Infall and accretion of substructure in hydrodynamical simulations of galaxy clusters

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy, April 2022.

Roan Haggar

Supervisors:	Prof. Meghan Gray	
	Prof. Frazer Pearce	
Examiners:	Prof. Justin Read	*(University of Surrey)*
	Dr. Emma Chapman	*(University of Nottingham)*
Submitted:	28th April 2022	
Examined:	9th June 2022	
Final version:	24th August 2022	

Contents

List of Tables

List of Figures

Abstract

The evolution of galaxies is heavily influenced by the environment in which they are found. Dense regions of the Universe can quench a galaxy's star formation by removing its gas, and transform a galaxy's morphology from being disk-dominated to bulge-dominated. Galaxy clusters represent the most extreme example of this: compared to the cosmic field, these highly dense environments contain a much greater fraction of red, elliptical galaxies. However, a subtlety of this is that a galaxy's evolution is not just driven by the environment in which it is currently found, but is also influenced by the environments through which it has previously passed. For example, a galaxy in the outskirts of a cluster may also have been quenched during pre-processing, meaning that it was accreted onto the cluster through a cosmic filament or as a member of a group, both of which can quench star formation. Alternatively, it could be a backsplash galaxy: one that has passed through the dense cluster environment in the past, but subsequently left and now resides in the cluster outskirts.

Pre-processing, backsplash galaxies, and the direct impact of a cluster environment are difficult to disentangle from each other, because observations cannot provide us with the full histories of galaxies. This makes observational studies of galaxy environment difficult, particularly nearby to clusters. For instance, it is not clear how common are backsplash galaxies around clusters, how their frequency varies between clusters, or whether we can identify which galaxies are indeed backsplash. It is also not fully understood what happens to galaxy groups nearby to clusters, and how a cluster can influence the impact of a group on its constituent galaxies. These questions are challenging to answer with observational data, but can be approached using cosmological simulations that complement the available observations.

In this thesis, we use data from The Three Hundred project, a suite of hydrodynamical simulations of large galaxy clusters, to study the environmental histories

of galaxies in and around clusters. We begin by establishing that these simulations are fit-for-purpose, by comparing them to equivalent dark matter-only simulations. We find that, compared to the hydrodynamical runs, our dark matter-only simulations underestimate the number density of galaxies in the central regions of both groups and clusters, for which we discuss several potential causes. This indicates that hydrodynamical simulations are necessary in studying cluster substructure, as the evolution of galaxy groups will be different in dark matter-only simulations.

Having established this, we then use these hydrodynamical simulations to examine how galaxy groups evolve as they approach, enter, and pass through a cluster. These galaxy groups become gravitationally unbound very quickly, losing most of their member galaxies less than 1 Gyr after entering a cluster. In fact, the overwhelming majority of groups do not survive a full passage through a cluster, meaning that any groups nearby to a cluster are almost certainly on their first infall towards the cluster centre.

We then investigate backsplash galaxies and find that, on average, over half of all galaxies between R_{200} and $2R_{200}$ from their host at $z = 0$ are backsplash galaxies. However, this fraction depends on the dynamical state of a cluster; dynamically relaxed clusters, which are isolated and accreting new material slowly, have a far greater fraction of backsplash galaxies in their outskirts. This backsplash population is mostly developed in the last few Gyr, and is dependent on the recent dynamical history of a cluster.

This work uses simulations to shed light on the different processes that galaxies can experience during their accretion onto a galaxy cluster. More importantly though, the findings from these simulations can be applied to real, observational studies. The dynamical state of a cluster is a measurable property, and so can be used to infer how 'contaminated' the population of infalling galaxies in an observed cluster's outskirts will be by backsplash galaxies. Furthermore, galaxy groups observed nearby to a cluster are on their first infall, and so will contain very few backsplash galaxies. Any galaxies that are members of a group inside of a cluster will have experienced the central region of a group, but have likely only joined the cluster very recently. Simulations will be a valuable tool to complement upcoming surveys like the WEAVE Wide-Field Cluster Survey and *Euclid* – due to begin in late 2022 and 2023 respectively – and will allow us to more deeply interpret this observational data, and infer the environmental histories of galaxies.

Acknowledgements

Firstly and most importantly, I would like to thank my supervisors, Meghan, for always being a constant source of ideas, suggestions, support and encouragement, and Frazer, for your clear, frank thoughts and opinions on everything, astronomy-related or otherwise. I've been incredibly fortunate to have two supervisors who have been so enthusiastic when my work has gone well, and supportive when things have gone not so well. There are also a wealth of others who have provided me with huge amounts of insight and assistance with my work, including (but not limited to) Ulli Kuchner, Alexander Knebe and Weiguang Cui. I'm also very grateful to Phil Parry for his tech support over the past few years, you're truly the engine that keeps everything in CAPT moving.

I'd like to thank my office mates in B112, for providing a constant distraction from work. A special thanks to Finlay, Charutha, Amy, CB and Matt, for providing distractions in the form of cryptic crosswords (CC?), treasure hunts, and a whole load of other nonsense. Without you, my PhD would have been much more productive, but much less fun.

Thanks to all the Inflativersers, including Jake, Rachel, Tom, Amy, CB, Kellie, Lizzie T and Jimi, for helping me to learn that science is much more enjoyable when you get to waffle on about it to other people. It's also been a pleasure getting to know every one of the other coffee-drinkers, wall-climbers, movie-watchers and pub-frequenters from CAPT – too many people to name, but a shout-out to my fellow fourth years (Leo and Sukhi love you guys). A special acknowledgement also to my housemates Mick and Jacob, for filling the past three years with air hockey, Dance Dance Revolution, darts, board games, The Circle™, vegan cat food, whisky sours, Bo'm Gear, spooky spiders, odd noises, and general silliness. My time in Nottingham has been further enriched by the other residents of Dunkirk, so thank you to Marc from Bargain Booze, Geoff from Geoff's, and particularly to Tony from The Boat for always being so welcoming on a Friday night.

Thank you also to my Mum and Dad for all their support, and to my brothers for their unwavering disinterest in my work.

Finally, I'd like to thank Liza, for making everything worthwhile.

Published Work

The papers that I have led or contributed to throughout this thesis are listed below, along with the co-authors of each piece of work:

i) **Haggar R.**, Gray M. E., Pearce F. R., Knebe A., Cui W., Mostoghiu R., Yepes G. (2020), *The Three Hundred Project: Backsplash galaxies in simulations of clusters*, MNRAS, 492, 6074

ii) Kuchner U., Aragón-Salamanca A., Pearce F. R., Gray M. E., Rost A., Mu C., Welker C., Cui W., **Haggar R.**, Laigle C., Knebe A., Kraljic K., Sarron F., Yepes G. (2020), *Mapping and characterisation of cosmic filaments in galaxy cluster outskirts: Strategies and forecasts for observations from simulations*, MNRAS, 494, 5473

iii) Knebe A., Gámez-Marín M., Pearce F. R., Cui W., Hoffmann K., de Petris M., Power C., **Haggar R.**, Mostoghiu R. (2020), *The Three Hundred Project: Shapes and radial alignment of satellite, infalling, and backsplash galaxies*, MNRAS, 495, 3002

iv) **Haggar R.**, Pearce F. R., Gray M. E., Knebe A., Yepes G. (2021), *The Three Hundred Project: Substructure in hydrodynamical and dark matter simulations of galaxy groups around clusters*, MNRAS, 502, 1191

v) Kuchner U., **Haggar R.**, Aragón-Salamanca A., Pearce F. R., Gray M. E., Rost A., Cui W., Knebe A., Yepes G. (2022), *An inventory of galaxies in cosmic filaments feeding galaxy clusters: galaxy groups, backsplash galaxies, and pristine galaxies*, MNRAS, 510, 581

vi) Contreras-Santos A., Knebe A., Pearce F., **Haggar R.**, Gray M., Cui W., Yepes G., De Petris M., De Luca F., Power C., Mostoghiu R., Nuza S., Hoeft

M. (2022), *The Three Hundred Project: galaxy clusters mergers and their impact on the stellar component of brightest cluster galaxies*, MNRAS, 511, 2897

vii) Cui W., Dave R, Knebe A., Rasia E., Gray M., Pearce F., Power C., Yepes G., Anbajagane D., Ceverino D., Contreras-Santos A., de Andres D., De Petris M., Ettori S., **Haggar R.**, Li Q., Wang Y., Yang X., Borgani S., Dolag K., Zu Y., Kuchner U., Cañas R., Ferragamo A., Gianfagna G. (2022), *The Three Hundred project: The Gizmo-Simba runs*, MNRAS, 514, 977

viii) **Haggar R.**, Kuchner U., Gray M. E., Pearce F. R., Knebe A., Yepes G., Cui W. (2022), *The Three Hundred project: Galaxy groups do not survive cluster infall*, MNRAS, submitted

Papers i), iv) and viii) are projects that I led; papers i) and iv) have both been published in journals, while paper viii) has been submitted and is currently under review.

Papers i), iv) and viii) form the basis of Chapter 5, Chapter 3 and Chapter 4 respectively. Material from each of these papers also makes up parts of Chapter 1, Chapter 2 and Chapter 6.

Chapter 1

Introduction

The properties of galaxies are strongly dependent on the cosmic environment in which they reside. A result of this is that the present-day environment of a galaxy is not the only factor at play, and the previous environments through which a galaxy has passed can also play a substantial role. Consequently, it is important to understand the full environmental histories of galaxies in order to understand their evolution.

In this chapter, we first introduce the ΛCDM cosmological model in Section 1.1. In Section 1.2 we describe how this model leads to the formation of galaxy clusters and large-scale structure, and in Section 1.3 we give an overview of galaxy evolution, and how it is affected by cosmic environment. Finally, in Section 1.4, we outline the aims and content of this thesis.

Contents

1.1 ΛCDM

The development of modern cosmology over the past century has resulted in a multitude of observational and theoretical leaps forward, which have transformed our understanding of the Universe. A major shift in our view of the Universe came in 1929, when Edwin Hubble found that galaxies further away from the Earth are receding at greater velocities (Hubble, 1929). This result provided evidence that the Universe is expanding, a result that had been derived two years earlier by Georges Lemaître (Lemaître, 1927), and has since become known as the 'Hubble–Lemaître law'. Whilst this idea was controversial at the time, it was bolstered in the 1960s by the discovery of the Cosmic Microwave Background (CMB), a high-redshift radio signal with a blackbody spectrum (Penzias & Wilson, 1965). This was a prediction of the expanding universe model: such a universe needed to have a beginning, which by this stage had acquired a nickname, the 'Big Bang'.

The Universe was small, hot and dense in its early years, and filled with fast-moving protons and electrons. When the Universe cooled, these particles were able to combine to form neutral atoms for the first time. This resulted in the decoupling of photons from matter, at a time known as the epoch of recombination. In the newly transparent Universe, photons were finally able to travel unimpeded – it is these photons that we observe today in the CMB. Modern estimates place this event approximately 380,000 years after the Big Bang, at a redshift of $z \sim 1000$ (Planck Collaboration et al., 2020).

One of the striking features of the CMB is its near-perfect isotropy. In all directions, the CMB spectrum is that of a blackbody, with a temperature of approximately 2.73 K. However, this temperature varies in different directions, albeit by by less than one part in 10^4 across the entire sky. Although they are incredibly small, these anisotropies are very important. After first being detected by the Cosmic Background Explorer (COBE) in 1992 (Smoot et al., 1992), the spatial resolution of CMB measurements was massively improved by the subsequent Wilkinson Microwave Anisotropy Probe (WMAP, Spergel et al., 2003) and *Planck* (Planck Collaboration et al., 2014) instruments. These incredibly detailed results allowed the average variation in CMB temperature over different angular distances to be characterised by a power spectrum, shown in Fig. 1.1.

The variations in CMB temperature are the result of small fluctuations in energy density in the early Universe: in fact, both the existence and magnitude of these variations had been predicted 20 years earlier as a consequence of the Big Bang

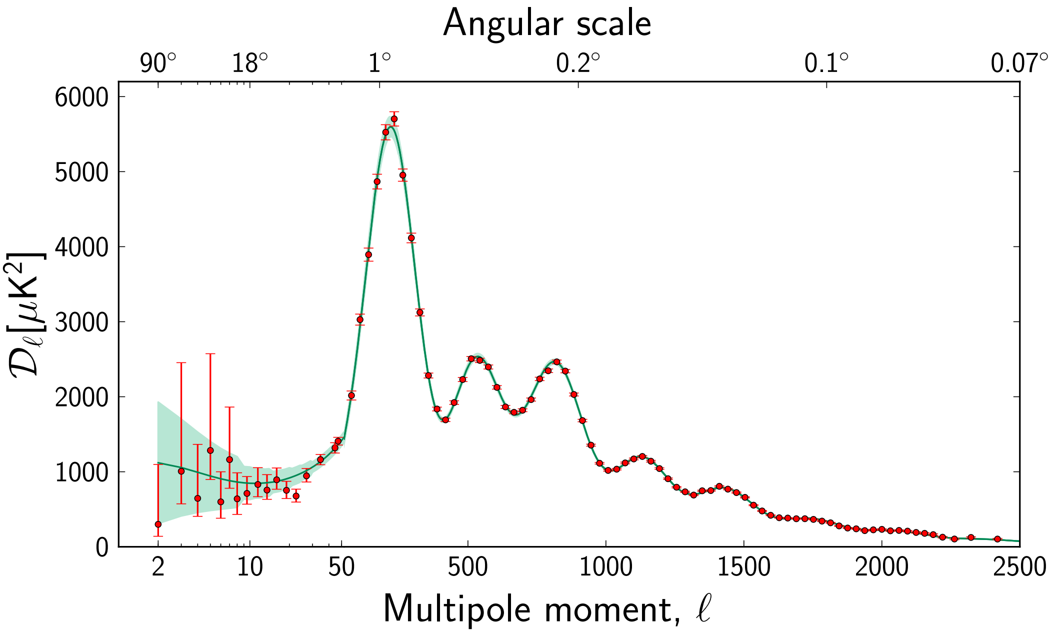

Figure 1.1: Angular power spectrum of the Cosmic Microwave Background temperature, measured by the *Planck* satellite. Figure taken from Planck Collaboration et al. (2014).

(Peebles & Yu, 1970; Zeldovich, 1972, for example). As well as providing qualitative evidence of the origins of our Universe, the magnitudes and positions of peaks in the CMB power spectrum can tell us more about the details of the our Universe, by allowing us to calculate numerous cosmological parameters such as the total density parameter, Ω_0, and the Hubble constant, H_0. For example, the ratio between the odd- and even-numbered peaks can be used to calculate Ω_b, the baryonic matter density parameter (Tegmark, 1996). CMB studies also tell us that the total matter density in our Universe, Ω_m, is about six times greater than the baryonic matter density, indicating that most of the matter in the Universe is not in the form of baryons (Planck Collaboration et al., 2016).

However, this is not a new idea. In 1933, Fritz Zwicky measured the velocity dispersion of galaxies in the Coma cluster, and used this to calculate the cluster's mass (Zwicky, 1933, 1937). On finding that this calculated mass was 400 times greater than the total observed mass of the constituent galaxies, he postulated that the cluster contained a huge amount of undetected mass, which he called *'dunkle Materie'* (dark matter).

Further demonstration of the existence of this dark matter came when Freeman (1970) and Rubin & Ford (1970) measured the rotation speeds of disk galaxies – that is, the orbital speeds of stars in a galaxy, as a function of the distance from the galactic centre. Assuming spherical symmetry (or circular symmetry for a disk galaxy), the orbital speed of a star, v_{r}, at a distance r from the galactic centre is given by

$$v_{\mathrm{r}} = \sqrt{\frac{GM_{<\mathrm{r}}}{r}} \, , \tag{1.1}$$

where G is the gravitational constant, and $M_{<\mathrm{r}}$ is the mass enclosed within the distance r. Different density profiles will therefore result in a different relationship between v_{r} and r. For example, the orbital speed around a compact or point mass (such as the planets around the Sun) scales as $v \propto r^{-1/2}$, as $M_{<\mathrm{r}}$ is a constant. Alternatively, if $M_{<\mathrm{r}} \propto r^3$, then v_{r} and r will be directly proportional. Based on the radial density of visible matter in galaxies, the orbital speeds of stars would be expected to decrease in the galactic outskirts. Instead, Freeman (1970) and Rubin & Ford (1970) found that v_{r} is constant in this region. While earlier work such as that of Babcock (1939) and Volders (1959) had found similar results, the ubiquity of flat rotation curves shown by Freeman (1970) and Rubin & Ford (1970) strongly indicated the presence of a large amount of unseen matter, impacting galactic dynamics.

Their independent detections showed that, as well as clusters, individual galaxies also contain large quantities of dark matter. This played a key role in constraining the distribution of dark matter within galaxies, and was later followed by gravitational lensing studies of clusters, which once again confirmed the large amount of diffuse dark matter within a cluster (Lynds & Petrosian, 1989; Tyson et al., 1990). Today, a multitude of similar techniques exist, which allow us to indirectly detect dark matter. Observational and theoretical studies have also investigated different classes of dark matter, including cold dark matter (CDM), warm dark matter (WDM) and self-interacting dark matter (SIDM) (Lovell et al., 2020; Stücker et al., 2022, for example). Evidence suggests that slow-moving cold dark matter is likely to make up the majority of structure in our Universe. However, to date no direct detections of dark matter have been made, and so for now the exact nature of this matter remains a mystery.

The final piece in this cosmological puzzle is a much more recent development. In the late 1990s, two independent studies (Riess et al., 1998; Perlmutter et al., 1999) used observations of Type Ia supernovae to constrain the energy composition of the Universe. Two important findings came from this: the matter density of the Universe is substantially less than the total energy density, and the acceleration of the Universe is expanding. This was the first evidence of a 'cosmological constant', also known as the vacuum energy or dark energy, and commonly denoted by Λ. Subsequent work, such as studies of the CMB power spectrum, have found that the vacuum energy density parameter, Ω_Λ, has a value of approximately 0.7 (Planck Collaboration et al., 2020). Incredibly, this indicates that roughly 70% of the energy content of our Universe is in the form of this dark energy, and yet the origins and details of this energy are still largely unexplained.

These ingredients – the CMB produced by a hot Big Bang, and the dark matter and dark energy contained within our Universe – are the basis of the Lambda cold dark matter (ΛCDM) model, commonly referred to as the 'standard model' of modern cosmology. This model forms the basis of many theoretical and observational studies, and can explain in detail many of the cosmological observations that we make about our Universe.

1.2 Cosmic structure formation and galaxy clusters

The Universe that we observe today is not homogeneous. On large scales it exhibits incredibly detailed structure, and the ability of the ΛCDM model to predict this structure is one of its greatest achievements.

1.2.1 Collapse and growth of dark matter haloes

The ΛCDM model shows that density fluctuations in the early Universe, as detected in the CMB, are amplified as the Universe ages. A region with an increased density will eventually collapse under gravity, while under-dense regions of the Universe will remain under-dense (Frenk et al., 1983; Frenk & White, 2012). This evolution is driven by dark matter, which dominates the relatively small amount of baryonic matter in the Universe. Eventually, dark matter collapses into small, dense, bound 'haloes' (White & Rees, 1978). These dark matter haloes are approximately spherically symmetric (although often with some ellipticity), and densest in the centre. Several empirical density profiles exist for these haloes, including the commonly used Navarro-Frenk-White (NFW) profile (Navarro et al., 1996). This describes the radial density, ρ, as a function of the distance from the halo centre, d, in terms of two free parameters, a characteristic density ρ_0, and the scale radius of the halo, R_s:

$$\rho(d) = \frac{\rho_0}{\frac{d}{R_\mathrm{s}}\left(1 + \frac{d}{R_\mathrm{s}}\right)^2} \,. \tag{1.2}$$

These dark matter haloes then grow, both through the gravitational accretion of diffuse dark matter, and through hierarchical structure formation – the merging of smaller haloes to form larger haloes. Subsequently, the gravitational pull of a halo attracts baryonic matter in the form of gas, which falls into the potential well of the halo. As this gas cools and condenses, it can collapse and form stars. Eventually, this leads to galaxies forming in the centres of these haloes, where the dark matter density is greatest (Springel et al., 2005), resulting in the high dark matter content of galaxies that is indicated by galactic rotation curves (Freeman, 1970; Rubin & Ford, 1970).

More exotic structure can also form in similar ways. For example, hierarchical structure formation means that dark matter haloes are not smooth, but instead contain 'subhaloes', smaller haloes that are imbedded within a larger halo. This substructure comes from the accretion of smaller haloes, which do not immediately merge and become fully incorporated into a larger halo, thus remaining in orbit around the host halo's centre (Taylor & Babul, 2004). Additionally, the collapse of dark matter can form structures of completely different shapes. Regions that fully collapse in two directions, but have had insufficient time to collapse in all three, will form long, extended structures, known as cosmic filaments (Codis et al., 2012). Other structures can be formed similarly, resulting in the rich array of large-scale structure in our Universe – a region that collapses in only one direction will form a cosmic wall (Bond et al., 1996), while gravitational collapse in all three directions forms nodes, which we discuss in the following section.

1.2.2 Galaxy groups, clusters, and the cosmic web

Cosmic filaments span the entire Universe, stretching tens of megaparsecs in length and marking out the boundaries of cosmic voids. These structures are ubiquitous in our Universe, and both simulations (detailed further in Section 2.1) and observations have detected vast networks of these filaments, known as the cosmic web (see Fig. 1.2). Filaments contain copious amounts of material; Ganeshaiah Veena et al. (2019) show that approximately two thirds of galaxies in the EAGLE hydrodynamical simulation are located within filaments.

Cosmic filaments meet each other at nodes, where we find the largest bound structures in the Universe, galaxy clusters. These are represented by huge dark matter haloes with masses up to $\sim 10^{15}\ M_{\odot}$. Galaxy clusters typically consist of a brightest cluster galaxy (BCG) located near to the centre, with subhaloes in the cluster containing hundreds of smaller satellite galaxies that make up the population of cluster members (Frenk & White, 2012).

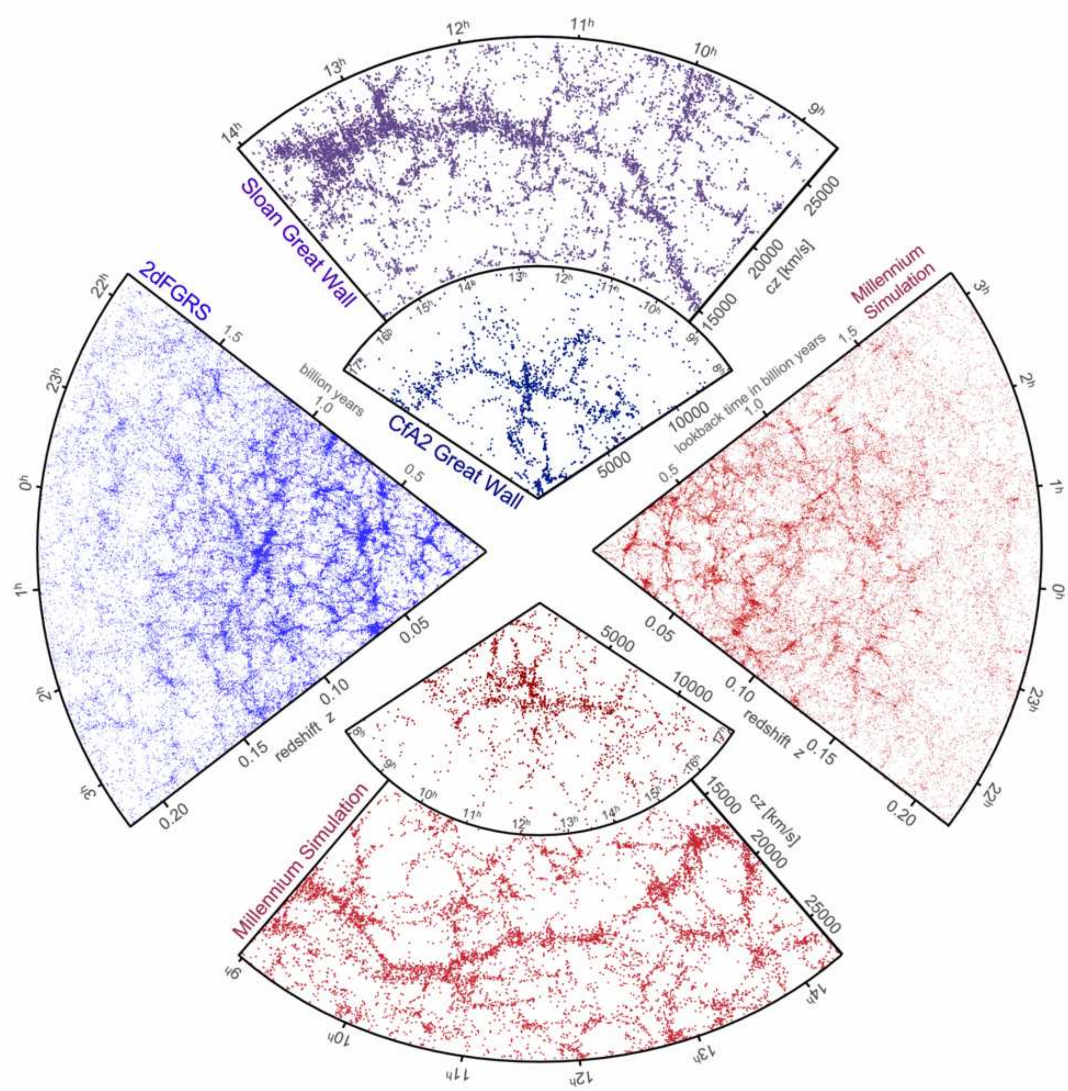

Figure 1.2: Composite figure showing observed distributions of galaxies from the Sloan Digital Sky Survey (SDSS, York et al., 2000) and the 2dF Galaxy Redshift Survey (2dFGRS, Colless et al., 2001), plus simulated distributions from the Millenium simulation (Springel et al., 2005). Similar distributions of galaxies can be seen in each panel, showing the large-scale structure that is traced by galaxies. Figure taken from Springel et al. (2006).

After their formation, further hierarchical growth allows clusters to continue increasing their size by accreting more dark matter haloes, and therefore more galaxies. These accreted structures can vary in size by several orders of magnitude: individual galaxies can enter a cluster, as well as galaxy groups. Groups are effectively low-mass clusters, with masses around 10^{13} $M_{\odot}$, and typically contain tens or hundreds of member galaxies (Benavides et al., 2020). The distinction between groups and clusters has historically been somewhat arbitrary. Observationally, both are often identified by finding overdensities of galaxies or gas (e.g. Tempel et al., 2016), and in simulations by finding large dark matter haloes or bound groups of galaxies (see Section 4.1 for some further discussion of this). Various ways of separating groups and clusters based on the number of members, temperature of intracluster/intragroup gas, and the dark matter halo mass have been suggested, but no widely-accepted, physically-motivated boundary separating groups and clusters exists. Despite this, some recent studies have attempted to find physical differences between the two: Paul et al. (2017) showed that different scaling relations apply to haloes with masses above and below $8 \times 10^{13} M_{\odot}$, and recommended this as an appropriate boundary.

A significant portion of this thesis (Chapter 4, and to a lesser extent Chapter 3), is dedicated to studying galaxy groups, and the role that these play in the assembly of clusters – we discuss the definition of groups that we adopt in Section 3.2.2 and Section 4.2. In extreme cases, clusters can grow rapidly by experiencing major cluster-cluster mergers involving thousands of galaxies (Moore et al., 1999; Contreras-Santos et al., 2022), but throughout this thesis we do not discuss these events in detail.

1.3 An overview of galaxy evolution

As galaxies form at the centres of dark matter haloes, the visible component of groups, clusters, and large-scale structure is dominated by galaxies.

Galaxies vary in terms of their size, shape and colour. One of the earliest, and still most well-known, classifications of galaxies is the Hubble sequence (Hubble, 1926), in which galaxies were split into two main morphological categories ('elliptical' and 'spiral', also known as early-type and late-type respectively). In the original Hubble sequence, lenticular galaxies were included as a subset of elliptical – they are now often seen as a separate class of galaxies, although usually still under the umbrella of 'early-type' (Liller, 1966). The spiral galaxies were also split into two further categories ('normal' and 'barred'), each of which was in turn split into subcategories based on properties such as the dominance of their central bulge. Other galaxies without regular structure were classed as 'irregular'.

The Hubble sequence has proven a robust means of classifying galaxies, with observations across a wide range of redshifts supporting the original sequence (Lee et al., 2013). However, in recent years some flaws have begun to emerge, as well as the need to focus on different properties. Although the Hubble sequence is defined based on galaxy morphology, we must also consider their sizes, colours, and star formation rates in order to understand these objects fully, and these properties are not included in the Hubble sequence. It is also now clear that these properties are not independent – for example, early-type galaxies are typically larger, redder, and have quenched star formation compared to late-type galaxies (see Conselice, 2014, for a detailed review of the relationships between galaxy properties). This has led to the common practice of separating galaxies based on their locations on a colour-magnitude diagram, as the 'blue cloud' and 'red sequence' form two distinct populations on these diagrams (e.g. Baldry et al., 2004).

Furthermore, it is now clear that galaxies are in fact incredibly diverse, and that this apparent bimodality does not tell the full story of galaxies' properties. Some galaxies have intermediate star formation rates, meaning that they are found between the blue cloud and red sequence, in the so-called 'green valley' (Gonçalves et al., 2012; Schawinski et al., 2014). Others, like post-starburst galaxies, are galaxies whose star formation has been quenched within the last ~ 1 Gyr, after having recently experienced a rapid burst of star formation (Zabludoff et al., 1996; Wild et al., 2020). Evidence suggests that the properties of these galaxies differ to other quenched galaxies; Sazonova et al. (2021) find that post-starburst galaxies are highly morphologically disturbed, indicating that they have recently experienced a major merger. Even more exotic classes of galaxies add to the complexity of this picture, such as active galaxies (Urry & Padovani, 1995), and jellyfish galaxies, whose gas has been stripped away into a long, extended tail (Cramer et al., 2019).

One of the greatest aims of extragalactic astronomy is to understand how galaxies form and evolve over the course of the Universe. Despite their slightly misleading names, it is now widely believed now that late-type galaxies evolve into early-type galaxies. This process is complex; Schawinski et al. (2014) use data from Galaxy Zoo (Lintott et al., 2008, 2011) to show that multiple different evolutionary pathways between late-type and early-type galaxies are required to explain the properties of galaxies in the present day. However, the evolution of galaxies can mostly be summarised into two main changes:

- Change in morphology: Galaxies change from disk-dominated to bulge-dominated.

Spiral arms are lost, and galaxies evolve towards a featureless elliptical shape.

- Quenching of star formation: Star formation stops, either due to removal of gas from the galaxy, or prevention of gas from collapsing to form stars. This eventually leads to a corresponding change in galaxy colour.

Because of the complex nature of galaxy evolution, there also exist other, more subtle evolutionary processes. For instance, the kinematics of galaxies also vary, giving us 'fast rotating' and 'slow rotating' early-type galaxies that can be studied observationally using integral-field spectroscopy (Cappellari, 2016). Nonetheless, morphological change and quenching are still the two most dramatic, and most visible, changes that a galaxy experiences throughout its lifetime.

1.3.1 Environmental effects

In the past 40 years it has become apparent that the processes that drive galaxy evolution do not occur at random, but are instead driven by the local environment of galaxies – their location within the cosmic web. An early study by Dressler (1980) revealed that cluster environments contain mostly early-type galaxies, whereas galaxies in field regions typically have late-type morphologies, with intermediate environments such as galaxy groups lying between these two. More recent studies have also confirmed that, compared to field galaxies, cluster galaxies have quenched star formation rates (Balogh et al., 1999; McNab et al., 2021) and lower gas fractions (Jaffé et al., 2015), across a large range of redshifts (van der Wel et al., 2007; Quadri et al., 2012). Fig. 1.3 shows how the fraction of red (quenched) galaxies depends on the mass and cosmic environment of the galaxies.

A wide range of mechanisms exist that are enhanced or suppressed in different environments, and so can explain the effect of environment on galaxies. Cosmological simulations show that galaxies in cluster environments experience enhanced ram pressure stripping: the removal of gas from a galaxy moving through a medium, caused by drag forces from that medium acting on the gas. This is due to the dense intracluster medium – the hot gas that fills the space between cluster galaxies – which can lead to infalling galaxies being almost entirely stripped of their halo gas. This can occur even in the outskirts of a cluster, resulting in the quenching of star formation in cluster galaxies (Zinger et al., 2018; Arthur et al., 2019). Observational evidence for ram pressure stripping includes cluster galaxies whose molecular gas reservoirs have been disturbed, meaning the gas is distributed asymmetrically with respect to the stellar component of the galaxy (Zabel et al., 2019). In extreme cases, this leads to the

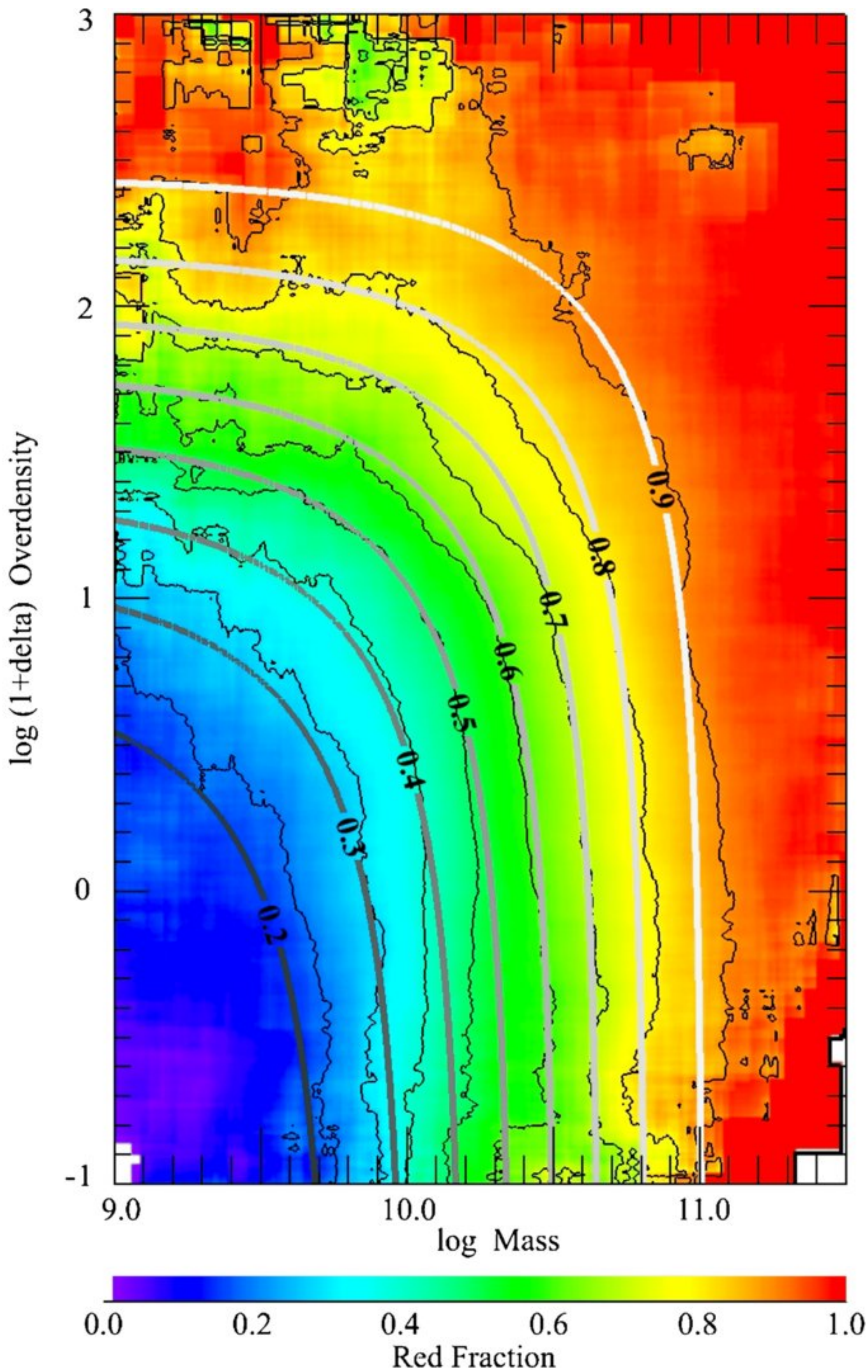

Figure 1.3: Fraction of red galaxies in SDSS, as a function of the galaxy stellar mass and local overdensity, a proxy for cosmic environment. A greater fraction of high-mass galaxies, and a greater fraction of those in dense environments, are red. Figure taken from Peng et al. (2010).

previously-mentioned jellyfish galaxies – Cramer et al. (2019) use Hα observations to map the ionised gas around a Coma cluster galaxy, and find a tail of ram pressure stripped gas several times longer than the galactic diameter. Ram pressure stripping is a rapid quenching process, but other slow mechanisms can similarly quench star formation. Galaxy starvation is one such example, in which any extended reservoir of gas around a galaxy is stripped away. This leaves the galaxy to slowly exhaust its supplies of gas over the next several Gyr, eventually halting the formation of new stars (Larson et al., 1980; Maier et al., 2016, 2019). Other processes, like shocks, can quench a galaxy by heating its gas and reducing the efficiency of star formation, without actually removing the gas from the galaxy (Alatalo et al., 2016).

As well as these processes that quench star formation, mechanisms that impact galaxy morphology are also affected by environment. For instance, although galaxy mergers are not common in clusters, galaxy groups have been shown to enhance the rate of mergers, due to their combination of a high galaxy number density, and low velocity dispersion[1] (Jian et al., 2012). Mergers drastically impact the evolution of galaxies by disturbing their morphology, but can also remove gas from a galaxy by triggering outflows and AGN. Galaxy harassment (Moore et al., 1996a) is another gravitational process, which is enhanced within clusters. This is driven by high-speed close encounters between galaxies, which can disturb the morphology of the galaxies. Boselli & Gavazzi (2006) provide a more extensive review of environment-dependent galaxy evolution.

1.3.2 Environmental history and pre-processing

A consequence of this connection between galaxies and their environments is that a galaxy's evolution is not just impacted by the environment in which it is currently found – it can also be affected by the environments through which it has previously passed. In particular, galaxies that are currently observed within a cluster can have experienced a range of different environments before entering the cluster. For instance, galaxies can fall into a cluster by passing along a relatively high-density cosmic filament, which can quench star formation, similarly to clusters (Kraljic et al., 2018; Laigle et al., 2018). In fact, compared to clusters, these filaments are likely to play a role in the evolution of many more galaxies. This is because, while clusters are rare objects, a very substantial fraction of galaxies in our Universe reside in cosmic

[1] The relative velocities of merging galaxies are usually < 500 km s^{-1} (Lotz et al., 2008; An et al., 2019), but dark matter haloes with masses greater than 10^{14} $M_{\odot}$ typically have velocity dispersions greater than 500 km s^{-1} (McClintock et al., 2019; Wetzell et al., 2021). Consequently, mergers are more likely in group-sized haloes, with masses less than 10^{14} $M_{\odot}$.

filaments (Ganeshaiah Veena et al., 2019). Despite this, the effects of filaments on galaxy evolution are actually far less understood than the effects of clusters, and so this remains an open field of research with many outstanding questions, although these questions are not addressed in detail within this thesis.

Regardless, it is still clear that galaxies entering a cluster through a filament are likely to experience different environmental effects to those being accreted from the field. Furthermore, galaxies that are members of an infalling galaxy group will have different histories to those that are infalling as isolated objects (White et al., 2010; Cybulski et al., 2014; Jaffé et al., 2016). The means by which a galaxy is affected before it enters a cluster are known collectively as 'pre-processing'.

Backsplash galaxies are a related concept to pre-processing. These are galaxies that have previously been accreted by a cluster at some point in their past, meaning that they have passed within a distance of R_{200} from the cluster's centre (where R_{200} is the radius within which the mean density of a cluster is equal to 200 times the critical density of the Universe). However, they have since left the cluster (passed outside of R_{200}), meaning that they are currently outside of the cluster, in the region that we hereafter refer to as the 'cluster outskirts'. These galaxies will therefore have experienced the effects of a cluster environment in their past, but are observed in the same locations as infalling galaxies, that are approaching a cluster for the first time (Gill et al., 2005; Haggar et al., 2020).

Pre-processing and backsplash galaxies are both problematic when trying to observationally study the impact of environment on galaxy evolution, as they act as contaminants, making it difficult to disentangle different effects and find causal relationships. Comparing the properties of cluster galaxies and field galaxies, differences between the two may have been caused by the cluster, or by pre-processing during the infall of galaxies towards the cluster, or by a combination of the two. Similarly, samples of infalling galaxies collected by surveys of cluster outskirts (which may have been pre-processed) are likely to be contaminated by the backsplash population, making it difficult to disentangle the effects of pre-processing and the effects of a cluster.

1.4 Aims and outline of this thesis

The numerous different evolutionary pathways of galaxies described in the previous section are fairly well-understood. However, an open problem is understanding the relative contribution of all of these processes to galaxy evolution, and in which cases different environmental processes dominate. For example, although much work has gone into understanding galaxy evolution within groups, it is still not clear exactly how a galaxy group is affected by approaching a cluster, and to what extent a cluster can dominate the effects of a group, and potentially disrupt its structure. Additionally, although previous work (Hester, 2006) has shown that processes such as ram pressure stripping are enhanced in group centres, little work has been carried out investigating how other processes are affected by a galaxy's position within a group, and how these processes are impacted by a group nearing a cluster environment. In this thesis, we aim to understand how a galaxy's position in the phase space of its host group affects its evolution, and how this phase space, and the more general group structure, is impacted by a group entering a large galaxy cluster.

We also aim to explore the connection between groups and other objects like backsplash galaxies – for example, whether groups preferentially contain many or few backsplash galaxies. On a more fundamental level, it is still not even clear what fraction of galaxies around a cluster are backsplash galaxies, and whether this fraction is universal, or differs between clusters. Again, some work has been carried out in this area (e.g. Gill et al., 2005), but mostly using dark matter-only simulations. Another aim of this thesis is to determine whether these simulations are useful in studying this cluster substructure, or whether full-physics hydrodynamical simulations are necessary.

Cosmological simulations are an excellent tool in exploring these questions, as they allow us to examine the full history of a cluster and determine which environments galaxies have experienced in their past. Indeed, we use cosmological simulations throughout this thesis to address the questions listed above: in Chapter 2 we first give an overview of simulations, before describing the simulations that are used throughout this thesis.

Questions on the environmental histories of galaxies are much more challenging to answer using observational data. Because of this, the overarching motivation behind this thesis is to determine how simulations can be used to interpret observational data. For instance, we can use simulations to find out which properties of galaxies or larger structures should be measured, in order to learn more about the environments through which galaxies have passed. In this thesis, we aim to understand which observable properties of a galaxy cluster can tell us about the contribution of backsplash galaxies to this cluster, and about the role that nearby group substructures play in the assembly of the cluster population.

After we introduce cosmological simulations in Chapter 2, we make some direct comparisons between two otherwise identical sets of hydrodynamical and dark matter-only simulations in Chapter 3, to find any systematic differences between the two that would impact this research. We settle on using the hydrodynamical simulations, which we use in Chapter 4 to study the changing distribution of galaxies in group-centric phase space, for infalling galaxy groups. In Chapter 5 we look at backsplash galaxies, and quantify the fraction of these around clusters, and how this depends on measurable cluster properties. Finally, we conclude and summarise our findings in Chapter 6.

Chapter 2

Cosmological simulations

Constraining the environmental histories of galaxies is not easy with observations, as these only provide a single snapshot of data, from which we are required to infer the history of these objects. Simulations, on the other hand, allow us to probe the full histories of galaxies, and learn which observable quantities can be used to constrain their cosmic environments.

In this chapter, we first describe some different simulation techniques in Section 2.1, and discuss how cosmological simulations are used as a tool in this field. We then introduce the simulation data that is used throughout this thesis in Section 2.2.

Contents

2.1 Simulations of cosmic structure

Small density perturbations can be modelled and described analytically. Such perturbations are described as 'linear', and satisfy the equation

$$|\delta| = \left| \frac{\rho - \bar{\rho}_{\mathrm{m}}}{\bar{\rho}_{\mathrm{m}}} \right| \ll 1 \, , \tag{2.1}$$

where ρ is the density of a perturbation, and $\bar{\rho}_{\mathrm{m}}$ is the mean matter density in that region of the Universe (Sahni & Coles, 1995). However, when the density contrast, δ, is no longer far less than one, analytic solutions can no longer be used to describe the evolution of these perturbations.

It is this that has motivated the development of cosmological simulations, which use numerical techniques to model the evolution of cosmic structure, based on some initial conditions. The earliest cosmological simulations were simple, N-body simulations. These involved modelling a cosmological volume ('box') by populating it with particles[1] of a given mass, whose dynamics are defined by the statistical properties of the Universe at high redshift, for instance, by a power spectrum describing the density perturbations on different length scales (Fidler et al., 2017). The particles are then typically evolved under Newtonian gravity – even in cluster-scale simulations, the speeds and distances involved are sufficiently small that full general relativistic equations of motion are not necessary (Dehnen & Read, 2011). Some other necessary approximations are usually included, such as finite time steps and gravitational softening (see also Section 2.1.4 and Barnes, 2012).

Some of the earliest N-body simulations, such as the work of Press & Schechter (1974) and White (1976), involved simulating hundreds of gravitational masses. By including no other physics beyond gravity, N-body simulations are, in effect, simulations of a box of dark matter. Because the matter density of our Universe is dominated by dark matter, the gravitational collapse of cosmic structure is dominated by the effects of dark matter, making these simulations a good approximation for our Universe (Jenkins et al., 1998; Borgani & Kravtsov, 2011). Consequently, the usage of N-body simulations, or dark matter-only simulations, has continued into the 21st century. The size of these simulations has continued to increase exponentially: 2005 saw the release of Millennium, the first 10^{10} particle simulation (Springel et al., 2005), which was followed just 12 years later by the 10^{12} particle Euclid Flagship

[1]In cosmological simulations, the 'particles' are not physical particles, but represent discretised packages of matter. The mass of these particles is set by the simulation resolution: higher-resolution simulations (i.e. lower particle mass) will be more physically accurate, but more computationally expensive.

simulation (Potter et al., 2017). Modern cosmological simulations can cover volumes of several Gpc^3, although smaller simulation volumes such as the $\sim 50\ \mathrm{Mpc}^3$ TNG50 simulation (Nelson et al., 2019) can give a detailed view of cosmic structure by allowing a higher resolution to be used.

2.1.1 Hydrodynamical simulations

Simulations that include baryons, as well as dark matter, have become increasingly popular in recent years. These allow the gravity-dominated structure of the Universe to be modelled, as well as the visible component of the Universe that is dominated by baryonic physics. These simulations have numerous benefits: they allow us to include the effects of baryons on cosmic structure, make predictions for future observational work, and model gas, stars and galaxies to study how they evolve. Unfortunately, these benefits come at the price of efficiency, as these simulations are far more computationally expensive than dark matter-only simulations. Because of this, many different computational methods are used, in order to maximise accuracy while minimising computational cost.

Perhaps the most physically motivated approach to modelling baryons and galaxies is with hydrodynamical simulations. As well as dark matter particles, particles of gas are included in these simulations. These only interact with dark matter particles via gravity, but experience hydrodynamical interactions with each other. Consequently, the gas particles are collisional, meaning the gas has non-zero viscosity. Collapsing clouds of this gas can therefore lead to features such as shocks, and so radiative gas cooling is usually included in these simulations to allow shock-heated gas to evolve. Once the gas has cooled, star formation is modelled by producing star particles. These new particles are often generated stochastically, with a small amount of the mass of a gas particle being converted into a new star particle (representing more than one physical star, due to the limited simulation resolution) (Somerville & Davé, 2015; Wechsler & Tinker, 2018).

