 331, L1
- Silk J., Rees M. J., 1998b, Astronomy & Astrophysics, 331, L1
- Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91
- Smoot G. F., et al., 1992, ApJ, 396, L1
- Snowden S. L., et al., 1997, The Astrophysical Journal, 485, 125
- Soltan A., 1982, Monthly Notices of the Royal Astronomical Society, 200, 115
- Sousbie T., 2011, Monthly Notices of the Royal Astronomical Society, 414, 350
- Spergel D. N., et al., 2007, ApJS, 170, 377
- Spitzer Lyman J., 1956, The Astrophysical Journal, 124, 20
- Springel V., 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

- Springel V., 2010, Monthly Notices of the Royal Astronomical Society, 401, 791
- Springel V., Hernquist L., 2003, Monthly Notices of the Royal Astronomical Society, 339, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Monthly Notices of the Royal Astronomical Society, 328, 726
- Springel V., Di Matteo T., Hernquist L., 2005, Monthly Notices of the Royal Astronomical Society, 361, 776
- Springel V., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 676
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, The Astrophysical Journal, 717, 289
- Stevens A. R. H., et al., 2019, Monthly Notices of the Royal Astronomical Society, 483, 5334
- Stocke J. T., Keeney B. A., Danforth C. W., Shull J. M., Froning C. S., Green J. C., Penton S. V., Savage B. D., 2013, The Astrophysical Journal, 763, 148
- Stott J. P., et al., 2012, Monthly Notices of the Royal Astronomical Society, 422, 2213
- Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, The Astrophysical Journal, 693, 1142
- Tanimura H., et al., 2019a, Monthly Notices of the Royal Astronomical Society, 483, 223
- Tanimura H., Aghanim N., Douspis M., Beelen A., Bonjean V., 2019b, Astronomy & Astrophysics, 625, A67
- Terrazas B. A., Bell E. F., Woo J., Henriques B. M. B., 2017, The Astrophysical Journal, 844, 170
- Terrazas B. A., et al., 2019, arXiv e-prints, p. arXiv:1906.02747
- Thob A. C. R., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 972
- Thompson T. A., Quataert E., Murray N., 2005, The Astrophysical Journal, 630, 167
- Trayford J. W., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2879
- Trayford J. W., et al., 2017, Monthly Notices of the Royal Astronomical Society, 470, 771

- Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, Monthly Notices of the Royal Astronomical Society, 470, 1121
- Tumlinson J., et al., 2011, Science, 334, 948
- Tumlinson J., et al., 2013, The Astrophysical Journal, 777, 59
- Tumlinson J., Peebles M. S., Werk J. K., 2017, ARA&A, 55, 389
- Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2014, Monthly Notices of the Royal Astronomical Society, 445, 794
- Turner M. L., Schaye J., Steidel C. C., Rudie G. C., Strom A. L., 2015, Monthly Notices of the Royal Astronomical Society, 450, 2067
- Turner M. L., Schaye J., Crain R. A., Rudie G., Steidel C. C., Strom A., Theuns T., 2017, Monthly Notices of the Royal Astronomical Society, 471, 690
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769
- Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, The Astrophysical Journal, 640, 691
- Vogelsberger M., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 1518
- Voit G. M., Balogh M. L., Bower R. G., Lacey C. G., Bryan G. L., 2003, The Astrophysical Journal, 593, 272
- Voit G. M., Kay S. T., Bryan G. L., 2005, Monthly Notices of the Royal Astronomical Society, 364, 909
- Walsh D., Carswell R. F., Weymann R. J., 1979, Nature, 279, 381
- Wang W., White S. D. M., Mandelbaum R., Henriques B., Anderson M. E., Han J., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2301
- Weinberger R., et al., 2017, Monthly Notices of the Royal Astronomical Society, 465, 3291
- Weinberger R., et al., 2018, Monthly Notices of the Royal Astronomical Society, 479, 4056
- Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O'Meara J. M., Peebles M. S., 2013, The Astrophysical Journal Supplement Series, 204, 17
- Werk J. K., et al., 2014, The Astrophysical Journal, 792, 8

- Werk J. K., et al., 2016, *The Astrophysical Journal*, 833, 54
- White S. D. M., Frenk C. S., 1991, *The Astrophysical Journal*, 379, 52
- White S. D. M., Rees M. J., 1978, *Monthly Notices of the Royal Astronomical Society*, 183, 341
- White S. D. M., Frenk C. S., Davis M., 1983, *ApJ*, 274, L1
- Wiersma R. P. C., Schaye J., Smith B. D., 2009a, *Monthly Notices of the Royal Astronomical Society*, 393, 99
- Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, *Monthly Notices of the Royal Astronomical Society*, 399, 574
- Wright E. L., et al., 1992, *ApJ*, 396, L13
- Zwicky F., 1933, *Helvetica Physica Acta*, 6, 110
- Zwicky F., 1937, *The Astrophysical Journal*, 86, 217
- de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2019, *Astronomy & Astrophysics*, 624, A48
- van de Voort F., Schaye J., 2012, *Monthly Notices of the Royal Astronomical Society*, 423, 2991
- van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, *Monthly Notices of the Royal Astronomical Society*, 414, 2458
- van de Voort F., Springel V., Mandelker N., van den Bosch F. C., Pakmor R., 2019, *Monthly Notices of the Royal Astronomical Society*, 482, L85