Methods such as smoothed-particle hydrodynamics (SPH) are commonly employed to better approximate the physical properties of gas (Wechsler & Tinker, 2018). SPH was first developed by Gingold & Monaghan (1977) and Lucy (1977), and allows fluids (such as gas) to be accurately modelled by discrete particles. In contrast to fixed-mesh 'Eulerian' methods (which calculate properties of the fluid at fixed locations across a simulation box), SPH is a 'Lagrangian' method, meaning that the coordinate system moves with the fluid itself. Each fluid particle is described by hydrodynamical properties such as its temperature and density, plus its position and velocity. Crucially though, in SPH the particle properties are interpolated between particles, meaning that the fluid properties in each region of the simulation are smooth, continuous, and dependent on the properties of multiple nearby particles. The power of SPH can be improved further by allowing this smoothing length to vary, such that it is shorter in denser regions of the simulation. This allows high-density regions to be modelled with high spatial precision, whilst also efficiently modelling large, low-density regions (see Rosswog, 2009; Price, 2012, for further details).

The gas/dark matter particle masses in cosmological simulations are large, typically between $10^6 - 10^9\ M_\odot$ (Vogelsberger et al., 2020), and so there are many important, complex processes in galaxy evolution that occur far below the resolution limit of these simulations. Such processes are referred to as 'subgrid' processes. Star formation relies on subgrid physics to determine the quantity and type of stars that are produced, as do black hole formation, magnetic fields, and feedback from AGN and supernovae. However, these subgrid processes are complex, and much of the physics governing them is not fully understood, making them challenging to implement. Analytic 'recipes', based off either theory or empirical data, are used to statistically describe these processes. For example, subgrid AGN feedback can involve heating the gas nearby to an AGN, and placing hot bubbles around the AGN to represent the effects of a collimated jet – the strength of these two processes depends on the accretion rate of the central black hole (Somerville & Davé, 2015).

2.1.2 Efficient modelling techniques

It is desirable to have simulations with high resolution, as they will model smaller structures like galaxies more accurately. It is also desirable to have large volumes, to provide a statistically significant sample of large objects like clusters and filaments. However, due to the computational cost of hydrodynamical models, large, high-resolution simulations can be prohibitively expensive. One solution to deal with this is 'zoom simulations'. These start with a large, low-resolution box that is simulated to $z = 0$. Then, the particles making up an object of interest (often a single halo) can be identified and traced back to their initial conditions, where they are split into a number of smaller particles. The whole simulation is then run again – this allows a particular object to be chosen and simulated at high-resolution, whilst also including the cosmic environment of this object, which is simulated at a lower resolution (Somerville & Davé, 2015). Zoom simulations are ideal for studies of structures like galaxies and clusters: while these simulations only allow pre-selected objects to be studied, they produce high-resolution models of these objects, at a fraction of the cost of a large cosmological simulation.

One example of zoom simulations is the FIRE simulations (Hopkins et al., 2014). This suite of hydrodynamical simulations consists of haloes from the size of dwarf galaxies to small galaxy groups, each simulated in very high-resolution with the surrounding structure also included. For example, the group halo in this suite, which has a mass of $10^{13}\ h^{-1}\ M_\odot$, has a dark matter particle mass of $1.6 \times 10^6\ h^{-1}\ M_\odot$, about 10 times better than the resolution of the $(60\ h^{-1}\ \mathrm{Mpc})^3$ box in which it was initially simulated (Kim et al., 2014). The NIHAO simulations are another suite of zoom simulations, over a similar mass range (Wang et al., 2015). On larger scales, several different zoom simulations have been used to study samples of tens or hundreds of galaxy clusters, by extracting them from large cosmological volumes and resimulating them. Examples include the Hydrangea simulations (Bahé et al., 2017), the MACSIS clusters (Barnes et al., 2017a) and The Three Hundred project (Cui et al., 2018), which are the simulations used throughout this thesis.

Instead of hydrodynamical simulations, various other approaches exist that are less physically motivated, but attempt to deal with baryonic physics by employing empirical models to save on computing power. These typically utilise dark matter-only simulations to simulate the cosmic structure found in hydrodynamical simulations, that are then post-processed to retroactively include the baryonic material. Semi-analytic models are an example of this. In these models, baryonic gas is added to numerically simulated dark matter haloes, and the gas is subsequently evolved using models of gas cooling and star formation to reproduce the evolution of galaxies (Benson et al., 2001; Baugh, 2006; Croton et al., 2016). Numerous semi-analytic codes now exist, which can reproduce the results from both hydrodynamical simulations and observations with impressive accuracy. For example, Fig. 2.1 (from Somerville & Davé, 2015) shows the galaxy stellar mass function from a selection of semi-analytic models, hydrodynamical simulations and observations, and demonstrates the agreement between these, particularly at low redshifts. Semi-analytic models do have limitations though, and care must be taken when studying their outputs – for instance, the model parameters are usually calibrated using hydrodynamical simulations or astronomical observations, and so it is not meaningful if a model simply returns the same calibration data that was used as an input.

Other empirical approaches include halo occupation models, in which haloes within a dark matter-only simulation are populated with galaxies statistically, based on a probability distribution that matches galaxies of given properties to corresponding haloes (Guo et al., 2016; Wechsler & Tinker, 2018). We discuss the advantages and

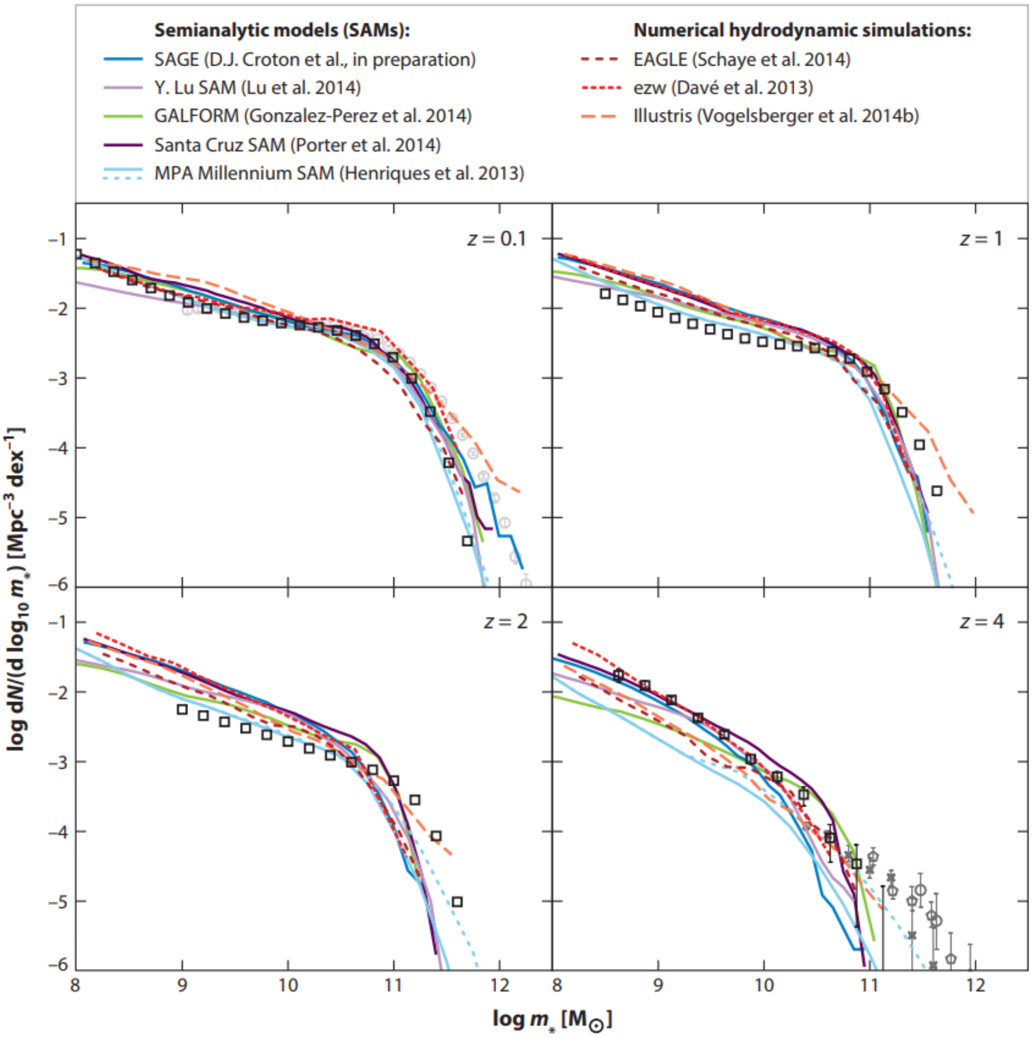

Figure 2.1: Galaxy stellar mass function, across four redshifts, calculated from five semi-analytic models, three hydrodynamical simulations, and eight observational studies. Semi-analytic models (Henriques et al., 2013; Gonzalez-Perez et al., 2014; Lu et al., 2014; Porter et al., 2014; Croton et al., 2016) are shown by solid lines. Hydrodynamical simulations (Davé et al., 2013; Vogelsberger et al., 2014; Schaye et al., 2015) are shown by dotted/dashed lines. Observational results (Baldry et al., 2008; Marchesini et al., 2010; Caputi et al., 2011; Bernardi et al., 2013; Moustakas et al., 2013; Muzzin et al., 2013; Duncan et al., 2014; Tomczak et al., 2014) are shown by individual shapes (squares, circles, etc.). All of the models show relatively good agreement, particularly at $z = 0$, although there are some systematic differences. Figure taken from Somerville & Davé (2015) – a full description of this figure is included in that work.

limitations of these empirical models, particularly semi-analytic models, in greater detail in Section 3.1.

2.1.3 Halo finders and merger trees

The output from hydrodynamical simulations is saved in discrete snapshots, storing the positions, velocities and masses of each particle at pre-defined time stamps. In order to extract useful properties of dark matter haloes and galaxies, this raw simulation data must then be post-processed.

An important stage of this involves using halo finders. These group particles together, in order to determine which particles belong to different dark matter haloes or subhaloes. Additional steps include associating particles of stars and gas with dark matter haloes, and calculating properties of the haloes such as their position, mass, and concentration. Most galaxy halo finders are based off of one of two principles. Density peak locators search for local peaks in a density field and denote these as halo centres. A spherical region is then expanded outwards from each point, collecting particles until the local density drops below a given threshold. Alternatively, an algorithm such as a friends-of-friends algorithm (Davis et al., 1985) can find groups of particles that are close together, either in physical space or phase space, and denote these as a bound halo (see Knebe et al., 2011b, for a far more extensive overview). Halo finders are imperfect, and can struggle to accurately pick out haloes in some situations, such as when a subhalo is close to the centre of a host (Muldrew et al., 2011), or when two similar sized haloes merge (Behroozi et al., 2015). However, their usage allows us to automatically pick out and study galaxies, groups and clusters from simulations.

Halo finders are used on static, single-snapshot data, but it is also useful to link haloes together in successive snapshots – by identifying two haloes in successive snapshots that correspond to the same physical object, we can study how an object evolves over time. Tree-builders are a useful tool for this, as their role is to track haloes and connect halo catalogues between different snapshots, typically by using a merit function to determine how many particles are shared by haloes in successive snapshots. This allows us to follow the position and properties of a galaxy moving through a simulation.

Using these methods, haloes in a simulation are assigned a progenitor (the same halo identified in a previous snapshot), and descendent (the same halo, in a subsequent snapshot). However, tree-builders also can track which haloes merge hierarchically to form larger haloes, and so keep track of galaxies at $z = 0$ that have multiple different progenitors (Srisawat et al., 2013). This allows us to track galaxy mergers, when smaller haloes are absorbed into the 'main branch' of a halo's merger tree.

2.1.4 Limitations of simulations

The development of cosmological simulations over the last few decades has allowed entirely new areas of astronomy and cosmology to be opened up. Unfortunately, the nature of numerical simulations means that they are not perfect. Limits on computational power mean that continuous (or almost continuous) quantities must be discretised in simulations. The masses of particles used in large cosmological simulations – typically millions of solar masses – mean that structures less massive than this, such as molecular clouds and globular clusters, cannot be resolved. Simulations also use a finite time step to evolve their particles, which in turn requires gravitational softening to be implemented. For example, Plummer softening (Plummer, 1911; Barnes, 2012) involves adding an extra term, ϵ, to the gravitational potential, Φ, such that the potential at a distance r scales as:

$$\Phi(r) \propto \frac{-1}{\sqrt{r^2 + \epsilon^2}} \,. \tag{2.2}$$

This places an upper limit on the force that can be applied to a particle, preventing very strong, short-range interactions between two very nearby particles from occurring. A consequence of this is that the interactions between particles become non-physical on length scales comparable to the softening length, which can inhibit the formation of structure on this scale. For instance, in a simulation with $\epsilon = 5$ kpc, thin galactic disks would be unable to form, as these typically have a thickness of ~ 1 kpc (see Mostoghiu et al., 2021b, for some further discussion on this).

All cosmological simulations are affected by these unavoidable limitations, but the variation between different simulations is also a potential problem. The results from cosmological simulations depend on the physics that is used – this includes the cosmological parameters, and the calibration and implementation of hydrodynamical effects (although this most strongly affects the properties of galaxies, which we do not directly study in detail during this thesis). Additionally, hydrodynamical simulations rely on the post-processing techniques described in Section 2.1.3, above. While previous studies have shown that most halo finding codes perform with similar effectiveness (Knebe et al., 2011b; Onions et al., 2012, for example), the properties of haloes in denser regions of the Universe, and especially inside of other haloes, are not always well-constrained. This is particularly relevant to Chapter 3 of this thesis, in which we discuss the impact of halo finders on the distributions of galaxies predicted by cosmological simulations.

2.1.5 Combining simulations with observations

Despite these limitations, simulations are an incredibly useful tool in astronomy, allowing us to build a theoretical analogue to the real, observed Universe. However, the true power of simulations comes by combining them with observations, and it is this combination that motivates the work in this thesis. As with simulations, the nature of extragalactic observational astronomy means that it also has several limitations. Cosmic structure evolves over $\sim$Gyr timescales, meaning that we cannot directly observe how structure changes. Furthermore, in the context of galaxy clusters we can generally only measure the positions of galaxies in two dimensions, and their speeds in one dimension (the line-of-sight).

By utilising simulations, we can interpret observations more deeply, and make inferences from the observations that would not be possible without the additional information provided by simulations. For example, mock observations can be produced from simulation data, allowing us to find observable quantities that correlate with non-observable quantities. Then, by making observations, we can imply properties of galaxies, clusters and large-scale structure that cannot be measured directly. This process is two-way, as observations are also invaluable when developing better simulations, and empirical results can be used to refine the physics, calibration, and subgrid processes within simulations.

In this thesis, we study the build-up of clusters and the environmental histories of galaxies, using simulations, but from an observational perspective. Previous studies have already begun to look at this problem – Kuchner et al. (2022) use the same simulations used throughout this thesis (detailed in Section 2.2) to take a thorough inventory of the fraction of group galaxies, backsplash galaxies and filament galaxies nearby to clusters. We will build on this by determining how future observational studies will be able to constrain the environmental histories of galaxies, with a particular focus on galaxies and galaxy groups that are being accreted by a cluster.

2.2 The Three Hundred project

This thesis utilises simulation data from The Three Hundred project[2] (Cui et al., 2018). In Section 2.2.1 we give a technical outline of these simulations, including the

[2] https://www.the300-project.org

motivation for using these simulations specifically. We then give an overview of some of the analysis conducted throughout this work in Section 2.2.2 and Section 2.2.3.

2.2.1 Simulation data

The Three Hundred dataset is a suite of 324 hydrodynamical zoom simulations of large galaxy clusters. These simulations were produced from the MDPL2 MultiDark simulation (Klypin et al., 2016)[3]. MDPL2 is a large, dark matter-only, cosmological simulation, consisting of a box with sides of comoving length 1 h^{-1} Gpc, containing 3840^3 particles each of mass 1.5×10^9 $h^{-1}M_\odot$, and using *Planck* cosmology ($\Omega_\mathrm{M} = 0.307$, $\Omega_\mathrm{B} = 0.048$, $\Omega_\Lambda = 0.693$, $h = 0.678$, $\sigma_8 = 0.823$, $n_\mathrm{s} = 0.96$) (Planck Collaboration et al., 2016)[4].

To generate The Three Hundred suite, the 324 most massive clusters at $z = 0$ were chosen from MDPL2. For each cluster, the particles within a spherical region of radius 15 h^{-1} Mpc ($\sim 10R_{200}$) from the cluster centre at $z = 0$ were traced back to their initial positions. These dark matter particles were split into dark matter and gas particles, of masses $m_\mathrm{DM} = 1.27 \times 10^9$ $h^{-1}M_\odot$ and $m_\mathrm{gas} = 2.36 \times 10^8$ $h^{-1}M_\odot$ respectively, set by the baryonic matter fraction of the Universe. Lower-resolution particles were used beyond 15 h^{-1} Mpc, to replicate any large-scale tidal effects on the cluster at a lower computational cost. Each cluster was then resimulated from its initial conditions with full baryonic physics. The box size and cosmology used for each of the cluster simulations in The Three Hundred are the same as those used in MDPL2.

These simulations have been carried out multiple times, using several different physics models and simulation codes – specifically, using the GadgetX, Gadget-MUSIC (Sembolini et al., 2013), and Gizmo-Simba (Cui et al., 2022) hydrodynamical codes. Additionally, three semi-analytic models have been used to analyse these same clusters from the MDPL2 simulation: SAG (Cora et al., 2018), SAGE (Croton et al., 2016) and Galacticus (Benson, 2012). This allows rigorous studies of galaxy properties and evolution to be made using these simulations, however throughout this thesis we choose to focus purely on the GadgetX simulations. This choice was made because we spend little of this thesis discussing galaxy properties, and so using multiple physics models would not add significantly to our work. Additionally, The Three Hundred suite includes several simulations using GadgetX, such as the high-resolution and dark matter-only runs used in Chapter 3, and so using the GadgetX runs allows us to include these additional simulations in our work.

[3]The MultiDark simulations are publicly available from the cosmosim database, https://www.cosmosim.org.

[4]The reduced Hubble constant, h, is defined such that $H_0 = h \times 100$ km s^{-1} Mpc^{-1}.

GadgetX is a modified version of the Gadget3 code, which is itself an updated version of the Gadget2 code, and uses a smoothed-particle hydrodynamics scheme to fully evolve the gas component of the simulations (Springel, 2005; Beck et al., 2016). The final dataset comprises of a mass-complete cluster sample from $M_{200} = 5 \times 10^{14}\ h^{-1}M_{\odot}$ to $M_{200} = 2.6 \times 10^{15}\ h^{-1}M_{\odot}$, where M_{200} is the mass contained within a sphere of radius R_{200}. The hydrodynamical clusters consist of dark matter and gas particles, but also contain stellar particles of variable masses, typically with $m_{\rm star} \sim 4 \times 10^{7}\ h^{-1}M_{\odot}$, produced by the stochastic star-formation that is implemented by GadgetX (Tornatore et al., 2007; Murante et al., 2010; Rasia et al., 2015). The Three Hundred dataset has been used to examine galaxy groups (Haggar et al., 2021), cosmic environment (Wang et al., 2018), cosmic filaments (Kuchner et al., 2020; Rost et al., 2021), backsplash galaxies (Haggar et al., 2020) and ram pressure stripping (Arthur et al., 2019; Mostoghiu et al., 2021a), as well as numerous other areas relating to galaxy clusters. A full, technical description of The Three Hundred dataset is available in Cui et al. (2018).

Comparison to other datasets

The Three Hundred is a state-of-the-art suite of galaxy cluster simulations, perfectly suited for the work contained within this thesis. One of its strengths is the large number of clusters it contains. This large number allows us to make statistically significant conclusions, even if a subsample of clusters is used, or if clusters with imperfect analysis are removed (as described later in Section 2.2.3). Additionally, this large number of clusters allows us to identify multiple examples of relatively rare events (see Chapter 4), and find correlations between different cluster properties (Chapter 5).

The 324 clusters in The Three Hundred sample make up a substantially larger sample than many other suites of hydrodynamical cluster simulations. For example, the Hydrangea (Bahé et al., 2017) and Cluster-EAGLE (Barnes et al., 2017b) suites only consist of 24 and 30 clusters, respectively, covering similar mass ranges to The Three Hundred. An advantage of having a smaller sample of clusters is that both Hydrangea and Cluster-EAGLE have better resolution than the simulations used in this work. However, the resolution of The Three Hundred suite is still sufficient to resolve huge numbers of galaxies, and is approximately four times better than the resolution of the MACSIS cluster sample (Barnes et al., 2017a), a similar-sized sample of large galaxy clusters (390 clusters in total). Finally, the high-resolution region surrounding The Three Hundred clusters extends to $\sim 10R_{200}$

from the cluster centre, further than the $5R_{200}$ high-resolution region surrounding each MACSIS cluster, allowing The Three Hundred to be used to study infalling objects and the outskirts of galaxy clusters, as we do in this thesis.

2.2.2 Galaxy identification and tree-building

The data for each cluster in The Three Hundred consists of 129 snapshots saved between $z = 16.98$ and $z = 0$. To identify the haloes and subhaloes in each cluster, each snapshot was processed using the AHF[5] halo finder (see Gill et al. (2004a) and Knollmann & Knebe (2009) for further details). AHF operates by identifying peaks in the matter density field, and returns the position and velocity of each halo and subhalo, as well as properties such as their radii, and their masses in gas, stars and dark matter. In this work, the word 'galaxy' is used to refer to all the components of an object in the hydrodynamical simulations, including its stellar and dark matter components. These galaxies can either be individual objects, or may be bound to a group. We use the word 'galaxy' to describe these in a general context, but in the specific context of the dark matter-only simulations used in Chapter 3, we refer to these objects as 'subhaloes' instead.

Throughout this work, we place a limit on the total (dark matter, gas and stars) mass of galaxies/subhaloes within the simulations of $M_{200} \geq 10^{10.5}\ h^{-1} M_{\odot}$. This corresponds to approximately 100 particles in the 15 h^{-1} Mpc high-resolution region surrounding each cluster. This limit is quite small, and galaxy properties such as their shapes (Tenneti et al., 2014; Chisari et al., 2017) and ram pressure (Steinhauser et al., 2016) cannot be reliably calculated with this few particles. However, throughout this thesis we do not focus on these properties of galaxies – instead, we use more basic properties such as their mass composition and positions, which can be found with as few as 100 particles (Onions et al., 2012). On the occasion that we do use other halo properties (such as halo concentrations in Section 4.2.1), we do so for group and cluster haloes. These generally contain > 1000 particles, which is easily sufficient to calculate the required quantities.

We also remove all objects from the hydrodynamical simulations that contain more than 30% of their mass in stars. These haloes are typically found very close to the centre of a larger halo, meaning that much of their dark matter has been stripped (evidence of this tidal stripping in The Three Hundred simulations is presented in Knebe et al. (2020)). The remnants of this process are very compact objects with high stellar mass fractions, whose properties (such as their radii and masses) are not

[5]http://popia.ft.uam.es/AHF

well-determined by our halo finder. Given this, and the fact that these haloes make up only 1% of all haloes within $5R_{200}$ of the clusters, we make the decision to remove these objects from our analysis.

In our hydrodynamical simulations, we also only use galaxies with a stellar mass $M_{\rm star} \geq 10^{9.5} M_{\odot}$. This is approximately equivalent to removing all galaxies with a luminosity $L < 10^8 L_{\odot}$, whilst not removing any with $L > 10^9 L_{\odot}$. However, it should be noted that we do not apply this stellar mass cut to the hydrodynamical simulations in Chapter 3, as this chapter includes direct comparisons between hydrodynamical and dark matter-only simulations. By not applying an absolute stellar mass cut in the hydrodynamical simulations, we keep all objects with a total mass above $10^{10.5}\ h^{-1} M_{\odot}$, and therefore ensure that the hydrodynamical and dark matter-only simulations are equivalent.

The halo merger trees were built using MERGERTREE, a tree-builder that forms part of the AHF package. For each halo in a given snapshot, this tree-builder calculates a merit function for each halo in all previous snapshots; specifically, MERGERTREE uses the merit function $M_{\rm i}$, as described in Table B1 of Knebe et al. (2013). This merit function is then used to identify a main progenitor, plus other progenitors, based on the number of particles that they share with the halo of interest. The version of MERGERTREE used in this work produces one-directional trees, not networks – a halo can have multiple progenitors, but cannot split into multiple descendants, instead being limited to just one.

Because it is able to skip snapshots, MERGERTREE is able to 'patch' over gaps in the merger tree, for example when a subhalo is near to the centre of its host halo and so is not easy to identify against the high background density (Onions et al., 2012). We also place a limit on the change in mass permitted between snapshots, such that no halo can more than double in dark matter mass between successive snapshots. This helps to prevent 'mismatches', caused by a subhalo located close to the centre of a larger halo being detected as the main halo (as shown in Behroozi et al., 2015, and detailed in the following section). Additional information on AHF and MERGERTREE can be found in Knebe et al. (2011b) and Srisawat et al. (2013).

2.2.3 Sub-sample of clusters

Throughout much of this work, we are interested in the evolution of galaxy clusters. We therefore require reliable simulation data at redshifts greater than zero, and so reliable tracking data at $z > 0$. However, in some cases the merit function used by MERGERTREE can incorrectly assign links between haloes in different snapshots.

This can lead to an apparent 'jump' in the position of a halo (in box coordinates), as well as a sudden change in its properties, due to one halo being incorrectly labelled as the progenitor of another. These mismatching events are uncommon, typically only affecting a small number of snapshots, and are fairly inconsequential when they affect individual galaxies. However, the merger trees of the cluster haloes can also be affected in this way, leading to a sudden change in the position of the main halo. Although such events are uncommon and typically only affect a small number of snapshots, they can be a major issue when tracking the times and positions of objects entering a cluster, and can result in many galaxies and groups being erroneously tagged as members of a cluster. These 'jumps' can also occur at late times (after $z = 1$), which is particularly problematic in this work, as most of the galaxy infalls take place at late times; for clusters that can be tracked back to $z = 5$, approximately 80% of infall events occur after $z = 1$.

These merger tree mismatches are especially common during a major merger between two haloes. Behroozi et al. (2015) show that various halo finders experience this same problem, where two merging haloes of similar size can be accidentally switched by a tree-builder, leading to the sizes and positions of haloes appearing to change suddenly and dramatically. Many of the clusters in The Three Hundred experience major mergers; a recent study, Contreras-Santos et al. (2022), discusses cluster mergers in The Three Hundred simulations in detail. In fact, we find that 59 of our 324 simulated clusters experience a change in position of $> 0.5R_{200}(z)$ between two snapshots after $z = 1$. Given that the typical time elapsing between snapshots at this redshift is ~ 0.3 Gyr, we find this distance to be non-physical and so likely due to these tree-builder issues.

In some cases, the tree-builder instead misses a link in the merger tree, causing a branch of the merger tree to end prematurely and the history of the halo before this link to be lost. For 17 clusters, the central cluster halo is affected in this way, and the evolution of the cluster halo cannot be tracked back further than $z = 0.5$. We choose to also remove these clusters from our analysis, in order to avoid affecting our results with clusters that do not have complete, reliable merger trees. Nine of these clusters also experience the halo mismatches described in the previous paragraph, resulting in a total of 67 clusters that we choose to remove from our sample.

While this is a large number of clusters to remove, an advantage of using a simulation suite such as The Three Hundred is that we still have a very large number of clusters remaining, easily a big enough sample to reach statistically significant results. Work is currently ongoing to improve the tree builder used in this analysis –

for instance, the factor of two change in dark matter mass described in Section 2.2.2 was introduced partly to reduce the jumps in cluster halo position. Future improvements to this tree builder include weighting the dark matter, gas, and star particles differently, as the population of stars at the centre of a halo is more resistant to processes such as stripping, and provides a good measure of a halo's position. This should reduce the need for removing clusters in future work. An alternative solution would be to manually add missing links to the merger tree, but we do not do this to avoid adding any bias or false information to our simulations. Additionally, the number of clusters available means that this is not necessary, and so we simply opt to remove these clusters from our analysis.

The remaining 257 clusters have M_{200} masses (dark matter, gas and stars, including subhaloes) ranging from $5 \times 10^{14}\ h^{-1}M_{\odot}$ to $2.6 \times 10^{15}\ h^{-1}M_{\odot}$, with a median value of $8 \times 10^{14}\ h^{-1}M_{\odot}$. Their radii ($R_{200}$) range from 1.3 h^{-1} Mpc to 2.3 h^{-1} Mpc, with a median of 1.5 h^{-1} Mpc. We use these clusters in the parts of this thesis that require cluster data from $z > 0$, but in the sections that only use data from $z = 0$, we instead use the full sample of 324 clusters.

Chapter 3

Hydrodynamical and dark matter-only simulations of clusters

Much of the work in this thesis revolves around using cosmological simulations to study the substructure of large galaxy clusters. Dark matter-only simulations are able to produce the cosmic structure of a ΛCDM universe at a much lower computational cost than hydrodynamical simulations, but it is not entirely clear how well smaller substructure is reproduced by dark matter-only simulations. In this chapter, we investigate this by directly comparing the substructure of galaxy clusters and of surrounding galaxy groups in hydrodynamical and dark matter-only simulations from The Three Hundred project. Dark matter-only simulations underestimate the number density of galaxies in the centres of groups and clusters relative to hydrodynamical simulations, and this effect is stronger in denser regions. We also look at the phase space of infalling galaxy groups, and show that dark matter-only simulations underpredict the number density of galaxies in the centres of these groups by about a factor of four. This implies that the structure and evolution of infalling groups may be different to that predicted by dark matter-only simulations, and so hydrodynamical simulations are crucial when studying the structure of clusters. Finally, we discuss potential causes for this underestimation, considering both physical effects, and numerical differences in the analysis. The content of this chapter has been published in Haggar et al. (2021).

Contents

3.1 Introduction

The mass composition of galaxy clusters is dominated by dark matter, which makes up over 80% of the mass of a typical cluster (Allen et al., 2011), and so the gravitational collapse of these structures is dominated by the effects of dark matter (Jenkins et al., 1998; Springel, 2005; Borgani & Kravtsov, 2011). Indeed, this dominance of gravitational effects over baryonic effects was partly the motivation behind the earliest numerical simulations of the non-linear collapse of cosmic structure, such as the work of Press & Schechter (1974), White (1976) and Gott et al. (1979). These each used N-body simulations of collisionless particles to study the build-up of structure, although many early simulations failed to produce adequate amounts of halo substructure. This was attributed to the 'over-merging' problem; dark matter subhaloes passing through a larger halo are heavily stripped, and pass below the simulation resolution limit (Frenk et al., 1996; Moore et al., 1996b). Subsequent work, such as that of Moore et al. (1998), was able to resolve this substructure, and confirmed that this over-merging was indeed responsible for the apparent lack of substructure in dark matter-only simulations.

Modern simulations, such as Millennium-XXL (Angulo et al., 2012), the Jubilee project (Watson et al., 2014) and the MultiDark simulations (Klypin et al., 2016) contain billions of cold dark matter particles in gigaparsec-scale volumes, and allow for detailed studies of the dark matter-dominated formation of large-scale structure. However, building hydrodynamical simulations of this scale requires huge amounts of computing power, and so various alternative approaches exist that are less physically motivated, but employ empirical models to save on computing power (as outlined in Section 2.1.2). Semi-analytic models and halo occupation models are two such examples, which use dark matter-only simulations as their starting point.

These alternatives to hydrodynamical models are generally successful, and recent models have been able to reproduce galaxy properties in impressive detail. For example, the GalICS 2.0 semi-analytic code created by Cattaneo et al. (2017) is able to reproduce both the galaxy stellar mass function and Tully-Fisher relation for galaxies in a (100 Mpc)^3 volume, which had previously been a significant issue for these models (Heyl et al., 1995; Baugh, 2006). Similarly, the semi-analytic model of Porter et al. (2014) (which was tuned using several different hydrodynamical simulations) can successfully reproduce the Fundamental Plane relation from observations of early-type galaxies (Djorgovski & Davis, 1987; Dressler et al., 1987). The increased speed of both halo occupation models and semi-analytic models over hydrodynamical simulations means that they are particularly useful for exploring large parameter spaces (Benson, 2010; Wechsler & Tinker, 2018), and can also be used to to study large cosmological volumes, and create samples of galaxies by generating mock observations (Eke et al., 2006; Frenk & White, 2012; Somerville & Davé, 2015). The power of this is demonstrated by Carretero et al. (2015), who have used a halo occupation model and the MICE simulations (Crocce et al., 2015) to produce mock observations that are being used by the upcoming *Euclid* mission[1]. Similarly, Knebe et al. (2018) have applied three distinct semi-analytic models to the MDPL2 MultiDark simulation (Klypin et al., 2016), generating the largest ever public mock galaxy catalogues.

The underlying assumption of these models is that the substructure of dark matter-only simulations is valid, compared to a more physical picture involving baryons. However, several studies have indicated regimes in which this may not be the case. Previous work has shown that the cumulative halo mass function is dependent on the baryonic processes that are present in a simulation (see Cui et al., 2012, 2014, for example). Similarly, van Daalen et al. (2011) study the matter power spectra of several hydrodynamical simulations from the OWLS project (Schaye et al., 2010), as well as the power spectrum of an equivalent dark matter-only simulation. They find that on length scales less than the typical size of clusters ($\lesssim 1$ Mpc), the power spectrum amplitude is suppressed in dark matter-only simulations, which is attributed to the cooling and collapsing of baryonic material into dense halo cores in hydrodynamical simulations. This steepened baryonic radial density profile can then cause the dark matter halo to contract, in the same way as was shown by Blumenthal et al. (1986), potentially leading to denser regions of dark matter in hydrodynamical simulations. Other work has found similar results on smaller scales. Jia et al. (2020) compare hydrodynamical and dark matter-only simulations of a $10^{14.8}\ h^{-1}M_{\odot}$ halo, at three mass resolutions. They find that the subhalo mass function is steeper in the hydrodynamical simulations, supporting the idea that the substructure in dark matter-only simulations is suppressed on small scales. Libeskind et al. (2010) show analogous results, but in even smaller (Local Group-sized) simulations.

Jia et al. (2020) also show that subhaloes are more concentrated in their hydrodynamical simulations, again confirming the mechanism of dark matter halo contraction described by Blumenthal et al. (1986). Indeed, other studies have found that halo density profiles are steepened by baryonic effects, leading to haloes being more concentrated in hydrodynamical simulations (Rasia et al., 2004; Lin et al., 2006).

[1]The MICE mock galaxy catalogue is publicly available from the CosmoHub database, https://cosmohub.pic.es.

Additionally, the central regions of dark matter haloes appear to be most strongly affected (Cui et al., 2014; Schaller et al., 2015). Such a mechanism is also supported by the work of Dolag et al. (2009). They show that subhaloes in radiative hydrodynamical simulations, which produce dense stellar regions in the centres of dark matter haloes, are more resistant to the stripping of gas and dark matter than haloes in non-radiative hydrodynamical simulations that lack these stellar cores.

In spite of this, the significance of the effect of baryons on dark matter haloes remains unclear, partly because of its dependence on the models that are implemented (Tissera et al., 2010; Cui et al., 2016). There does exist some disagreement within the literature, with some studies instead finding less substructure in hydrodynamical simulations, which is often attributed to increased tidal disruption in hydrodynamical simulations (Zhu et al., 2016; Garrison-Kimmel et al., 2017; Richings et al., 2020). The effect of baryons at different halo masses is also unclear; Chua et al. (2017) show that the subhalo mass function is shallower in hydrodynamical simulations than in dark matter-only simulations, in contrast to several other studies, including the recent work of Jia et al. (2020). Other studies show that the presence of baryons simply does not have a strong effect on substructure. In Bahé et al. (2019), the fraction of galaxies being accreted by clusters that survive to redshift $z = 0$ is only weakly dependent on whether baryons are included, although they explain that this may be due to the sub-grid physics implemented within their simulations.

In this chapter, we investigate how cosmic structure and substructure are affected by including baryons in cosmological simulations, to establish whether hydrodynamical simulations are necessary for studying cluster substructure, as we do in the following chapters. We begin by studying galaxies in galaxy clusters, and go on to investigate the outskirts of clusters. We particularly focus on galaxy groups in these outskirts, as a significant fraction of the galaxies that are accreted by clusters join the cluster as members of a galaxy group. Berrier et al. (2009) use dark matter-only simulations to find that 30% of galaxies of virial mass[2] greater than $10^{11.5}\ h^{-1}M_{\odot}$ have joined a cluster as part of a group, and 12% as part of a group of six or more galaxies. They also show that these fractions are slightly greater when a lower galaxy mass limit is used, in agreement with subsequent studies (Choque-Challapa et al., 2019). There is, however, some disagreement in this figure; some studies have found the fraction of infalling galaxies within groups to be as low as 10% (Arthur et al., 2017). Others have found that groups can make up almost half of infalling galaxies (McGee et al., 2009), although it is important to note that this variation is partly

[2]Berrier et al. (2009) use the definition of virial mass laid out by Bryan & Norman (1998).

down to the way in which groups are defined, which varies between different works (we discuss this in greater detail in Section 4.1). Studies of groups in cluster outskirts are therefore crucial in learning about the growth of clusters, and the histories of galaxies in cluster environments.

In Section 3.2 we introduce the analysis used in this chapter. In Section 3.3 we present our results, and in Section 3.4 we discuss some of the causes and effects of the differences we find between the two types of simulations. Finally, we summarise our findings in Section 3.5.

3.2 Methodology

The production of galaxy halo catalogues and merger trees for the work in this chapter follows the processes detailed in Section 2.2, as is also the case for Chapter 4 and Chapter 5. However, one distinction is that there is no limit applied to the stellar mass of galaxies in the hydrodynamical simulations used in this chapter, unlike the subsequent chapters which only include galaxies with $M_{\rm star} \geq 10^{9.5} M_{\odot}$. In doing so, all haloes with a total mass above $10^{10.5}\ h^{-1} M_{\odot}$ are kept in the hydrodynamical simulations, making them directly comparable to the dark matter-only simulations, which also include all haloes above this mass limit. Note that, by using a mass cut rather than a particle number cut, any bias towards the hydrodynamical simulations is removed, as the gas/star particles have lower masses than the dark matter particles, and so lower mass objects can be found in these simulations.

3.2.1 Dark matter-only simulations

In addition to the hydrodynamical simulations that make up The Three Hundred dataset, for the work in this chapter the 324 clusters have also been resimulated using only dark matter, to allow us to make direct comparisons between hydrodynamical and dark matter-only simulations. The same dark matter particle masses were used as in the original MDPL2 simulation, but the simulations run using the GadgetX code, as opposed to Gadget-2, which was used in MDPL2 (Klypin et al., 2016). These clusters were hence evolved from the same initial conditions as their hydrodynamical counterparts, and using the same simulation code and analysis. This allows us to make like-for-like comparisons between two simulations of the same clusters, showing the effects of baryonic physics on the dynamics of clusters. Finally, one of the clusters from The Three Hundred sample has been simulated two further times, in order to investigate resolution effects. We hereafter refer to this cluster by

Table 3.1: Parameters for the four classes of simulations used in this work. Columns 2-5 represent the mass of dark matter particles, m_{DM}, and of gas particles, m_{gas}, in the central high-resolution region, the Plummer equivalent gravitational softening length for both gas and dark matter particles, $\epsilon_{\mathrm{DM,gas}}$, and the Plummer equivalent gravitational softening length for star particles, $\epsilon_{\mathrm{stars}}$ (see Plummer (1911) and Barnes (2012) for a more detailed description of this softening). The bottom two rows only apply to 'CLUSTER_0002', which was used to test resolution effects.

Simulation	m_{DM} / $10^8\ h^{-1}M_\odot$	m_{gas} / $10^8\ h^{-1}M_\odot$	$\epsilon_{\mathrm{DM,gas}}$ / h^{-1} kpc	$\epsilon_{\mathrm{stars}}$ / h^{-1} kpc
Hydrodynamical	12.7	2.36	6.5	5
Dark matter-only	15	–	6.5	–
Reduced softening	12.7	2.36	6.5	1
High-res DM-only	1.88	–	3.25	–

its ID, 'CLUSTER_0002'. This cluster has been simulated hydrodynamically with a shorter gravitational softening length for stellar particles, and in dark matter-only with a factor of eight increase in resolution. These additional simulations are detailed in Section 3.3.2, and a summary of the particle data is given in Table 3.1.

3.2.2 Group identification

In this thesis, we identify galaxy groups in the simulations by considering each galaxy, and determining how many other galaxies in the same snapshot are associated with it. If the galaxy has four or more other galaxies associated with it, we take it to be the host halo of a group. Galaxies are defined as being associated with this 'group host' (and hence a member of the group) if they satisfy the same criteria as Han et al. (2018) and Choque-Challapa et al. (2019) use to define galaxy groups; the total (dark matter, gas, and stars) mass of a galaxy must be less than that of its group host, and the galaxy must satisfy the criterion below:

$$\frac{v^2}{2} + \Phi\left(r\right) < \Phi\left(2.5R_{200}^{\mathrm{grp,h}}\right) . \tag{3.1}$$

Here, v is the relative velocity of a galaxy with respect to its group host, $\Phi(r)$ is the gravitational potential due to the group host at a distance r from its centre, and $R_{200}^{\mathrm{grp,h}}$ is the radius of the group host halo. It is important to note that this is different to the radius of the host cluster in each simulation, which is subsequently referred to by $R_{200}^{\mathrm{clus,h}}$. Any galaxies that are less massive than their group host and that satisfy Eq. 3.1 are taken to be bound members of this group. Although we hereafter refer to these group members as being 'bound' to their host group, it is important to note that this definition is not technically equivalent to gravitational binding. Previous work (e.g. Behroozi et al., 2013) has shown that halo particles can be gravitationally balanced against the Hubble flow out to $\sim 4R_{200}$ from the halo centre. However, Eq. 3.1 places an artificial radial limit on groups, so that galaxies can only be found as far as $2.5R_{200}^{\rm grp,h}$ from the centre of the group – group members must have a total specific energy less than an object with zero velocity at $r = 2.5R_{200}^{\rm grp,h}$.

This outer limit is the same as was used by Han et al. (2018): their choice was motivated by the work of Mamon et al. (2004), who showed that backsplash galaxies are typically found out to approximately $2.5R_{200}$ from their host halo, but rarely any further. By setting this as the outer limit of a group, we include almost all galaxies that are on bound orbits around the group (having passed through its central halo at least once), whilst excluding galaxies that have not entered the group halo before. Furthermore, the relative velocity term in Eq. 3.1 means that only slow-moving galaxies at large distances are included as group members. Galaxies moving at greater velocities are excluded from the group, as these are likely 'fly-by' galaxies or 'renegade subhaloes' (Knebe et al., 2011c), which happen to be passing near to the group, but are not bound to it.

If a halo has four or more galaxies associated with it that each have a smaller total mass (including dark matter, gas, and stars) than the halo, we define this as a group, with the halo being the 'group host' halo. Very small groups with fewer members than this are common, but for the mass constraints that we put in place in Section 2.2.2, a collection of $\gtrsim 5$ associated galaxies is typically required to define a group (see Tully, 2015, for example). Additionally, we only study groups with 50 members or fewer (including the host object) that each satisfy the mass constraints given in Section 2.2.2, in order to exclude major cluster-cluster mergers.

3.2.3 Infalling groups

Throughout this chapter, we consider galaxy groups in two regimes. In Section 3.3.1, we identify galaxy groups located in the region between $[R_{200}^{\rm clus,h}, 5R_{200}^{\rm clus,h}]$ from the cluster centre at $z = 0$, which we refer to as the 'cluster outskirts' throughout this chapter. Then, in Section 3.3.2, we instead identify infalling galaxy groups, at all redshifts, again using the same methods as Han et al. (2018) and Choque-Challapa et al. (2019). Using the halo merger trees, we first identify all objects that have just fallen into the cluster, taken from the whole history of the cluster; that is, we find galaxies that were at a distance greater than $R_{200}^{\rm clus,h}$ from the cluster centre in one snapshot, but are within $R_{200}^{\rm clus,h}$ in the following snapshot. We refer to these objects as the 'infalling' galaxies.

For each infalling galaxy, we consider the first snapshot in which it has passed within $R_{200}^{\mathrm{clus,h}}$, and use the method in Section 3.2.2 to determine if any galaxies are bound to it at this time, thereby making the infalling galaxy a group host halo. This galaxy, and any galaxies that are bound to it, then make up the infalling group. A schematic is given in Fig. 3.1, showing the configuration of an infalling group. Haloes that are on a second (or subsequent) infall are excluded, so that our sample of groups consists only of those entering the cluster for the first time. These repeat infallers make up only 13% of the infalling galaxies, and less than 1% of the bound groups that we identify.

3.3 Results

3.3.1 Radial density profiles of clusters and groups

Fig. 3.2 shows the number density profile of galaxies that are bound (according to the criterion in Eq. 3.1) to a host cluster. For consistency, we use the same criteria to define galaxies bound to the cluster as we use for defining group member galaxies, but we note that this bound population of galaxies represents almost all galaxies in the cluster; over 99% of galaxies within R_{200} of a cluster are gravitationally bound. Note also that the distances from the cluster centre, r, are given in units of the cluster radii $R_{200}^{\mathrm{clus,h}}$ and so are normalised between clusters, but the number densities are given in units of h^3 Mpc^{-3}. For each cluster we generate a kernel density estimation (KDE) of the distribution of galaxies. Using a KDE with an optimised bandwidth provides a smoothed distribution, and removes most of the effects of bin selection that can impact a histogram. We then average the KDEs across all clusters. As this analysis only requires data from $z = 0$, we use data from all 324 clusters.

Clusters in the dark matter-only simulations have a deficit of subhaloes in their central regions, relative to the hydrodynamical cluster simulations. The deficit decreases with increasing distance from the cluster centres, and the two profiles are indistinguishable outside of $2.1R_{200}^{\mathrm{clus,h}}$, as shown by the residual in the bottom panel of Fig. 3.2. This is equal to the ratio of the galaxy number density profile in the hydrodynamical clusters (given by ρ_{hydro}) to the density profile in the dark matter-only simulations (ρ_{DM}). Consequently, in regions where this residual is equal to one, the number density of galaxies is equal in the two types of simulation. The error bounds on this residual come from the uncertainty in the mean density profiles (the dark shading in the top panel). This residual demonstrates that the substructure of hydrodynamical simulations is only reproduced by dark matter-only simulations in

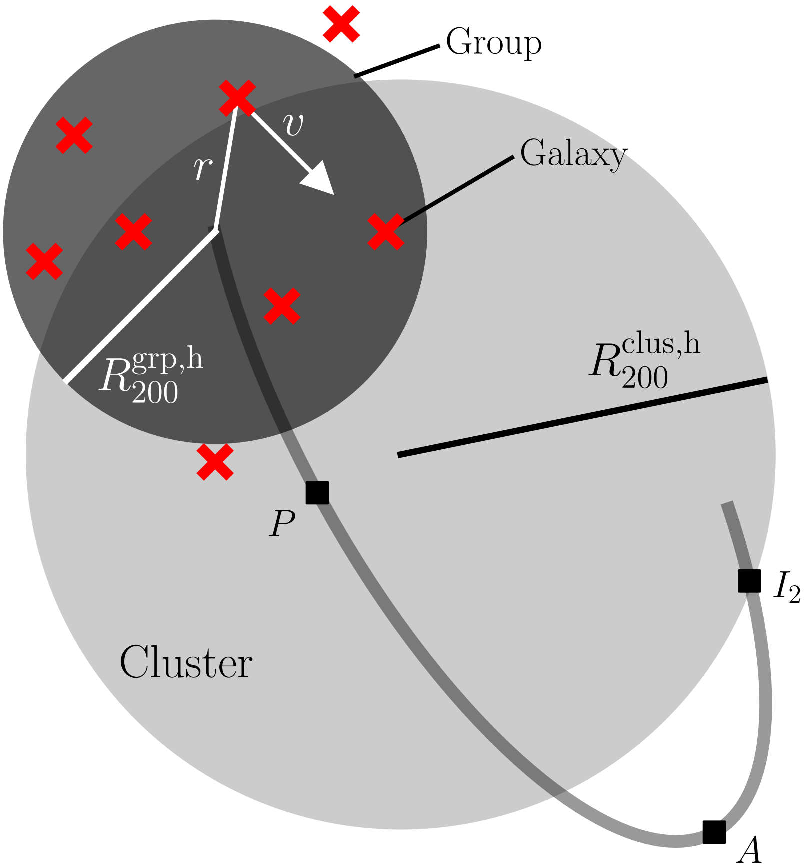

Figure 3.1: Schematic of a galaxy group halo (dark circle) falling into a cluster (light circle). The centre of the host group has just passed within $R_{200}^{\mathrm{clus,h}}$ of the cluster. Red crosses represent galaxies that are members of this group; note that these are not limited to be within R_{200} of the cluster or the group at infall, but are just defined based on Eq. 3.1. The position, r, and velocity, v, of one galaxy relative to its host group are also labelled. It is this configuration that is used in the second part of this chapter (Section 3.3.2). The subsequent path of this group through the cluster is shown by the thick, grey line, and the black squares on this line represent the moments of pericentre, apocentre, and second infall of the group (marked P, A and I_2 respectively). These are not used in this chapter, but feature extensively in Chapter 4 of this work.

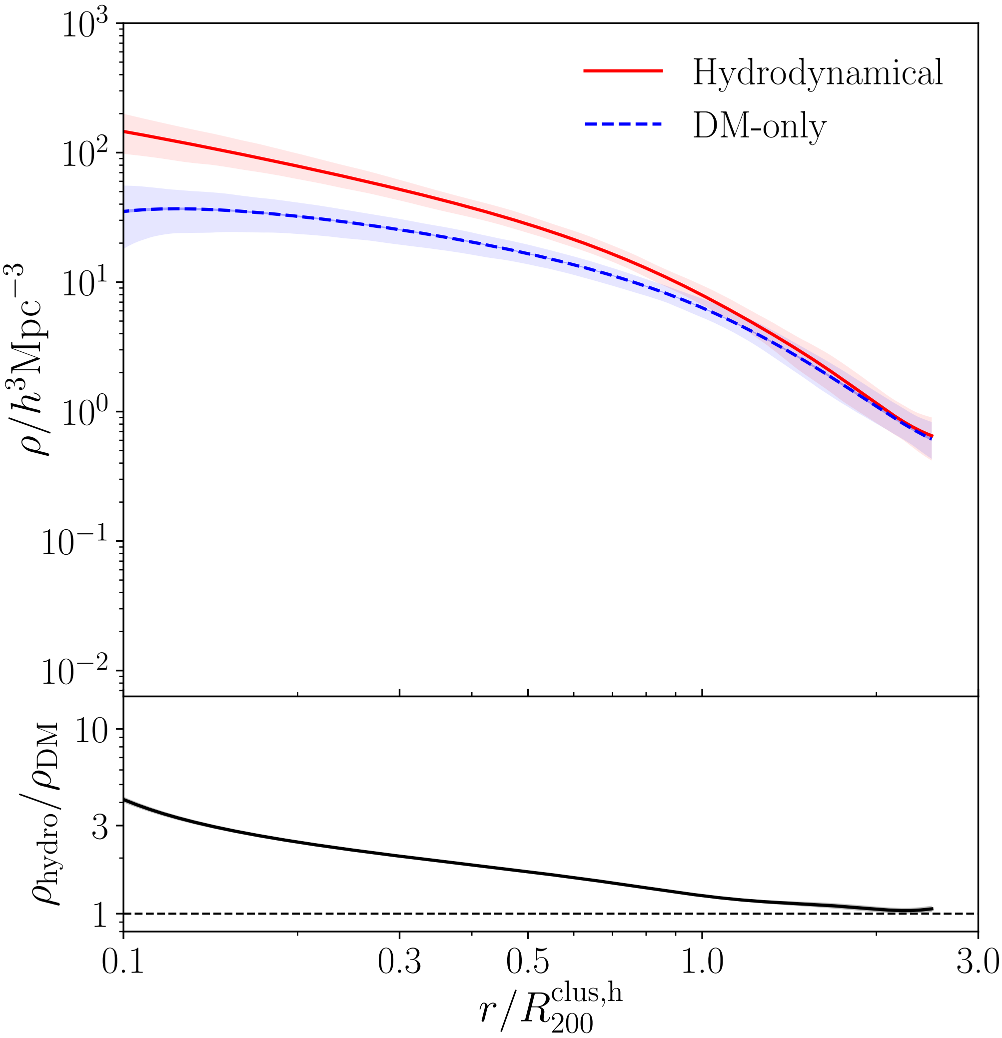

Figure 3.2: Radial number density, ρ, of galaxies gravitationally bound to clusters, for 324 hydrodynamical clusters and dark matter-only clusters at $z = 0$ (top panel). Light shaded regions show the 1σ spread between clusters, and dark shaded regions represent 1σ uncertainty in the mean radial number density profile, although these are mostly too small to be seen. Bottom panel shows fractional residuals, which we define in the main text.

the outer regions of clusters.

Some of the difference between the number densities of galaxies in the cluster outskirts can be explained by the inclusion of backsplash galaxies, which have previously passed within $R_{200}^{\mathrm{clus,h}}$ of a cluster, but now exist beyond this radius in the cluster outskirts. If we exclude these backsplash galaxies from our analysis, we instead find that the number density profiles of the hydrodynamical and dark matter-only clusters agree at radii beyond $1.4R_{200}^{\mathrm{clus,h}}$. This indicates that backsplash galaxies are less likely to survive the passage through a cluster in dark matter-only simulations, as they contribute less to the number density of galaxies in these simulations. Indeed, we find that in the dark matter-only simulations, an average of 45% of galaxies in the radial region $[R_{200}^{\mathrm{clus,h}}, 2R_{200}^{\mathrm{clus,h}}]$ are backsplash galaxies, compared to 51% in the hydrodynamical simulations (which we discuss in detail in Chapter 5[3]). As this difference in backsplash fraction is small, it can only explain a small part of the number density deficit in dark matter-only simulations. At cluster distances greater than $1.4R_{200}^{\mathrm{clus,h}}$ the difference in number density between the two simulations is less than 15%, even when including backsplash galaxies. Inside this radius, the increased deficit of galaxies in the dark matter-only simulations cannot be fully explained by these missing backsplash galaxies.

Unlike the mass density profiles of these clusters (Mostoghiu et al., 2019), the number density profiles in Fig. 3.2 are not well-described by an NFW profile (Navarro et al., 1996), particularly in the cluster outskirts. This is potentially because of the boundness criteria we employ, which place strict limits on the velocities of galaxies in the outer regions of the clusters. However, in the radial region $[0.2, 2.0]R_{200}^{\mathrm{clus,h}}$, the cluster number density profiles are well-described by an Einasto profile (Einasto, 1965; Navarro et al., 2004). Specifically, we use the form of Springel et al. (2008),

$$\rho\left(r\right) = \rho_0 \exp\left(-\frac{2}{\alpha}\left[\left(\frac{r}{r_0}\right)^{\alpha} - 1\right]\right), \tag{3.2}$$

where ρ_0, r_0 and α are free parameters describing the profile. The values of these parameters for the radial density profiles of clusters in both the hydrodynamical and dark matter-only simulations are given in Table 3.2. The parameter α describes the curvature of the profile, and its greater value for the dark matter-only clusters demonstrates how this profile is shallower near to the centre of a cluster, but drops off equally steeply at greater radii.

[3]Our work in Chapter 5, which is also published in Haggar et al. (2020), gives a slightly different value for the average backsplash fraction in hydrodynamical simulations. This is due to the fact that, in Chapter 5, we also include a stellar mass cut in the galaxies.

Table 3.2: Parameters for the best-fit Einasto profiles (Eq. 3.2), for the hydrodynamical and dark matter-only simulations of the number density profiles of galaxy clusters (Fig. 3.2) and galaxy groups in the cluster outskirts (Fig. 3.3). For clarity, we have not included these fits in the relevant figures.

	Simulation	ρ_0 / h^3 Mpc^{-3}	r_0 / R_{200}^{h}	α
Clusters	Hydrodynamical	11.709 ± 0.006	0.8339 ± 0.0003	0.640 ± 0.001
	DM-only	5.479 ± 0.003	1.0747 ± 0.0004	0.878 ± 0.001
Groups	Hydrodynamical	33.43 ± 0.04	0.5303 ± 0.0002	0.931 ± 0.001
	DM-only	7.46 ± 0.01	0.7615 ± 0.0003	1.302 ± 0.002

Fig. 3.3 shows data equivalent to that in Fig. 3.2, except that instead of the density profiles of the clusters, it gives the mean radial number density profile of galaxy groups located in the cluster outskirts, between $[R_{200}^{\mathrm{clus,h}}, 5R_{200}^{\mathrm{clus,h}}]$ from the cluster centre, at $z = 0$. Groups that have between five and 50 galaxies (including the group host galaxy) that each satisfy our mass criteria are included. Here, we generate a KDE for each individual galaxy group, and average these across all groups in the whole sample of 324 clusters. Throughout our analysis of the 'group members', we exclude the 'host galaxy' at the centre of each group's host halo.

As is the case for the clusters in Fig. 3.2, groups in the dark matter-only simulations also have a deficit of subhaloes, relative to groups in the hydrodynamical simulations. This deficit is strongest in the central regions, and as shown by the residuals in the bottom panel in Fig. 3.3, the two profiles agree within uncertainties in the group outskirts, beyond $2.1R_{200}^{\mathrm{grp,h}}$. The apparent spike in the residual at $r \approx 2.5R_{200}^{\mathrm{grp,h}}$ is due to small number statistics, as there are very few galaxies at this distance from the group centres. In the hydrodynamical simulations, we find a very small dependence of the number density profile on the distance of the groups from the cluster centre – the number density of galaxies is approximately 20% lower in groups within $2R_{200}^{\mathrm{clus,h}}$ of the cluster centre, compared to those beyond this distance. We find no significant systematic dependence on the cluster distance in the dark matter-only simulations. The shapes of the radial number density profiles for both the hydrodynamical and dark matter-only groups are different to those of the profiles of the clusters. Generally the cluster profiles are flatter, with a much shallower decrease in number density beyond R_{200}. However, in the radial region $[0.3, 2.0]R_{200}^{\mathrm{grp,h}}$, the group number density profiles can also be well-described by an Einasto profile – the parameters of the best-fit profiles are given in Table 3.2. As is also the case for the cluster profiles in Fig. 3.2, the change in the slope is sharper in the dark matter-only groups.

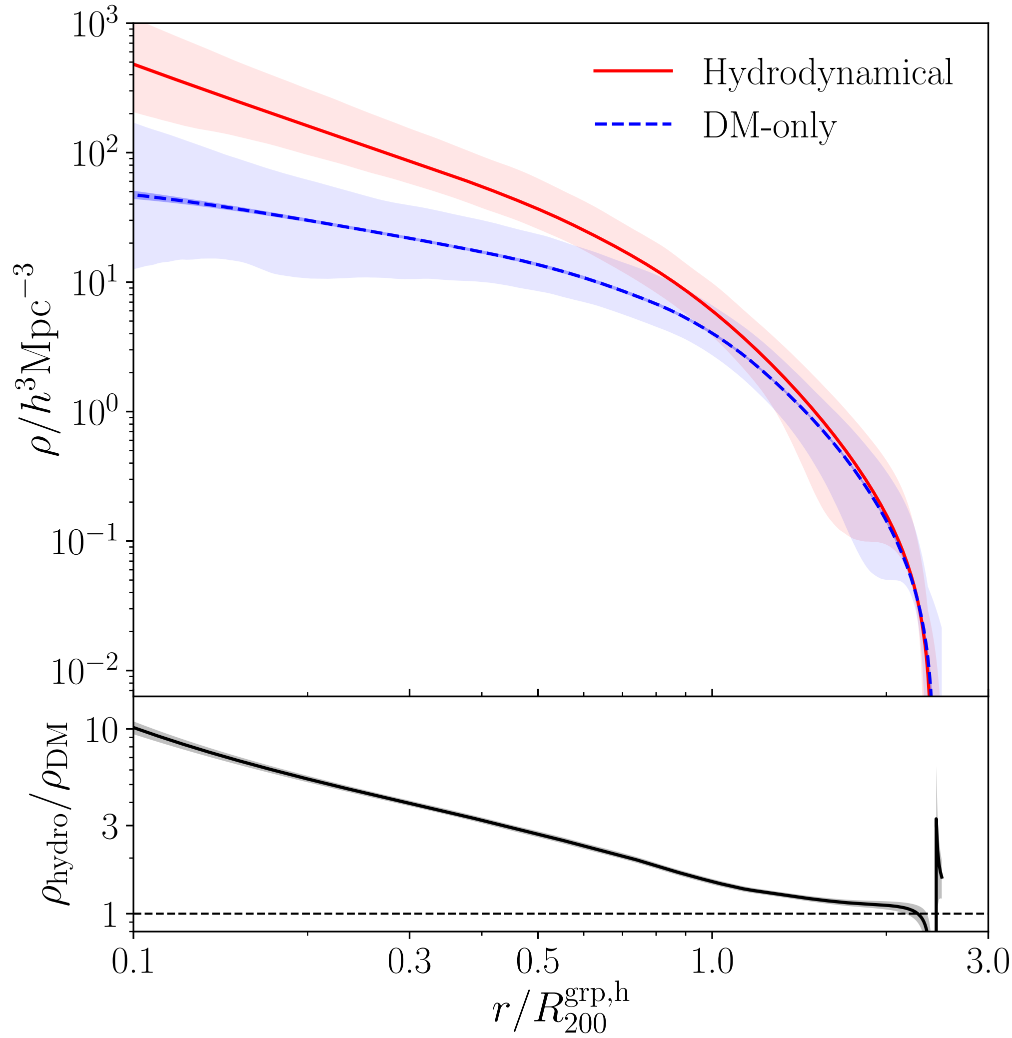

Figure 3.3: Radial number density, ρ, of galaxies in groups located in cluster outskirts at $z = 0$ (top panel). Includes all groups in the radial range $[R_{200}^{\mathrm{clus,h}}, 5R_{200}^{\mathrm{clus,h}}]$, consisting of between five and 50 galaxies. Light shaded regions show the 1σ spread between groups, and dark shaded regions represent 1σ uncertainty in the mean radial number density profile, although these are mostly too small to be seen. Bottom panel shows fractional residuals.

Masses and radii of groups

We have also investigated whether the radial number density profiles of groups in the cluster outskirts are dependent on the mass of the group host haloes, $M_{200}^{\mathrm{grp,h}}$. The median mass of the host halo for groups in the cluster outskirts is $10^{13.5\pm0.4}\ h^{-1}M_{\odot}$ in the dark matter-only simulations, and $10^{13.2\pm0.4}\ h^{-1}M_{\odot}$ in the hydrodynamical simulations (these error bars represent the 1σ spread of the data). The range of $M_{200}^{\mathrm{grp,h}}$ is approximately two orders of magnitude; the range of group masses is $[10^{12.3}, 10^{14.5}]\ h^{-1}M_{\odot}$ in the dark matter-only simulations, and $[10^{12.0}, 10^{14.2}]\ h^{-1}M_{\odot}$ in the hydrodynamical simulations, for the groups that we have selected (with between five and 50 members). We generally find that the number density of galaxies is less in larger groups, and that the number density profiles are flatter. In large groups the radial number density profile becomes closer to the cluster profile, particularly in the inner regions of the group. Splitting the galaxy groups into three categories based on halo mass, we find that the variation in radial number density between the most massive and least massive groups is relatively small, typically less than a factor of three.

Although the average mass of group haloes is slightly greater in our dark matter-only simulations, this is not enough to fully explain the significantly flatter radial density profile in these simulations. The difference in average group mass between our hydrodynamical and dark matter-only simulations is 0.3 dex, which is small compared to the variation in group mass within each simulation (approximately 2 dex). It is therefore not enough to account for the difference in number density between the simulations, which is a factor of 10 in the central regions of the groups. We discuss the dependence of the radial number density on the group halo mass in more detail in Appendix A.

Naturally, given that groups in the dark matter-only simulations have slightly greater masses, we also find that these groups have greater average radii than in the hydrodynamical simulations. This is shown in Fig. 3.4; the median radius, $R_{200}^{\mathrm{grp,h}}$, for groups of between five and 50 members is $0.51^{+0.17}_{-0.13}\ h^{-1}$ Mpc in the outskirts of the dark matter-only clusters, compared to $0.41^{+0.17}_{-0.10}\ h^{-1}$ Mpc in the hydrodynamical simulations (these error bars also represent the 1σ spread of the data). This difference in median radius is equivalent to a 92% increase in the median volume of these groups, meaning that although the number density of galaxies in the group outskirts is the same in both data sets, this would result in a greater number of galaxies in the outskirts of dark matter-only groups.

Indeed, we find that the median number of group members outside of $R_{200}^{\mathrm{grp,h}}$ is

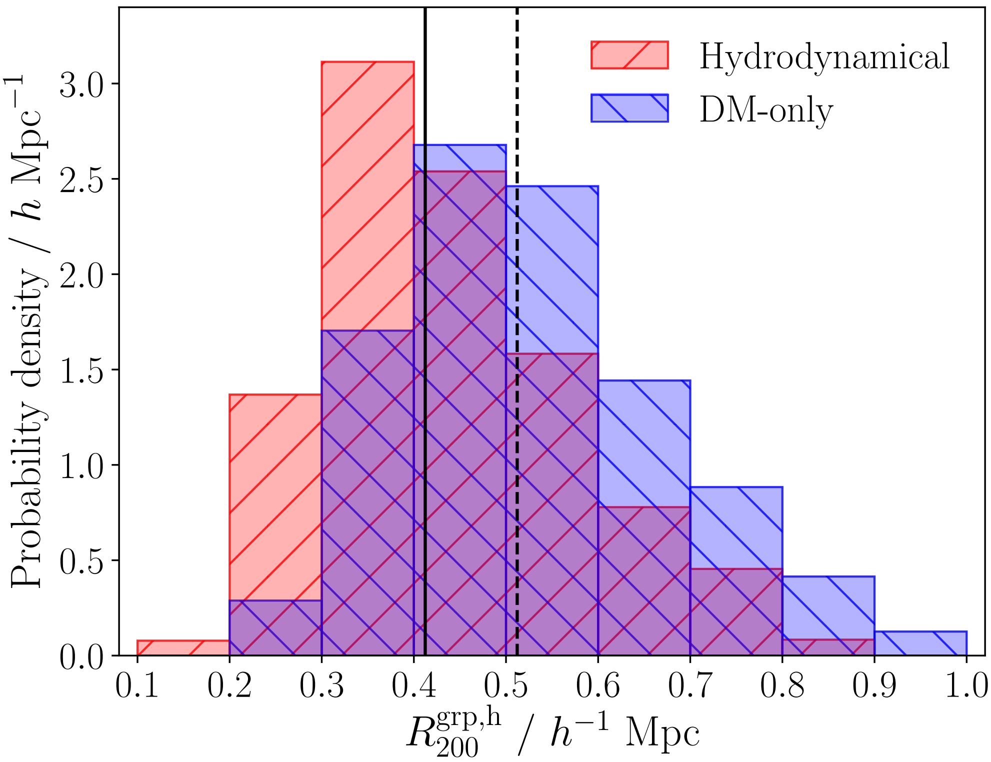

Figure 3.4: Histogram of radii, $R_{200}^{\rm grp,h}$, of groups with between five and 50 members, located between $R_{200}^{\rm clus,h}$ and $5R_{200}^{\rm clus,h}$ from a cluster, in hydrodynamical and dark matter-only simulations. Solid and dashed vertical lines show the median group radius in the hydrodynamical and dark matter-only simulations, respectively.

2^{+4}_{-2} in the hydrodynamical simulations, and 3^{+4}_{-2} in the dark matter-only, demonstrating this increase, although the difference is much smaller than the spread in the data. The median total number of group members is the same in the hydrodynamical and dark matter-only simulations (8^{+11}_{-3} and 8^{+9}_{-3} respectively). The difference in the average group halo radius is not seen in the radii of the host clusters, $R_{200}^{\rm clus,h}$; on average, there is less than a 1% variation in the radius of each cluster between the hydrodynamical and dark matter-only simulations.

Fig. 3.4 also shows that some of the groups, particularly in the hydrodynamical simulations, have very small radii ($< 200\ h^{-1}$ kpc). This is due to the fact that we do not apply a lower mass limit to the group host haloes (besides the $10^{10.5}\ h^{-1}M_{\odot}$ mass limit that is applied to all haloes). Despite this, even the smallest groups that we identify have $R_{200}^{\rm grp,h} \approx 160\ h^{-1}$ kpc, $M_{200}^{\rm grp,h} \approx 10^{12}\ h^{-1}M_{\odot}$, and typically contain approximately five galaxies, which we consider large enough to still represent physical galaxy groups. In fact, certain galaxy groups, such as Hickson Compact Groups (Hickson, 1982), can contain similar numbers of galaxies to this within an even smaller radius, sometimes less than 50 kpc (Barton et al., 1996).

3.3.2 Phase space of infalling groups

The results from Fig. 3.2 show that the inclusion of baryonic material affects the substructure in galaxy clusters, and Fig. 3.3 shows that this effect is even stronger in galaxy groups located in the cluster outskirts. As described in Section 3.1, a significant fraction of galaxies within clusters have previously been members of a group, that has since been accreted by a cluster (Berrier et al., 2009; McGee et al., 2009; Arthur et al., 2017). Indeed, in our hydrodynamical simulations we find that over the history of a cluster, an average of $(14.2 \pm 0.2)\%$ of galaxies that enter $R_{200}^{\rm clus,h}$ do so as part of a bound group (this error represents the uncertainty in the mean; throughout most of this work we instead quote the spread in the data). For the dark matter-only clusters, $(6.2 \pm 0.1)\%$ of subhaloes are accreted as members of groups; this lower fraction is expected, given the lower number density of group members in the dark matter-only simulations, shown by Fig. 3.3. The fraction of galaxies accreted in hydrodynamical groups is similar to other work that uses similar sized hydrodynamical clusters (Arthur et al., 2017), and is in line with the typical mass fraction found in subhaloes (Gao et al., 2011). The greater fraction that is found in some other work (e.g. McGee et al., 2009) is likely caused by their use of different mass limits for galaxies and the group host (note that we use the same limit for these) – we discuss this in more detail in Section 4.1.

The accretion of galaxy groups is therefore an important part of the growth of galaxy clusters, and this has motivated a wide range of studies into the properties of these groups. For example, Choque-Challapa et al. (2019) use dark matter-only simulations to identify groups using the same method as this work, and then look at the phase space of subhaloes in groups at the time when groups enter the cluster. They go on to look at the evolution of these groups by examining which subhaloes remain bound to the group, how this depends on the position and velocity of group members, and when subhaloes become unbound from their group host.

However, as we have shown, the structure of groups in dark matter-only simulations is different to the substructure predicted by more physically motivated hydrodynamical simulations. Fig. 3.5 demonstrates this further, by showing the phase space distribution of bound satellite galaxies in infalling groups in our sample of 257 hydrodynamical (left panel) and dark matter-only (right panel) clusters described in Section 2.2.3. We again note that here we are looking at groups at the moment when they enter a galaxy cluster (i.e. the first snapshot at which the group host is inside

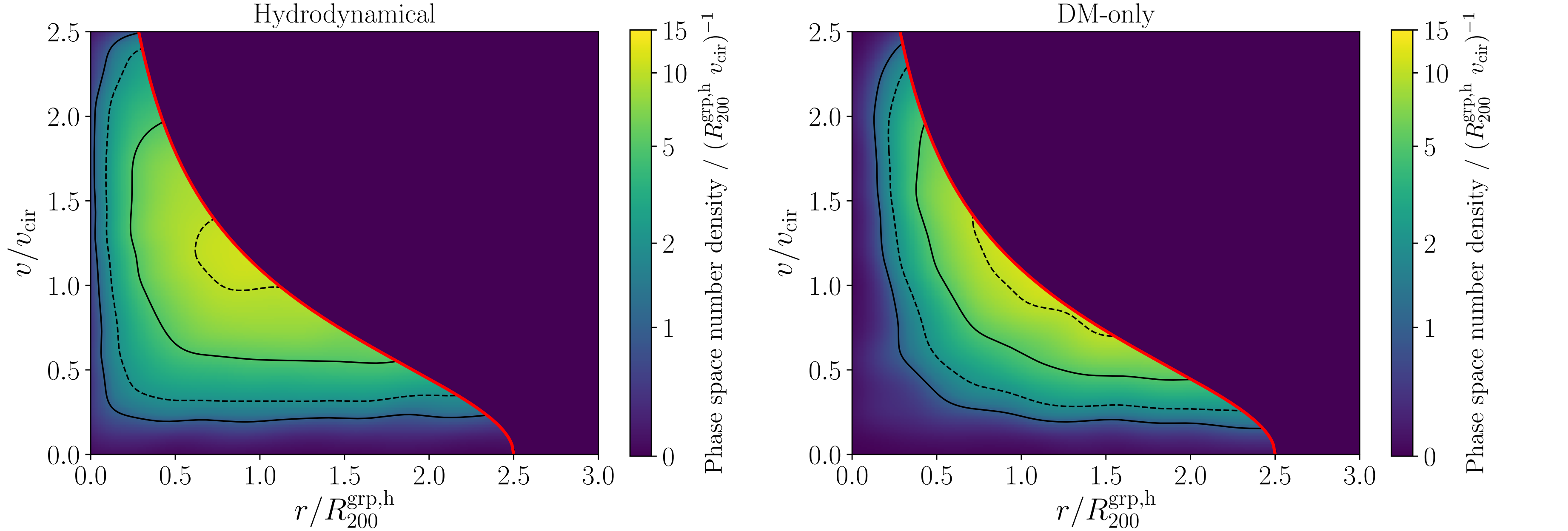

Figure 3.5: Average phase space distribution of galaxies/subhaloes in groups, at the time of infall, for hydrodynamical (left) and dark matter-only (right) cluster samples, where lighter colours represent regions of phase space that contain more galaxies. All groups with between five and 50 members are used, across all clusters. Phase space density is the number of galaxies expected in one group, in a square interval of the phase space, of size $[R_{200}^{\rm grp,h}, v_{\rm cir}]$, where $v_{\rm cir}$ is the circular orbital velocity at a distance of $R_{200}^{\rm grp,h}$ from the group centre. Contours are at densities of [1, 2, 5, 10] $(R_{200}^{\rm grp,h} v_{\rm cir})^{-1}$ in both plots. Red line shows the boundness criterion given by Eq. 3.1, meaning that galaxies below this line are bounded to their group host. Galaxies above this line are not group members, and so are excluded from this figure.

$R_{200}^{\mathrm{clus,h}}$), as opposed to Section 3.3.1, in which we look at all groups in the cluster outskirts. The phase space consists of the radial distance of a galaxy from its host group halo, in units of $R_{200}^{\mathrm{grp,h}}$, and its velocity relative to the group halo, in units of v_{cir}, the circular orbital velocity at $r = R_{200}^{\mathrm{grp,h}}$. This figure shows a 2D KDE of the stacked phase space data for all infalling groups across the cluster samples, and is normalised by the total number of infalling groups in each sample. As the typical size of a group (~ 8 members) is small, using a KDE with an optimised bandwidth allows the mean distribution of galaxies to be clearly seen.

The phase space of group members in our dark matter-only simulations is in agreement with the distribution found by Choque-Challapa et al. (2019), despite their use of lower mass clusters ($\sim 10^{14}\ h^{-1}M_{\odot}$), and a lower limit on satellite masses ($\sim 10^{7.8}\ h^{-1}M_{\odot}$). The greatest concentration of subhaloes in this phase space is close to the line representing the boundness criterion, spread between approximately $0.7R_{200}^{\mathrm{grp,h}}$ and $1.5R_{200}^{\mathrm{grp,h}}$, with the maximum located at $r = 1.1R_{200}^{\mathrm{grp,h}}$, $v = v_{\mathrm{cir}}$. We also find that there are very few subhaloes near to the central regions of groups, particularly with low velocities. Only 9% of dark matter-only groups contain at least one subhalo with $r < 0.3R_{200}^{\mathrm{grp,h}}$ and $v < v_{\mathrm{cir}}$ (excluding the host halo at the centre of each group). This is a region that previous work, such as that of Choque-Challapa et al. (2019), has also shown to contain few satellites in dark matter-only simulations.

However, carrying out this analysis with the hydrodynamical simulations, as shown by the left panel of Fig. 3.5, gives a different distribution of galaxies. The most prominent region of high phase space density does not extend as far from the group centres, as it reaches from $0.6R_{200}^{\mathrm{grp,h}}$ to $1.1R_{200}^{\mathrm{grp,h}}$, with a maximum at $r = 0.9R_{200}^{\mathrm{grp,h}}$, $v = 1.2v_{\mathrm{cir}}$. The central regions are also more populated with galaxies; 33% of hydrodynamical groups have at least one galaxy in the central, low-velocity region described in the previous paragraph.

The differences between the phase spaces are demonstrated more clearly in Fig. 3.6. This plot shows the fractional difference between the phase space density of the hydrodynamical and dark matter-only simulations, shown in the two panels of Fig. 3.5. In the regions marked as 'Hydro. excess' in Fig. 3.6, the colour represents the fractional excess of the number density in the hydrodynamical simulations, relative to the dark matter-only. Similarly, in the 'DM-only excess' regions, the quantity plotted is the density excess in the dark matter-only phase space, relative to the hydrodynamical. For example, a value of 1.5 in the 'DM-only excess' region would mean that the phase space density in this region is 150% greater in the dark matter-only groups than in the hydrodynamical groups (i.e. 2.5 times the magnitude). The dashed line

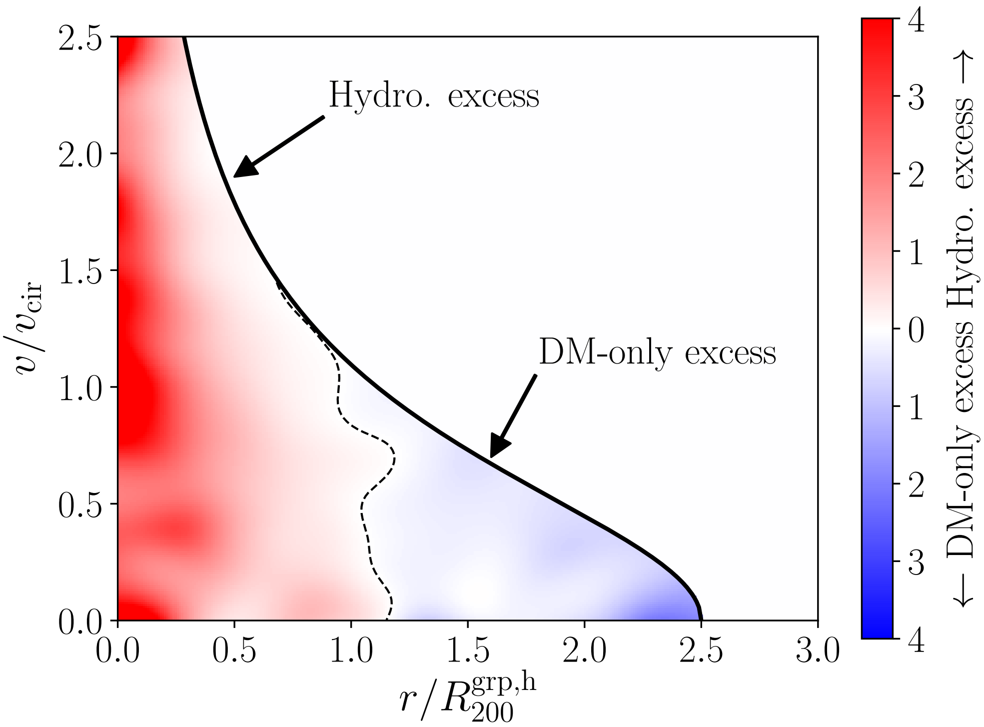

Figure 3.6: Fractional difference of the phase space distributions of group members in hydrodynamical and dark matter-only simulations, shown in Fig. 3.5. The values represent the fractional difference of the greater density value relative to the lower value, and the colour represents whether the hydrodynamical or dark matter-only simulations have an excess of galaxies in this region. These two regimes are separated by the dashed contour at zero, such that hydrodynamical groups have a greater density of galaxies to the left of the line, and dark matter-only groups to the right.

marks a contour where the phase space densities are the same, and clearly divides the phase space into two distinct regions. At greater distances from the centres of groups (to the right of this contour), dark matter-only simulations over-predict the abundance of galaxies in this region of phase space. Meanwhile, closer to the centres of groups, dark matter-only simulations under-predict the numbers of galaxies.

Fig. 3.6 shows that there are more galaxies beyond $1.1R_{200}^{\mathrm{grp,h}}$ in infalling dark matter-only groups, regardless of the relative velocity of galaxies. This appears to contradict the results of Section 3.3.1 and Fig. 3.3, which show that the number density of galaxies in the outskirts of groups is the same in dark matter-only and hydrodynamical simulations. However, because of the larger $R_{200}^{\mathrm{grp,h}}$ values for dark matter-only groups, these groups have approximately twice the volume of hydrodynamical groups, which will lead to twice as many galaxies in a given radial region, in units of $R_{200}^{\mathrm{grp,h}}$. Indeed, in most of the phase space with $r > R_{200}^{\mathrm{grp,h}}$, the relative excess of dark matter-only subhaloes is approximately equal to one.

Fig. 3.6 also shows the excess of galaxies in group centres in the hydrodynamical simulations. As discussed earlier in this section, dark matter-only groups are less likely to contain galaxies at very small radii and with low velocities. This is shown by the large hydrodynamical density excess: in the region with $r < 0.3R_{200}^{\mathrm{grp,h}}$ and $v < v_{\mathrm{cir}}$, the average excess is 2.8, corresponding to almost four times as many galaxies in this region of phase space in the hydrodynamical simulations. However, there is a similar excess for all galaxies in these central regions, regardless of their relative velocities.

Resolution effects

As described in Section 2.2.1 and Table 3.1, we have also simulated one of the clusters (referred to as 'CLUSTER_0002') from THE THREE HUNDRED sample two more times, with a shorter gravitational softening for stellar particles, and in dark matter-only at a resolution eight times greater than the standard dark matter-only clusters. These four simulations of a single cluster, using the same initial conditions and cosmological code, are shown in Fig. 3.7. The top two panels show the dark matter distribution of CLUSTER_0002 in the two hydrodynamical runs, and the bottom two panels show the dark matter-only simulations.

We find that decreasing the softening length in the hydrodynamical simulation has no significant effect on either the radial profiles in Section 3.3.1, or the phase space of members in infalling groups (Section 3.3.2), for this cluster. Conversely, comparing the original and high-resolution dark matter-only runs shows that a change

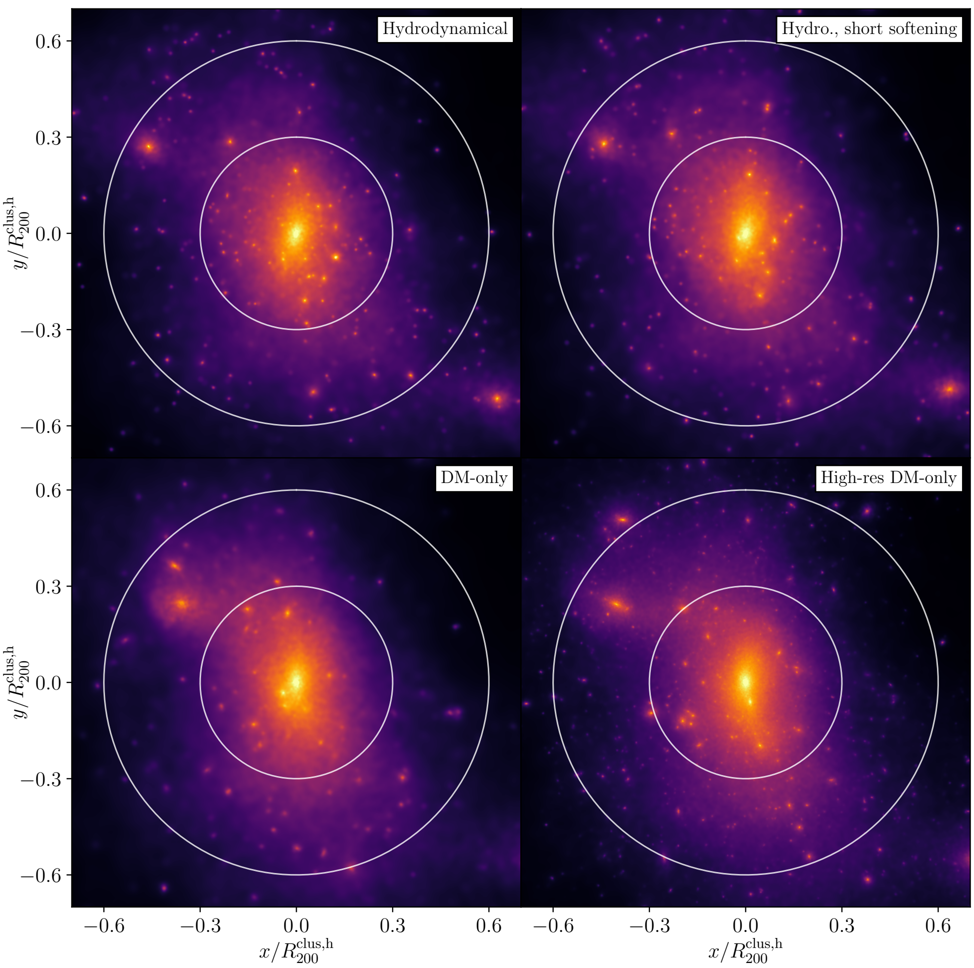

Figure 3.7: Dark matter distribution of inner region of CLUSTER_0002, which was simulated four times. Top-left: the original hydrodynamical run. Top-right: hydrodynamical run with a reduced gravitational softening length for stars. Bottom-left: original dark matter-only run. Bottom-right: high-resolution dark matter-only run. Overlaid circles are at $0.3R_{200}^{\mathrm{clus,h}}$ and $0.6R_{200}^{\mathrm{clus,h}}$ from the cluster centre; the value of $R_{200}^{\mathrm{clus,h}}$ varies by approximately 1% between the simulations. Note the visibly greater number of compact dark matter haloes within $0.3R_{200}^{\mathrm{clus,h}}$ in the panels representing the hydrodynamical runs.

Table 3.3: Number of galaxies within $0.5R_{200}^{\mathrm{clus,h}}$ of the centres of the clusters in Fig. 3.7, and total number of galaxies within $0.5R_{200}^{\mathrm{grp,h}}$ of groups around these clusters. Note that using a high-resolution dark matter-only simulation does increase these quantities relative to the regular dark matter-only simulation, but less so than using hydrodynamical simulations. Changing the gravitational softening length has almost no effect.

Simulation	$N(r < 0.5R_{200}^{\mathrm{clus,h}})$	$N(r < 0.5R_{200}^{\mathrm{grp,h}})$
Hydrodynamical	264	35
Reduced softening	258	35
Dark matter-only	86	14
High-res DM-only	198	23

in resolution does have some effect on our results. For example, the radii, $R_{200}^{\mathrm{grp,h}}$, of the groups in the cluster outskirts are on average 11% smaller in the high-resolution dark matter-only simulation, compared to the original dark matter-only simulation. However, due to the relatively small sample of groups in the outskirts of this one cluster, the difference is not significant ($\sim 1\sigma$).

Table 3.3 shows the number of galaxies within $0.5R_{200}^{\mathrm{clus,h}}$ of the centre of the cluster shown in Fig. 3.7, and the total number of galaxies found within $0.5R_{200}^{\mathrm{grp,h}}$ of group centres in the cluster outskirts, for each of the four simulations. This too is somewhat affected by increasing the resolution; the high-resolution dark matter-only simulation contains more subhaloes in the inner region of the cluster, and more subhaloes in the centres of surrounding groups. However, these increases that are caused by the greater dark matter resolution are less than the difference between the hydrodynamical and dark matter-only simulations. Moreover, the difference in resolution between the hydrodynamical and dark matter-only simulations is much less than a factor of eight (the difference between the standard and high-resolution dark matter-only simulations), so this slight difference in resolution cannot fully account for the differences between the hydrodynamical and dark matter-only simulations seen throughout this work. Finally, we note that slightly more groups are found in the cluster outskirts in the high-resolution dark matter-only simulation. When this is accounted for, the average number of galaxies within $0.5R_{200}^{\mathrm{grp,h}}$ of the centre of a group is actually unaffected by resolution.

This demonstrates that our findings are an effect of the inclusion of baryonic material, rather than a resolution effect, and the differences between the four simulations of this same cluster are in fact visible in Fig. 3.7. The main features of the cluster are similar in each of the simulations, although they show some variation, as these plots focus on the central cores of the clusters ($r < 0.7R_{200}^{\mathrm{clus,h}}$). Nevertheless, there are visibly more compact dark matter haloes in the central regions of the top two panels (showing the hydrodynamical runs), particularly very close to the cluster centres. There is very little difference in the substructure produced when increasing the resolution or decreasing the softening length.

3.4 Discussion & implications

The difference in the distribution of galaxies between hydrodynamical and dark matter-only simulations clearly has implications for the study of galaxy groups and clusters via simulations. This difference could have three main causes. The inclusion of gas and baryonic physics in the simulations may result in different substructure forming, both within the cluster itself, and in groups in the cluster outskirts. Alternatively the difference could be due to systematic issues with the halo finder (in our case, AHF), which then impact the halo catalogue that it produces. Finally, the presence of baryonic material may alter the properties of dark matter haloes, making them more likely to survive in certain environments.

Tidal stripping can cause the mass of a dark matter halo to decrease on entering a larger halo, whilst having a more minor effect on the baryonic material at the centre of such haloes (Smith et al., 2016). Previous work (Knebe et al., 2011a) has also indicated that the ratio of stellar mass to halo mass is greater for backsplash galaxies, which have experienced the environmental effects of a cluster in their past, indicating the stripping of their dark matter haloes (which also agrees with our findings in Chapter 5). Furthermore, tidal effects were partly responsible for the over-merging problem, as seen in early simulations (Moore et al., 1998). If these tidal effects are enhanced in dark matter-only simulations relative to hydrodynamical simulations, this could result in a drop in the number density of galaxies in denser regions, causing the deficit of galaxies seen in the centres of groups and clusters.

However, there are a number of issues with this. As tidal effects are a result of the gravitational potential of host groups and clusters, they are present in hydrodynamical simulations, as well as other stripping mechanisms such as ram pressure stripping (Arthur et al., 2019). This would cause gas to be stripped from galaxies in groups and clusters, as well as dark matter, potentially leaving a fully stripped galaxy core that consists mostly of stellar material. Such objects would not be found in a dark matter-only simulation, and so would indeed lead to a deficit of galaxies. However, as described in Section 2.2.2, we remove all galaxies from the hydrodynamical simulations that have more than 30% of their mass contained in stellar particles, corresponding to about 1% of all objects within $5R_{200}^{\mathrm{clus,h}}$ of the clusters. This would remove these heavily stripped galaxies, meaning that they would be absent in both the hydrodynamical and dark matter-only simulations.

Furthermore, we note that the deficit in subhaloes in dark matter-only groups is similar at all velocities in the group centre. This includes high-velocity galaxies which are likely near pericentre of a radial orbit, moving quickly from the group outskirts into its dense central region, as well as low-velocity galaxies, which are on roughly circular, bound orbits near the middle of the group. Assuming an inside-out formation history of groups (van der Burg et al., 2015) implies that these low-velocity galaxies joined their host group at an earlier time in its history, when it was less massive, and have since settled into virialised orbits. Consequently, we would expect these low-velocity galaxies to be less affected by tidal effects than those entering the group at a later time. This is not the case, as we see that galaxies are more prevalent in the hydrodynamical simulations, regardless of their velocities. Additionally, the right panel of Fig. 3.5 shows that very low-velocity central galaxies ($r < 0.3R_{200}^{\mathrm{grp,h}}$, $v < 0.5v_{\mathrm{cir}}$) are almost completely absent in dark matter-only simulations, despite these galaxies being less affected by tidal effects. This indicates that tidal effects are not playing a strong role in changing the number density of galaxies in group and cluster centres. However, the large deficit of extremely high-velocity haloes ($v \gtrsim 2v_{\mathrm{cir}}$) in the dark matter-only simulations could indeed be a result of tidal effects. Similarly, the (small) deficit of backsplash galaxies around our dark matter-only clusters could also be due to tidal effects, as backsplash galaxies are known to follow highly radial orbits and so experience strong tidal forces (Knebe et al., 2020).

In spite of this, previous work has found results analogous to ours, that have been largely attributed to tidal forces. Libeskind et al. (2010) compare two simulations of a system similar to the Local Group, simulated from the same initial conditions in both dark matter-only, and with full hydrodynamics. They too find that subhaloes are more concentrated in the centres of host haloes in their hydrodynamical simulation. The reason provided for this is that the dense baryonic region at the centre of hydrodynamical haloes restricts the tidal stripping of the dark matter halo. This leads to stronger dynamical friction in the hydrodynamical simulation, resulting in these objects being dragged towards the centre of their host, and increasing the galaxy number density. While this may partially explain the results in our work, it would also lead to a drop in the total masses of galaxy haloes in the dark matter-only simulations. In fact, we find that the cumulative halo mass functions for our hydrodynamical and dark matter-only clusters do not differ by more than 25%, between $M_{200} = 10^{10.5}\ h^{-1}M_{\odot}$ and $M_{200} = 10^{15}\ h^{-1}M_{\odot}$. This indicates that, although the mechanism described by Libeskind et al. (2010) may partially contribute to the results in this work, the effect does not seem to be strong enough to fully explain the differences in density that we find in groups and clusters.

An alternative explanation for the trends seen in this work is a numerical effect, that the inclusion of baryonic material will alter the effectiveness of halo finders that are used in simulations. The initial step of most galaxy halo finders relies on locating a dense halo centre, and then expanding from this to determine the extent of the galaxy halo. Consequently, the presence of a dense region of star and gas particles at the middle of a dark matter halo will benefit halo finders, as this will help a halo satisfy the conditions to become a seed for the halo finder to use. However, this non-physical explanation is unlikely to be the only cause of the difference between the simulations. Our third explanation, that the dense central region of baryonic material can also have a physical effect on the dark matter, is strongly supported by both this work and previous studies in the literature. As described by Blumenthal et al. (1986) and van Daalen et al. (2011), this dense baryonic region can increase the steepness of the dark matter density profile in a halo centre. This would indeed enhance the ability of a halo finder to detect the halo, but because of a physical difference in the simulations, not simply a numerical effect. The initial step in AHF involves using a refined grid to locate peaks in the density field (Knollmann & Knebe, 2009), and so if these peaks are less sharp in a dark matter-only simulation, detection of galaxies near to the centres of groups will be more challenging. These subhaloes may instead be included as part of the group halo, potentially contributing to the greater sizes of group haloes in dark matter-only simulations, as shown in Fig. 3.4.

In principle, the idea that haloes are easier to pick out in hydrodynamical simulations could be tested, by identifying individual particles in haloes in the hydrodynamical simulations, and examining where these particles were found in the dark matter-only simulations. If these particles are not found in dark matter-only haloes identified by AHF, but are found in an overdensity that we could identify by eye, it would indicate that the halo finder has failed to identify the halo. While this investigation is beyond the scope of this work, it would be a useful study to understand whether all substructure is affected by hydrodynamics in the same way, and to understand the relative contributions of the causes listed above in different regions of the simulations.

We stress that, due to the nature of most halo finders, this effect is not unique to the halo finder used in this work. The visible differences between the simulations in Fig. 3.7 show that the baryonic material is also having a physical impact on the dark matter, and so this is not purely a numerical effect. Similar effects would likely be observed with many other halo finders used widely in cosmological simulations. For example, Knebe et al. (2011b) examine 16 different halo finders in a dark matter-only cosmological volume, and find that the cumulative halo mass functions predicted by these agree to within $\sim 10\%$ over nearly four orders of magnitude, from $M_{200} = 2 \times 10^{11}\ h^{-1}M_{\odot}$ to $M_{200} = 10^{15}\ h^{-1}M_{\odot}$. Onions et al. (2012) find a similar result in a lower mass regime, instead studying the effectiveness of 11 halo finders at detecting subhaloes within a Milky Way-sized dark matter halo. They too find a variation of approximately 10% in the cumulative halo mass function, this time between $M_{200} = 6 \times 10^{6}\ h^{-1}M_{\odot}$ and $M_{200} = 10^{10}\ h^{-1}M_{\odot}$. They also demonstrate that this result still holds in the dense, central regions of the galaxy halo.

For the specific example of infalling galaxy groups that we have investigated in Section 3.3.2, we show that the tools typically used to analyse simulation data lead to different views of the composition of galaxy groups depending on the inclusion of baryonic material, which is not immediately obvious. This difference in the composition of groups will affect conclusions relating to their evolution. For example, Choque-Challapa et al. (2019) determine, amongst other results, the fraction of galaxies that become unbound from infalling groups in dark matter-only simulations, and how this depends on the position/velocity of galaxies relative to their group host. Our work adds a caveat to results such as these: we show that dark matter-only simulations underestimate the fraction of central, low-velocity galaxies, which are more likely to remain bound to infalling groups.

In Chapter 4, we build on the work in Section 3.3.2, using hydrodynamical simulations to investigate the infall of these galaxy groups, to study how their dynamical properties change over time.

3.5 Conclusions

In this chapter we have examined the differences in galaxy cluster substructure produced by hydrodynamical and dark matter-only simulations. We make this comparison by using a suite of hydrodynamical and dark matter-only simulations, to obtain two cluster samples that use the same cosmology, initial conditions, simulation codes and analysis. We then use the specific example of the phase space of infalling galaxy groups to investigate how the analysis of cluster simulations could be affected by these differences. Our findings are summarised below.

- Apart from the outskirts of galaxy groups and clusters, where the number density of bound galaxies is below ~ 1 Mpc^{-3}, dark matter-only simulations underestimate the radial number density profiles of galaxies in clusters, and in groups located in cluster outskirts. It is only at distances beyond $\sim 2R_{200}$ from the centres of groups/clusters that the profiles are indistinguishable.

- Closer to the centres of groups and clusters, the deficit of galaxies in dark matter-only simulations increases. At $r = 0.1R_{200}$ in clusters, the number density of galaxies is four times greater in hydrodynamical simulations. At the same distance in groups (scaled by R_{200}), the average number density is 10 times greater in hydrodynamical simulations.

- In galaxy groups that are entering a cluster, the deficit of galaxies in dark matter-only simulations is particularly pronounced when considering galaxies close to the group centre, with either high or low velocities relative to their group host. In some regions of the position-velocity phase space, there are up to five times as many galaxies in an average hydrodynamical group.

- The presence of a dense region of baryonic material in the centres of hydrodynamical haloes, and the increased central density of dark matter caused by this, means that galaxies produce a sharper peak in the density field within hydrodynamical simulations. The increased prominence of over-densities, and hence the increased contrast of galaxies against their host group, makes them easier for halo finders to detect in hydrodynamical simulations. The discrepancies between the two simulation types that we describe are a consequence of this.

The results from Section 3.3.2 show that infalling galaxy groups appear less compact when using dark matter-only simulations, compared to more physically motivated hydrodynamical simulations. This will affect the evolution of these groups, and the fate of group members after the infall of their host group, as previous work has shown that compact groups appear more likely to survive cluster infall (Choque-Challapa et al., 2019). We investigate the survival of galaxy groups during cluster infall in the following chapter of this thesis.

However, this work has wider implications for cosmological simulations. We show that the use of dark matter-only simulations, either as an approximation of a full-physics simulation, or as a framework for techniques such as semi-analytic modelling or halo occupation models, may need to be adjusted. Many semi-analytic models already account for similar effects, by including ‘orphan galaxies’ in their catalogues; these are subhaloes that have been heavily stripped or disrupted as they approach the centre of their host halo, and so can no longer be located in the simulation (Contini & Kang, 2015; Pujol et al., 2017; Cora et al., 2018). Similar methods have been used in halo occupation models, to find orphan galaxies and to adequately populate clusters with satellite galaxies (Carretero et al., 2015; Guo et al., 2016).

Despite this, such methods are not widely used in dark matter-only simulations, and as we show, this can potentially lead to different conclusions regarding the structure and composition of groups and clusters. As a minimum, the caveat that dark matter-only simulations produce incomplete halo catalogues in dense cosmological regions must be included. Further to this, corrections need to be made to work involving dark matter-only simulations, to account for the fact that observational surveys of groups and clusters will find a greater population of central galaxies than predicted by these simulations. These potential issues with dark matter-only simulations help to motivate our use of hydrodynamical simulations in the remainder of this thesis.

Chapter 4

Evolution of galaxy groups during cluster infall

Galaxy groups play an important role in the assembly of clusters, but it is not clear how the dynamical properties and structure of groups are affected when they interact with a large cluster, or whether all group members necessarily experience the same evolutionary processes. In this chapter, we use the hydrodynamical simulations from The Three Hundred project to study the properties of groups passing through a cluster. We find that galaxies become gravitationally unbound from groups very quickly, less than 1 Gyr after entering a cluster. Most groups quickly mix with the cluster satellite population, with just 8% of infalling group haloes later leaving the cluster, although nearly half of these have lost all of their member galaxies by this stage. The position of galaxies in group-centric phase space is also important – only galaxies near to the centre of a group ($r \lesssim 0.7R_{200}$) remain bound once a group is inside a cluster, and slow-moving galaxies in the group centre are likely to be tidally disrupted, or merge with another galaxy. This work will help future observational studies to better constrain the environmental history of group galaxies. For instance, groups observed nearby to clusters have likely approached very recently, meaning that their galaxies will not have experienced a cluster environment before. The content of this chapter forms a paper (Haggar et al. 2022) that has been submitted and is currently under review.

Contents

4.1 Introduction

Galaxies can enter a cluster via numerous different environments, resulting in the pre-processing of galaxies. For example, Kuchner et al. (2022) find that 45% of cluster galaxies are accreted via cosmological filaments which, like clusters, can quench star formation. This results in degeneracy, as it is not immediately clear whether cluster galaxies have been quenched by the cluster itself, or are quenched due to pre-processing.

As galaxies can also enter clusters as members of galaxy groups, these are another contributor to pre-processing, although the exact degree of groups' contribution is debated. Some simulations (McGee et al., 2009; Han et al., 2018) and observations (Dressler et al., 2013) find that close to half of all cluster members have been accreted as members of galaxy groups, while others (Arthur et al., 2019) find a much lower fraction. There are multiple explanations for this. For example, previous studies have shown that this fraction depends on the stellar mass of the accreted galaxies (De Lucia et al., 2012), and in Section 3.3.2 we showed that using hydrodynamical or dark matter-only simulations can also have an impact. Additionally, the definition of a 'galaxy group' is not standardised, and different definitions can lead to different conclusions. Various studies identify group members as galaxies that lie within the radius of a host group halo (Arthur et al., 2019; Donnari et al., 2021), that satisfy a boundness criterion (Han et al., 2018; Choque-Challapa et al., 2019), or by using a Friends-of-Friends algorithm (Benavides et al., 2020), all of which can result in different selections of group members. Furthermore, Berrier et al. (2009) find that, although 30% of cluster members are accreted via group haloes, half of these 'groups' only contain two or three galaxies. Clearly, the minimum (and maximum) size of what constitutes a group is also an important consideration.

Both theoretical and observational studies have shown that the effects of a group environment on the evolutionary processes in galaxies can be enhanced even further when a group enters a cluster. Galaxy mergers (Benavides et al., 2020) and gas removal (Pallero et al., 2019; Kleiner et al., 2021) are common in infalling groups, and multiple studies have connected this galaxy evolution to the external forces acting on a group, such as the effects of large-scale structure and clusters. Vijayaraghavan & Ricker (2013) used cosmological simulations to show that mergers, ram-pressure stripping, and tidal truncation of galaxy haloes are all enhanced further when their groups enter clusters, for a variety of reasons – for example, the intra-group medium is shocked during a group-cluster merger, increasing its density and thus increasing the ram-pressure stripping of the group members. Similar mechanisms have been described in previous works, such as Mauduit & Mamon (2007), who showed that galaxies passing through shocked regions in cluster-cluster mergers also experience increased ram pressure. In a related observational study, Roberts & Parker (2017) found that dynamically relaxed groups, which are typically isolated and slowly-growing, contain a smaller fraction of star-forming galaxies than unrelaxed groups. Again, this indicates that the processing of galaxies in groups is dependent on the disturbance of these groups by the larger environment in which they are located (see also Gouin et al., 2021).

All of this means that galaxies that have joined clusters as members of a group have experienced different evolutionary processes to those that have joined as individuals. Bahé et al. (2019) use the Hydrangea suite of hydrodynamical simulations (Bahé et al., 2017), and find that only $\sim 50\%$ of infalling group members survive to $z = 0$ once they have entered a cluster – in contrast, they find that more than 90% of galaxies that have not experienced any pre-processing survive to $z = 0$. This survival fraction is higher than in some other studies, although much of the prior work in this field uses dark matter-only simulations (e.g. Gill et al., 2004b), in which substructure can be more easily stripped (Smith et al., 2016). The results of Bahé et al. (2019) show that group members are particularly strongly influenced within clusters, and that they can be very heavily disturbed during accretion onto a cluster. Moreover, previous work has hinted that galaxy groups themselves can be heavily disrupted when entering a cluster. Gonzalez-Casado et al. (1994) showed that tidal forces from clusters can rapidly increase the internal energy of infalling groups, by up to a factor of 10 for the smallest groups, allowing these groups to be disrupted. Furthermore, Choque-Challapa et al. (2019) find that, using dark matter-only simulations and a similar group definition as is used in this work, over 90% of group members become unbound after this group enters a cluster, and that these galaxies quickly form part of the cluster population of galaxies.

However, beyond this, there is little work that has examined in detail how the dynamics of galaxy groups evolve when they are accreted by a cluster, particularly with large numbers of clusters in hydrodynamical simulations. While previous studies have looked at the overall disruption of groups that enter a cluster and the subsequent ‘post-processing’ of their constituent galaxies, we do not currently have a detailed understanding of the timescales over which groups change, and how the evolutionary processes that galaxies experience are affected by the group dynamics (Cohn, 2012; Bahé et al., 2019).

In this chapter, we use the same hydrodynamical simulations used in the previous chapter to study the evolution of groups as they enter galaxy clusters, and the processes that galaxies in these groups experience in their subsequent passage through the cluster halo. Specifically, we look at how the phase space of groups evolves: that is, how the positions and speeds of galaxies change, relative to the group that they are bound to. We make comparisons between groups before and after they pass through a cluster, to find the cumulative effect that a cluster has on the dynamics and structure of galaxy groups. Then, we look at the fates of group galaxies, categorising them based on the processes they experience in the several Gyr after entering a cluster (such as mergers and stripping), and how this depends on the structure of groups. Finally, we discuss how this theoretical work can help observational studies.

This chapter is structured as follows: In Section 4.2 we introduce some definitions, and the methods that we use to analyse groups. In Section 4.3 we show how the internal dynamics of groups change as they pass through a cluster, and in Section 4.4 we focus on the state of galaxies and groups after passing through a cluster. Finally, we summarise our findings in Section 4.5.

4.2 Methodology

In this chapter, we identify infalling groups in the same way as described in Section 3.2.2 and Section 3.2.3, by finding galaxies that are 'bound' to a host halo at the moment of cluster infall, according to Eq. 3.1. Similarly to in Chapter 3, we place a lower limit of $M_{200} \geq 10^{10.5}\ h^{-1}M_{\odot}$ on the total mass of haloes and subhaloes, and require that they contain more than 30% of their mass in stars, as described previously in Section 2.2.2. Additionally, in this chapter and Chapter 5, we also include a lower limit on the stellar mass of these objects, of $M_{200} \geq 10^{9.5}\ h^{-1}M_{\odot}$. From Section 4.3.2 onwards, we look at groups that are entering a cluster for the first time, and subsequently pass through the cluster – Fig. 3.1 in the previous chapter shows a schematic diagram of this.

We study galaxy groups with between five and 50 members (including the host object) that each satisfy our mass constraints. Groups of this size are considered to be small or intermediate sized groups (Tully, 2015), but are large enough to provide an environment that can strongly impact galaxy evolution (Hester, 2006). These groups have an average mass at cluster infall of $10^{13.5\pm0.4}\ h^{-1}M_{\odot}$ (median and 1σ spread). This means that the typical mass ratio between the group and cluster is roughly $1:20$, although this varies across the range of group and cluster masses, from approximately $1:5$ to $1:100$. The upper limit of 50 members on the group size has been imposed to ensure that major cluster-cluster mergers are not included in this study: we are primarily interested in the impact of clusters on the smaller group structures they accrete, rather than violent $1:1$ mergers that can completely disrupt a cluster's structure. The work of Contreras-Santos et al. (2022) uses The Three Hundred simulations to study major mergers, and so we refer the reader to this study for a detailed discussion of these events.

4.2.1 Tidal radius of groups

Subhaloes passing through a larger halo can experience strong tidal stripping, and group-sized haloes can often lose a large fraction of their mass due to stripping from a cluster (Muldrew et al., 2011; Bahé et al., 2019). Similarly, galaxies can be tidally stripped from these groups (Gonzalez-Casado et al., 1994; Choque-Challapa et al., 2019), although the extent of this stripping varies between different studies. For instance, Vijayaraghavan et al. (2015) find that the central regions of galaxy groups are largely unaffected by a cluster potential, and are only disrupted by dynamical friction after several Gyr.

The tidal radius of a group or dark matter halo is an effective way to predict and explain tidal stripping. Generally, the tidal radius is defined as the distance from a smaller object at which the self-gravity of that object is less than the tidal force due to a larger object. However, the tidal radius actually comes from a tidal tensor, which is a more complex property that depends on multiple factors, including the density profiles of the structures involved and the orbit of a satellite around a larger object (Read et al., 2006). Because of this, the tidal radius can be calculated differently for different scenarios, to best describe the tidal forces acting on an object (see van den Bosch et al., 2018, for a detailed summary). Perhaps the simplest example is the Roche limit, the tidal radius of a point mass that is being tidally influenced by another point mass. More physically motivated scenarios such as an extended subhalo within a larger extended halo (as is used in this work) require more complex descriptions.

Calculating a tidal radius is complicated further by the fact that subhalo properties are often poorly defined by a halo finder, and can be strongly dependent on the distance of a subhalo from the group centre. Muldrew et al. (2011) test the ability of AHF and another halo finder, SUBFIND (Springel et al., 2001), to recover subhalo properties. They find that AHF performs better at identifying all the particles of a subhalo, and thus constrains the subhalo mass more effectively. However, for subhaloes within $\sim 0.5R_{\rm vir}$ ($\sim 0.7R_{200}$)[1], both halo finders underestimate the number of particles in the subhalo. This is due to the greater background density near the centre of a large halo, which makes the overdensities of subhaloes less pronounced. In particular, the density in the outskirts of a subhalo will be comparable to the local density of the host halo, meaning that the edge of the subhalo will not be clear. This results in some of the subhalo particles being mistakenly labelled as belonging to the main halo, truncating the subhalo. Because of this, it is challenging to predict the mass, and therefore the radius, of subhaloes in these regions. Furthermore, in our work we wish to combine the data from multiple galaxy groups (of different sizes) in multiple galaxy clusters (also of different sizes). It is therefore convenient to have an expression for the group tidal radius that is independent of the cluster or group size, and solely depends on the separation between these two.

We define the tidal radius of an infalling subhalo by adapting the descriptions in Klypin et al. (1999) and van den Bosch et al. (2018). Specifically, they give the tidal radius in terms of a function, $f(d)$, whose value is the minimum of two expressions:

$$R_{\rm t} = d\left(\frac{M^{\rm grp,h}_{\rm <d}(R_{\rm t})}{M^{\rm clus,h}_{\rm <d}(d)}\frac{1}{2-f(d)}\right)^{\frac{1}{3}}, \tag{4.1}$$

$$f(d) = \min\left[1,\ \left.\frac{{\rm d}(\ln M^{\rm clus,h}_{\rm <d})}{{\rm d}(\ln d)}\right|_{d}\right]. \tag{4.2}$$

Here, $R_{\rm t}$ is the tidal radius of the group, d is distance from a group to the cluster centre, and M are the radial enclosed mass profiles of the group and the cluster. We assume the radial density of the dark matter haloes follow an NFW profile (Navarro et al., 1996), as given previously in Eq. 1.2. The concentration of a halo, c, is equal to the ratio between R_{200} and the scale radius of the halo, $R_{\rm s}$:

$$R_{\rm s} = \frac{R_{200}}{c}. \tag{4.3}$$

Integrating the NFW profile, Eq. 1.2, gives the enclosed mass in a sphere of radius d:

[1] Muldrew et al. (2011) use the definition of virial radius presented in Bryan & Norman (1998). For their cosmology, the mean density of a halo within the virial radius is $101\rho_{\rm crit}$, where $\rho_{\rm crit}$ is the critical density of the Universe. Hence, for the clusters used in their work and ours, $R_{\rm vir} = R_{101} \approx 1.3R_{200}$, although it is important to note that this conversion depends on the concentrations and density profiles of dark matter haloes.

$$M_{<\mathrm{d}} = 4\pi\rho_0 R_\mathrm{s}^3 \left[\ln\left(\frac{d+R_\mathrm{s}}{R_\mathrm{s}}\right) + \frac{R_\mathrm{s}}{d+R_\mathrm{s}} - 1\right] . \tag{4.4}$$

This can then be used to rewrite Eq. 4.2. For a general NFW profile, $f(d) = 1$ in the region $d \lesssim 2.2R_\mathrm{s}$. However, $f(d) < 1$ outside of this region, and so must be calculated for each subhalo. Solving the derivative in the expression of $f(d)$ gives:

$$f(d) = \min\left[1, \left(\frac{x^2}{x+1}\frac{1}{(x+1)\ln(x+1) - x}\right)\right] . \tag{4.5}$$

Here, we define the quantity x to make the equations in this section more easily readable:

$$x = \frac{d}{R_\mathrm{s}} , \tag{4.6}$$

where d is the distance from the halo centre, and R_s is the scale radius of the halo, as used to define the NFW profile in Eq. 1.2.

Using the fact that $M_{<\mathrm{d}}(d = R_{200}) = M_{200}$, we can produce an expression for M_{200} using Eq. 4.4. Substituting this into Eq. 4.4 gives

$$M_{<\mathrm{d}} = M_{200}\left[\ln(1+x) - \frac{x}{1+x}\right] \times \left[\ln(1+c) - \frac{c}{1+c}\right]^{-1} . \tag{4.7}$$

This expression can then be substituted into the equation for tidal radius, Eq. 4.1, for the cluster enclosed mass, $M_{<\mathrm{d}}^{\mathrm{clus,h}}(d)$, and for the mass enclosed within the tidal radius of a group, $M_{<\mathrm{d}}^{\mathrm{grp,h}}(R_\mathrm{t})$. This gives the expression for tidal radius below,

$$\begin{aligned}\frac{R_\mathrm{t}}{R_{200}^{\mathrm{grp,h}}} = {} & \frac{d}{R_{200}^{\mathrm{clus,h}}}\left(\frac{1}{2-f(d)}\right)^{\frac{1}{3}} \\ & \times \left(\frac{\left[\ln\left(1+\frac{C_\mathrm{c}d}{R_{200}^{\mathrm{clus,h}}}\right) + \left(1+\frac{C_\mathrm{c}d}{R_{200}^{\mathrm{clus,h}}}\right)^{-1} - 1\right]}{\left[\ln(1+C_\mathrm{c}) - \frac{C_\mathrm{c}}{1+C_\mathrm{c}}\right]}\right)^{-\frac{1}{3}} \\ & \times \left(\frac{\left[\ln\left(1+\frac{C_\mathrm{g}R_\mathrm{t}}{R_{200}^{\mathrm{grp,h}}}\right) + \left(1+\frac{C_\mathrm{g}R_\mathrm{t}}{R_{200}^{\mathrm{grp,h}}}\right)^{-1} - 1\right]}{\left[\ln(1+C_\mathrm{g}) - \frac{C_\mathrm{g}}{1+C_\mathrm{g}}\right]}\right)^{\frac{1}{3}} ,\end{aligned} \tag{4.8}$$

where C_c and C_g are the concentrations of the cluster and group haloes, respectively, and $f(d)$ is given by Eq. 4.5. Finally, we take the halo concentrations to be constant for all of the clusters, and all of the groups. Specifically, we set the value of $C_{\rm c}$ equal to the median value for our clusters, $C_{\rm c} = 3.9$, and $C_{\rm g}$ equal to the median value for our groups, $C_{\rm g} = 4.4$. Approximating these concentrations as constant has a small effect because Eq. 4.8 is not strongly dependent on them. For a group at a distance $d = 0.2R_{200}^{\rm clus,h}$ from the cluster centre, the value of $R_{\rm t}$ varies from its median value by 20% across the full range of cluster concentrations (from $C_{\rm c} = 2.3$ to $C_{\rm c} = 7.7$). At greater distances from the cluster centre, this variation is even smaller. Similarly, the 1σ deviation[2] in $C_{\rm g}$, between 2.6 and 6.9, leads to a variation in $R_{\rm t}$ of less than 18%. This variation is used as the uncertainty in the tidal radii that we calculate later, in Section 4.3.2 and Fig. 4.4.

By making these assumptions, we are able to reach an expression for the tidal radius of a group in units of $R_{200}^{\rm grp,h}$ that depends only on the distance from the group to the cluster centre. As $R_{200}^{\rm grp,h}$ of a group can change over the course of infall, the tidal radius could be scaled by this changing group radius. However, we instead choose to scale the tidal radius by $R_{200}^{\rm grp,h}$ at the moment of cluster infall, to allow us to stack groups and study their evolution more clearly – this is explained in further detail in Section 4.3.2.

Eq. 4.8 is an ideal form of the tidal radius for our analysis, as it allows us to calculate the average tidal radius for all groups in a radial bin across many clusters. This form of the tidal radius may also be useful in future studies, both observational and theoretical, that wish to stack substructure on multiple different size scales.

4.3 Phase space evolution

Much of the work in this chapter revolves around studying the phase space of galaxies within galaxy groups, as the groups enter and pass through a cluster, and how the distribution of galaxies within this phase space changes over time. This chapter builds on the analysis of the phase space of galaxies in infalling groups from Chapter 3. In that chapter, the phase space consists of the radial distance of a galaxy from its host group halo in terms of the group halo radius, $R_{200}^{\rm grp,h}$, and the galaxy's velocity relative to the group halo, in units of $v_{\rm cir}$, the circular orbital velocity at $r = R_{200}^{\rm grp,h}$. It is important to stress that this work involves looking at the phase space of galaxies relative to their host group, not the cluster (as has been done by numerous previous studies, such as Jaffé et al., 2015; Arthur et al., 2019, for example). This method can provide detailed information, by showing both the spatial and velocity distribution

[2]There are a small number of groups with highly concentrated haloes, so we use the 1σ spread in $C_{\rm g}$ to avoid skewing our data.

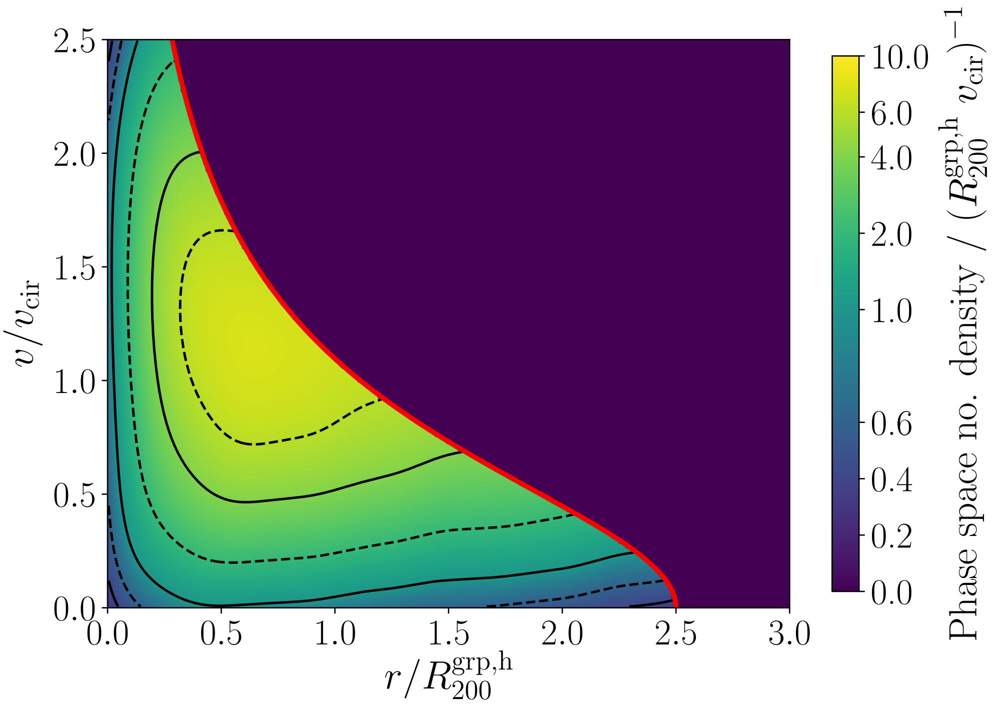

Figure 4.1: Distribution of galaxies in group phase space, for groups at moment of infall into galaxy cluster. Data for all groups from all 257 clusters that are used in this analysis are shown, stacked together. Lighter colours represent regions of phase space with more galaxies – the maximum value is at $r = 0.65R_{200}^{\rm grp,h}$, $v = 1.15v_{\rm cir}$, representing the region of this phase space in which group members are most likely to be found. The red line represents the boundness criterion for galaxies, Eq. 3.1; galaxies above this line are not considered group members, and so are excluded from this figure. Contours are at densities of [0.2, 0.4, 0.6, 1, 2, 4, 6] $(R_{200}^{\rm grp,h} v_{\rm cir})^{-1}$.

of galaxies in groups, and telling us how the speed and acceleration of galaxies differ in different regions of the group.

Fig. 4.1 shows the average distribution of galaxies in phase space, for an infalling group – a group that has just passed within $R_{200}^{\rm clus,h}$ of the cluster centre for the first time, as shown in Fig. 3.1. Similarly to in Chapter 3, we produce a smoothed distribution of galaxies using a 2D KDE with an optimised bandwidth. This plot is very similar to the left panel of Fig. 3.5, which also shows the distribution of galaxies in infalling hydrodynamical groups. The two are nevertheless slightly different, due to the inclusion of a stellar mass cut in this chapter, which was not used in the previous chapter (see Section 2.2.2). In the remainder of this section, we examine how this phase space changes as a group passes through a cluster.

Overall, we identify 1340 infalling groups across the 257 clusters, with a median size of 8^{+7}_{-3} members (median and 1σ spread). This indicates that, although we permit groups to contain up to 50 members, groups of this size are rare compared to the large number of smaller groups – only 8% of the groups contain more than 20 members. These groups enter the cluster over a wide range of redshifts, with a median value of $z_{\rm infall} = 0.4^{+0.6}_{-0.3}$.

Our 1340 groups represents an average of 5.2 accreted groups per cluster – this might appear to be a small number, but it is important to note that this is not the entire accreted group population, as this only accounts for intermediate size groups (with between five and 50 members). Berrier et al. (2009) demonstrate that about half of galaxy groups contain only two or three members, and such groups are not included in our analysis. If we do include these smaller groups, we find that approximately 14% of $z = 0$ cluster galaxies in our simulations were accreted as members of a group, comparable to the results from other studies presented in Section 4.1 (see also Section 3.3.2).

4.3.1 Groups beyond the cluster outskirts

Before studying groups passing through clusters, we first study how this phase space changes in groups that are not under the influence of a cluster, and are located far from the cluster centre (greater than $3R_{200}^{\rm clus,h}$ from the cluster). This can then be used as a control, showing how the distribution of galaxies changes for a group evolving secularly, as an (approximately) isolated system. Fig. 4.2 shows the direction and rate at which galaxies in groups move around this phase space, for groups between $3R_{200}^{\rm clus,h}$ and $10R_{200}^{\rm clus,h}$ from the centre of a cluster, between redshifts of $z = 0.1$ and $z = 0$. We assume that these groups are isolated, as they are sufficiently far from a cluster that they are not subject to its strongest effects. We did not study groups at greater cluster distances because the resolution of the simulations decreases outside of this distance. This figure includes only galaxies that were bound to a group (i.e. that lay below the thick red line) at $z = 0.1$, which we then follow until $z = 0$ (about 1.3 Gyr).

Throughout this section, in order to study how the speeds and positions of group galaxies change, we examine the changes of these properties for bound group members, relative to $R_{200}^{\rm grp,h}$ and $v_{\rm cir}$ of their host group measured at a previous time. In Fig. 4.2 and the following figures in Section 4.3.2 we show how the phase space of groups changes over time. In these plots, the direction of arrows shows the average direction that galaxies in this region are moving in phase space, and darker arrows mean that the galaxies are moving across the phase space more quickly. For example, a galaxy going from $[1.0R_{200}^{\rm clus,h}, 0.5v_{\rm cir}]$ to $[2.0R_{200}^{\rm clus,h}, 1.5v_{\rm cir}]$ in 2 Gyr would be represented by an arrow located at $[1.5R_{200}^{\rm clus,h}, 1.0v_{\rm cir}]$, pointing at a 45° angle to the top-right, with a colour of ~ 0.71. Note that the horizontal and vertical axes in Fig. 4.2 are dimensionless, as they have been normalised to prior values of $R_{200}^{\rm grp,h}$ and $v_{\rm cir}$, and so we describe the distance moved across this phase space in a given time with the term 'virial units per Gyr'.

The positions and velocities of the galaxies in Fig. 4.2 are scaled relative to $R_{200}^{\rm grp,h}$ and $v_{\rm cir}$ of each group at $z = 0.1$. Some regions of phase space do not contain any arrows because of a lack of data, indicating that almost no galaxies were found in this region across all the groups – for example, there are no galaxies in the top-right of Fig. 4.2, because they were all below the red line just a short time previously. Despite this, some galaxies are found above the line, because they have become unbound between $z = 0.1$ and $z = 0$.

The phase space of these groups is not in equilibrium, and bound galaxies in the centres of these groups appear to be moving downwards on this plot (i.e. losing speed but remaining a similar distance from the group centre). This indicates that energy is being dissipated during their orbits. Dynamical friction is strongest in the group centres, and so this is likely responsible for the loss of energy during these orbits. Fig. 2 in Arthur et al. (2019) shows analogous behaviour to this for the phase space of a galaxy cluster: subhaloes move horizontally in phase space when approaching the centre of their host halo, then move sharply downwards when they are near to the halo centre, resulting in the apparent 'spiral' motion of galaxies in Fig. 4.2.

This trend could also potentially be explained by the destruction of some inner galaxies by mergers before they have time to leave the group centre. However, we find that this is not the case, as the majority (82%) of galaxies within $0.5R_{200}^{\rm grp,h}$ of the group centre survive to $z = 0$ without merging into another halo or being heavily stripped (see Section 4.4.3 for further discussion of the fates of group members). Furthermore, if we remove these galaxies from our analysis, there is a negligible change in the trends in Fig. 4.2.

4.3.2 Groups passing through clusters

To study groups falling into clusters, the phase space diagrams that we present are instead scaled relative to $R_{200}^{\rm grp,h}$ and $v_{\rm cir}$ of each group at $z_{\rm infall}$, the moment of cluster infall. We scale the positions and speeds of galaxies by these values, even in subsequent snapshots after $z_{\rm infall}$. This approach is not perfect, because the radius and circular velocity of a host group halo changes as the group approaches the centre of a cluster, likely due to tidal stripping. Despite this, we choose to measure these

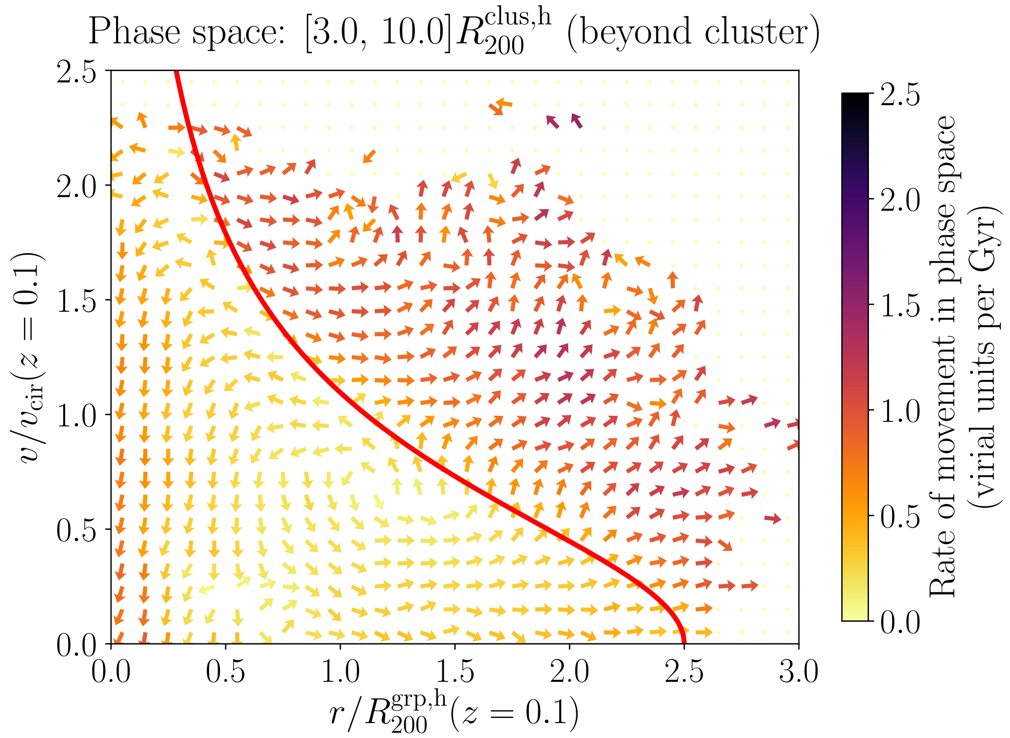

Figure 4.2: Motion of group members in phase space of host group, for groups located beyond the influence of a cluster. Thick red line shows boundness criterion, providing an approximate measure of galaxies that have become unbound from their group. Colours of arrows represent the rate at which galaxies are moving in this phase space. This plot shows stacked data for 2769 groups, located between $3R_{200}^{\mathrm{clus,h}}$ and $10R_{200}^{\mathrm{clus,h}}$ from the centre of a cluster, between $z = 0.1$ and $z = 0$. This shows how galaxies move in the phase space of groups when the group is not affected by the external environment. All galaxies lie below the bounded line at $z = 0.1$, but some move above the red line and become unbound, although many remain bound to the group. The motion of the bound galaxies follows a characteristic pattern, rather than being in random directions.

properties only at the moment of infall because, in the central regions of a large halo, the mass and radius of a subhalo are not well-defined; due to the high background density in the centre of the cluster halo, it can be challenging for a halo finder to identify the overdensity of a subhalo. Consequently, near the centre of a cluster, the mass and radius of a group ($M_{200}^{\mathrm{grp,h}}$ and $R_{200}^{\mathrm{grp,h}}$) are not reliable (Muldrew et al., 2011). Scaling by the values of $R_{200}^{\mathrm{grp,h}}$ and v_{cir} at z_{infall} allows us to visualise how the absolute values of the distance and speed of galaxies relative to their groups are changing. This means that galaxies lying below the line of boundness after infall are not strictly bound to the group, but the approach still provides a good approximation.

In this section we consider groups that are entering the cluster for the first time, and so have not previously experienced a cluster potential. We also only include groups at times between their first infall and their first apocentric passage (the turnaround in their cluster orbit). It is important to note that this is not necessarily the true 'first apocentre' of an orbit, as haloes are not accreted onto clusters in perfectly radial orbits. Instead, they have some tangential component to their velocity, meaning that some haloes will pass an apocentre before their entry to the cluster (Ghigna et al., 1998; Tollet et al., 2017). However, as the focus of this chapter is on the evolution of groups after their cluster infall, we will hereafter refer to the first apocentric passage post-infall as the 'first apocentre'. Finally, we do not separate groups by redshift – for example, some of these groups have passed their first pericentre by $z = 0$, but some have not and so are absent from the post-pericentre analysis.

Fig. 4.3 shows how the phase space of these groups changes as they pass through a cluster. We find that the behaviour of groups as they enter and pass through a cluster can be approximately split into two main phases, with the group dynamics changing suddenly as a group makes its closest approach to the cluster centre, as shown by the two panels in this figure. The left panel of Fig. 4.3 shows groups on their infall, moving from the cluster outskirts towards their first pericentric passage, near to the cluster centre. Generally, the galaxies in these groups move upwards on this plot, showing an increase in their velocity relative to their host group. These data are for galaxies that are bound to groups at the moment of infall (z_{infall}), but some of these move above the red line and so become unbound from their host group. Similarly to in Fig. 4.2, the direction of arrows shows the average direction that galaxies are moving in phase space, for galaxies in this region of phase space. It is important to note that the arrows in the left panel ('pre-pericentre') are pointing vertically upwards, with a very small horizontal component. This shows that, although these galaxies have a large change in speed, their distance to the group centre does not change very much; galaxies within $R_{200}^{\mathrm{grp,h}}$ remain within $R_{200}^{\mathrm{grp,h}}$.

This behaviour is different for groups that have passed the pericentre of their orbit, shown in the right panel of Fig. 4.3. This panel shows data for groups at snapshots when they have passed pericentre, but have not yet reached their first apocentre, and so are receding from the cluster centre. Groups are also only included here at stages of their orbit when they are between the cluster centre and $R_{200}^{\mathrm{clus,h}}$, to allow us to compare the two panels in Fig. 4.3. Instead of increasing their velocity, most galaxies in these post-pericentre, receding groups keep a fairly constant relative velocity, and instead move horizontally on this plot, becoming spatially separated from the centre of their host group. This behaviour is stronger for galaxies that are above the boundness line, meaning they have become unbound from the group – these move to greater distances from the group centre, often with an accompanying slight increase in relative speed. Galaxies that are still bound to a group instead experience a drop in their relative speed, as well as an increase in separation from the group centre.

In summary, Fig. 4.3 shows that there are two phases of evolution for a galaxy group passing through a large cluster. First, galaxies are given a kinetic energy kick, increasing their speed relative to their host group. This rapid boost in kinetic energy is manifested after the group passes pericentre, which typically occurs ~ 0.5 Gyr after entering the cluster, by being converted into potential energy as the galaxies recede from the group centre.

Tidal effects and dynamical friction

In Fig. 4.4 we break down the results from Section 4.3.2 into individual steps, separating the infalling groups into bins based on their cluster-centric distance, both before and after passing pericentre. This gives a much more detailed view of how this phase space changes over the average course of a group through a cluster. We note that each panel does not represent an identical sample of groups, as most groups will not have a snapshot in all of these radial bins, and so these data represent the evolution of all groups that are found in this radial range. If we instead select only groups that have passed through each of these bins, there is only a minimal impact on our results, but large amounts of noise are introduced due to the small number of groups.

In each panel, the tidal radius (based on the approximations detailed in Section 4.2.1) for a group in the centre of this bin is also marked, in units of the group

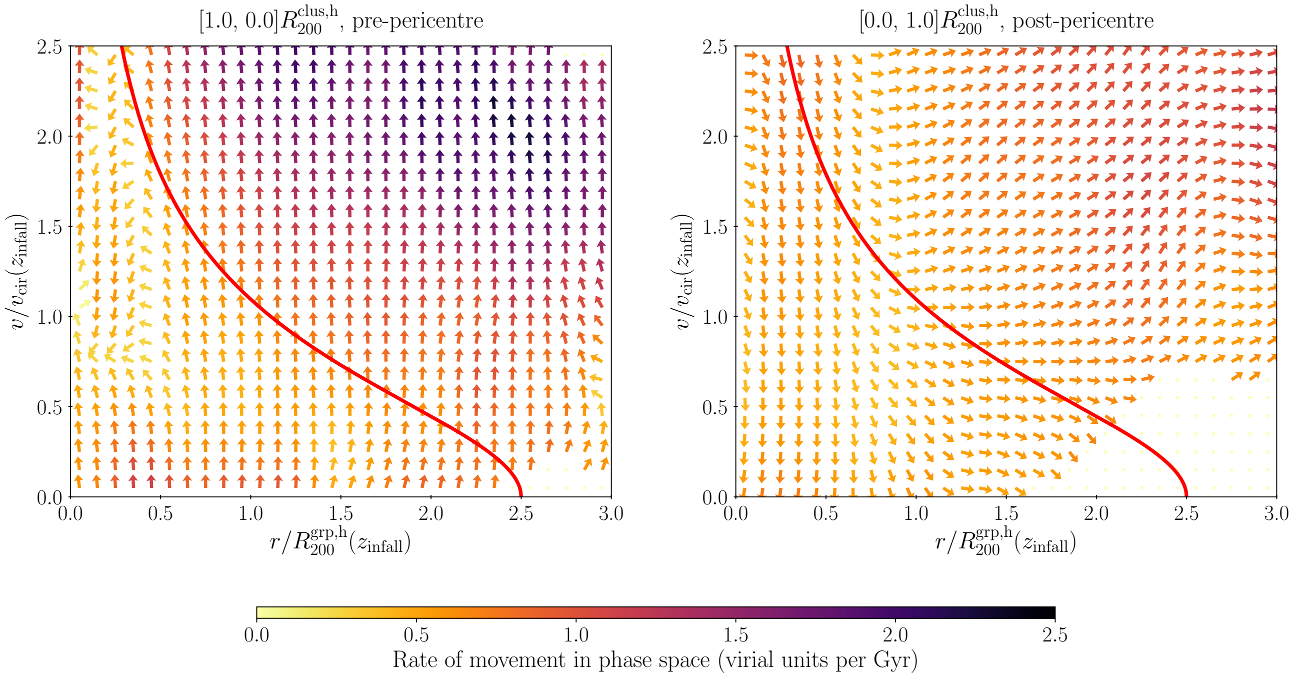

Figure 4.3: Same as Fig. 4.2, but showing motion of group members in phase space of host group, as the group moves through a galaxy cluster. Data shown are for galaxies that are bound to groups at the moment of infall, for groups on their first passage through the cluster. Red line shows boundness criterion at moment of infall, and so provides an approximate measure of galaxies that have become unbound from their group. The two panels show data for groups in two different stages of their orbit around the cluster: left panel shows data for groups before reaching their first pericentric passage of the cluster, moving between $R_{200}^{\mathrm{clus,h}}$ and cluster centre, and the right panel shows groups moving between the cluster centre and $R_{200}^{\mathrm{clus,h}}$, which have passed their pericentre and are now receding from the cluster, moving towards their first apocentric passage. For pre-pericentre groups, the bulk motion of the galaxies is upwards, representing an increase in their group-centric speed, but little change in the spatial separation of galaxies from their host group. In contrast, for groups that have passed pericentre and are now receding from the cluster, group members are moving approximately horizontally in phase space, increasing their distance from the group to which they were previously bound.

radius at infall. The closer a group is to the cluster centre, the stronger the effect of the cluster will be, and this is demonstrated by the decrease and subsequent increase of the tidal radius as groups pass through the cluster. Interestingly, across the eight panels, the tidal radius appears to mark a transition, such that the group dynamics evolve differently within the tidal radius, compared to beyond the tidal radius. Outside the tidal radius, galaxies first experience a kinetic kick and then recede from the group centre, as detailed in the previous section. However, inside the tidal radius, galaxies generally behave in a way similar to that seen in the centres of isolated groups in Fig. 4.2 – they mostly move downwards on these plots, showing a decrease in speed.

Physically, this distinction indicates how the dynamics in some regions of the group are dominated by the group itself, whilst others are dominated by the effects of the cluster. As described in Section 4.3.1, galaxies in isolated groups experience dynamical friction due to the group's halo (Vijayaraghavan et al., 2015). This is particularly strong in the dense central regions of the group, where dynamical friction will cause galaxies to slow down and spiral inwards, dominating over the effect of the cluster. However, beyond the tidal radius, tidal effects from the cluster dominate this dynamical friction, meaning that the movement of galaxies in phase space is dictated by the cluster, not the group. The change in the tidal radius means that the two phases of group evolution are clearer in the outskirts of a group, as the dynamics of these regions are dominated by the cluster for much of the group's journey. Conversely, galaxies in the group centres ($r < 0.5R_{200}^{\mathrm{grp,h}}$) decrease in velocity at almost all times, as they are almost always within the tidal radius. The only exception to this is in the very deepest parts of the cluster (such as in panel (d) in Fig. 4.4). Dekel et al. (2003) showed that at the very centre of a dark matter halo, tidal forces can become fully compressive – this could explain why all group galaxies change their orbits around their host groups, with their speeds increasing and their distances either remaining the same or decreasing.

Finally, as the change from an increase in v to an increase in r is dependent on the tidal radius, this switch in behaviour is not instantaneous as it might appear to be in Fig. 4.3. Once a group reaches a distance of approximately $0.3R_{200}^{\mathrm{clus,h}}$ beyond pericentre (panel (f) in Fig. 4.4), the motion of galaxies away from the group begins in the centre, and then spreads throughout the group as it once again dominates over the cluster. Eventually, for groups that are long past pericentre (panel (h) of Fig. 4.4), all galaxies are either decreasing in relative speed, or their speed is staying the same. All the galaxies remaining in groups at this stage are also moving away from the

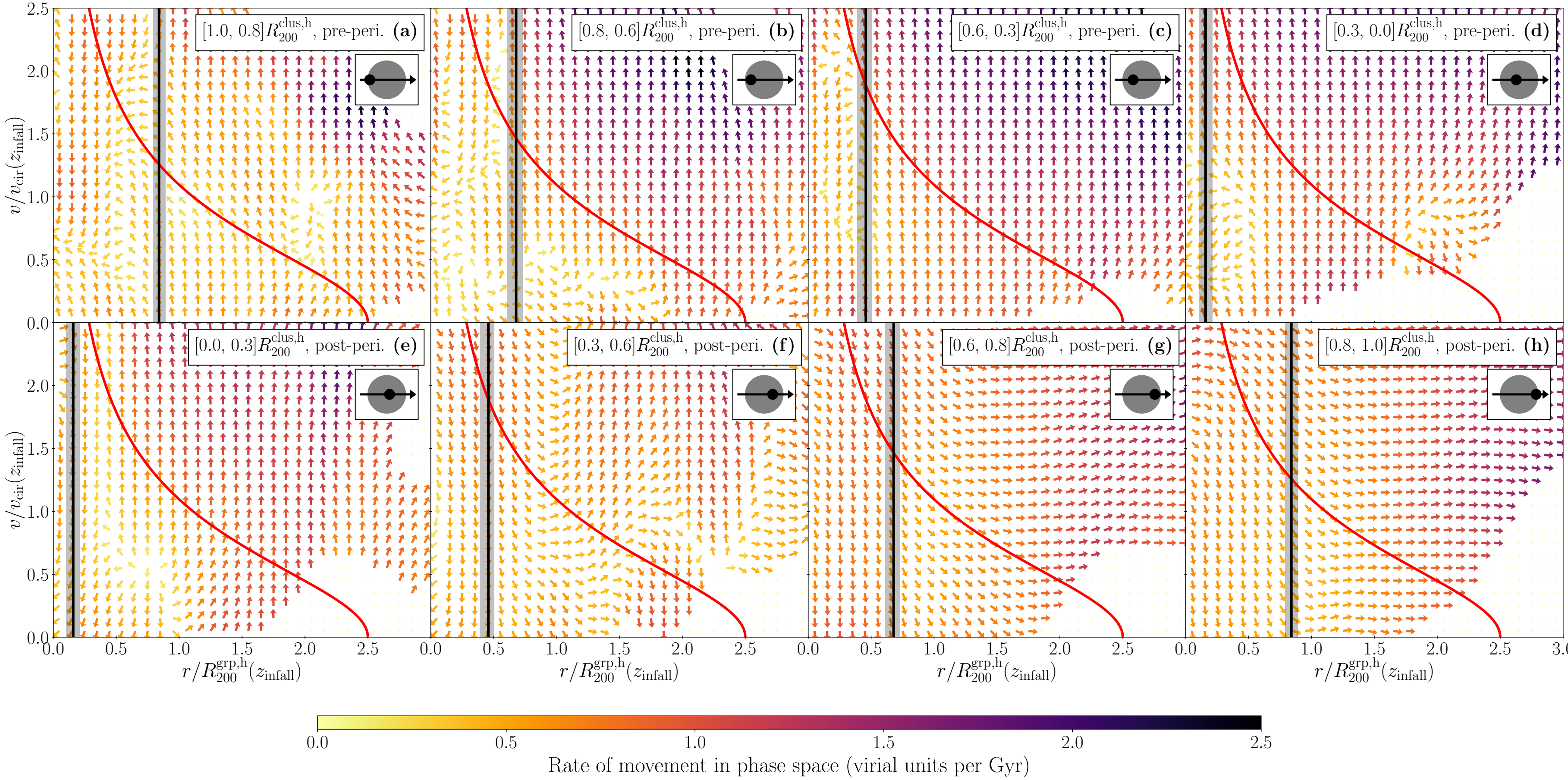

Figure 4.4: Same as Fig. 4.3, but showing motion of group members in phase space of host group at multiple stages of the group's passage through a cluster *(caption continued on next page)*.

Figure 4.4: *(caption continued from previous page)* Each panel shows groups at different stages of their journey, showing the groups as they enter a cluster, approach the cluster centre, and recede from the cluster out to a distance of $R_{200}^{\rm clus,h}$. Top row (panels (a)-(d)) shows groups before reaching pericentre, binned by their distance from the cluster centre. Bottom row ((e)-(h)) shows groups after passing pericentre, before reaching their first apocentre. Data are shown for groups on their first infall only. The red line shows boundness criterion at moment of infall, and so provides an approximate measure of galaxies that have become unbound from their group. Vertical black line on each plot represents the tidal radius for a group in this radial bin, and the shaded region represents the variation of this radius due to the 1σ spread in the concentrations of clusters and groups – these have median halo concentrations of $C_{\rm c} = 3.9^{+1.5}_{-1.1}$ and $C_{\rm g} = 4.4^{+2.5}_{-1.8}$, respectively. Colours of arrows represent the rate at which galaxies are moving in this phase space, with darker arrows indicating that galaxies are moving at a greater rate in this phase space. A small schematic is shown in the top-right of each panel, showing the point at which the groups (small black circle) are on their journey through the cluster halo (large grey circle). The transition between the two phases of group evolution shown in Fig. 4.3 can be seen, as well as the corresponding movement of the tidal radius discussed in Section 4.3.2.

group centre, towards the bottom-right of the phase space, which is characteristic of galaxies approaching the apocentre of a bound orbit around a group.

To help visualise this behaviour, Appendix B shows how group galaxies move around phase space for a single example group as it passes through a cluster, clearly showing the two main phases of group evolution.

4.4 Groups after cluster infall

The results in Section 4.3 show how the dynamics of galaxy groups change as they pass through a cluster. In this section, we discuss the differences in the properties of a group before and after it passes through a cluster, in order to understand how distinguishable are these two classes of groups.

4.4.1 Orbits of galaxy groups

The data used in Section 4.3 are for groups on their first passage through a cluster, but not all of these groups will follow the same path. Just as some galaxies that are accreted by a cluster can become 'backsplash galaxies' (which we discuss in Chapter 5), some groups will pass through the cluster and exit $R_{200}^{\rm clus,h}$ again, becoming 'backsplash groups' that can enter the cluster for a second time. Others will 'stick' to the cluster, remaining bound and not leaving $R_{200}^{\rm clus,h}$ once they have entered.

We find that, across the 257 clusters used in this work, most groups (92%) that fall into a cluster do not leave it again. By $z = 0$, only 42% of the groups that enter a cluster have reached their first turnaround (apocentre) in their cluster orbit, while 58% do not reach this stage. These groups do not reach apocentre for multiple reasons; either they have merged with the cluster halo before reaching apocentre rather than remain a substructure of the cluster, they have been heavily stripped by the cluster and so fall below the resolution limit before reaching apocentre, or they have simply not had time to reach apocentre by $z = 0$. Of the groups that do reach the apocentre of their orbit, 20% have left the cluster after entering $R_{200}^{\mathrm{clus,h}}$, while 80% reach apocentre within $R_{200}^{\mathrm{clus,h}}$ of the cluster centre, and so remain 'stuck' to the cluster potential. We hereafter refer to these as 'backsplash groups' and 'sticky groups', respectively. Overall, just 8% of all infalling groups go on to leave the cluster again. Finally, 81% of the backsplash groups in our sample later experience a second cluster infall, and 19% are still outside of the cluster at $z = 0$.

The paths that groups can take through a cluster can be described in terms of the distance from a group to the cluster centre at pericentre and apocentre. Interestingly, we find that the distance of a group halo from the cluster centre at pericentre is very consistent, regardless of the group's later behaviour. Groups that do not reach apocentre have a median pericentric distance of $0.36^{+0.14}_{-0.09}R_{200}^{\mathrm{clus,h}}$ from the cluster centre, which is very similar to the pericentre of groups that do later reach apocentre: backsplash groups and sticky groups have median pericentric distances of $0.38 \pm 0.13R_{200}^{\mathrm{clus,h}}$ and $0.36^{+0.11}_{-0.08}R_{200}^{\mathrm{clus,h}}$, respectively. This justifies our decision to normalise the figures throughout this chapter by the group radius at infall, as most groups pass well within $0.7R_{200}^{\mathrm{clus,h}}$, where Muldrew et al. (2011) showed that subhalo sizes cannot be reliably measured.

This shows that most groups take a similar trajectory into clusters, passing by the cluster centre at a similar distance. However, the subsequent orbits of these groups can vary dramatically, with groups reaching a wide range of apocentric distances, and some not being tracked to reach their apocentre at all. By definition, the post-infall apocentric cluster distances of backsplash groups and sticky groups are very different. Backsplash groups have a median apocentric distance of $1.16^{+0.29}_{-0.12}R_{200}^{\mathrm{clus,h}}$, and sticky groups of $0.63^{+0.22}_{-0.16}R_{200}^{\mathrm{clus,h}}$, which correspond to median orbital eccentricities of 0.53 ± 0.12 and $0.25^{+0.19}_{-0.15}$, respectively.

4.4.2 Removal of galaxies from groups

The sample of backsplash groups that exit a cluster and then re-enter allow us to directly compare how a single passage through a cluster permanently affects the properties of a group. Comparing the same sample of groups at the moment of first infall and second infall means that the groups are in approximately the same configuration (at a distance of $\sim R_{200}^{\mathrm{clus,h}}$, falling towards the cluster).

Overall, we find that groups on a second infall contain far fewer galaxies, when compared to groups infalling for the first time. On their first infall, the median size of these groups was 6^{+3}_{-1} members (note that this is slightly smaller than the value of 8^{+7}_{-3} quoted in Section 4.3, which includes groups that do not exit and re-enter the cluster). By their second infall, the median size of these same groups is 2 ± 1 members. In fact, 46% of the groups that are infalling for the second time contain only one member. Physically, these objects are not actually groups at all: a 'group' with one member instead represents a single galaxy that has no other galaxies bound to it, having previously been the central galaxy in a group. This shows that, in a single passage through a cluster, almost all galaxies become unbound from groups. Often this process completely disrupts a group, resulting in no galaxies remaining bound together.

Similarly, the dark matter haloes of these groups are heavily stripped during their passage through the cluster. At first infall, the median radius, $R_{200}^{\mathrm{grp,h}}$, of a group was $0.51^{+0.15}_{-0.10}$ h^{-1} Mpc. By their second infall, these same groups had a median radius of $0.32^{+0.10}_{-0.07}$ h^{-1} Mpc. Similarly, the median mass of these groups, $M_{200}^{\mathrm{grp,h}}$, decreases by a factor of three in this time, from $10^{13.2}$ $h^{-1}M_{\odot}$ to $10^{12.7}$ $h^{-1}M_{\odot}$, consistent with the decrease in the number of galaxies. The infall mass of groups that later have a second infall is slightly smaller than the average mass of all groups, $10^{13.5}$ $h^{-1}M_{\odot}$ (Section 3.2.3) but, as we discuss in Section 4.4.3, we still consider these to be a representative sample of all infalling groups. This factor of three drop in mass is comparable to the results from other previous studies which have found that dark matter subhaloes are heavily stripped; Muldrew et al. (2011) found that a halo passing through the centre of a cluster has approximately half of its mass stripped away, and Taylor & Babul (2004) used semi-analytic models to show that subhaloes on orbits similar to our groups lose $> 40\%$ of their mass with each pericentric passage of a cluster. Some studies find even more extreme evidence of this removal of dark matter: Smith et al. (2016) used hydrodynamical simulations to show that a cluster halo can strip away $\sim 80\%$ of the dark matter in galaxy-sized subhaloes.

4.4.3 The fates of group galaxies

We can investigate the removal of galaxies from groups further, by comparing these groups at different stages of their infall and journey through a cluster. As shown in Section 4.3, the speed of galaxies relative to their host group increases before they have reached pericentre of their cluster orbit, and their group-centric distance increases post-pericentre. Therefore, although groups become spatially separated after pericentre, it is not clear when the galaxies become unbound from these groups.

For backsplash groups that also have a second infall, their member galaxies are removed from their host group very quickly. Of those galaxies that are bound to a group at first cluster infall (i.e. that satisfy Eq. 3.1), $60^{+20}_{-35}\%$ are no longer bound to the group by the first pericentre, $76^{+24}_{-26}\%$ are removed by apocentre, and $89^{+11}_{-29}\%$ by the second infall into the cluster (median and 1σ spread for backsplash groups). These numbers are almost identical for backsplash groups that do not have a second infall.

For groups that reach apocentre but do not leave the cluster ('sticky groups'), $75 \pm 25\%$ of previously bound galaxies are no longer group members at pericentre, and $73^{+27}_{-33}\%$ at apocentre. Although it appears that the number of unbound galaxies drops slightly between pericentre and apocentre, this can actually be explained by the fact that the number of particles (and thus the mass and radius) in a subhalo is artificially suppressed in the centre of a large halo, as we describe in Section 4.2.1. This apparently lower mass can have the effect of making more galaxies appear to be unbound.

However, although these galaxies are no longer members of the group, this is not necessarily because they have become gravitationally unbound from their host group. In this section, we analyse the final fates of these galaxies after their group enters a cluster. To do this, we separate the galaxies' states into five categories:

- Bound: galaxy is still bound to its host group, according to Eq. 3.1.
- Unbound: galaxy does not satisfy Eq. 3.1, and so is no longer bound to its host group.
- Disrupted: no descendent of a group member has not been found by the halo finder, typically because its dark matter halo has been heavily stripped.
- Merged with group: galaxy has been absorbed by the halo of its host group, effectively merging with the brightest group galaxy.
- Other merger: merging with another, larger object. Galaxy has either been absorbed by the cluster halo (effectively merging with the brightest cluster galaxy), or absorbed by another galaxy of greater mass than itself.

The 'disrupted' galaxies in our sample represent a class of objects that have physical similarities. However, because of the nature of the simulations and tree-builder that we use in this work, the branches of their merger trees are cut off prematurely, meaning that they appear to have no descendent halo in the simulations and so their final fate cannot be determined. Before their branches end, the dark matter masses of these galaxies are changing rapidly – in their final ten snapshots before they are removed from the merger tree, 76% of these galaxies experience at least one drop of $> 30\%$ in their halo mass between two snapshots (~ 0.3 Gyr), and 37% have a measured drop of $> 40\%$. However, MERGERTREE does not allow for the dark matter mass of an object to change by more than a factor of two between snapshots (see Section 2.2.2 for an explanation of this). Consequently, if a galaxy's dark matter halo mass drops by $> 50\%$ between snapshots, this change will not be recorded, no descendent for the halo will be added to the catalogue, and this branch in the merger tree will end. Despite this heavy stripping of dark matter, very few of the disrupted galaxies violate the mass limits that are imposed in Section 2.2.2; if we remove these mass limits, the median final mass of these galaxies before their merger tree ends is $\log_{10}(M_{200}/h^{-1}M_{\odot}) = 11.3^{+0.7}_{-0.6}$, with a median stellar mass of $\log_{10}(M_{\rm star}/h^{-1}M_{\odot}) = 10.4^{+0.5}_{-0.4}$, and a ratio between these of $0.14^{+0.12}_{-0.08}$. Consequently, few of these galaxies are removed from the merger trees due to violating these imposed mass limits.

Fig. 4.5 shows the status of group member galaxies as their host group passes through a cluster. These data are averaged across all groups that become backsplash groups and then have a second cluster infall, meaning that we have data for their entire passage through a cluster. Overlaid as solid, dashed, dot-dashed and dotted lines are the boundaries between the coloured regions when all groups are included. For example, this indicates the states of galaxies at pericentre for all groups that reach their first pericentre, regardless of what subsequently happens to the group. Similarly, the apocentre data show the fates of all galaxies in groups at apocentre, whether or not this apocentre is outside of the cluster. These data closely follow the data for groups that have a second infall, showing that these second infallers are representative of the entire group sample. We therefore only discuss these groups that later have a second infall, allowing us to make comparisons of the same sample of groups at different stages of their orbit.

As stated above, only approximately 40% of group members are still members of the group at the pericentric passage of the cluster centre (note that here we use the mean behaviour of each group, as opposed to the median used earlier on in this section, and so the quantities differ slightly). However, of the 61% of galaxies that are no longer group members at pericentre, only 45% have become unbound from their

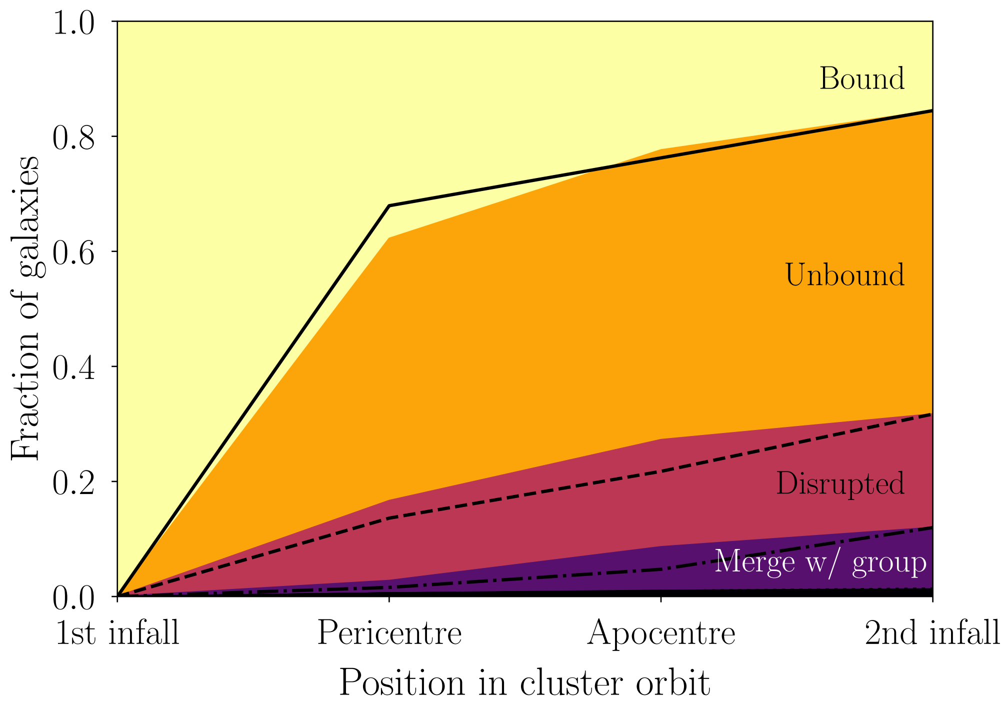

Figure 4.5: Status of galaxies that were bound to groups at cluster infall, as their host group passes through a cluster and begins its second infall. These data are averaged across all groups that become backsplash groups, and experience a second infall. All galaxies are bound at first infall, by definition. Areas representing galaxies that have become bound, unbound, disrupted, or merged with the group halo are labelled. The small, black region represents other mergers, which is unlabelled for clarity. Solid/dashed/dot-dashed/dotted lines show the boundaries between these regions for all groups that reach this stage of their orbit, regardless of whether they go on to reach apocentre or have a second infall.

host group, while 16% have experienced one of the other fates described above. As these groups exit the cluster and re-enter, the number of galaxies becoming unbound increases slightly (to 53%), but the number of galaxies leaving the group for another reason doubles, to 32%, showing that these other processes are more important after a group's initial infall.

Fig. 4.5 represents all group members at infall, however Fig. 4.4 shows that galaxies in different regions of the group phase space will experience different processes, and so the likelihood of each outcome is not the same for all galaxies in a group. Accordingly, we also find that the evolution and fates of group galaxies is strongly dependent on their position within the phase space of their host group. Fig. 4.6 and Fig. 4.7 show the evolution of members of groups that pass through and re-enter a cluster, in the bottom-left[3] ($r < 0.5R_{200}^{\rm grp,h}$, $v < 0.5v_{\rm cir}$) and bottom-right ($r > 0.8R_{200}^{\rm grp,h}$) of the phase space shown in Fig. 4.1. These represent the slow-moving galaxies deep within the group's potential well, and loosely-bound galaxies in the group outskirts, respectively.

Clearly, galaxies in the central (Fig. 4.6) and outer (Fig. 4.7) regions of a group have vastly different evolutionary histories. Slow-moving galaxies in the centres of groups almost never become unbound from the group – instead, the majority of them are disrupted by the time the group re-enters the cluster, although a sizeable fraction (17%) of them merge with the group halo. Dynamical friction likely plays a role in this, by causing these galaxies to spiral in towards the group centre, making them likely to merge with their host group's halo. This is in contrast to the outskirts of the groups, where the vast majority of group members become unbound from the group almost immediately after the group enters the cluster, and only a small fraction are heavily disrupted. In both cases, the black lines on the figures show that galaxies in other infalling groups experience similar evolution, although slightly more galaxies are disrupted in groups that become backsplash groups.

It is also important to consider that these four stages in the group orbit – infall, pericentre, apocentre and second infall – are not equally spaced in time. For the groups shown in Fig. 4.5 (backsplash groups with a second infall), pericentre, apocentre and the second infall occur an average of 0.5 ± 0.2 Gyr, $2.6^{+0.4}_{-0.7}$ Gyr, and $3.5^{+1.2}_{-1.0}$ Gyr after the first infall, respectively. Consequently, not only do most of the unbound galaxies leave the group between infall and pericentre, but this process takes place in just ~ 0.5 Gyr, compared to the ~ 2 Gyr between pericentre and apocentre.

[3]This selection specifically examines slow-moving galaxies near the group centre, as fast-moving galaxies near the group centre exhibit different behaviour. We elaborate on this in Section 4.4.3.

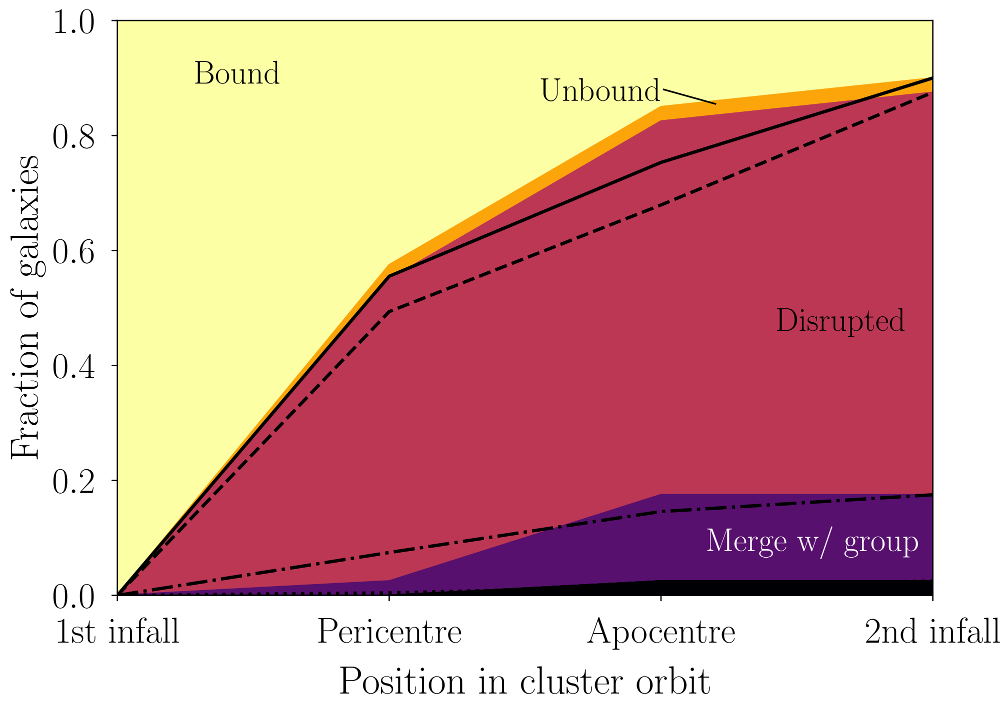

Figure 4.6: Same as Fig. 4.5, but for slow-moving galaxies in the centre of groups ($r < 0.5R_{200}^{\mathrm{grp,h}}$, $v < 0.5v_{\mathrm{cir}}$). Again, the small 'other mergers' region is unlabelled for clarity. Galaxies in this region are much more likely to become heavily disrupted and have an incomplete merger tree, although a substantial fraction merge with the group halo, mostly between pericentre and apocentre.

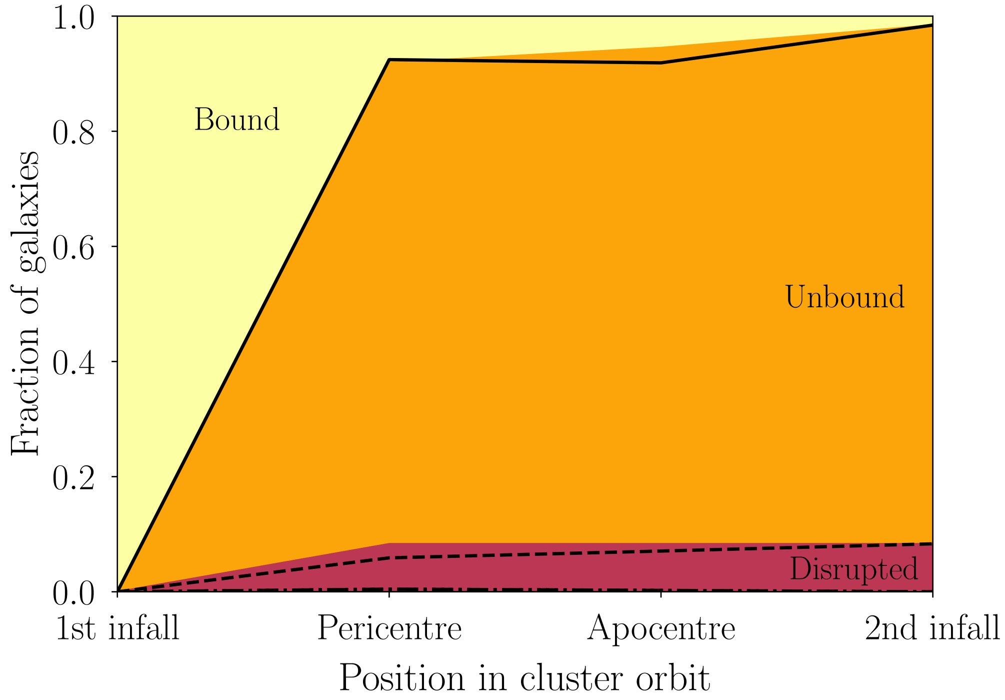

Figure 4.7: Same as Fig. 4.5, but for galaxies in the outskirts of groups ($r > 0.8R_{200}^{\rm grp,h}$). The 'group mergers' and 'other mergers' regions are unlabelled for clarity. Galaxies in the outskirts of the groups are highly likely to become unbound from their host group, which usually happens between infall and pericentric passage.

This means that galaxies are actually removed from groups more suddenly than it appears in Fig. 4.5, Fig. 4.6 and Fig. 4.7, as the four stages are equally spaced in these three figures. To demonstrate this further, these three figures are reproduced in Appendix C, with the horizontal axis scaled to show the average elapsed time between these stages of the orbits.

Finally, we note that these timescales are redshift dependent: the time for a galaxy entering a cluster to reach pericentre at $z = 0$ can typically range from $1 - 2$ Gyr (see Fig. B1 in Tollet et al., 2017, for further details), but the time taken decreases at higher redshifts. Our method consequently returns an average infall-to-pericentre time of less than 1 Gyr, because we stack data from groups at numerous different redshifts (for some additional discussion of cluster crossing times, see Contreras-Santos et al., 2022).

Galaxy fates across group phase space

Finally, we can take a more general approach to this by looking at the phase space of the infalling groups, to determine the typical fates of galaxies at the second cluster infall, as a function of their initial position in this phase space. Fig. 4.8 shows how common different outcomes are for group members, as a function of their relative position and speed at cluster infall; this is in effect a generalisation of Fig. 4.6 and Fig. 4.7. For example, in the bottom-left region of the phase space, there is a high density of 'disrupted' galaxies, showing that galaxies here during infall later became disrupted, in agreement with Fig. 4.6.

The top-left and top-right panels of Fig. 4.8 show a substantial decrease in the number of galaxies that remain bound to a group outside of $r \sim 0.7R_{200}^{\mathrm{grp,h}}$ from the group centre. This indicates that, for almost all groups, virtually all galaxies outside of this radius are removed. Similarly to in Section 4.3.2, the tidal stripping of groups can explain this sharp cut. According to Eq. 4.8, a tidal radius of $0.7R_{200}^{\mathrm{grp,h}}$ corresponds to a group that is approximately $0.7R_{200}^{\mathrm{clus,h}}$ from the cluster centre. This distance is the maximum typical pericentric distance that we find for groups in our sample – almost all groups (95%) have a pericentric passage of $r \leq 0.7R_{200}^{\mathrm{clus,h}}$. Consequently, almost all groups will have had a tidal radius of $R_{\mathrm{t}} = 0.7R_{200}^{\mathrm{grp,h}}$ at some point in their orbit, but not all groups will have experienced a tidal radius less than this. This explains why some galaxies remain in the groups within $0.7R_{200}^{\mathrm{grp,h}}$, but none remain beyond this distance.

Generally, only galaxies near to the group centre with high velocities remain as bound group members. These are on longer, eccentric orbits – galaxies with

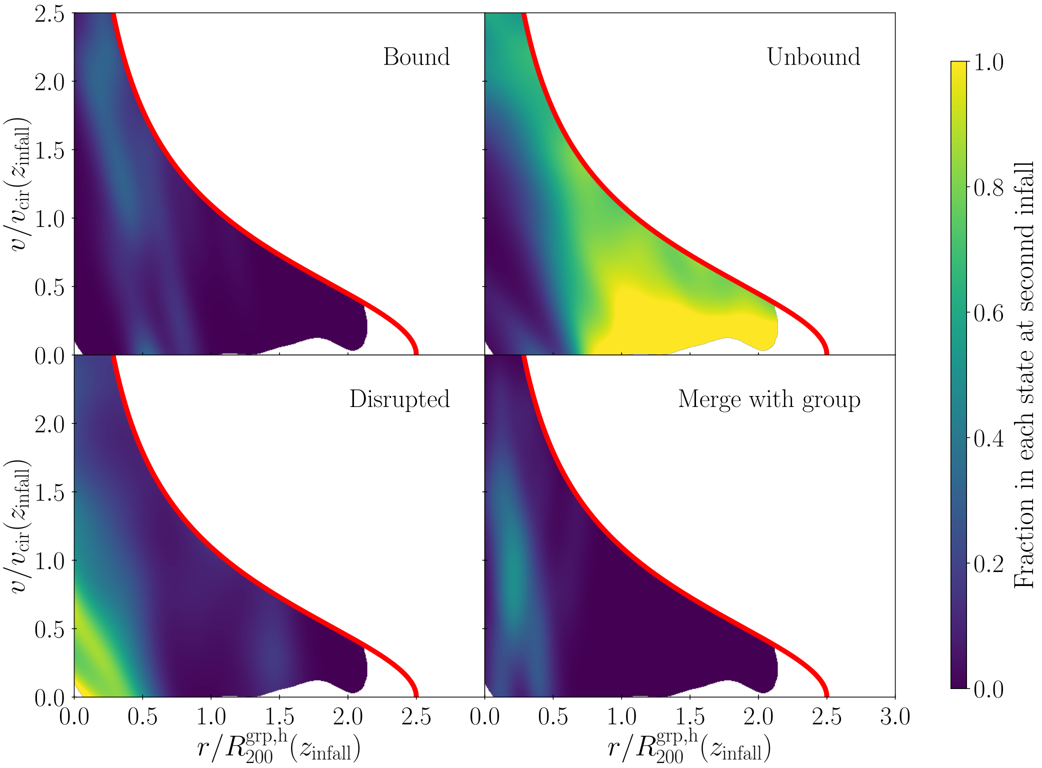

Figure 4.8: Of the galaxies that are bound to a group at its first infall, each panel shows the fraction of these in each state at the moment of second infall. Phase space is defined by the position/speeds of the galaxies at the first infall. Top-left panel shows the fraction of galaxies that remain bound to the group. Top-right panel shows the fraction that become unbound from the group. Bottom-left shows the fraction that are 'disrupted'. Bottom-right shows the fraction that merge with the group halo. Lighter colours represent regions of the phase space with a greater number of galaxies. White regions either represent the 'unbound' region, or regions where the number of galaxies is very low.

lower velocities spend more time nearer the group centre, and so are more likely to merge with the group, or to be disrupted. Furthermore, the bottom-left and bottom-right panels show that the disrupted galaxies inhabit different parts of phase space, compared to those that later merge with the group halo. Disrupted galaxies have large amounts of their dark matter stripped in a short period of time: for two-thirds of these galaxies, in the snapshot immediately after they are 'disrupted', more than 50% of their dark matter particles appear either in the halo of their host group or (less often) their host cluster. This disruption by a larger halo is similar to how galaxy harassment can occur in clusters (Moore et al., 1996a). However, the galaxies in the centre of these disrupted haloes do not immediately become associated with the group halo – if this were the case, these objects would be tagged as merging with the group halo, which they are not. This implies that a tidal disruption is occurring, in which large amounts of material are removed from the galaxy, forming a substructure in the group such as a tidal stream. This substructure will most likely merge with the group halo at some later time (Moore et al., 1998), effectively making this process a merger with the group halo, but over a longer time period.

Disruption is more likely for galaxies in the centres of groups that are slow-moving at the moment of infall, while galaxies with greater speeds are somewhat more likely to merge with the group halo. One explanation for this lies in the left panel of Fig. 4.3, showing pre-pericentre groups. Before a group reaches pericentre, galaxies in the group centre with high speeds move downwards in phase space, indicating that their speed is decreasing due to dynamical friction, and they are slowly spiralling into the group centre where they merge. Slow-moving galaxies are instead moving upwards on this plot, indicating that they are experiencing strong, accelerating forces that can disrupt their structure. Additionally, high-speed galaxies are on radial, eccentric orbits, meaning that they pass the group centre infrequently. In contrast, a low speed and low group-centric distance indicates that a galaxy is on a small, circular orbit, and so will be able to make multiple orbits of the group in a short period of time, providing more opportunities for heavy stripping by the central group galaxy.

The different outcomes for galaxies from different regions of phase space can be seen more clearly in Fig. 4.9. This combines the four panels from Fig. 4.8, to show the region in infall phase space from which the greatest fraction of galaxies have each of the four fates shown above. Only one contour is shown for each galaxy fate, but this shows that each region of phase space is responsible for a different population of galaxies in the future. These results broadly agree with the findings of

Choque-Challapa et al. (2019), who use dark matter-only simulations to study the fates of galaxies in infalling groups as a function of their position in group phase space. Among other results, they show that outside of $r \sim 0.8R_{200}^{\mathrm{grp,h}}$ there is a sharp increase in the fraction of members becoming unbound from their host group, and that galaxies lying near to the boundness line are more likely to be stripped from their groups.

Observational analogues

It might appear that this preferential removal of outer group members could lead to the formation of very dense, centrally-concentrated groups such as Hickson Compact Groups (Hickson, 1982) in and around clusters. As Fig. 4.4 shows though, the galaxies that remain bound to a group do not remain in the same region of this phase space. Instead, the remaining galaxies are redistributed throughout the group until they follow a similar distribution to that shown in Fig. 4.1. These group remnants are no more centrally concentrated than the infalling groups.

Most importantly, Fig. 4.5 and the top-left panel of Fig. 4.8 show that, of the galaxies that are bound to a group when it approaches a cluster, almost none are still bound to a group after just a single crossing of the cluster ($\sim$ 2 Gyr later). Typically, only a very small number of galaxies remain in a group, and so the remnant 'groups' are usually either single galaxies, or galaxy binaries. Groups with five or more members are extremely unlikely to have previously experienced a cluster environment; across all of our simulations, less than 1% of such groups entering a cluster after $z = 0.1$ have previously passed through a cluster. Because of this, galaxy groups nearby to a cluster (i.e. in the cluster outskirts, just outside of $R_{200}^{\mathrm{clus,h}}$) are very unlikely to contain backsplash galaxies, which as we show in Chapter 5, make up a substantial fraction of the galaxies surrounding a cluster (see also Gill et al., 2005). Instead, these groups represent a population of galaxies that are unprocessed by their host clusters, but have experienced group effects in their past.

The fact that galaxy groups observed nearby to clusters are very likely to be on their first approach to the cluster has important implications for observational studies of galaxy evolution and environmental pre-processing. Additionally, we can infer greater detail about the histories of the galaxies in these groups. For example, cluster galaxies that are currently not in groups are unlikely to have previously experienced the dense, central regions of a group, as galaxies in group centres are much more likely to remain in their groups. Similarly, galaxies associated with groups inside clusters have almost certainly previously passed through the group centre, even if

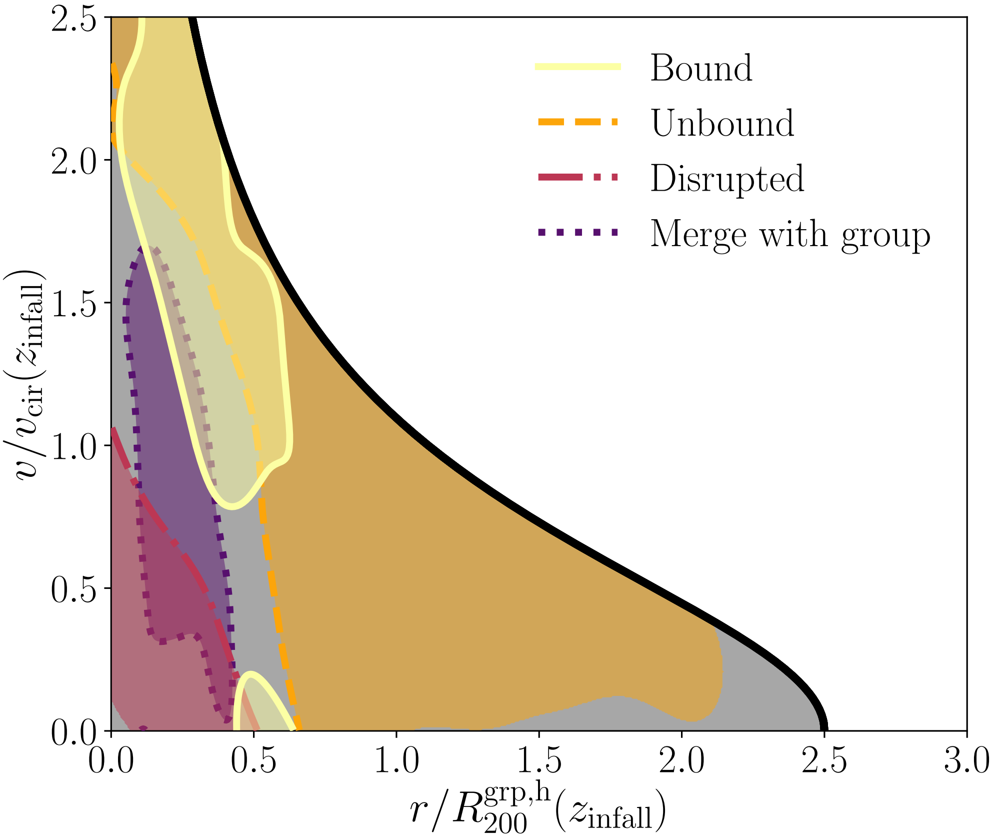

Figure 4.9: Overlay of four panels from Fig. 4.8, to aid with comparison of different phase space distributions, for galaxies bound to groups (below the thick line) at cluster infall. The highlighted region corresponding to each galaxy fate shows the region of phase space that is most likely to produce these galaxies. Contour surrounding each region is placed at a value equal to half of the maximum, from each panel in Fig. 4.8. Grey regions either represent areas of phase space that contain few galaxies, or are not an important region for any of these four classes of galaxies. From this, it is clear that galaxies in different regions of phase space later experience different environments, and different evolutionary processes.

they now reside in the group outskirts. This means that they will have experienced the most extreme environmental impacts of the group. Hester (2006) showed that, in groups of a similar size to those used in this work ($M_{200}^{\mathrm{grp,h}} = 10^{13} M_{\odot}$), a disk galaxy with a dark matter mass of $10^{11} M_{\odot}$ at $r = 0.75 R_{200}^{\mathrm{grp,h}}$ will have $\sim 20\%$ of its disk gas removed, but if this galaxy passes within $r = 0.25 R_{200}^{\mathrm{grp,h}}$ of the group centre, it can lose approximately 90% of its gas. They attribute this to the stronger ram pressure stripping that takes place in group centres.

4.5 Conclusions

In this chapter we build on the previous chapter of this thesis, in which we studied how galaxy groups are modelled differently in hydrodynamical and dark matter-only simulations. Here, we use the hydrodynamical simulations to study the evolution of intermediate sized galaxy groups (five to 50 members with stellar masses $M_{\mathrm{star}} \geq 10^{9.5} M_{\odot}$) in the vicinity of large galaxy clusters, and specifically from the time after the groups pass within R_{200} of the cluster. We begin by studying the positions and speeds of galaxies relative to their host group in order to characterise how this 'phase space' of the group changes over time, before studying the fates of group members after the passage of their group through a cluster. Our findings are summarised below.

- On entering a cluster, galaxy groups typically pass within $0.6 R_{200}^{\mathrm{clus,h}}$ of the cluster centre. Most of these groups remain permanently bound to the cluster, although a small fraction ($\sim 10\%$) reach an apocentric distance outside of the cluster's radius, $R_{200}^{\mathrm{clus,h}}$.

- The dynamics of these groups change in two phases. First, the member galaxies increase their speeds relative to the group centre, often becoming gravitationally unbound. Then, after the group passes the pericentre of its cluster orbit (which typically occurs after ~ 0.5 Gyr in the cluster), the distances of galaxies from the group centre increases.

- The majority of galaxies bound to a group at its first cluster infall are no longer in the group after a full passage through the cluster. Many of these galaxies become either unbound from the group, heavily stripped, or merge with the brightest group galaxy, and the fate of a galaxy depends strongly on its location within the group at the infall time.

- Consequently, the overwhelming majority ($> 99\%$) of groups that enter a cluster are doing so for the first time in their histories. In observations, groups that are seen just outside of a cluster are very unlikely to have previously experienced a cluster environment.

Although the composition and structure of simulated galaxy groups is dependent on the physical models that are used, the results from this work still allow us to make conclusions about groups that can be applied to observational work. Groups that are observed nearby to clusters are almost certainly recent infallers, particularly groups with low velocity dispersions, as galaxies in groups become gravitationally unbound almost immediately after entering a cluster. Furthermore, any galaxies that are observed in a group that is inside a cluster will have previously passed through the group centre, and so will be severely stripped by the tidal forces and ram pressure of their host group.

In addition to the approach taken in this chapter, which draws conclusions on galaxy groups that can be applied observationally, work remains to be done on the dynamics of these groups. In future work (beyond the scope of this thesis), we plan to study the dynamics of these groups in greater detail. For example, the binding energy-angular momentum phase space, and the orbital parameters of galaxies, can tell us about the anisotropy of group members' orbits (Wojtak et al., 2008; Lotz et al., 2019, for example), which can in turn be used to describe how virialised is a group. In the future, we plan to study the time evolution of these dynamical parameters. This work could be further extended by using other definitions of galaxy groups (like those described in Section 4.1) – for example, by defining group members as all galaxies inside $R_{200}^{\mathrm{grp,h}}$. However, our results are unlikely to be sensitive to this: Fig. 4.8 and Fig. 4.9 show that in most regions of the group, high-speed galaxies (those lying near the boundness condition) are most likely to become unbound, similarly to galaxies in the group outskirts ($r > R_{200}^{\mathrm{grp,h}}$). Consequently, the high-speed group members that would be introduced by removing the velocity limit in our group definition (Eq. 3.1) would generally experience the same fates as the galaxies in the group outskirts, which would be excluded based on other group definitions.

This work further strengthens the case that galaxy groups provide a unique way to study galaxy evolution, and particularly pre-processing. As they are almost all first-infallers, groups in the outskirts of clusters will have experienced no cluster processing, and will have a very low contamination by backsplash galaxies. Consequently, pre-processing effects will dominate over any effects from the cluster in these structures, and so studying these objects in more detail will allow us to further disentangle the environmental effects of clusters, and the effects of other cosmic environments. Processes such as gas removal and morphological changes in these galaxies will exclusively have occurred pre-infall in groups or cosmic filaments, and so ultimately the properties of galaxies in groups will help inform us of the role that environment plays in driving galaxy evolution.

Chapter 5

Backsplash galaxies in the outskirts of clusters

In the outer regions of a galaxy cluster, galaxies may be either falling into the cluster for the first time, or have already passed through the cluster centre at some point in their past. This previous passage through a cluster can occur as an individual galaxy or, as we showed in the previous chapter, as a member of a previously bound group. In this chapter, we investigate the 'backsplash population' of galaxies; those that have passed within R_{200} of the cluster centre at some time in their history, but are now outside of this radius. These are contrasted with infalling galaxies, which are found in the cluster outskirts, and have never been inside a cluster. We find that, on average, over half of all galaxies between R_{200} and $2R_{200}$ from their host cluster at $z = 0$ are backsplash galaxies, but that this fraction is dependent on the dynamical state of a cluster, as dynamically relaxed clusters have a greater backsplash fraction. We also find that this population is mostly developed at recent times ($z \lesssim 0.4$), and is dependent on the recent history of a cluster. Finally, we show that the dynamical state of a given cluster, and thus the fraction of backsplash galaxies in its outskirts, can be predicted based on observational properties of the cluster. The content of this chapter has been published in Haggar et al. (2020).

Contents

5.1 Introduction

It is well established that there is a relative excess of early-type galaxies in cluster environments (Dressler, 1980), as well as a deficit of star-forming galaxies (Thomas et al., 2010; Patel et al., 2011). However, these galaxies can pass through several different environments during their lifetime, each of which can have an impact on the galaxy's properties. For instance, galaxies entering a cluster can experience pre-processing in group or filament environments (see also Section 1.3.2.

Clusters are not static objects, and do not smoothly accrete matter throughout their history. Some of the first work in which this idea was investigated was that of Fillmore & Goldreich (1984), who numerically studied the gravitational collapse of collisionless matter haloes, and the paths that particles take on their first and subsequent infalls. More specifically, the presence of a 'splashback radius' in galaxy clusters indicates that material can leave a cluster and then re-enter at a later stage. Theoretical work on this radius includes that of Adhikari et al. (2014), Diemer & Kravtsov (2014) and More et al. (2015), who each describe the splashback radius as the distance from a cluster centre at which accreting matter first reaches the apocentre of its orbit, and show that this radius physically corresponds to the distance at which the cluster density profile drops most steeply. They then proceed to identify and study the splashback radius in simulations of clusters. Observational studies have also confirmed the presence of a splashback radius, although the detected splashback radii appear to consistently take smaller values than predicted by simulations. Both More et al. (2016) and Baxter et al. (2017) stack the number density profiles of galaxies around large numbers of SDSS clusters, finding a sharp drop in the profiles in the cluster outskirts, corresponding to the splashback radius. Similarly, weak lensing (Chang et al., 2018) and S-Z measurements (Shin et al., 2019; Zürcher & More, 2019) have been used to measure the radial mass density profiles of clusters, finding a steep drop in their density, as simulations predict should be the case at the splashback radius.

Closely tied to the splashback radius are 'backsplash galaxies', a population of galaxies that have fallen into a cluster, but have overshot the cluster centre and have passed back beyond a certain distance from the cluster centre. Typically, distances of either R_{200}, or a definition of the virial radius of the cluster (such as that of Bryan & Norman (1998)) are used to define this 'backsplash population'. At the present day these galaxies are outside of the cluster, either receding from the cluster centre, or on a second (or subsequent) infall.

It is particularly important to note the distinction between the definitions for splashback and backsplash, and that these two are not interchangeable; the splashback radius is based on the radial density profile of a cluster (More et al., 2015) and – for a spherical system – clearly separates infalling material from matter orbiting in the potential of the halo. However, a backsplash galaxy refers to an individual object that has simply left the cluster, having previously been within a given distance of its centre. Consequently, the splashback radius does not necessarily include all backsplash galaxies. For example, as no assumptions are made about the boundness of backsplash galaxies, they could travel far beyond the splashback radius, rather than remaining on bound orbits. This would be analogous to the 'renegade subhaloes' identified by Knebe et al. (2011c) in simulations of the Local Group, which were associated with a host halo, but entered a different host at a later time.

Gill et al. (2005) identify backsplash galaxies in simulations of clusters, and show that the outskirts of a cluster contain a significant population of these galaxies. These galaxies will therefore have experienced the effects of a cluster environment in their past, but are found in the same locations as infalling galaxies when observed. Because of this, samples of infalling galaxies collected by surveys of cluster outskirts are likely to be contaminated by this backsplash population, making it difficult to disentangle the effects of pre-processing and the effects of a cluster on a population of galaxies.

Backsplash galaxies are not easy to observationally differentiate from infalling galaxies, although the two are potentially distinguishable through kinematics. Gill et al. (2005) show that the infall speeds of backsplash galaxies are generally lower than those entering the cluster for the first time, and Pimbblet (2011) observationally identified a population of cluster galaxies with line-of-sight velocities predicted by Gill et al. (2005), although their 'backsplash population' is likely affected by interlopers. Simulations are a vital tool for studying a backsplash population, as they allow us to examine the full history of a cluster and determine the fraction of galaxies that are indeed backsplash.

In Section 5.2 we introduce the analysis used in this chapter, including our definitions of backsplash galaxies and cluster dynamical state. In Section 5.3 we present our results, and discuss which cluster properties affect the population of backsplash galaxies. We then summarise our findings in Section 5.4.

5.2 Methodology

5.2.1 Backsplash population

The definition of a 'backsplash galaxy' is somewhat subjective. A common definition is one based purely on the present-day locations of galaxies; the backsplash population consists of galaxies that have passed within the virial radius of a cluster at some previous time, but are now found outside of the cluster, at some distance $D > R_{\rm vir}$ from the cluster centre (Gill et al., 2005; Bahé et al., 2013), although the 'virial radius' used in this definition is also open to interpretation. Other work uses a definition based on the dynamics of galaxies. For example, Haines et al. (2015) place no radial distance constraints on their backsplash galaxies, and instead take all galaxies that have passed through the pericentre of their orbital path but have yet to reach apocentre (and hence have an outwards radial velocity) to be backsplash galaxies. However, by this definition a significant portion of their backsplash galaxies are within the virialised region of the cluster, and galaxies that have passed through the cluster centre and are on a second infall are exempt from this definition.

We adopt a definition similar to that of Gill et al. (2005), based on the orbital history of each galaxy relative to R_{200}, which is the radius we use as the extent of the cluster. For the remainder of this chapter, we use R_{200} to refer to the radius of the cluster, rather than $R_{200}^{\rm clus,h}$ as in the previous chapters. We categorise each galaxy in or around a cluster into one of three groups, based on their radial distance to the cluster centre at $z = 0$, $D_{z=0}$, and their minimum distance to the cluster centre at any time in their history, $D_{\rm min}$:

- $D_{z=0} = D_{\rm min}$ or $D_{\rm min} > R_{200}$:
  The infalling population: Galaxies that are on their first infall towards the cluster. These are either on approximately radial paths (giving $D_{z=0} = D_{\rm min}$), or are members of infalling groups that are yet to reach R_{200}, which can lead to $D_{z=0} > D_{\rm min} > R_{200}$.

- $D_{\rm min} < D_{z=0} < R_{200}$:
  The cluster population: Galaxies within the radius of the cluster (taken to be R_{200}). These are the 'normal' satellite galaxies, which we consider to be members of the cluster. Many of these are on bound orbits, although galaxies that are on paths heading out of the cluster can also be included in this definition.

- $D_{z=0} > R_{200}$, $D_{\rm min} < R_{200}$:

The backsplash population: Galaxies that have previously fallen through the cluster, but have now exited the cluster and exist beyond R_{200} at $z = 0$. These can either be receding from the cluster centre, or on a subsequent infall towards the cluster centre.

Our definition deviates slightly from that of Gill et al. (2005), who instead define backsplash galaxies relative to a larger radius, $R_{\rm vir} \sim R_{100} \sim 1.4R_{200}$. However, they also note that 90% of the backsplash galaxies they identify pass within $0.5R_{\rm vir}$ of the cluster centre, meaning that by considering R_{200}, we are unlikely to neglect a large fraction of these galaxies.

Furthermore, this definition of backsplash galaxies applies to clusters we study at $z = 0$. If instead we are interested in the backsplash galaxies of a cluster observed at a redshift $z_{\rm obs} > 0$, we adjust the definition by replacing $D_{z=0}$ with the radial distance from the cluster centre at $z_{\rm obs}$, and by replacing $D_{\rm min}$ with the minimum distance a galaxy has passed to the cluster centre at any redshift $z \geq z_{\rm obs}$. Specifically, in this chapter we focus on the fraction of all galaxies in the radial region $[R_{200}, 2R_{200}]$ (which we explain below in Section 5.3) that are members of the backsplash population, and so have previously been within $R_{200}(z)$ of the cluster centre, where $R_{200}(z)$ is the radius of a cluster at a redshift z.

5.2.2 Dynamical state

In this chapter, we study how the fraction of backsplash galaxies around a cluster varies, for clusters in different dynamical states. The dynamical state is a property of a large dark matter halo, such as a cluster halo, which describes how disturbed or relaxed a galaxy cluster is – isolated, slowly-growing clusters are generally more dynamically relaxed, while clusters experiencing dynamically violent processes will be unrelaxed. Numerous measures of dynamical state exist, but Cui et al. (2018) describe three parameters that are used to determine the dynamical state of each of the clusters in The Three Hundred simulations. These parameters are:

- Centre of mass offset, $\Delta_{\rm r}$: The offset of the centre of mass of the cluster from the density peak of the cluster halo, as a fraction of the cluster radius R_{200}.
- Subhalo mass fraction, $f_{\rm s}$: Fraction of the cluster mass contained in subhaloes.
- The virial ratio, η: A measure of how well a cluster obeys the virial theorem, based on its total kinetic energy, T, its energy from surface pressure, $E_{\rm s}$, and its total potential energy, W. It is defined as $\eta = (2T - E_{\rm s}) \,/\, |W|$.

Further description of each of these dynamical state parameters is available in Cui et al. (2018), and more comprehensive details in Cui et al. (2017).

Cui et al. (2018) describe a cluster as being dynamically relaxed if it satisfies $\Delta_{\mathrm{r}} < 0.04$, $f_{\mathrm{s}} < 0.1$ and $0.85 < \eta < 1.15$, and denote it as unrelaxed if it does not satisfy all of these. In order to obtain a continuous, non-binary measure of dynamical state, we combine these three parameters into a single measure of dynamical state, the so-called 'relaxation' of a cluster, χ_{DS}:

$$\chi_{\mathrm{DS}} = \sqrt{\frac{3}{\left(\frac{\Delta_{\mathrm{r}}}{0.04}\right)^2 + \left(\frac{f_{\mathrm{s}}}{0.1}\right)^2 + \left(\frac{|1-\eta|}{0.15}\right)^2}}\,. \tag{5.1}$$

Note that for a cluster to be most relaxed, we require Δ_{r} and f_{s} to be minimised, and $\eta \rightarrow 1$ (Cui et al., 2017) – physically, this corresponds to a symmetrical, virialised cluster, with little substructure. $\chi_{\mathrm{DS}} = 1$ corresponds approximately to the Cui et al. (2018) definition of dynamical state, such that all of the clusters they denote as 'dynamically relaxed' have $\chi_{\mathrm{DS}} > 1$.

5.3 Results & discussion

In Fig. 5.1, we show the distribution of $D_{z=0}$ and D_{min} for each halo and subhalo in our simulations, stacking the data from the 257 selected clusters. On average, 90% of the backsplash haloes at $z = 0$ are found between R_{200} and $2R_{200}$, and 99.8% between R_{200} and $3R_{200}$. This indicates that the region just outside of the cluster – specifically, the radial range $[R_{200}, 2R_{200}]$ – is the region where backsplash galaxies make the most important contribution to the total population of galaxies. Consequently, although we find a small number of backsplash galaxies outside of this region, for most of this chapter we characterise the prevalence of backsplash galaxies around a cluster using the fraction of backsplash galaxies inside this region.

In total, we find 27114 galaxies in the radial range $[R_{200}, 2R_{200}]$, of which 15811 have previously passed within R_{200} and are members of the backsplash population, corresponding to a mean backsplash fraction of 58%. This is consistent with the work of Gill et al. (2005), who found a backsplash fraction of 50%, albeit in a slightly different radial range and using dark matter-only simulation data (which, as we showed in Section 3.3.1, results in a suppressed backsplash fraction).

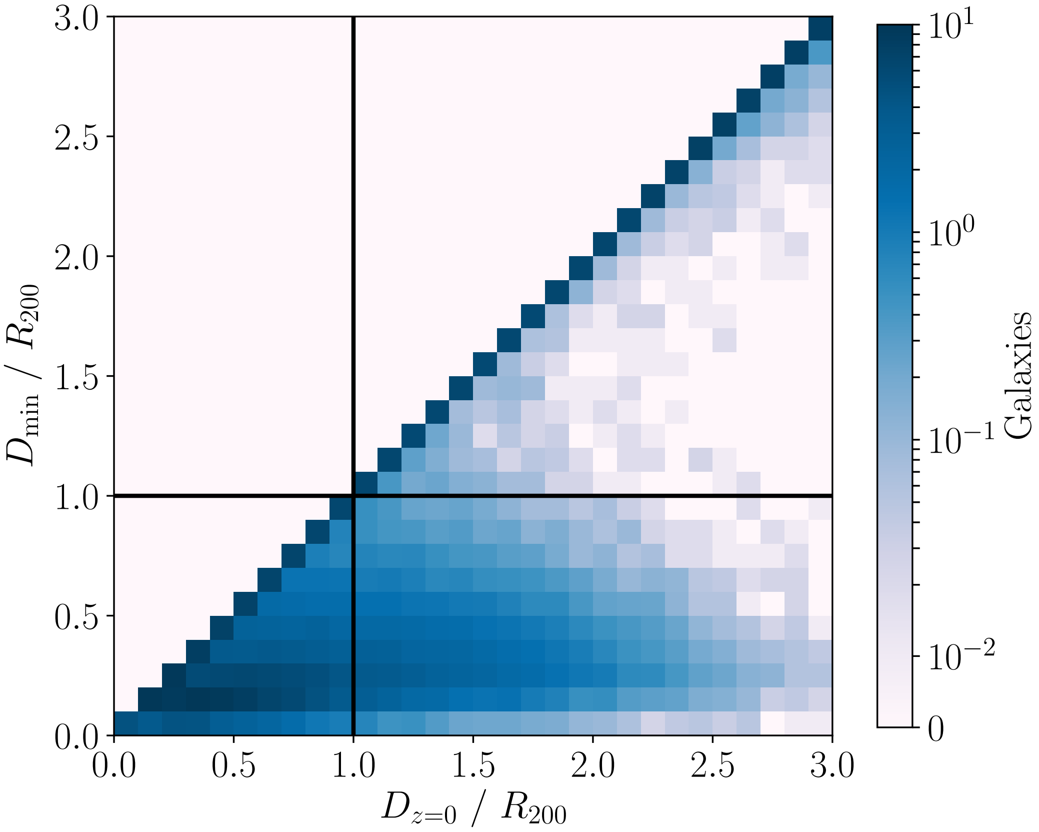

Figure 5.1: Histogram, showing $z = 0$ and minimum cluster distances of galaxy population (using galaxy haloes with $M_{200} \geq 10^{10.5} M_\odot$, $M_{\rm star} \geq 10^{9.5} M_\odot$ and $M_{\rm star} < 0.3 M_{200}$) averaged across 257 clusters. Note the characteristic large number of objects along the line $D_{\rm min} = D_{z=0}$, corresponding to infalling galaxies. Backsplash galaxies are located in the bottom-right of this phase space.

5.3.1 Dynamical state parameters

Fig. 5.2 shows the distribution of the dynamical state, $\chi_{\rm DS}$, of the 257 clusters we have selected for use in this work. The 67 clusters we remove from our original sample of 324 are slightly biased towards lower values of $\chi_{\rm DS}$, however they still cover most of the range of relaxation values. This is a result of the fact that the highly relaxed clusters (with greater values of $\chi_{\rm DS}$) are less likely to fail the selection criteria we detail in Section 2.2.3. Based on the values of $\chi_{\rm DS}$, we split our sample into three groups, allowing the third of clusters that are most relaxed ($\chi_{\rm DS} > 1.030$) and least relaxed ($\chi_{\rm DS} < 0.619$) to be compared. Each of these groups contains 86 clusters. The unrelaxed clusters have slightly greater average values of M_{200} and R_{200}, however the difference is small compared to the spread of these quantities across the whole sample of clusters.

In addition to the stacked data in Fig. 5.1, we also calculate the backsplash fraction, F, for each of the 257 clusters individually. Fig. 5.3 shows the backsplash fraction for each cluster against its relaxation parameter, and shows that clusters that are more relaxed have a greater fraction of backsplash galaxies. A smaller centre of mass offset, $\Delta_{\rm r}$, smaller fraction of mass in subhaloes, $f_{\rm s}$, and a virial ratio, η, closer to one, all of which are indicative of a relaxed cluster, each result in a larger backsplash population.

The third of clusters that are least relaxed (with $\chi_{\rm DS} < 0.619$) have a median backsplash fraction of $F = 45^{+15}_{-20}\%$ (1σ spread). The most relaxed third ($\chi_{\rm DS} > 1.030$) have a backsplash fraction $F = 69^{+9}_{-11}\%$. Of the three dynamical state parameters, $f_{\rm s}$ correlates most strongly with the backsplash fraction, followed by $\Delta_{\rm r}$. Although there is a weaker relationship between F and $|1 - \eta|$, a relationship does indeed exist. We note that, when considering $f_{\rm s}$, part of the correlation between this parameter and the backsplash fraction is likely caused by the backsplash galaxies themselves; the movement of a large number of galaxies from within R_{200} to the cluster outskirts will reduce the amount of substructure within R_{200}, therefore causing the fraction of mass contained in subhaloes, $f_{\rm s}$, to decrease.

There is also a very weak correlation between the backsplash fraction and M_{200}, in which the less massive clusters have a marginally higher backsplash fraction, although this can be fully accounted for by the fact that the relaxed clusters have a slightly lower average mass. We find no significant correlation between the backsplash fraction and R_{200}. We therefore conclude that more dynamically relaxed clusters have greater backsplash populations, and that the fraction of mass in subhaloes and the centre of mass offset of the cluster are specific properties that we expect to strongly

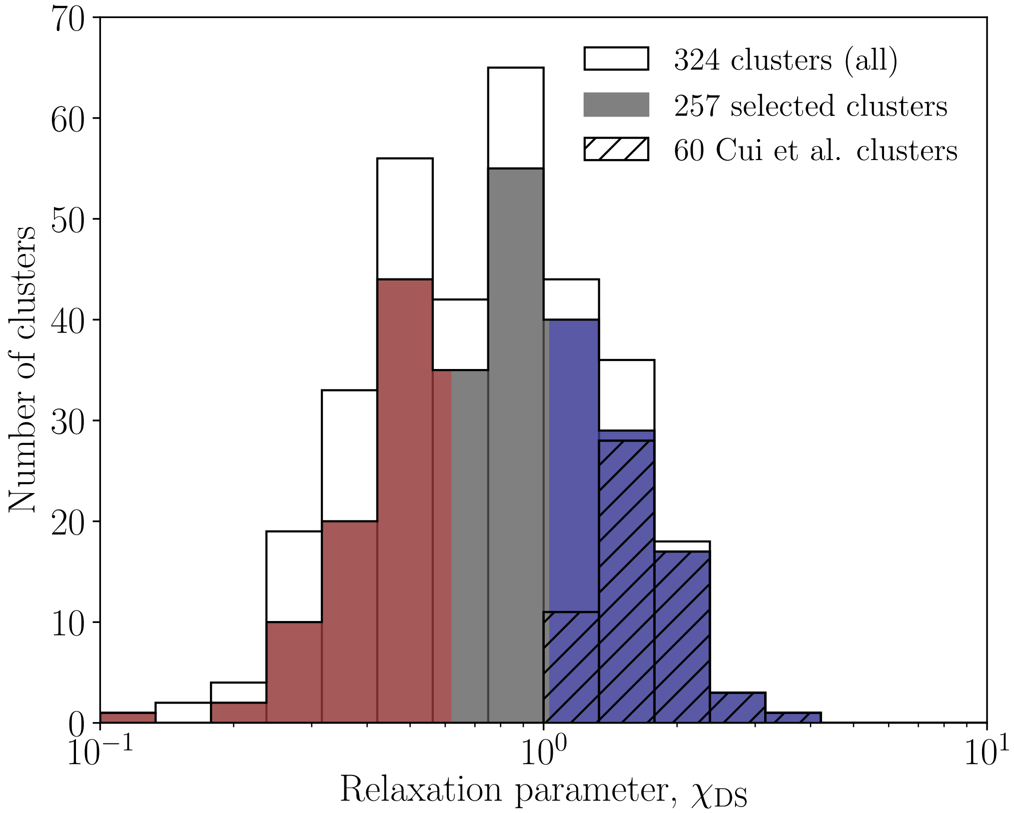

Figure 5.2: Distribution of the 'relaxation' of each cluster, $\chi_{\rm DS}$, with the height of each bar showing the number of clusters in this range. The full sample of 324 clusters is shown by the white bars, and the overlaid filled bars show our selected sample of 257 clusters. The regions $\chi_{\rm DS} < 0.619$ ('unrelaxed' clusters) and $\chi_{\rm DS} > 1.030$ ('relaxed' clusters) are highlighted in red and blue, respectively. Our sample consists of 86 'relaxed' clusters, 86 'unrelaxed' clusters, and 85 with $0.619 < \chi_{\rm DS} < 1.030$. The hatched bars represent the clusters that are dynamically relaxed according to Cui et al. (2018); by definition, these have $\chi_{\rm DS} > 1$.

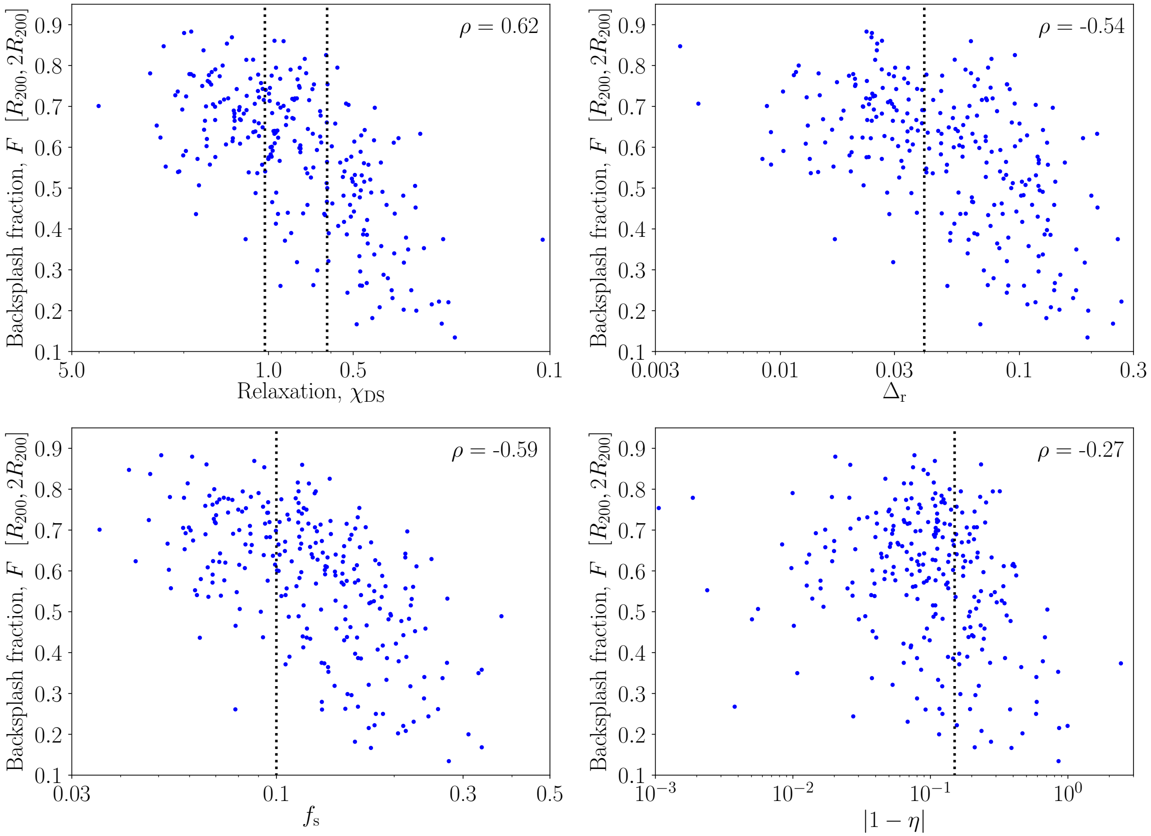

Figure 5.3: 'Relaxation', $\chi_{\rm DS}$, of each of the selected 257 clusters, against backsplash fraction in the radial range $[R_{200}, 2R_{200}]$ at $z = 0$, shown in top-left plot. The variation of backsplash fraction with each individual parameter is also shown. Note the reversed horizontal axis on the top-left plot; in each of these plots, the 'more relaxed' clusters are on the left. The vertical dashed lines in the top-left plot show the boundaries between our relaxed, unrelaxed, and intermediate clusters, and the vertical lines in the other three plots show the boundary values for relaxation given by Cui et al. (2018) (see also Section 5.2.2). The Spearman's rank correlation coefficient, ρ, for each plot is inset, showing the tighter correlation achieved by combining the three dynamical state parameters.

affect this.

Finally, we find that there is very little dependence of the backsplash fraction on the mass of galaxies. Separating the galaxies into stellar mass bins of width 0.5 dex, we find that the median backsplash fraction does not change for galaxies with stellar masses between $10^{9.5}\ h^{-1}M_{\odot}$ and $10^{11}\ h^{-1}M_{\odot}$. Considering only galaxies with stellar masses above $10^{11}\ h^{-1}M_{\odot}$, the median backsplash fraction of clusters appears to drop slightly, although this drop is not statistically significant due the far smaller number of these high-mass galaxies present in cluster outskirts. However, the backsplash fraction of clusters does depend on the total masses of galaxies – that is, the mass including the galaxy halo mass. We find that galaxies with a low halo mass are more likely to be members of the backsplash population. For example, on average 69% of galaxies with total masses in the range $[10^{10.5}, 10^{11}]\ h^{-1}M_{\odot}$ are backsplash galaxies, compared to 43% of those with mass in the range $[10^{12}, 10^{12.5}]\ h^{-1}M_{\odot}$. Similarly, galaxies with a greater ratio of stellar mass to total mass are also more likely to be backsplash galaxies.

We expect the halo mass and stellar mass of these galaxies to be closely linked (Moster et al., 2010). As stellar material experiences very little stripping between infall and leaving a cluster in these simulations, we therefore deduce that the dark matter haloes of backsplash galaxies have been tidally stripped during their passage through the cluster. This means that the apparent bias towards low-mass haloes becoming backsplash galaxies is due to the stripping of the haloes around backsplash galaxies, rather than due to a strong dependence on the halo mass of galaxies at infall. Previous studies have shown that galaxy-sized dark matter haloes are heavily stripped when passing through a larger halo (Taylor & Babul, 2004; Smith et al., 2016, for example), as we discussed in detail in Chapter 4, and particularly in Section 4.4.2.

Another potential explanation for this is dynamical friction, as previous work has shown that the location of the splashback feature in simulations of dark matter haloes is dependent on the mass of subhaloes being considered. Specifically, both Adhikari et al. (2016) and More et al. (2016) find that the splashback feature for haloes of greater masses is found at smaller distances. However, for the rest of this chapter, we continue to focus on the effect of cluster properties on the backsplash population, rather than galaxy properties.

5.3.2 Evolution of backsplash fraction

By examining how the dynamical state parameters for each cluster vary over time, we are able to examine the stability of these parameters. We find that the parameters are only stable over a relatively short timescale, as the dynamical state of a cluster in the $z = 0$ snapshot is uncorrelated to its dynamical state in snapshots before $z = 0.5$; that is, the Spearman's rank correlation coefficient between $\chi_{\rm DS}$ at $z = 0$ and $z > 0.5$ is zero. Consequently, we infer that these measures of dynamical state are dependent on the recent history of the cluster, rather than being an inherent property of the cluster that has been present since its formation. However, as Fig. 5.3 shows, the backsplash fraction is correlated with the dynamical state of a cluster. We therefore expect that the backsplash population must also be established over these relatively short timescales.

Fig. 5.4 shows that (with quite a large spread), the median backsplash fraction of each cluster is zero at $z = 1.7$, and reaches half its present day value at $z = 0.6$. Our definition of backsplash galaxies at $z > 0$ is as we describe in Section 5.2.1, and we define the backsplash fraction at a redshift $z_{\rm obs}$ as the fraction of galaxies in the radial region $[R_{200}(z_{\rm obs}), 2R_{200}(z_{\rm obs})]$ that have previously passed within $R_{200}(z)$. Consequently, if a cluster was viewed at $z > 0$, this is the backsplash fraction that would be observed, based on its radius at this time. The particularly large scatter in the data at high redshifts is due mostly to our measure of backsplash fraction; if only a small ($\lesssim 10$) number of galaxies are present in the outskirts of a cluster, then the presence of just one backsplash galaxy can dramatically change the value of F.

Fig. 5.5 shows the median backsplash fraction plotted for the relaxed and unrelaxed clusters separately. Note that the clusters are selected by dynamical state at $z = 0$, and the same clusters are then studied at each previous redshift. Consequently, the 'relaxed' sample of clusters at $z > 0$ are not necessarily those that are most relaxed at this redshift. We see that the backsplash fractions of the relaxed and unrelaxed samples agree at times before approximately $z = 0.4$. However, after this time the backsplash fraction of the relaxed clusters keeps growing, while the backsplash fraction of the unrelaxed cluster sample does not. In fact, the average backsplash fraction of the unrelaxed clusters at $z = 0$ is almost the same as it was at $z = 0.4$. This indicates that the backsplash fraction is very much dependent on the recent history of a cluster, as the two types of cluster have only become distinguishable since $z = 0.4$.

Finally, we examine how the current backsplash population has evolved. Fig. 5.6 shows the fraction of the current backsplash galaxies that were also backsplash galaxies at previous redshifts. Note the distinction between this and Fig. 5.4, as Fig. 5.6 considers only the $z = 0$ backsplash galaxies, and does not include galaxies that were members of the backsplash population at $z > 0$ but are within the cluster at $z = 0$.

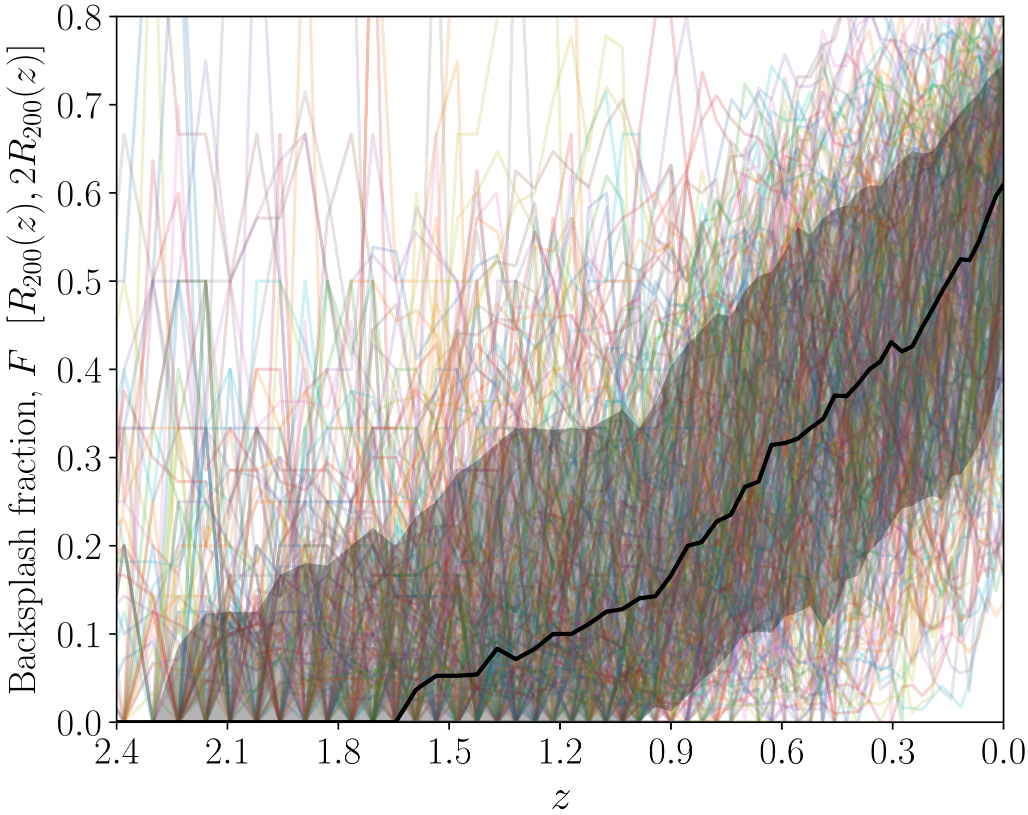

Figure 5.4: Evolution of backsplash fraction for 257 clusters, shown in colour. The median backsplash fraction plotted in black, and the shaded region shows the 68% bounds. The backsplash fraction starts to be established at $z \approx 1.7$, and increases continuously until $z = 0$.

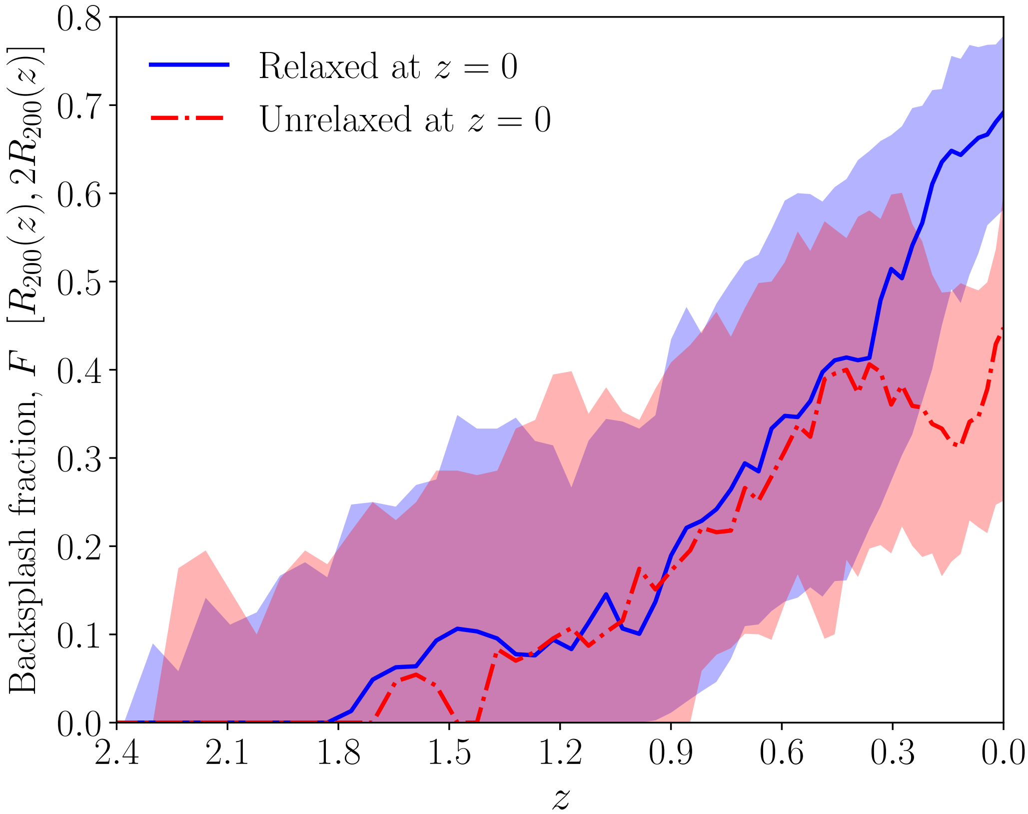

Figure 5.5: Median backsplash fraction over time, for 86 most relaxed ($\chi_{\rm DS} > 1.030$) and 86 least relaxed ($\chi_{\rm DS} < 0.619$) clusters. Note that the clusters are selected by dynamical state at $z = 0$, and the same clusters are then studied at each redshift. These clusters have built up their backsplash populations differently since $z = 0.4$, but were identical before that.

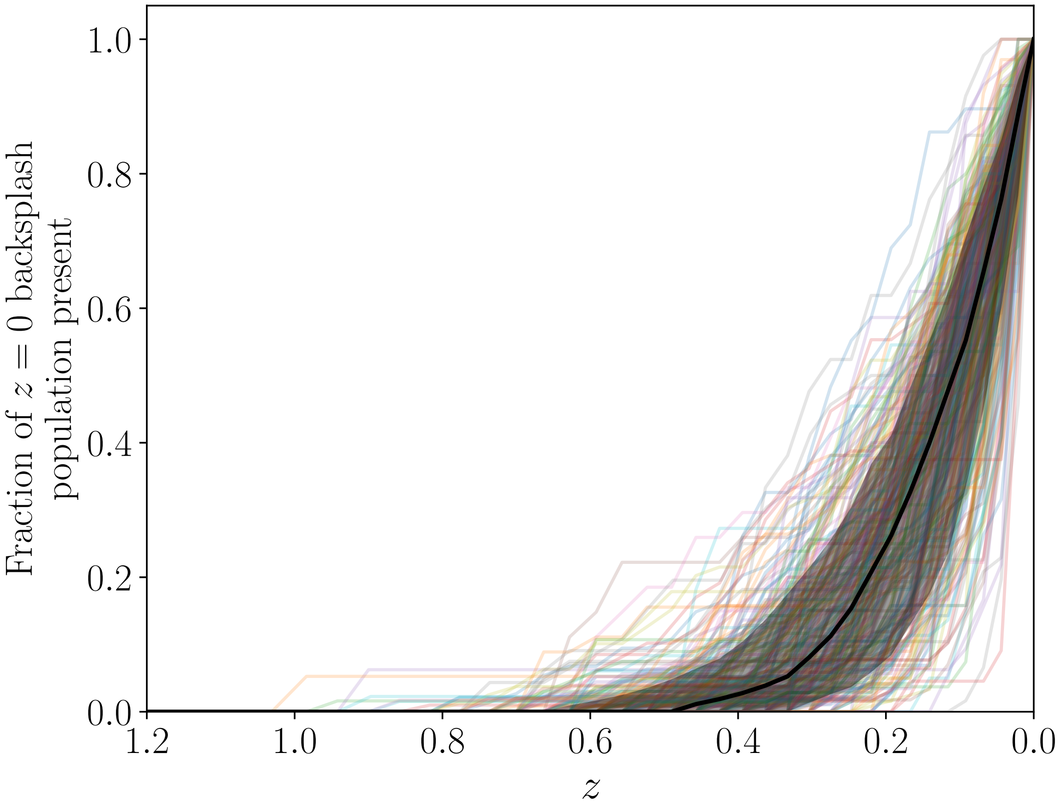

Figure 5.6: Fraction of backsplash population at $z = 0$ that are also members of the backsplash population at previous snapshots, for 257 clusters. Consequently, this shows the redshift at which the galaxies left the cluster, and passed outside of R_{200}. Median is shown in black. These times at which galaxies leave the cluster are independent of the cluster's dynamical state. Most of the current backsplash galaxies have only been in the backsplash population since approximately $z = 0.1$.

For a typical cluster, the present-day backsplash all become members of the backsplash population after $z = 0.5$, and half of the backsplash population is only built up at very late times ($z < 0.1$). This is the case for both the relaxed and unrelaxed cluster samples. This implies that there is a dynamic population of backsplash galaxies – a significant number of galaxies are joining and leaving the backsplash population, resulting in the overall backsplash fraction increasing relatively slowly, compared to the very rapid growth seen in Fig. 5.6.

Note also that in Fig. 5.6 we consider the time since a galaxy most recently left its host cluster, meaning if a backsplash galaxy has passed through a cluster twice (and so is on its third infall), we take the time since it left the cluster for the second time (i.e. the time since it was last within R_{200}). Backsplash galaxies that have passed through the cluster only once make up 90% of the backsplash population between R_{200} and $2R_{200}$, as the typical time to cross a cluster of diameter 4 Mpc is ~ 2 Gyr,

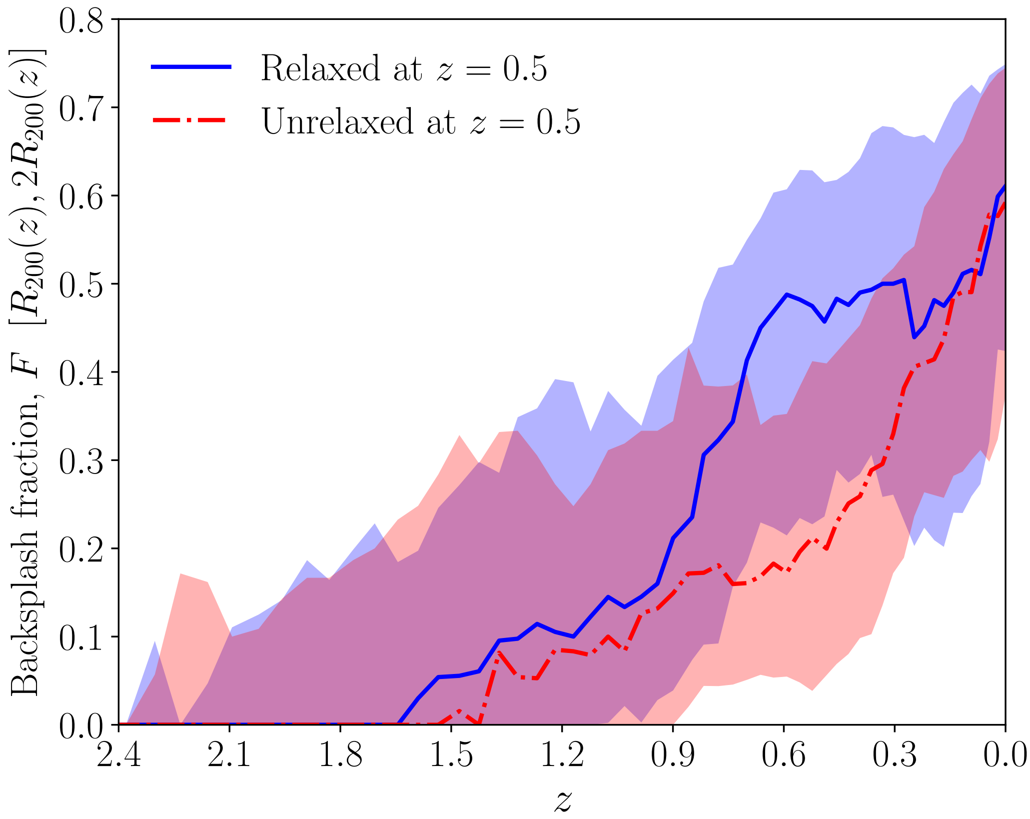

Figure 5.7: Median backsplash fraction over time, for clusters separated into two groups based on their values for $\chi_{\rm DS}$ at $z = 0.5$. Backsplash fractions are for 86 most relaxed ($\chi_{\rm DS}(z = 0.5) > 0.910$) and 86 least relaxed ($\chi_{\rm DS}(z = 0.5) < 0.586$) clusters. As we make the distinction in dynamical state at $z = 0.5$, the relaxed/unrelaxed samples of clusters are different to the samples used in Fig. 5.5.

meaning that only backsplash galaxies that enter the cluster at very early times are able to pass through the cluster a second time. A crossing time of 2 Gyr is consistent with Fig. 5.6, as this period corresponds approximately to the time between $z = 0.2$ and $z = 0$, which is the time over which most of the current backsplash population is built up.

The recent build-up of the backsplash population also gives further support to the idea that the observed backsplash fraction of a cluster is strongly dependent on its recent history. This is corroborated by Fig. 5.7, which shows how the backsplash fractions of clusters evolve, when the clusters are separated by dynamical state at $z = 0.5$. We see comparable behaviour to Fig. 5.5; the backsplash populations grow at similar rates, but that of the unrelaxed clusters plateaus after $z = 0.9$, whilst F continues to grow for the relaxed clusters. Note that the backsplash fractions of these two cluster samples are the same at $z = 0$, indicating that the dynamical state of a cluster at $z = 0.5$ does not affect its backsplash population at $z = 0$.

5.3.3 Role of mergers

An interpretation of the dynamical state is that it represents a measure of the formation history and growth of a cluster. For example, Wen & Han (2013) determine the dynamical state of clusters based on observable quantities, and describe how large amounts of substructure in clusters (which they use as a measure of dynamical state) can be produced by major merger events. Contreras-Santos et al. (2022) use data from The Three Hundred to show that $\chi_{\rm DS}$ is strongly affected by the merger history of a cluster. Similarly, we find that the more relaxed clusters are those whose formation time is earlier – that is, they are currently going through a phase of slow accretion, after an earlier phase of fast accretion.

$z_{\rm form}$ is the redshift at which M_{200} is equal to half its value at $z = 0$, as defined in Mostoghiu et al. (2019). For the relaxed sample of clusters, the average formation redshift is $z_{\rm form} = 0.66^{+0.17}_{-0.15}$, and for the unrelaxed sample, $z_{\rm form} = 0.33^{+0.13}_{-0.09}$. This shows that the unrelaxed sample consists of clusters that have accreted much of their mass in recent times, potentially through an event such as a major merger, and have consequently had a rapid recent growth in M_{200} and R_{200}. Fig. 5.8 demonstrates that $\chi_{\rm DS}$ contains information about the formation history of a cluster, back to at least $z = 1$. However, this does appear to contradict our previous result, that when comparing relaxed and unrelaxed clusters based on their backsplash fractions, the two types of cluster are indistinguishable before $z \gtrsim 0.4$.

Fig. 5.9 shows the change in R_{200}, M_{200}, $\chi_{\rm DS}$ and backsplash fraction, F, for a single cluster, as an example of this process. Between $z = 1.0$ and $z = 0.5$, the cluster mass increases by approximately a factor of four, and its radius increases by 50%, indicating a period of rapid growth. In approximately the same period, $\chi_{\rm DS}$ drops from a maximum value of 1.44 (indicating a relaxed cluster) to 0.37 (unrelaxed), and the backsplash fraction in the outskirts of this cluster drops from 39% to 7%. This rapid increase in mass is followed by a period of near-constant cluster mass (during which the backsplash fraction and relaxation increase), and then a small merger event at $z = 0.1$, which causes F and $\chi_{\rm DS}$ to drop again.

To determine whether these periods of rapid mass accretion are responsible for the suppression of a backsplash population, we stack the mass evolution profiles and backsplash fraction histories for a large sample of clusters, as well as the evolution of their dynamical states, shown in Fig. 5.10. Specifically, we find 74 instances where the mass of a cluster increases by at least a factor of three within 10 snapshots ($\sim$ 3 Gyr) after $z = 1$, and stack these 74 events. For each individual event, we select the window of 10 snapshots in which the increase in mass is greatest, and shift

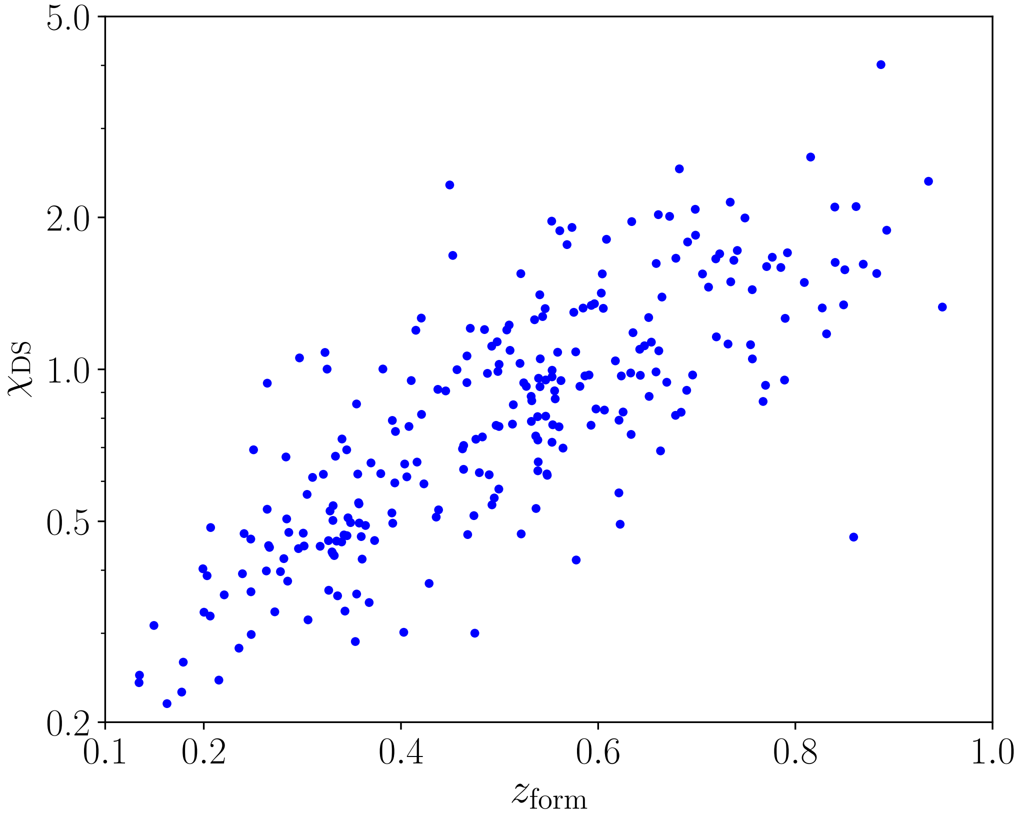

Figure 5.8: Relaxation at $z = 0$, $\chi_{\rm DS}$, against formation redshift, $z_{\rm form}$, as defined in Mostoghiu et al. (2019), for the 257 clusters. It is clear from this that the more dynamically relaxed clusters are those that accrete much of their mass at early times ($z \gtrsim 0.5$).

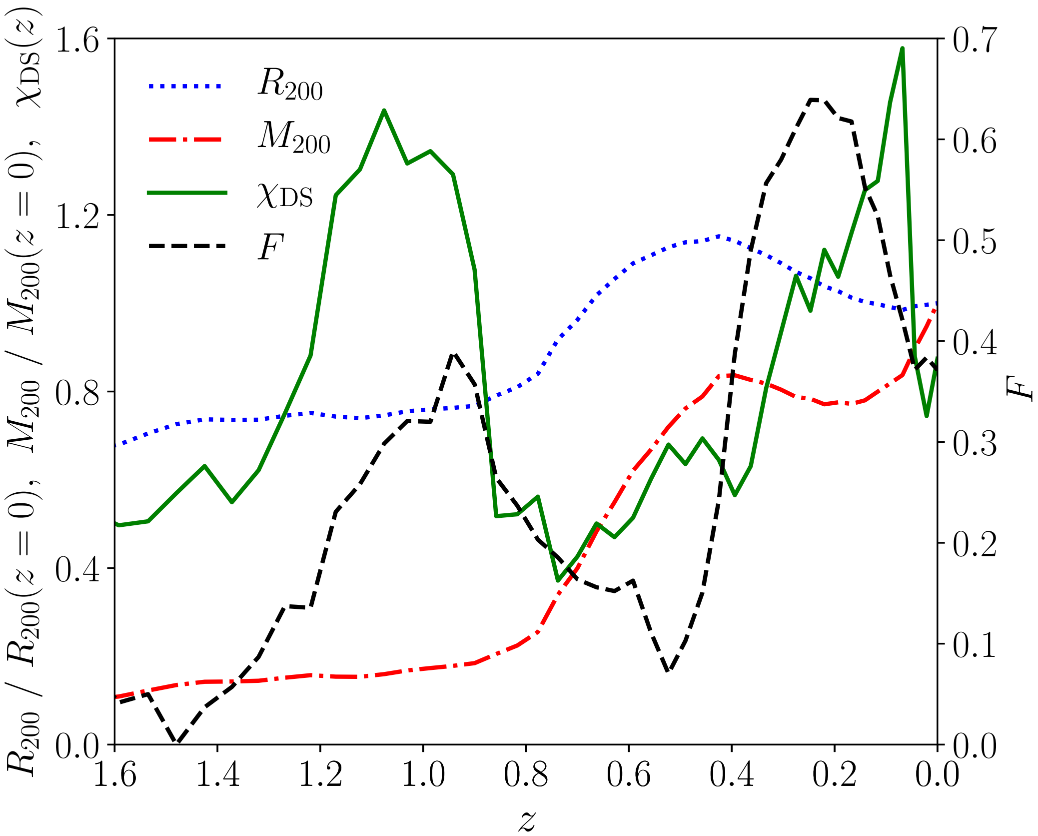

Figure 5.9: Evolution of the backsplash fraction, cluster radius, relaxation parameter and cluster mass (including subhalo masses) for one example cluster (with $\chi_{\rm DS}(z=0) = 0.88$). Note the period of rapid mass accretion between $z = 1.0$ and $z = 0.5$, which corresponds approximately with a decrease in backsplash fraction, and with a period in which the cluster is less relaxed.

the snapshot numbers, s, such that this window corresponds to the range between $s_{\mathrm{merger}} - 10$ and s_{merger}, as demonstrated in Fig. 5.10. In doing so, we assume that the time elapsed between each snapshot is identical, which is approximately the case for snapshots at $z \leq 1$.

Across these similar events, the median cluster mass increases in 10 snapshots by approximately a factor of four, from $0.18^{+0.12}_{-0.06} M_{200}(z=0)$ to $0.78^{+0.22}_{-0.33} M_{200}(z=0)$, either due to a series of merger events, or a rapid period of smooth accretion. In this same period, the median backsplash fraction drops from $26^{+26}_{-18}\%$ to $15^{+12}_{-10}\%$, and the median relaxation parameter decreases from $1.06^{+0.61}_{-0.43}$ to $0.41^{+0.25}_{-0.14}$. The median cluster radius, R_{200}, also increases in this period, by a factor of 40%. This confirms that the backsplash fraction in the outskirts of a cluster is reduced during and immediately after undergoing a merger or period of rapid accretion, and that such periods also place the cluster into an unrelaxed state, although there is still a large spread in the backsplash fraction between clusters. Fig. 5.10 also demonstrates how the reduced backsplash fraction returns within ~ 10 snapshots to the value we would expect if a merger had not taken place, indicating that only very recent periods of rapid growth will cause the present-day backsplash population to be suppressed.

Finally, we see from Fig. 5.10 that the dynamical state returns to its original value over a longer timescale than the backsplash fraction (> 10 snapshots). This difference in timescales over which χ_{DS} and F are sensitive to merger events explains the result from Fig. 5.8, which shows that χ_{DS} correlates with z_{form} for $z \leq 1$, despite F only being dependent on the history of the cluster for $z \leq 0.4$.

5.3.4 Radial backsplash profiles

To further investigate the effect of a sudden increase in cluster mass (and hence cluster radius), we also examine the radial profiles of the backsplash population. The median radial profiles of the relaxed and unrelaxed cluster samples are given in Fig. 5.11 – this plot shows the fraction of galaxies at a given radius that are backsplash galaxies. Note that the backsplash population is almost entirely contained within $2R_{200}$ for the unrelaxed clusters, but the relaxed clusters typically have backsplash galaxies present at distances up to $2.5R_{200}$ from the cluster centre – a similar trend is shown in Fig. 5 of Kuchner et al. (2022). This is also consistent with other work on the radial dependence of backsplash fractions. For example, Bahé et al. (2013), who use the same definition for backsplash fraction as us, study the radial backsplash fractions of clusters and groups using the GIMIC suite of hydrodynamical simulations. Although the clusters they examine have lower masses than our sample, they describe the radial

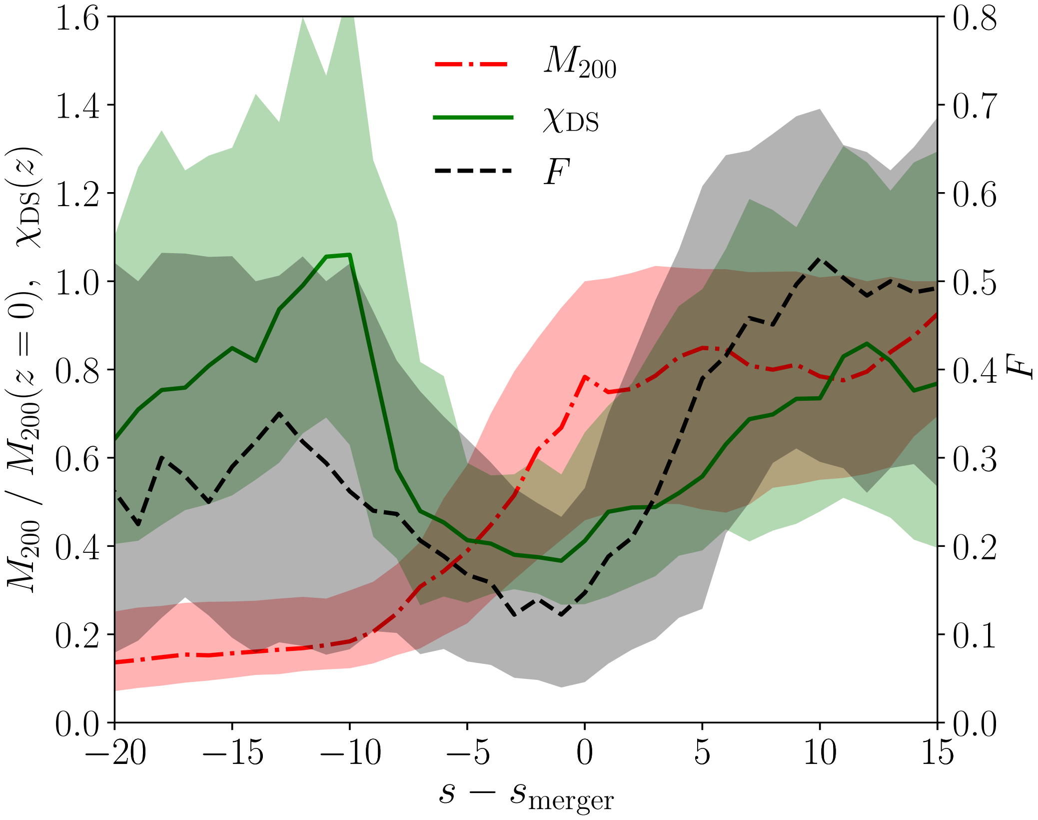

Figure 5.10: Stacked mass profiles, dynamical state profiles and backsplash fraction profiles, for 74 merger events in which there was a factor of three increase in mass within 10 snapshots at $z < 1$. Snapshot number, s, is adjusted for each event relative to the snapshot at which we define the merger event to be finished, $s_{\rm merger}$, such that the factor of three mass increase occurs between $s = s_{\rm merger} - 10$ and $s = s_{\rm merger}$.

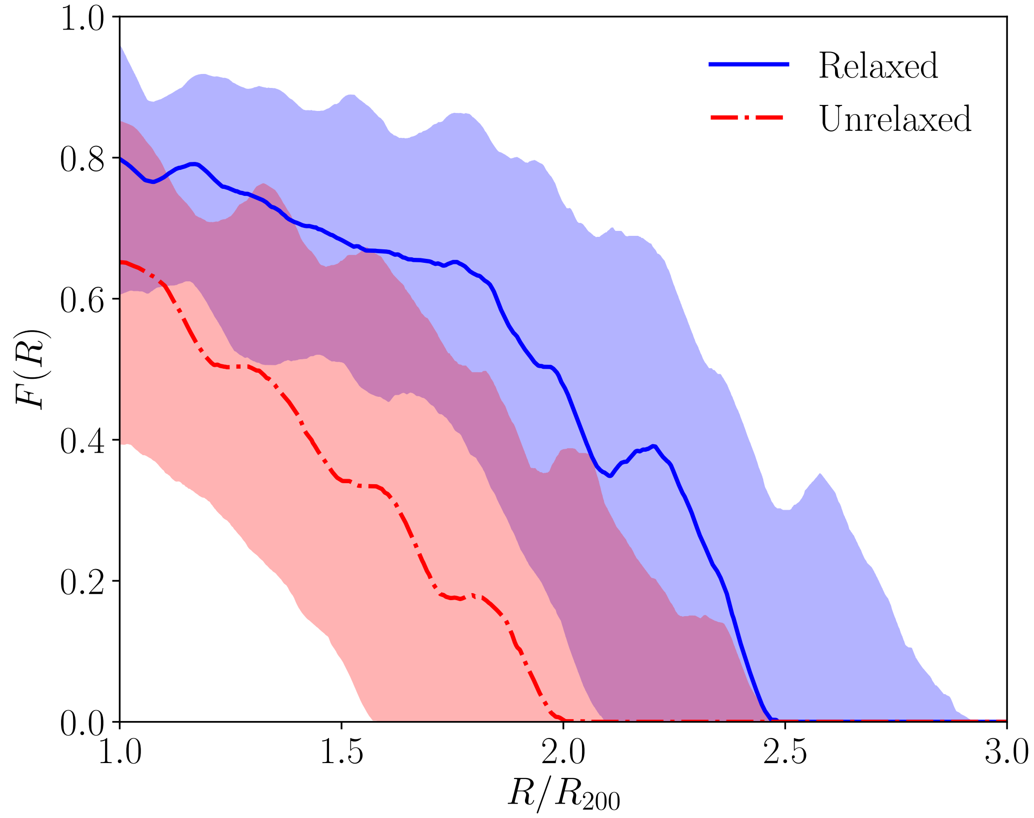

Figure 5.11: Median backsplash fraction as a function of radius, for relaxed and unrelaxed clusters at $z = 0$. Unrelaxed clusters have almost all of their backsplash galaxies containined within R_{200}, but they extend to greater distances in relaxed clusters.

backsplash fraction of one cluster with mass $M_{200} = 10^{15.2} M_{\odot}$, a typical mass for our sample of clusters. The backsplash fraction of this cluster agrees with our 'relaxed' radial profile, and drops below 5% at $R = 2.75R_{200}$.

This radial dependence shows why the backsplash fraction is lower in the unrelaxed clusters. As this sample have experienced a rapid increase in radius, the region we call the 'cluster outskirts' ($[R_{200}, 2R_{200}]$) has been pushed out to greater distances, and insufficient time has passed to allow this region to be populated with backsplash galaxies. Furthermore, the mass of the unrelaxed clusters has increased rapidly at a time shortly before $z = 0$, meaning that the potential well of the cluster is deeper, and so more difficult for galaxies to escape from and become backsplash. Consequently, the backsplash populations of the unrelaxed clusters are found at lower radial distances from the cluster, and fewer backsplash galaxies are present.

Previous work has hinted at this dependence of the backsplash population on the dynamical state of a cluster, by studying the splashback radius of clusters in different dynamical states, and accreting material at different rates. Numerous studies of the splashback feature in N-body simulations (Diemer & Kravtsov, 2014; Diemer et al., 2017) and in models of collapsing dark matter haloes (Adhikari et al., 2014; More et al., 2015) have found that the ratio between the splashback radius, $R_{\rm sp}$, and R_{200} is smaller in rapidly-accreting, unrelaxed clusters. Diemer & Kravtsov (2014) go on to explain that these clusters also typically form at later times, as indicated by Fig. 5.8. This reduction in $R_{\rm sp}$ relative to R_{200} appears to be analogous to our findings, that the backsplash population in unrelaxed clusters does not extend as far from the cluster centre as it does in relaxed clusters, resulting in a lower backsplash fraction around these clusters.

Generally, it appears that for a particularly large backsplash fraction to build up, a cluster must remain in a relaxed, stable state for extended period of time, to allow a significant amount of infalling galaxies to pass through and join the backsplash population.

5.3.5 Observational analogues for backsplash

As shown in Fig. 5.3, along with the general relaxation parameter $\chi_{\rm DS}$, the fraction of mass in subhaloes, $f_{\rm s}$, also correlates with backsplash fraction; relaxed, early-forming clusters with high backsplash fractions have a lower fraction of mass in subhaloes ($f_{\rm s} = 0.08^{+0.03}_{-0.02}$, compared to $0.19^{+0.05}_{-0.04}$ for the unrelaxed sample). This is consistent with the work of Wu et al. (2013), who show that clusters with earlier formation times have a lower fraction of their mass stored in subhaloes.

Directly measuring the total fraction of mass within subhaloes of a cluster with good accuracy is non-trivial. However, the luminous material within galaxies can be detected, and because The Three Hundred clusters use full-physics hydrodynamics, we are able to consider which properties of the clusters can be determined, based on observable quantities. Specifically, we are interested in which measures of the dynamical state can be determined. As the distribution of satellite galaxies in a cluster is a result of the distribution of subhaloes, we use the total stellar mass of galaxies, which is detectable by cluster surveys, as a proxy for the fraction of mass in subhaloes.

We make a radial cut at $3R_{200}$ around each cluster, and project the galaxy positions along a line-of-sight. We then take the total stellar mass of all galaxies found in the radial region $[0.2R_{200}, R_{200}]$, given by $M_{\rm star,200}$, and divide this by the total stellar mass within $0.2R_{200}$ of the cluster centre, $M_{\rm star,BCG}$. Note that we keep the same constraints on galaxies as are used throughout this work, such that we only consider galaxies with $M_{\rm star} \geq 10^{9.5} M_{\odot}$. The total stellar mass within $0.2R_{200}$

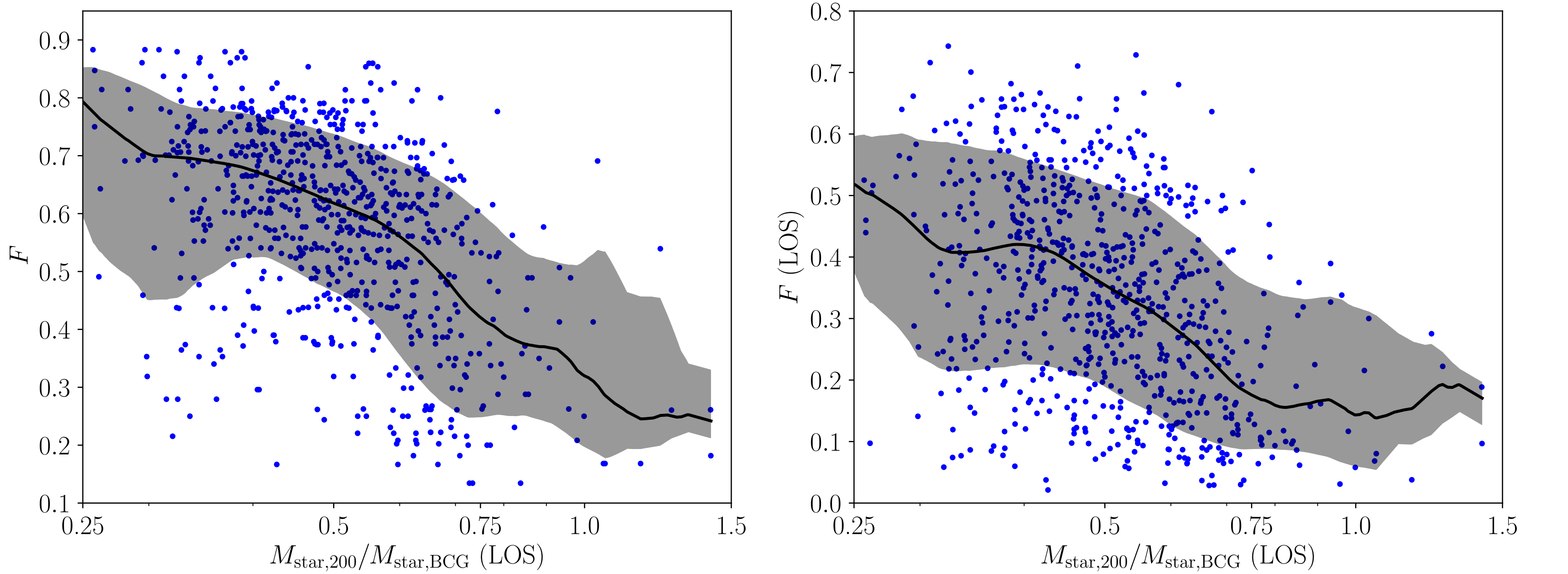

Figure 5.12: Ratio of stellar mass between $0.2R_{200}$ and R_{200}, $M_{\mathrm{star,200}}$, to stellar mass within $0.2R_{200}$, $M_{\mathrm{star,BCG}}$, against backsplash fraction, F, for 257 clusters. Left panel shows values of F calculated using full 3D data on each cluster, and hence is the backsplash fraction in the radial region $[R_{200}, 2R_{200}]$. Right panel shows backsplash fractions for a line-of-sight projection; that is, the fraction of galaxies in a 2D projected annulus between $[R_{200}, 2R_{200}]$ that have previously passed within R_{200} in 3D space. In both plots, $M_{\mathrm{star,200}}$ and $M_{\mathrm{star,BCG}}$ are found from line-of-sight projections at $z = 0$. Three orthogonal lines-of-sight are used for each cluster, which have different line-of-sight stellar mass ratios and line-of-sight backsplash fractions, but the same intrinsic 3D backsplash fraction. Consequently, 771 data points are used in each panel. The median backsplash fraction as a function of the stellar mass ratio is also shown for each plot.

corresponds approximately to that of the brightest cluster galaxy (BCG), which is usually within $0.1R_{200}$ of the cluster centre, and whose brightness provides a measure of the total cluster mass (Lin & Mohr, 2004). Using this ratio as a measure of the fraction of total mass in subhaloes (and hence, as a measure of dynamical state) is in line with the work of Wen & Han (2013), who use the steepness of the cluster's radial brightness profile as a measure of dynamical state, and describe how the light of relaxed clusters tends to be dominated by the stellar material of a single, very luminous BCG.

We find that for all of the clusters except for three, $M_{\rm star,200}$ has a value between $0.25M_{\rm star,BCG}$ and $1.5M_{\rm star,BCG}$. Fig. 5.12 shows the variation of backsplash fraction with this ratio of stellar masses, both in absolute terms, and with the $z = 0$ galaxy positions projected along the line-of-sight onto an observational plane, such that the line-of-sight backsplash fraction represents the fraction of galaxies in an observed annulus between R_{200} and $2R_{200}$ that are backsplash galaxies.

We see that clusters with a low stellar mass between $0.2R_{200}$ and R_{200}, relative to the stellar mass within their inner region, have a greater backsplash fraction than clusters with large populations of satellite galaxies containing large amounts of stellar material. For example, for clusters with $M_{\rm star,200} = 0.5M_{\rm star,BCG}$, the median backsplash fraction is $62^{+12}_{-15}\%$, or $35^{+16}_{-17}\%$ as measured along the line-of-sight. However, for clusters with $M_{\rm star,200} = M_{\rm star,BCG}$, we find that $F = 32^{+20}_{-12}\%$, equivalent to $14^{+17}_{-8}\%$ along the line-of-sight. This agrees with the trend observed in Fig. 5.3, where clusters with less mass contained in subhaloes (and hence more mass in the central cluster region, compared to in satellite galaxies) have greater populations of backsplash galaxies. The lower line-of-sight backsplash fractions are also as expected, due to the presence of greater-distance interlopers which are less likely to be members of the backsplash population.

We therefore conclude that this observable quantity, $M_{\rm star,200}/M_{\rm star,BCG}$, can act as a proxy for both the intrinsic backsplash fraction of a cluster, and for the backsplash contamination of its observed outskirts, indicating a method that cluster surveys would be able to use in order to account for the backsplash fraction when studying clusters. However, it is important to note that, due to the large spread in these data, this method would be best applied to large ensemble studies of many clusters, rather than to explain trends observed within individual clusters.

Clearly, in calculating these backsplash fractions, we assume that the radius of each cluster, R_{200}, is known exactly. However, the ratio of stellar masses is not strongly dependent on the value of R_{200} which is used. All of the clusters have radii between 1.3 h^{-1} Mpc and 2.3 h^{-1} Mpc, with a median R_{200} of 1.5 h^{-1} Mpc. If, instead, we assume this median value as the radius of each cluster, then the impact on the observational ratios shown in Fig. 5.12 is relatively small; 95% of the clusters experience a $< 20\%$ change in the ratio between $M_{\rm star,200}$ and $M_{\rm star,BCG}$, corresponding to a negligible difference in the inferred backsplash fraction, and further demonstrating that this stellar mass ratio is a potential means for estimating the backsplash fraction of clusters.

5.4 Conclusions

In this chapter, we have determined the fraction of galaxies in the outskirts of a cluster that are members of the 'backsplash population': galaxies that have travelled along a path through the centre of a cluster and now reside in its outskirts, either receding from the cluster centre or on a subsequent infall. We have studied the time-dependence and radial-dependence of this population, discussed physical processes that could impact the prevalence of these objects based on our definition of dynamical state, and proposed observable quantities that could reveal the backsplash population of clusters in future surveys. Our findings are summarised below:

- Across the 257 clusters we consider, 58% of galaxies found in the radial region $[R_{200}, 2R_{200}]$ around a cluster are backsplash galaxies. However, there is a large variation in this fraction between clusters; 95% of the clusters have a backsplash fraction between 21% and 85%.

- Clusters that are dynamically relaxed have a higher fraction of backsplash galaxies. Across our sample, approximately 70% of galaxies (between R_{200} and $2R_{200}$) in the third of clusters we deem 'most relaxed' are backsplash, compared to 45% in the 'least relaxed' third of clusters. For the least relaxed decile, this fraction drops even further, to a median value of approximately 30%.

- 50% of the backsplash galaxies at the present time only become backsplash galaxies (i.e. leave the region $R < R_{200}$) after $z = 0.1$, and less then 10% of the current backsplash galaxies have been continuous members of the backsplash population since before $z = 0.3$. Consequently, the $z = 0$ backsplash population is strongly dependent on the recent history of a cluster. In particular, clusters that have experienced a large increase in mass (and hence R_{200}) at recent times, either through rapid accretion or a major merger, have a suppressed backsplash fraction. Typically, we find that a cluster increasing in mass by a factor of three over ~ 3 Gyr will experience greater than a factor of two drop in backsplash fraction over the same period. We find that the clusters with a large increase in mass at late times (that is, a lower formation redshift, z_{form}) are less relaxed, and so contain a lower backsplash fraction.

- The backsplash galaxies in clusters we identify as 'unrelaxed' are mostly found within $2R_{200}$, whilst the backsplash population in 'relaxed' clusters extends to distances of $2.5R_{200}$. Almost no backsplash galaxies exist beyond $3R_{200}$.
- Measuring the stellar mass in a galaxy cluster relative to the stellar mass in its central region can allow the backsplash fraction of a cluster to be estimated, with an absolute uncertainty of approximately 10%. Clusters with outer regions containing large amounts of stellar material typically have a backsplash fraction of 30%, which rises to 60% for clusters dominated by a bright BCG. The backsplash fraction as measured by an observer – the fraction of galaxies in a projected annulus that are members of the backsplash population – can also be inferred by this measure, although not with the same precision. Typically, this line-of-sight backsplash fraction is approximately a factor of two less than the absolute backsplash fraction.

Our findings demonstrate that backsplash galaxies are likely to have a significant impact when studying the history of galaxies in cluster environments, and make it challenging to disentangle the effects of pre-processing and of a cluster environment. It should be noted that this work does not address whether an individual galaxy can be identified as a backsplash galaxy or an infalling galaxy. However, as properties such as galaxy luminosities in various bands are available for The Three Hundred simulations, we will be able to investigate this in future work, as we discuss in detail in Section 6.3.1. Nevertheless, we show that the backsplash population can be statistically accounted for within cluster surveys, allowing the contamination of an infalling sample of galaxies by backsplash galaxies to be quantified, and thus allowing corrections to be made to radial profiles of galaxy properties in cluster environments.

As we identify backsplash galaxies based on whether they have entered R_{200} at a given redshift, 'pseudo-evolution' of the cluster radius (that is, evolution in M_{200} due to the evolution in the critical density of the Universe) could result in galaxies being mistakenly tagged as backsplash galaxies (Diemand et al., 2007). Although this is possible, it is unlikely to be the case for two reasons. Firstly, as we showed in Section 5.3.5, the backsplash fraction is not sensitive to a moderate change in R_{200} – the change we use here is similar to the change in a cluster's radius between $z = 1$ and

$z = 0$ due to pseudo-evolution (Diemer et al., 2013). Secondly, as R_{200} grows due to pseudo-evolution, any infalling galaxies that are mistakenly tagged as having entered the cluster would likely have entered the cluster anyway a very short time later. As Fig. 5.1 shows, few backsplash galaxies make 'grazing passages' of the cluster (i.e. with $D_{\min} \sim R_{200}$), meaning that, although the infall time of the backsplash galaxies might be altered slightly, they are unlikely to be incorrectly labelled as backsplash.

We emphasise that it is important to distinguish between backsplash galaxies – which have been the focus of this chapter of this thesis – and the splashback radius, which has been discussed in other recent work (More et al., 2015). The splashback radius is typically used to refer to an outer radius of a cluster, beyond which material is not expected to be virialised. The backsplash population instead refers to individual galaxies that have passed through the cluster, which can be difficult to identify observationally, but can be studied using simulations. As shown in this work, backsplash galaxies can be found at far greater distances ($\sim 2.5R_{200}$) than the typical splashback radius of a cluster ($\sim 1.5R_{200}$).

This indicates that $R_{\rm sp}$ does not necessarily represent a hard boundary, outside of which all objects are infalling, in the same way that R_{200} does not represent such a boundary. This is in agreement with the work of Diemer (2017), who shows that the splashback radius typically contains the apocentre of approximately 87% of particles within a dark matter halo, and so some particles can indeed pass beyond it. However, by considering bound haloes rather than single particles, and by including hydrodynamical effects, we have shown that a significant fraction of galaxies are also expected to pass through the cluster and travel back out to relatively distant regions.

Other potential contributions to the apparent discrepancy between splashback radius and backsplash galaxies include different typical infall speeds of unbound gas and bound galaxy haloes (Bahé et al., 2013), which could also be indicated by the presence of an accretion shock around clusters at a similar distance to the splashback radius (Arthur et al., 2019). Anisotropy in the backsplash population could also explain how backsplash galaxies can be found beyond the splashback radius; a dependence of backsplash galaxies on large-scale structure around clusters could result in these galaxies being present at greater distances in different angular regions of the cluster outskirts. This would mean that a spherical splashback radius is still an insufficient method to determine the region of a cluster's influence, and a more physically motivated cluster boundary would likely be more complex and involve multiple different variables.

Chapter 6

Summary and conclusions

Throughout this thesis, we have used the hydrodynamical simulations from The Three Hundred project to study the environmental histories of galaxies, with a particular focus on galaxy groups and backsplash galaxies. This work began with a comparison between the hydrodynamical and dark matter-only runs of this suite of simulations, to understand how the inclusion of baryonic physics can impact the dense central regions of groups and clusters. This also allowed us to verify that hydrodynamical simulations are crucial in studying these dense structures. We then examined galaxy groups being accreted onto clusters, investigating how the dynamics of these groups are impacted by passing through a cluster potential, and the impact that this has on their constituent galaxies. Finally, we identified backsplash galaxies in the outskirts of the simulated clusters – galaxies that have entered a cluster, and subsequently left again. We found that these make up a large fraction of galaxies in the cluster outskirts, and that this fraction is predictable, based on the dynamical state of a cluster.

In this chapter, we summarise these results in detail, and discuss future implications of the work in this thesis.

Contents

6.1 Summary of results

In this section, we summarise the work from Chapter 3, Chapter 4 and Chapter 5.

6.1.1 Hydrodynamical and dark matter-only simulations

We begin this thesis with a study comparing hydrodynamical simulations of galaxy clusters, to equivalent dark matter-only simulations. Many previous studies of large-scale structure and galaxy clusters have utilised dark matter-only simulations, due to their simplicity and lower computational cost relative to hydrodynamical simulations. Prior work has found that the structure and substructure predicted by these simulations differs: dark matter-only simulations have suppressed matter power spectra on sub-Mpc scales (van Daalen et al., 2011), and the shallower subhalo mass functions (Libeskind et al., 2010; Jia et al., 2020).

Despite this, dark matter-only simulations are still often used for studies of galaxy groups and clusters, as well as the starting conditions for techniques such as semi-analytic models (although these can account for the different substructure predicted by dark matter-only simulations by including 'orphan galaxies'). However, in Chapter 3, we investigated how well the substructure in clusters and infalling galaxy groups is approximated purely by a dark matter-only simulation, when compared to a physically-motivated hydrodynamical simulation. We found that:

- Compared to hydrodynamical simulations, dark matter-only simulations under-predict the number of subhaloes in group-sized haloes ($M_{200} \sim 10^{13}\ h^{-1}M_{\odot}$), and cluster-sized haloes ($M_{200} \sim 10^{15}\ h^{-1}M_{\odot}$). This under-prediction is a function of the local density, where denser regions have a greater deficit of substructure in dark matter-only simulations. In the innermost regions of a cluster ($r < 0.1R_{200}$), the average number density of galaxies is four times greater in hydrodynamical simulations, and in the centres of groups, the number density is 10 times greater. Further away from the halo centres, the deficiency of subhaloes in the dark matter-only simulations decreases.

- This same behaviour is seen in galaxy groups in the cluster outskirts, and groups falling into a cluster. Furthermore, it is not strongly affected by numerical effects, such as the mass resolution or gravitational softening length of the simulations.

- We did not find convincing evidence that this effect can be explained purely by differences in tidal stripping between the two classes of simulations, and instead attribute these 'missing subhaloes' in the dark matter-only simulations to the mechanism described by Blumenthal et al. (1986) and van Daalen et al. (2011). A dense region of baryonic material in the centre of a hydrodynamically-simulated halo can enhance the central density of dark matter in the halo. The result of this is that, in hydrodynamical simulations, subhaloes near to the centres of larger haloes will be more pronounced, and thus easier for a halo finder to detect. It is well-established that halo finders are less effective at detecting subhaloes against a high background density (Muldrew et al., 2011; Onions et al., 2012, for example), and so this effect is particularly important in group and cluster centres.

This chapter helped to confirm the need for hydrodynamical simulations when studying cosmic structure on these scales, and that making conclusions based solely on dark matter-only simulations may result in details of the evolution of this structure being missed.

6.1.2 Dynamical evolution of galaxy groups falling into clusters

After the initial work on simulations of galaxy groups in Chapter 3, we then study galaxy groups in hydrodynamical simulations in greater detail in Chapter 4. Galaxy groups play a crucial role in cluster assembly, as a large fraction of the galaxies joining a cluster do so as part of a smaller group (Berrier et al., 2009; McGee et al., 2009). Additionally, these groups can have a strong effect on the evolution of galaxies, by providing an environment for pre-processing to take place, via galaxy mergers and gas stripping (Jian et al., 2012). However, it is not entirely clear how the groups themselves evolve after entering a cluster, and whether all group members are processed in the same way.

In Chapter 4 we examined how the properties of groups change as a group approaches, and subsequently passes through, a cluster. We focused particularly on the dynamical evolution of these groups, looking at how the spatial distribution of galaxies changes, as well as the distribution in speeds of group members. Then, once the evolution of the groups in these simulations had been constrained, we looked at the processes that galaxies are statistically likely to go through after they are accreted onto a cluster as a member of a group.

- On falling into a cluster, most groups have a pericentric passage at a distance of $r < 0.6R_{200}^{\mathrm{clus,h}}$ from the cluster centre, after which the majority of the groups remain within the cluster. Of the groups that enter the cluster and are tracked until their first turnaround, the apocentres of $\sim 80\%$ of these are within $R_{200}^{\mathrm{clus,h}}$ of the cluster centre, meaning that the groups do not leave the cluster after entering. These 'sticky groups' are contrasted with 'backsplash groups', whose apocentres are outside of the cluster, meaning that they can re-enter a cluster for a second time.

- During this passage through the cluster, we find that there are two main phases in the dynamical evolution of groups. Between the moment of infall and the first pericentre (which typically occurs ~ 0.5 Gyr later), the group members increase their speeds relative to their host group, indicating an increase in kinetic energy due to the cluster's potential. Then, after the group passes pericentre, the galaxies increase their distance to the group centre, becoming spatially separated and destroying the group substructure that entered the cluster.

- The majority of galaxies bound to a group at cluster infall become gravitationally unbound from the group during its passage through a cluster, with many of these becoming unbound by the first pericentric passage. A smaller, but still large, fraction merge with the brightest group galaxy, or become very heavily stripped and have their material accreted onto the group's dark matter halo. Galaxies in different regions of the group phase space experience drastically different processes: while galaxies in the group outskirts are likely to remain undisturbed and simply join the cluster satellite population, those in the group centres are likely to merge with the group. Only fast-moving galaxies nearby to the group centre at infall typically remain as bound group members.

Because of the findings in this chapter, we are able to draw conclusions about infalling galaxy groups that can be applied to observational studies. Groups that are observed in the vicinity of a cluster are almost certainly entering the cluster for the first time, as after a single passage through a cluster, almost all members of a group are lost via one of the mechanisms described above. Groups seen nearby to a cluster will not have previously experienced stripping by a cluster environment, and will likely contain few backsplash galaxies (galaxies that have been inside a cluster in their past). This complements the findings of Chapter 5 perfectly, which are summarised below in Section 6.1.3.

6.1.3 Estimating the prevalence of backsplash galaxies

In Chapter 5 we investigated backsplash galaxies in greater detail, building further on the overall aim of this thesis, which is to understand whether we can use simulations to observationally constrain the environmental histories of galaxies. Backsplash galaxies are a prime example of why this is important, as they are galaxies that have been previously processed by a cluster, but now exist in the cluster outskirts, where we would expect to find galaxies entering for the first time. Consequently, they act as a 'contaminant', as they are difficult to disentangle from galaxies that have been pre-processed by groups and filaments.

Backsplash galaxies have previously been shown to make up about 50% of the galaxies in the outskirts of a cluster (Gill et al., 2005), but few studies exist that have investigated the dependence of this fraction on properties of the clusters. However, the closely-related splashback radius has been shown to be greater in dynamically-relaxed, slowly-accreting galaxy clusters (Diemer & Kravtsov, 2014; More et al., 2015), and so in Chapter 5 we built on this by comparing the backsplash fraction in relaxed and unrelaxed clusters. Additionally, we looked at how this fraction changed over time, and what cluster properties could be used to estimate the number of backsplash galaxies around a real, observed cluster. Our findings are summarised below:

- On average, 58% of galaxies in the radial region $[R_{200}, 2R_{200}]$ from the centre of a cluster are backsplash galaxies, which is similar to (but slightly higher than) the 50% given by Gill et al. (2005). This difference can be explained by the fact that Gill et al. (2005) use dark matter-only simulations, which we showed in Chapter 3 to contain slightly fewer backsplash galaxies than the hydrodynamical simulations that we use. There is a large scatter in this average backsplash fraction; around some clusters, more than 80% of galaxies are backsplash galaxies, while other clusters have backsplash fractions of less than 20%. This region ($[R_{200}, 2R_{200}]$) contains most of the backsplash galaxies: the fraction of backsplash galaxies is lower beyond this radius, and almost no backsplash galaxies exist beyond $3R_{200}$.
- We introduce a new measurement of cluster dynamical state, $\chi_{\rm DS}$, combining three measures previously used in Cui et al. (2017). Using this measure, we find that dynamically relaxed clusters have a higher fraction of backsplash galaxies in their outskirts, compared to unrelaxed clusters. If we split the sample of clusters into thirds based on this relaxation parameter, approximately 70% of

galaxies (between R_{200} and $2R_{200}$) in the 'most relaxed' third of clusters are backsplash galaxies. This fraction drops to just 45% in the 'least relaxed' third of clusters. Of the three parameters used to calculate $\chi_{\rm DS}$, the subhalo mass fraction, $f_{\rm s}$, correlates most strongly with the backsplash fraction, and the virial ratio, η, the least strongly. $\chi_{\rm DS}$ correlates with the backsplash fraction more strongly than any of these individual parameters, or any combination of two parameters.

- The backsplash population is already well-established several Gyr before the present day: at $z = 0.7$, a lookback time of approximately 7 Gyr, the average backsplash fraction around a cluster is already 25%. However, many of these galaxies later re-join the cluster, meaning that the present-day backsplash population was only built up at recent times. Half of the $z = 0$ backsplash galaxies only exited the cluster after $z = 0.1$, and typically none of these galaxies have been outside of the cluster since before $z = 0.5$.

- Because the backsplash population is only developed at recent times, the $z = 0$ backsplash fraction is strongly dependent on the recent history of a cluster. Clusters with recent periods of rapid growth, and so recent formation times, have a lower backsplash fraction. This increase in mass can come from a major cluster-cluster merger, or a period of rapid accretion, but generally also leads to the cluster entering a more unrelaxed state – this correlation between dynamical state and cluster growth is explored further in Contreras-Santos et al. (2022).

We conclude Chapter 5 by giving an example of how an observable property of a cluster could be used to predict the fraction of backsplash galaxies surrounding it. Clusters with a more central concentration of stellar mass are generally more dynamically relaxed, and so will have a greater fraction of backsplash galaxies in their outskirts.

6.2 Connecting simulations with observations

Studying the environments of galaxies throughout their full histories, rather than just at the present-day, is vital in order to understand the impact of environment on galaxy evolution. Backsplash galaxies will bias and skew studies of the impact of clusters on galaxy properties, and make it appear that galaxies far from the cluster have been processed, when many of these are actually just ex-cluster members. Despite this, the past environments of galaxies and the contribution of backsplash galaxies is frequently overlooked in observational studies of clusters. While this is occasionally down to carelessness, more often it is simply because of the difficulty in identifying backsplash galaxies, which makes the backsplash contamination both unpredictable and unavoidable.

Combining the findings from throughout this thesis gives some powerful insights that can be used to solve this problem, and constrain the environmental histories of galaxies observed in and around large galaxy clusters. Measures of dynamical state can tell us what portion of galaxies are ex-cluster members, and identifying groups in the outskirts of clusters can show us galaxies that are likely on their first infall to the cluster. Although these two pieces of knowledge do not allow us to determine the environmental history of a single galaxy, they allow us to draw conclusions about the histories of an ensemble of galaxies, even with a sample of just a single cluster.

In this section, we illustrate the potential to constrain the environments of galaxies around an observed cluster, by using an example cluster taken from The Three Hundred suite. We use cluster_0002, the same cluster used in Section 3.3.2 and Fig. 3.7 to study resolution effects in our simulations. Although we do not develop detailed mock observations in this thesis[1], we can consider the properties of this cluster that could be measured, and use these as a first-order demonstration of how accurately a real, observed cluster could be interpreted. Fig. 6.1 shows cluster_0002 in four different views taken from the simulation: in dark matter, gas, stars, and the bound haloes and subhaloes detected by the halo finder. The view of the cluster in stars is the most obvious choice of observation, although the distribution of dark matter can be mapped using techniques such as weak lensing (Jauzac et al., 2016; Tam et al., 2020), while the gas in clusters and their constituent galaxies can be detected using X-ray and Hα measurements (Ge et al., 2016; Sun et al., 2021).

X-ray measurements can also be used to constrain the dynamical states of clusters (Yuan & Han, 2020; Zenteno et al., 2020), and although we do not use mock X-ray maps in this work, we know from the simulations that cluster_0002 is a relaxed cluster. Specifically, it has a relaxation parameter $\chi_{\mathrm{DS}} = 1.1$, putting it in the 'relaxed' third of clusters described in Section 5.3.1. This information can also be determined observationally by the high central concentration of its stellar mass, which is a sign of a dynamically relaxed cluster (Wen & Han, 2013). In fact, approximately two thirds of the stellar mass in this cluster is contained within galaxies in the inner $0.2R_{200}$ of the cluster. As we show in Chapter 5, a relaxed cluster with

[1]Several previous and ongoing projects have used data from The Three Hundred project to produce mock observations, including X-ray maps (De Luca et al., 2021) and images of strong lensing (Li et al. in prep.).

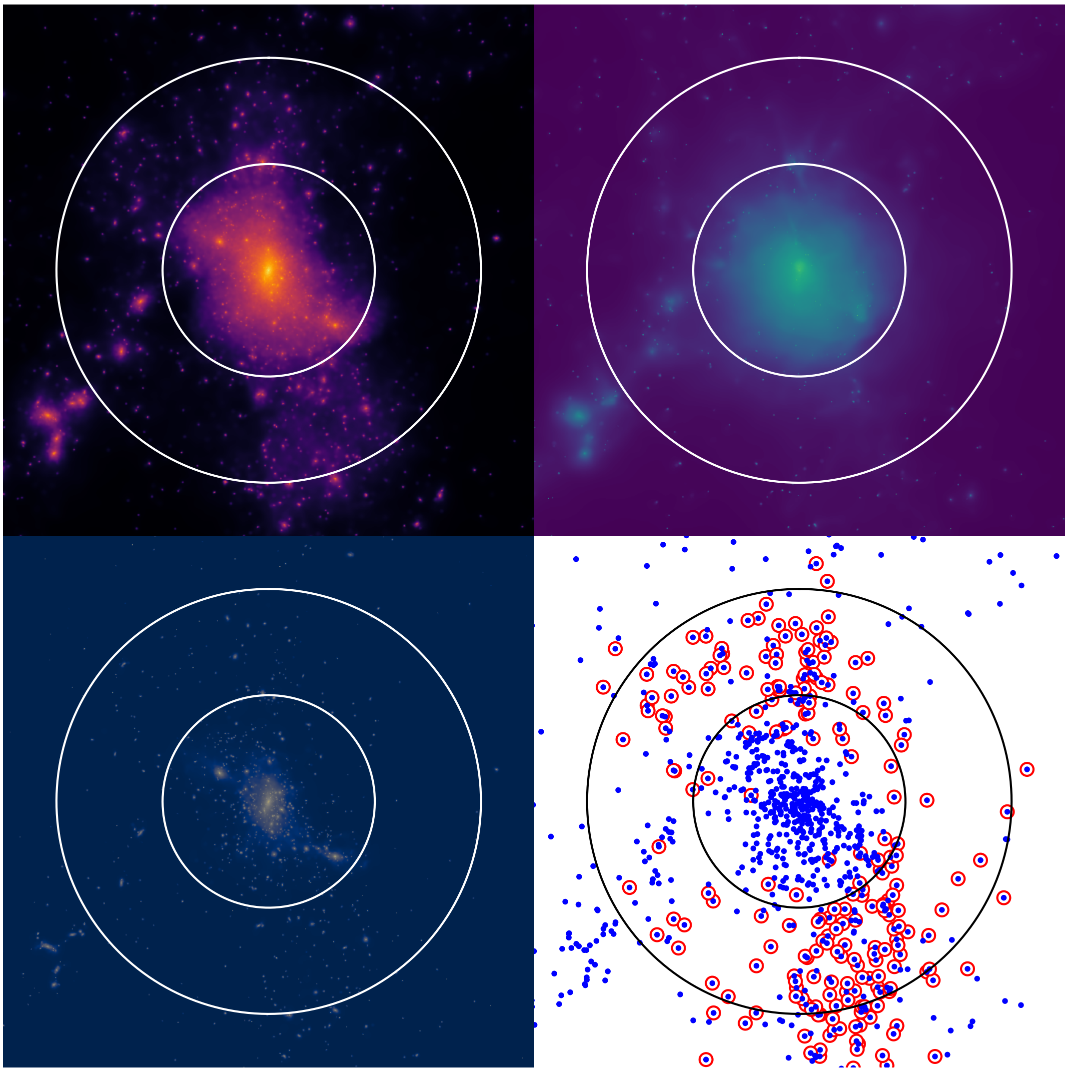

Figure 6.1: One cluster from our simulation suite, CLUSTER_0002, shown in four different views. The top-left, top-right, and bottom-left panels show the cluster in dark matter, gas, and stellar particles, respectively. Lighter colours represent a greater density in each of these panels. The bottom-right panel shows the distribution of haloes and subhaloes around this cluster, represented by blue dots. Backsplash galaxies are indicated by a red circle surrounding the dot. The two large circles on each panel indicate distances of $R_{200}^{\rm clus}$ and $2R_{200}^{\rm clus}$ from the cluster's centre. Note that some backsplash galaxies appear to be inside of the cluster (due to projection effects), and that some lie outside of $2R_{200}^{\rm clus}$.

stars concentrated this strongly in the centre should be surrounded by a sizeable backsplash population: Fig. 5.12 allows us to predict that $65^{+11}_{-15}\%$ of galaxies between $[R_{200}, 2R_{200}]$ from the cluster centre are backsplash galaxies (1σ uncertainty), or $40^{+14}_{-18}\%$ of galaxies when accounting for projection effects.

Additionally, in the left/bottom-left of this cluster, there are several large haloes, representing galaxy groups in the cluster outskirts (these are most clearly visible in the dark matter plot, but can be seen in the other panels of Fig. 6.1). In Chapter 4, we determined that these groups are almost certainly on their first infall towards the cluster centre. Consequently, despite the prediction that a large fraction of the galaxies in the cluster outskirts are backsplash galaxies, we can also predict that the galaxies in these groups are not members of the cluster backsplash population.

Because we are using simulated data here, we are able to check how successful these predictions are, by highlighting the backsplash galaxies in the bottom-right panel of Fig. 6.1 (circled in red). Indeed, as predicted, a large fraction of the galaxies between $[R_{200}, 2R_{200}]$ from the cluster centre are backsplash galaxies, but there is a lack of backsplash galaxies on the left of the cluster, where the infalling galaxy groups are found. This worked example demonstrates that we can not only predict the contamination of 'infalling cluster galaxies' by backsplash, but also successfully predict that this contamination will be different in different angular regions of the cluster.

This is just one example cluster, and although the methods laid out in this thesis work well in this example, there will likely be some cases where these predictions do not work as well. This is particularly apparent in the right panel of Fig. 5.12, which shows the very large spread in backsplash fraction of clusters. Part of the reason for this is the dependence on the line-of sight; if a large galaxy group is approaching a cluster along the line-of-sight, the first-time infalling galaxies will be disproportionately located along the line-of-sight, rather than in the projected cluster outskirts. Consequently, the fraction of backsplash galaxies in the projected outskirts is likely to be greater than predicted using the methods in this work.

However, in the context of future surveys where observational data from many clusters will be stacked together, this approach allows the environmental histories of a large ensemble of galaxies to be understood in far greater detail than has been possible before. While it is only a first-order solution, being able to isolate regions of observed clusters containing small amounts of backsplash contamination will allow us to finally draw unbiased conclusions on the impact of cluster environments on galaxy evolution. These findings can also form the basis of more complex methods, that will allow us to constrain galaxies' histories even better. We discuss some of these approaches in the following section.

6.3 Future work

There are two natural ways to extend the work in this thesis. Firstly, by investigating a greater variety of cosmic environments, for example, cosmic filaments. Using simulations, we could further constrain the fraction of galaxies that are affected by these, and how this differs between clusters. Alternatively, rather than focusing on large ensembles of galaxies and stacked cluster data, we can study individual galaxies, to learn about the environmental histories of individual galaxies, or smaller groups of galaxies.

6.3.1 Individual backsplash galaxies and machine learning

The work in Chapter 5 involves making a statistical estimation of backsplash fraction, from an ensemble of galaxies around a cluster. We showed in Section 6.2 how this estimate of the backsplash fraction could be refined, by using our knowledge of galaxy groups to correctly predict which regions around a cluster contained few backsplash galaxies. However, this could be taken even further by considering whether individual backsplash galaxies can be identified in astronomical observations, based on either observable properties of a cluster, or the galaxies themselves. Critically, this would allow us to not only quantify the contamination of the cluster outskirts by backsplash galaxies, but reduce this contamination by removing these galaxies from our observations.

We have recently begun a project investigating this, also using The Three Hundred simulations, with the hope that we can identify individual galaxies as backsplash, based on their observable properties. It is not immediately clear which galaxy or cluster properties are best for identifying backsplash galaxies, and so to solve this problem we have used machine learning techniques, which we train on simulated data from The Three Hundred project. Machine learning allows us to explore relationships between variables, with little prior knowledge of the strength or nature of the relationships between these variables.

Specifically, we are using a random forest classifier (Breiman, 2001) to identify backsplash galaxies, which is a relatively simple machine learning technique. For its input data, it uses a set of 'elements', each of which has a number of 'features', and is assigned to a 'class': in our case, each galaxy around a cluster is an element, which has numerous scalar properties ('features') such as its cluster-centric distance and mass, and falls into one of two classes, 'backsplash' or 'not backsplash'. With this data, a single decision tree is constructed – first, the data is split into two subsets, by choosing a boundary value in one of the features that splits the data into two classes as cleanly as possible. Each of these subsets is then split again based on another feature, and this process repeats until a decision tree has been built that separates the data into numerous subsets, each of which (mostly) consists of elements of a single class. A random forest classifier then repeats this process, by sampling the full set of input elements (for example, by bootstrap sampling), and generating a new, independent tree. These trees can then be combined to act as an ensemble, and predict which class an element will fall into, based on its features (see Piotrowska et al., 2022, for a descriptive summary of random forest classifiers).

Indeed, our preliminary work has shown that, by training a model of this type on simulated data from The Three Hundred project, we can successfully predict whether individual galaxies are infalling for the first time, or are backsplash galaxies. This model assigns a probability of being a member of the backsplash population, P, to each galaxy – KDEs representing the distribution of this value for backsplash and infalling galaxies are shown in Fig. 6.2.

Importantly, the input features for this predictor are all observable properties of a galaxy (like its projected distance from the cluster centre, line-of-sight velocity and stellar mass), rather than those that can only be determined in a simulation (such as a galaxy's 3D position) meaning that, although we have tested the model on simulated data, it could also be run on observational data of real clusters. This therefore represents an important proof of concept: using fairly rudimental observable properties of galaxies, we can extract additional information from observations, and classify galaxies by their environmental histories. The first iteration of this model has been able to take a sample of galaxies in the outskirts of a cluster (which are approximately 50% backsplash galaxies), and produce a sample of infalling galaxies with a purity of approximately 90%. Similarly, samples of backsplash galaxies can be constructed, with a purity of about 80% (Fig. 6.2).

Some features are of more importance to the classifier than others. Backsplash galaxies were best identified based on three observables: they are gas-poor, found nearby to R_{200} of the cluster, and have low line-of-sight speeds. This is in agreement with other studies in this area: Borrow et al. (2022) also find that gas content and cluster-centric distance are important in distinguishing backsplash and infalling galaxies. We have also extended this model to include other features – for example,

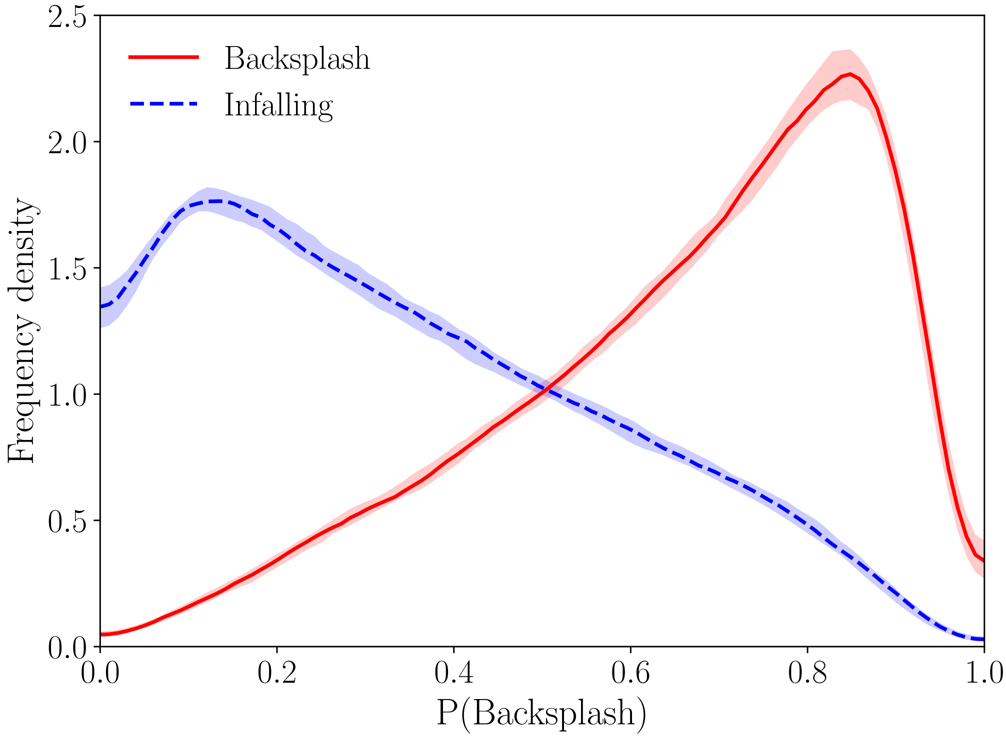

Figure 6.2: Random forest classifier returns a probability, P, that each galaxy is a backsplash galaxy, where $P = 1$ indicates that the classifier is 100% certain that the galaxy is backsplash. This plot shows the distribution of P for backsplash galaxies and non-backsplash (infalling) galaxies, located at a projected distance between R_{200} and $2R_{200}$ from the cluster centre. By only selecting galaxies with a value of P above or below a certain boundary, we can create pure samples of galaxies. For example, keeping only galaxies with $P > 0.8$ will return a sample made up predominantly of backsplash galaxies.

the stellar mass to halo mass ratio can be used to identify backsplash galaxies (as we also described briefly in Section 5.3.1). The angle between a galaxy's projected motion across the sky and the projected positional vector towards the cluster centre is also a strong indicator of backsplash galaxies – first-infallers are almost all on radial paths towards the cluster centre, but backsplash galaxies can be moving tangentially or away from the centre (this result was also shown in Knebe et al., 2020). These features are more challenging to measure, but could potentially be used – for example, the distortion of jellyfish galaxies can reveal the direction they are moving in the plane of the sky.

The next stage in this work is to improve this model further by including more properties of galaxies and galaxy clusters, such as different measures of cluster dynamical state, in order to construct large samples of galaxies that have had the same environmental histories. Findings such as those from Kuchner et al. (2022) could also be built into the model. In that paper, it is shown that backsplash galaxies have a characteristic distribution in their locations around a cluster (relative to their infall angle). This distribution depends on the dynamical state of the cluster, and whether the backsplash galaxies are receding from the cluster, or on a second infall – this is shown in Fig. 6.3, adapted from Kuchner et al. (2022). Using information like this could help to predict whether the amount of backsplash galaxies is different in different angular regions of a cluster, as we did in Section 6.2, above. Finally, we will improve the model further by incorporating more simulated data of galaxy properties, including the other hydrodynamical simulations from The Three Hundred, as well as the semi-analytic models that have been run on this same dataset.

Ultimately, this work could be applied to real, observational data, to tackle the same problems that we have approached throughout this thesis. Much of the work for this project is at an advanced stage, meaning that this machine learning model will likely be applied to data from cluster surveys in the near future, such as the WEAVE Wide-Field Cluster Survey (WWFCS, Kuchner et al. in prep.), due to begin in late 2022. Crucially though, as well as quantifying the contamination of a cluster's outskirts, this method would allow us to build samples of infalling and backsplash galaxies. This will be a huge help in studying galaxy evolution: it will not only solve the problem of having impure samples of infalling galaxies, but will also allow us to build samples of galaxies that were processed by clusters several Gyr ago, which could be studied to help understand the long-term impact of environment on galaxy properties.

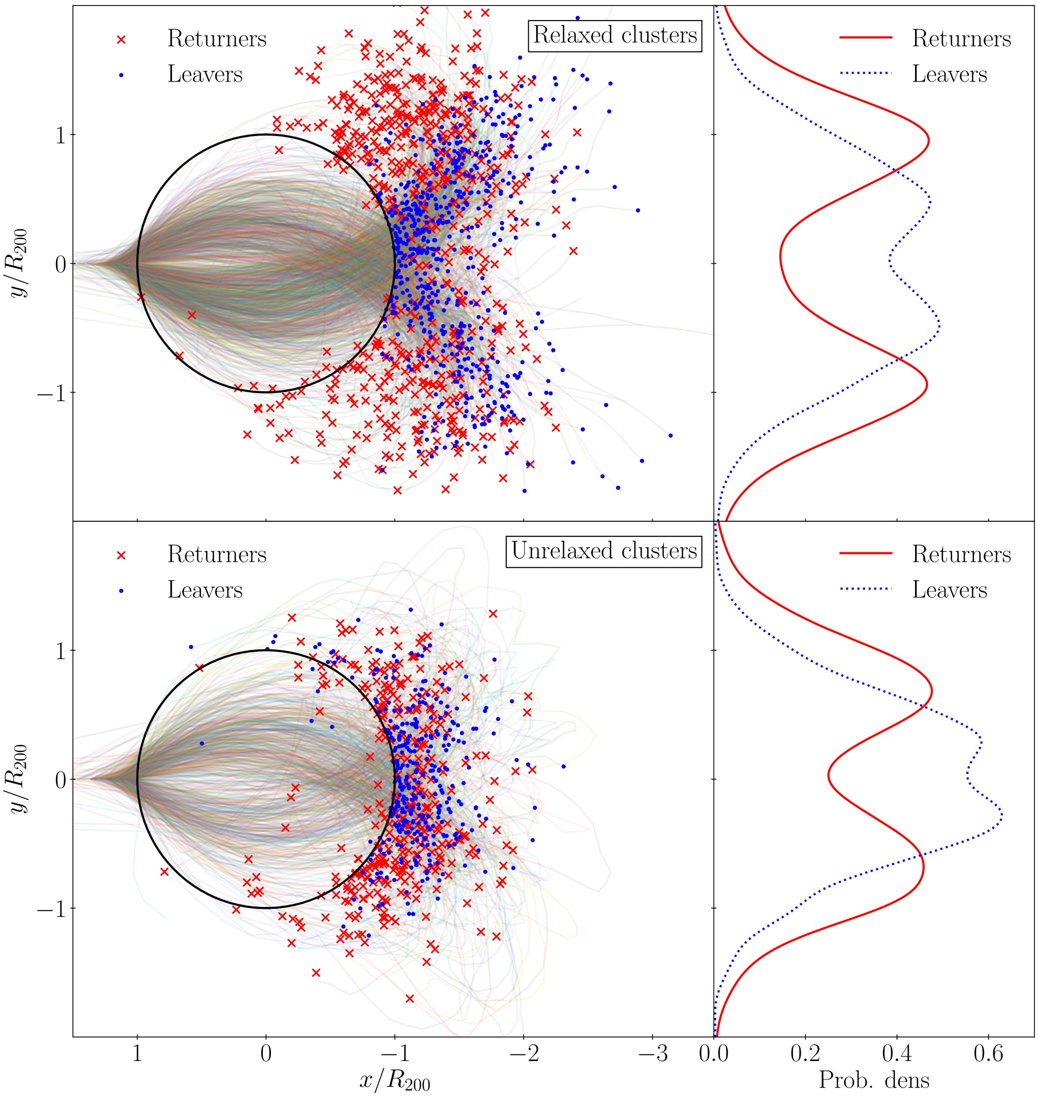

Figure 6.3: Adapted from Fig. 5 in Kuchner et al. (2022). Left panels show paths of backsplash galaxies passing through a cluster (with R_{200} shown by the large black circle), and their final positions at $z = 0$. Paths are rotated so that all galaxies enter the cluster from the same angle. Top row shows data for relaxed galaxy clusters, and bottom panel for unrelaxed clusters. Backsplash galaxies are split based on whether they are approaching the cluster on a second infall ('returners'), or receding from its centre ('leavers'). A KDE showing the distribution of the y-coordinates of the galaxies' positions is shown to the right of each plot. The receding backsplash galaxies are distributed differently to the second infallers, and the distributions are different for relaxed/unrelaxed clusters. See Kuchner et al. (2022) for a more detailed discussion of this figure.

6.3.2 Identifying galaxies processed by filaments

The work in Chapter 5 and the preliminary work in Section 6.3.1 show that we can make statistical estimates of the fraction of backsplash galaxies around clusters, and identify individual backsplash galaxies, based on measurable properties of galaxies, groups and clusters. This will help us understand how much of galaxy evolution can be attributed to environmental processing by clusters, and how much is due to other environments or secular evolution. To build on this further, we also plan to extend this work by applying the same methods to different classes of galaxies and environments, to study how environments other than clusters have impacted galaxies in their past.

Cosmic filaments are an example of one such environment, and in the future, we want to understand if the same methods used in this work can be applied to filaments. Close to half of all galaxies exist in a filament environment (Tempel et al., 2014; Libeskind et al., 2018), but their exact contribution to galaxy evolution is not clear, although it is now understood that filament galaxies are redder (Kraljic et al., 2018; Laigle et al., 2018), more massive (Malavasi et al., 2017) and more elliptical (Kuutma et al., 2017) than field galaxies. However, as well as present-day filament galaxies, it is important to identify galaxies that have previously passed through a filament. Without this, it is difficult to quantify filaments' exact contribution to galaxy evolution. To build on the work in this thesis, we aim to learn whether cluster properties can allow us to predict the fraction of cluster galaxies that have been accreted via a cosmic filament, and whether we can identify (using machine learning or other techniques) which individual galaxies have passed through a filament.

Previous work hints that this is a possibility. Gouin et al. (2021) use simulations to show that dynamically relaxed groups and clusters have a lower connectivity, meaning that they have fewer filaments connecting them to the cosmic web. On the other hand, rapidly-growing, unrelaxed clusters are strongly connected to their environment. This is in agreement with the work of Darragh Ford et al. (2019), who show that clusters that have experienced a recent major merger are more strongly connected to their environment. The connectivity of a cluster is likely to affect how many of its galaxies have been accreted along filaments, and so we hope to constrain the role of filaments in the build-up of the cluster population by using either a cluster's dynamical state, or direct measurements of its connectivity.

Kuchner et al. (2022) also show that filaments can interact with other environments, further complicating their role in the pre-processing of galaxies. For example, they show that the majority ($\sim 90\%$) of galaxy groups around a cluster are found in filaments. This kind of connection could potentially be exploited through our machine learning model described above, allowing us to deduce whether a galaxy has previously been, or is currently, inside of a filament. With this information, we hope to develop a model that can identify the full environmental history of galaxies, including whether they have passed through a cluster, group, or filament.

Finally, as well as looking at the cosmic environments of galaxies, simulations allow us to find exactly which processes occur in these environments (as we investigated briefly in Section 4.4.3). For example, early-type galaxies in groups can form through galaxy mergers, but mergers are less common in clusters, meaning that these early-type galaxies must form through other processes, either before or after entering a cluster. It is not entirely clear how common mergers are in filaments, but by using simulations, we can study the fraction of cluster galaxies whose evolution is in fact due to filaments. Applying these findings to observational data will then tell us which environments are actually responsible for driving galaxy evolution.

6.4 Upcoming complementary observations

In the next several years, numerous new observational surveys will provide incredibly rich, detailed data on galaxy clusters. These surveys will provide observations of galaxies in the centres of clusters, but also in the cluster outskirts, allowing us to probe many of the types of galaxies discussed in this thesis, such as galaxies in infalling groups, backsplash galaxies, and filament galaxies.

One example survey is the WEAVE Wide-Field Cluster Survey (WWFCS, Kuchner et al. in prep.). WEAVE (WHT Enhanced Area Velocity Explorer[2]) is a new spectroscopic instrument for the 4.2m William Herschel Telescope (WHT[3]). It consists of nearly 1000 fibres, that can each be configured to take optical spectroscopy of a different object over a two degree field of view (Dalton et al., 2014; Hughes et al., 2020). The WWFCS will use the WEAVE instrument to study approximately 20 nearby ($0.04 \leq z \leq 0.07$) galaxy clusters, out to distances of $5R_{200}$ from the cluster centres. This will provide spectroscopic measurements for several thousand galaxies in each of these clusters, down to stellar masses of $\sim 10^9 M_\odot$.

The WWFCS will complement the theoretical work presented in this thesis perfectly, as it will provide unprecedented spectroscopic coverage of the outskirts of galaxy clusters. From this data, it will be possible to identify cosmic filaments around these clusters (Kuchner et al., 2020), as well as galaxy groups, which we know to be a

[2] https://www.ing.iac.es//confluence/display/WEAV
[3] https://www.ing.iac.es/astronomy/telescopes/wht/

source of pre-processed galaxies entering the cluster for the first time. Additionally, the spectroscopic data will allow detailed galaxy properties to be measured, such as their stellar masses and line-of-sight velocities. This data will be ideal for feeding into the backsplash classifier we plan to build (Section 6.3.1), but will also help in identifying galaxies with similar velocities that are likely members of the same galaxy group. In addition to this detailed survey of a relatively small number of clusters, larger upcoming surveys like *Euclid* (Euclid Collaboration et al., 2019) will provide excellent cluster statistics, by imaging $> 10^5$ galaxy clusters.

As well as these optical surveys, the near future will bring impressive new X-ray observations. These will play a crucial role in our understanding of galaxy clusters, as X-ray observations of galaxy clusters are very well suited to studying their dynamical states. For instance, Ge et al. (2016) use *XMM-Newton*[4] measurements to quantify the amount of substructure, and the offset of the BCG from the cluster centre, in four clusters. They were also able to measure the cluster radii to a precision of less than 10%. Similarly, Zenteno et al. (2020) use *Chandra*[5] data to quantify the dynamical state parameters of a larger sample of clusters, which they then sort into relaxed and unrelaxed samples.

Future X-ray missions such as *Athena* (Barcons et al., 2017) will build on the currently available X-ray measurements, performing surveys close to two orders of magnitude faster than both *XMM-Newton* and *Chandra* (Nandra et al., 2013). Far deeper observations will consequently be available, allowing us to probe the fainter outer regions of clusters. One of the science goals of *Athena* is to detect and characterise the warm-hot intergalactic medium (WHIM), the gas that traces out cosmic filaments between clusters, allowing us to directly detect filament galaxies, as well as measure properties of clusters such as their connectivity (Nandra et al., 2013). Additionally, *Athena* aims to map the density and velocity of the hot intracluster medium, potentially allowing features such as cluster substructure and the splashback radius to be measured. These will be vital in understanding the dynamical states of clusters which, as we showed in Chapter 5, can provide us deeper insights on the histories of galaxies in and around these clusters.

The near future of extragalactic astronomy and cluster physics is exciting – huge upcoming observational studies, combined with recent jumps forward in simulations, mean that we are in a position to transform our understanding of galaxy evolution and cosmic environment. Upcoming measurements will give us new information about

[4] https://www.esa.int/Science_Exploration/Space_Science/XMM-Newton_overview
[5] https://chandra.harvard.edu

cluster properties, such as their dynamical states. Optical detection of structures that are associated with clusters, such as groups and filaments, will allow us to map out the cosmic environments around clusters. Finally, simulation data can be used to infer how and when the cluster galaxy population was built up, and the environmental histories of these galaxies, based on observed quantities. The work in this thesis will help to combine these approaches, constrain the environmental processes that galaxies have passed through, and finally understand the true impact of cosmic environment on galaxy evolution.

Bibliography

Adhikari S., Dalal N., Chamberlain R. T., 2014, J. Cosmology Astropart. Phys., 2014, 019

Adhikari S., Dalal N., Clampitt J., 2016, J. Cosmology Astropart. Phys., 2016, 022

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

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

An S.-H., Kim J., Moon J.-S., Yoon S.-J., 2019, ApJ, 887, 59

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

Arthur J., et al., 2017, MNRAS, 464, 2027

Arthur J., et al., 2019, MNRAS, 484, 3968

Babcock H. W., 1939, Lick Observatory Bulletin, 498, 41

Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017

Bahé Y. M., et al., 2017, MNRAS, 470, 4186

Bahé Y. M., et al., 2019, MNRAS, 485, 2287

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

Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945

Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, ApJ, 527, 54

Barcons X., et al., 2017, Astronomische Nachrichten, 338, 153

Barnes J. E., 2012, MNRAS, 425, 1104

Barnes D. J., Kay S. T., Henson M. A., McCarthy I. G., Schaye J., Jenkins A., 2017a, MNRAS, 465, 213

Barnes D. J., et al., 2017b, MNRAS, 471, 1088

Barton E., Geller M., Ramella M., Marzke R. O., da Costa L. N., 1996, AJ, 112, 871

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Baxter E., et al., 2017, ApJ, 841, 18

Beck A. M., et al., 2016, MNRAS, 455, 2110

Behroozi P. S., Loeb A., Wechsler R. H., 2013, J. Cosmology Astropart. Phys., 2013, 019

Behroozi P., et al., 2015, MNRAS, 454, 3020

Benavides J. A., Sales L. V., Abadi M. G., 2020, MNRAS, 498, 3852

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

Benson A. J., 2012, New Astron., 17, 175

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

Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, MNRAS, 436, 697

Berrier J. C., Stewart K. R., Bullock J. S., Purcell C. W., Barton E. J., Wechsler R. H., 2009, ApJ, 690, 1292

Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27

Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603

Borgani S., Kravtsov A., 2011, Advanced Science Letters, 4, 204

Borrow J., Vogelsberger M., O'Neil S., McDonald M. A., Smith A., 2022, arXiv e-prints, p. arXiv:2205.10376

van den Bosch F. C., Ogiya G., Hahn O., Burkert A., 2018, MNRAS, 474, 3043

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

Breiman L., 2001, Machine Learning, 45, 5

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

van der Burg R. F. J., Hoekstra H., Muzzin A., Sifón C., Balogh M. L., McGee S. L., 2015, A&A, 577, A19

Cappellari M., 2016, ARA&A, 54, 597

Caputi K. I., Cirasuolo M., Dunlop J. S., McLure R. J., Farrah D., Almaini O., 2011, MNRAS, 413, 162

Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646

Cattaneo A., et al., 2017, MNRAS, 471, 1401

Chang C., et al., 2018, ApJ, 864, 83

Chisari N. E., et al., 2017, MNRAS, 472, 1163

Choque-Challapa N., Smith R., Candlish G., Peletier R., Shin J., 2019, MNRAS, 490, 3654

Chua K. T. E., Pillepich A., Rodriguez-Gomez V., Vogelsberger M., Bird S., Hernquist L., 2017, MNRAS, 472, 4343

Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320

Cohn J. D., 2012, MNRAS, 419, 1017

Colless M., et al., 2001, MNRAS, 328, 1039

Conselice C. J., 2014, ARA&A, 52, 291

Contini E., Kang X., 2015, MNRAS, 453, L53

Contreras-Santos A., et al., 2022, MNRAS, 511, 2897

Cora S. A., et al., 2018, MNRAS, 479, 2

Cramer W. J., Kenney J. D. P., Sun M., Crowl H., Yagi M., Jáchym P., Roediger E., Waldron W., 2019, ApJ, 870, 63

Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, MNRAS, 453, 1513

Croton D. J., et al., 2016, ApJS, 222, 22

Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, MNRAS, 423, 2279

Cui W., Borgani S., Murante G., 2014, MNRAS, 441, 1769

Cui W., et al., 2016, MNRAS, 458, 4052

Cui W., Power C., Borgani S., Knebe A., Lewis G. F., Murante G., Poole G. B., 2017, MNRAS, 464, 2502

Cui W., et al., 2018, MNRAS, 480, 2898

Cui W., et al., 2022, arXiv e-prints, p. arXiv:2202.14038

Cybulski R., Yun M. S., Fazio G. G., Gutermuth R. A., 2014, MNRAS, 439, 3564

van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

Dalton G., et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. p. 91470L (`arXiv:1412.0843`), doi:10.1117/12.2055132

Darragh Ford E., et al., 2019, MNRAS, 489, 5695

Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645

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

De Luca F., De Petris M., Yepes G., Cui W., Knebe A., Rasia E., 2021, MNRAS, 504, 5383

De Lucia G., Weinmann S., Poggianti B. M., Aragón-Salamanca A., Zaritsky D., 2012, MNRAS, 423, 1277

Dehnen W., Read J. I., 2011, European Physical Journal Plus, 126, 55

Dekel A., Devor J., Hetzroni G., 2003, MNRAS, 341, 326

Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859

Diemer B., 2017, ApJS, 231, 5

Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1

Diemer B., More S., Kravtsov A. V., 2013, ApJ, 766, 25

Diemer B., Mansfield P., Kravtsov A. V., More S., 2017, ApJ, 843, 140

Djorgovski S., Davis M., 1987, ApJ, 313, 59

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

Donnari M., et al., 2021, MNRAS, 500, 4004

Dressler A., 1980, ApJ, 236, 351

Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, ApJ, 313, 42

Dressler A., Oemler Augustus J., Poggianti B. M., Gladders M. D., Abramson L., Vulcani B., 2013, ApJ, 770, 62

Duncan K., et al., 2014, MNRAS, 444, 2960

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Eke V. R., Baugh C. M., Cole S., Frenk C. S., Navarro J. F., 2006, MNRAS, 370, 1147

Euclid Collaboration et al., 2019, A&A, 627, A23

Fidler C., Tram T., Rampf C., Crittenden R., Koyama K., Wands D., 2017, J. Cosmology Astropart. Phys., 2017, 043

Fillmore J. A., Goldreich P., 1984, ApJ, 281, 1

Freeman K. C., 1970, ApJ, 160, 811

Frenk C. S., White S. D. M., 2012, Annalen der Physik, 524, 507

Frenk C. S., White S. D. M., Davis M., 1983, ApJ, 271, 417

Frenk C. S., Evrard A. E., White S. D. M., Summers F. J., 1996, ApJ, 472, 460

Ganeshaiah Veena P., Cautun M., Tempel E., van de Weygaert R., Frenk C. S., 2019, MNRAS, 487, 1607

Gao L., Frenk C. S., Boylan-Kolchin M., Jenkins A., Springel V., White S. D. M., 2011, MNRAS, 410, 2309

Garrison-Kimmel S., et al., 2017, MNRAS, 471, 1709

Ge C., Wang Q. D., Tripp T. M., Li Z., Gu Q., Ji L., 2016, MNRAS, 459, 366

Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, MNRAS, 300, 146

Gill S. P. D., Knebe A., Gibson B. K., 2004a, MNRAS, 351, 399

Gill S. P. D., Knebe A., Gibson B. K., Dopita M. A., 2004b, MNRAS, 351, 410

Gill S. P. D., Knebe A., Gibson B. K., 2005, MNRAS, 356, 1327

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

Gonçalves T. S., Martin D. C., Menéndez-Delmestre K., Wyder T. K., Koekemoer A., 2012, ApJ, 759, 67

Gonzalez-Casado G., Mamon G. A., Salvador-Sole E., 1994, ApJ, 433, L61

Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, MNRAS, 439, 264

Gott J. R. I., Turner E. L., Aarseth S. J., 1979, ApJ, 234, 13

Gouin C., Bonnaire T., Aghanim N., 2021, A&A, 651, A56

Guo H., et al., 2016, MNRAS, 459, 3040

Haggar R., Gray M. E., Pearce F. R., Knebe A., Cui W., Mostoghiu R., Yepes G., 2020, MNRAS, 492, 6074

Haggar R., Pearce F. R., Gray M. E., Knebe A., Yepes G., 2021, MNRAS, 502, 1191

Haines C. P., et al., 2015, ApJ, 806, 101

Han S., Smith R., Choi H., Cortese L., Catinella B., Contini E., Yi S. K., 2018, ApJ, 866, 78

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373

Hester J. A., 2006, ApJ, 647, 910

Heyl J. S., Cole S., Frenk C. S., Navarro J. F., 1995, MNRAS, 274, 755

Hickson P., 1982, ApJ, 255, 382

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

Hubble E. P., 1926, ApJ, 64, 321

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hughes S., et al., 2020, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 114477R (arXiv:2012.07734), doi:10.1117/12.2560331

Jaffé Y. L., Smith R., Candlish G. N., Poggianti B. M., Sheen Y.-K., Verheijen M. A. W., 2015, MNRAS, 448, 1715

Jaffé Y. L., et al., 2016, MNRAS, 461, 1202

Jauzac M., et al., 2016, MNRAS, 463, 3876

Jenkins A., et al., 1998, ApJ, 499, 20

Jia J.-Y., Gao L., Qu Y., 2020, Research in Astronomy and Astrophysics, 20, 174

Jian H.-Y., Lin L., Chiueh T., 2012, ApJ, 754, 26

Kim J.-h., et al., 2014, ApJS, 210, 14

Kleiner D., et al., 2021, A&A, 648, A32

Klypin A., Gottlöber S., Kravtsov A. V., Khokhlov A. M., 1999, ApJ, 516, 530

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Knebe A., Libeskind N. I., Knollmann S. R., Martinez-Vaquero L. A., Yepes G., Gottlöber S., Hoffman Y., 2011a, MNRAS, 412, 529

Knebe A., et al., 2011b, MNRAS, 415, 2293

Knebe A., Libeskind N. I., Doumler T., Yepes G., Gottlöber S., Hoffman Y., 2011c, MNRAS, 417, L56

Knebe A., et al., 2013, MNRAS, 435, 1618

Knebe A., et al., 2018, MNRAS, 474, 5206

Knebe A., et al., 2020, MNRAS, 495, 3002

Knollmann S. R., Knebe A., 2009, ApJS, 182, 608

Kraljic K., et al., 2018, MNRAS, 474, 547

Kuchner U., et al., 2020, MNRAS, 494, 5473

Kuchner U., et al., 2022, MNRAS, 510, 581

Kuutma T., Tamm A., Tempel E., 2017, A&A, 600, L6

Laigle C., et al., 2018, MNRAS, 474, 5437

Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692

Lee B., et al., 2013, ApJ, 774, 47

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Libeskind N. I., Yepes G., Knebe A., Gottlöber S., Hoffman Y., Knollmann S. R., 2010, MNRAS, 401, 1889

Libeskind N. I., et al., 2018, MNRAS, 473, 1195

Liller M. H., 1966, ApJ, 146, 28

Lin Y.-T., Mohr J. J., 2004, ApJ, 617, 879

Lin W. P., Jing Y. P., Mao S., Gao L., McCarthy I. G., 2006, ApJ, 651, 636

Lintott C. J., et al., 2008, MNRAS, 389, 1179

Lintott C., et al., 2011, MNRAS, 410, 166

Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2008, MNRAS, 391, 1137

Lotz M., Remus R.-S., Dolag K., Biviano A., Burkert A., 2019, MNRAS, 488, 5370

Lovell M. R., Hellwing W., Ludlow A., Zavala J., Robertson A., Fattahi A., Frenk C. S., Hardwick J., 2020, MNRAS, 498, 702

Lu Y., et al., 2014, ApJ, 795, 123

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

Lynds R., Petrosian V., 1989, ApJ, 336, 1

Maier C., et al., 2016, A&A, 590, A108

Maier C., Ziegler B. L., Haines C. P., Smith G. P., 2019, A&A, 621, A131

Malavasi N., et al., 2017, MNRAS, 465, 3817

Mamon G. A., Sanchis T., Salvador-Solé E., Solanes J. M., 2004, A&A, 414, 445

Marchesini D., et al., 2010, ApJ, 725, 1277

Mauduit J. C., Mamon G. A., 2007, A&A, 475, 169

McClintock T., et al., 2019, MNRAS, 482, 1352

McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937

McNab K., et al., 2021, MNRAS, 508, 157

Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996a, Nature, 379, 613

Moore B., Katz N., Lake G., 1996b, ApJ, 457, 455

Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, ApJ, 499, L5

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

More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36

More S., et al., 2016, ApJ, 825, 39

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

Mostoghiu R., Knebe A., Cui W., Pearce F. R., Yepes G., Power C., Dave R., Arth A., 2019, MNRAS, 483, 3390

Mostoghiu R., et al., 2021a, MNRAS, 501, 5029

Mostoghiu R., et al., 2021b, A&A, 652, A10

Moustakas J., et al., 2013, ApJ, 767, 50

Muldrew S. I., Pearce F. R., Power C., 2011, MNRAS, 410, 2617

Murante G., Monaco P., Giovalli M., Borgani S., Diaferio A., 2010, MNRAS, 405, 1491

Muzzin A., et al., 2013, ApJ, 777, 18

Nandra K., et al., 2013, arXiv e-prints, p. arXiv:1306.2307

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., et al., 2004, MNRAS, 349, 1039

Nelson D., et al., 2019, MNRAS, 490, 3234

Onions J., et al., 2012, MNRAS, 423, 1200

Pallero D., Gómez F. A., Padilla N. D., Torres-Flores S., Demarco R., Cerulo P., Olave-Rojas D., 2019, MNRAS, 488, 847

Patel S. G., Kelson D. D., Holden B. P., Franx M., Illingworth G. D., 2011, ApJ, 735, 53

Paul S., John R. S., Gupta P., Kumar H., 2017, MNRAS, 471, 2

Peebles P. J. E., Yu J. T., 1970, ApJ, 162, 815

Peng Y.-j., et al., 2010, ApJ, 721, 193

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

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

Pimbblet K. A., 2011, MNRAS, 411, 2637

Piotrowska J. M., Bluck A. F. L., Maiolino R., Peng Y., 2022, MNRAS, 512, 1052

Planck Collaboration et al., 2014, A&A, 571, A1

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2020, A&A, 641, A1

Plummer H. C., 1911, MNRAS, 71, 460

Porter L. A., Somerville R. S., Primack J. R., Johansson P. H., 2014, MNRAS, 444, 942

Potter D., Stadel J., Teyssier R., 2017, Computational Astrophysics and Cosmology, 4, 2

Press W. H., Schechter P., 1974, ApJ, 187, 425

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

Pujol A., et al., 2017, MNRAS, 469, 749

Quadri R. F., Williams R. J., Franx M., Hildebrandt H., 2012, ApJ, 744, 88

Rasia E., Tormen G., Moscardini L., 2004, MNRAS, 351, 237

Rasia E., et al., 2015, ApJ, 813, L17

Read J. I., Wilkinson M. I., Evans N. W., Gilmore G., Kleyna J. T., 2006, MNRAS, 366, 429

Richings J., et al., 2020, MNRAS, 492, 5780

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

Roberts I. D., Parker L. C., 2017, MNRAS, 467, 3268

Rosswog S., 2009, New Astron. Rev., 53, 78

Rost A., et al., 2021, MNRAS, 502, 714

Rubin V. C., Ford W. Kent J., 1970, ApJ, 159, 379

Sahni V., Coles P., 1995, Phys. Rep., 262, 1

Sazonova E., et al., 2021, ApJ, 919, 134

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

Schawinski K., et al., 2014, MNRAS, 440, 889

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

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

Sembolini F., Yepes G., De Petris M., Gottlöber S., Lamagna L., Comis B., 2013, MNRAS, 429, 323

Shin T., et al., 2019, MNRAS, 487, 2900

Smith R., Choi H., Lee J., Rhee J., Sanchez-Janssen R., Yi S. K., 2016, ApJ, 833, 109

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

Somerville R. S., Davé R., 2015, ARA&A, 53, 51

Spergel D. N., et al., 2003, ApJS, 148, 175

Springel V., 2005, MNRAS, 364, 1105

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

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

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

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

Srisawat C., et al., 2013, MNRAS, 436, 150

Steinhauser D., Schindler S., Springel V., 2016, A&A, 591, A51

Stücker J., Angulo R. E., Hahn O., White S. D. M., 2022, MNRAS, 509, 1703

Sun M., et al., 2021, Nature Astronomy, 6, 270

Tam S.-I., et al., 2020, MNRAS, 496, 4032

Taylor J. E., Babul A., 2004, MNRAS, 348, 811

Tegmark M., 1996, in Bonometto S., Primack J. R., Provenzale A., eds, Dark Matter in the Universe. p. 379 (arXiv:astro-ph/9511148)

Tempel E., Stoica R. S., Martínez V. J., Liivamägi L. J., Castellan G., Saar E., 2014, MNRAS, 438, 3465

Tempel E., Kipper R., Tamm A., Gramann M., Einasto M., Sepp T., Tuvikene T., 2016, A&A, 588, A14

Tenneti A., Mandelbaum R., Di Matteo T., Feng Y., Khandai N., 2014, MNRAS, 441, 470

Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775

Tissera P. B., White S. D. M., Pedrosa S., Scannapieco C., 2010, MNRAS, 406, 922

Tollet É., Cattaneo A., Mamon G. A., Moutard T., van den Bosch F. C., 2017, MNRAS, 471, 4170

Tomczak A. R., et al., 2014, ApJ, 783, 85

Tornatore L., Borgani S., Dolag K., Matteucci F., 2007, MNRAS, 382, 1050

Tully R. B., 2015, AJ, 149, 54

Tyson J. A., Valdes F., Wenk R. A., 1990, ApJ, 349, L1

Urry C. M., Padovani P., 1995, PASP, 107, 803

Vijayaraghavan R., Ricker P. M., 2013, MNRAS, 435, 2713

Vijayaraghavan R., Gallagher J. S., Ricker P. M., 2015, MNRAS, 447, 3623

Vogelsberger M., et al., 2014, Nature, 509, 177

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42

Volders L. M. J. S., 1959, Bulletin of the Astronomical Institutes of the Netherlands, 14, 323

Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83

Wang Y., et al., 2018, ApJ, 868, 130

Watson W. A., Iliev I. T., Diego J. M., Gottlöber S., Knebe A., Martínez-González E., Yepes G., 2014, MNRAS, 437, 3776

Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435

van der Wel A., et al., 2007, ApJ, 670, 206

Wen Z. L., Han J. L., 2013, MNRAS, 436, 275

Wetzell V., et al., 2021, arXiv e-prints, p. arXiv:2107.07631

White S. D. M., 1976, MNRAS, 177, 717

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

White M., Cohn J. D., Smit R., 2010, MNRAS, 408, 1818

Wild V., et al., 2020, MNRAS, 494, 529

Wojtak R., Łokas E. L., Mamon G. A., Gottlöber S., Klypin A., Hoffman Y., 2008, MNRAS, 388, 815

Wu H.-Y., Hahn O., Wechsler R. H., Behroozi P. S., Mao Y.-Y., 2013, ApJ, 767, 23

York D. G., et al., 2000, AJ, 120, 1579

Yuan Z. S., Han J. L., 2020, MNRAS, 497, 5485

Zabel N., et al., 2019, MNRAS, 483, 2251

Zabludoff A. I., Zaritsky D., Lin H., Tucker D., Hashimoto Y., Shectman S. A., Oemler A., Kirshner R. P., 1996, ApJ, 466, 104

Zeldovich Y. B., 1972, MNRAS, 160, 1P

Zenteno A., et al., 2020, MNRAS, 495, 705

Zhu Q., Marinacci F., Maji M., Li Y., Springel V., Hernquist L., 2016, MNRAS, 458, 1559

Zinger E., Dekel A., Kravtsov A. V., Nagai D., 2018, MNRAS, 475, 3654

Zürcher D., More S., 2019, ApJ, 874, 184

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

Zwicky F., 1937, ApJ, 86, 217

Appendices

Appendix A

Dependence of number density profile on group mass

In Section 3.3.1, we briefly discuss the effect of group host halo mass on the radial number density of galaxies within the group, for groups in the outskirts of a cluster (as shown in Fig. 3.3).

In both the hydrodynamical and dark matter-only simulations, the range of group halo masses is approximately two orders of magnitude. To investigate the effect of group mass on the radial density profiles of the groups, we split each sample of groups into three mass bins, each containing approximately equal numbers of groups. We then compare the radial number density profiles of groups in each bin, for the hydrodynamical and dark matter-only simulations.

This is shown in Fig. A.1. We find that the difference in radial number density between the high-mass and low-mass bins is less than a factor of 2.5 at all radii in the dark matter-only simulations. In the hydrodynamical simulations, the maximum difference between the two mass bins is a factor of four, but only in the very centres of the groups. This is a significant difference, but is less than the difference between the hydrodynamical and dark matter-only simulations, which reaches a maximum of a factor of 10. It is therefore unlikely to affect the overall conclusions of this work.

Most importantly, when comparing mass bins between the simulations, the same trend is seen as in Fig. 3.3, for each of the three mass bins. As Fig. A.2 shows, when considering either low, medium or high-mass groups in the cluster outskirts, the number density of galaxies, ρ, is consistently greater in the hydrodynamical simulations. The difference is slightly larger in the smallest groups, but even in groups with high masses, the number density of galaxies in the centre of hydrodynamical groups is seven times greater than in dark matter-only groups. This is shown by the fractional residual at the bottom of each plot – these are calculated in the same way as is described in Section 3.3.1. The same trend is seen when comparing like-for-like mass bins (i.e. when comparing groups with halo masses in the range $[10^{13.0}, 10^{13.5}]$ $h^{-1}M_{\odot}$ in both the hydrodynamical and dark matter-only simulations).

This shows that the mass of a group does impact the number density of galaxies within the group. However, it also demonstrates that this effect is substantially smaller than the difference caused by the inclusion of baryonic material within the simulations.

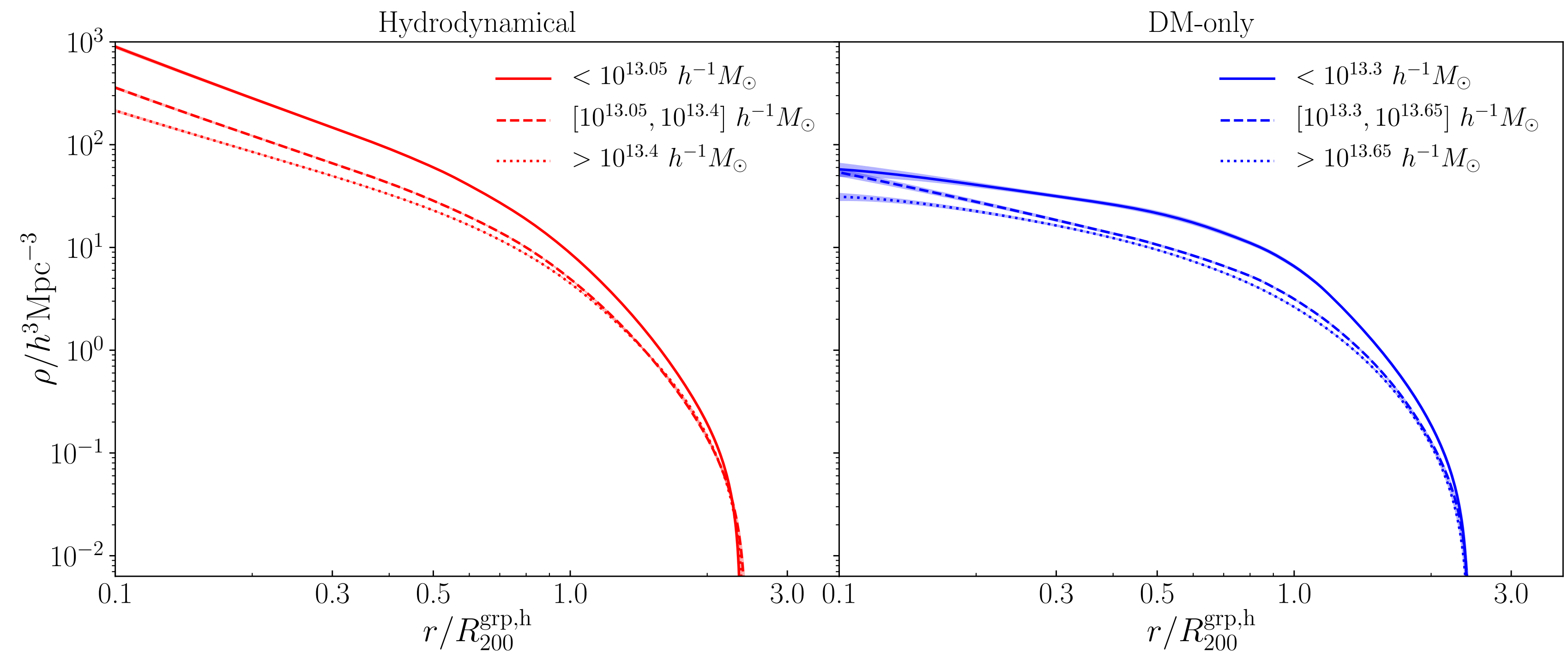

Figure A.1: Radial number density of galaxies in groups (as in Fig. 3.3), for hydrodynamical and dark matter-only simulations, split by group host halo mass. For each class of cluster simulations, we split the groups into three mass bins, with approximately equal numbers of groups in each bin; these mass bins are shown in the legend. Shaded regions represent the uncertainty in the average density profile. For clarity, the spread of the data for each sample of groups is not shown.

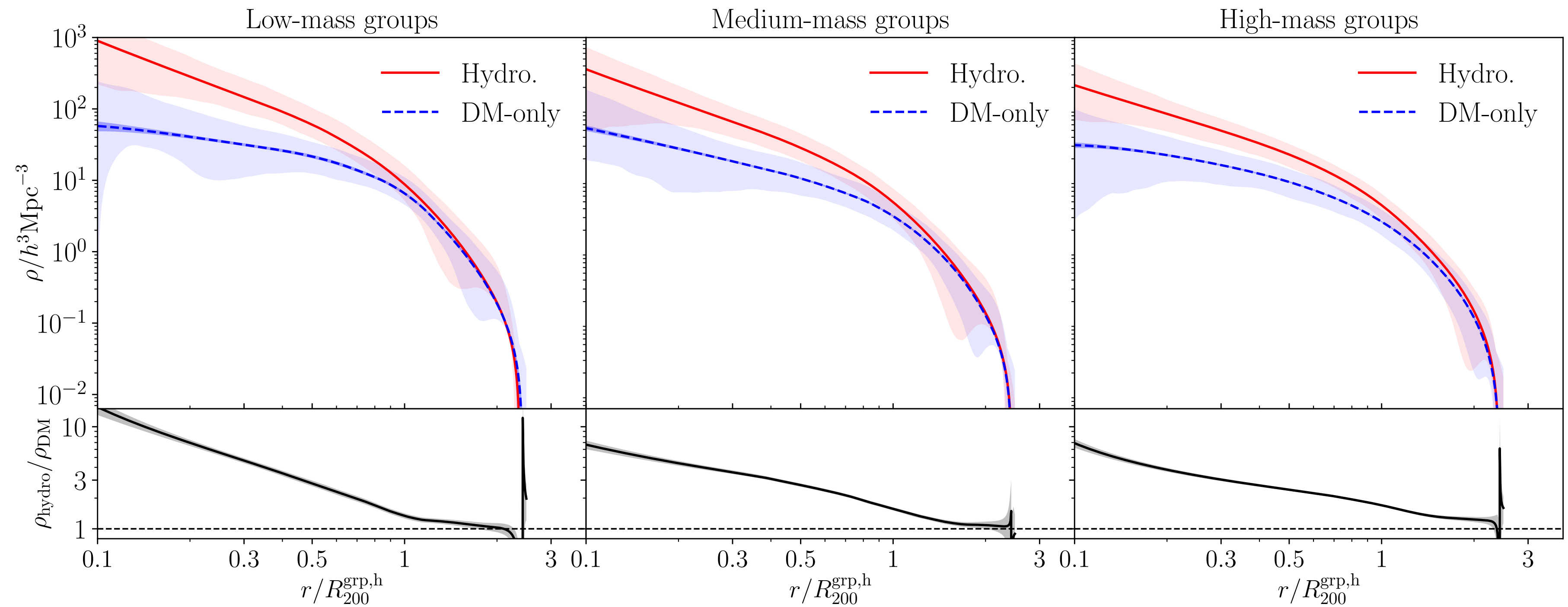

Figure A.2: Radial number density of galaxies in groups located in cluster outskirts at $z = 0$ (top panels). These plots show the same data as Fig. 3.3, but split into three mass bins, each containing approximately one third of the groups in the radial range $[R_{200}^{\mathrm{clus,h}}, 5R_{200}^{\mathrm{clus,h}}]$ around a cluster. For hydrodynamical simulations, the low, medium and high-mass groups are those with a group host halo mass, $M_{200}^{\mathrm{grp,h}}$, in the ranges ($< 10^{13.05}\ h^{-1}M_\odot$), ($[10^{13.05}, 10^{13.4}]\ h^{-1}M_\odot$) and ($> 10^{13.4}\ h^{-1}M_\odot$), respectively, as shown in Fig. A.1. For dark matter-only simulations, the mass bins are ($< 10^{13.3}\ h^{-1}M_\odot$), ($[10^{13.3}, 10^{13.65}]\ h^{-1}M_\odot$) and ($> 10^{13.65}\ h^{-1}M_\odot$). Shaded regions represent 1σ uncertainty in the mean radial number density profile, although these are mostly too small to be seen. Bottom panel shows fractional residuals (the ratio of the hydrodynamical and dark matter-only profiles, as defined in Section 3.3.1).

Appendix B

Evolution of a single galaxy group in phase space

Fig. B.1 shows an example of one galaxy group entering, passing through, and then re-entering a cluster, as a demonstration of the process discussed throughout Chapter 4. The right column shows the dark matter halo of the cluster, represented by the grey circle, and the paths of the galaxies in a group as the group passes through the cluster. The system is rotated so that the path of the group is in the plane of the page.

In the left column, the changing position of each group member in phase space is shown (i.e. its changing position and speed relative to the host group). Each line in phase space represents the path taken by one galaxy through phase space, as the galaxy has followed the path through the cluster shown in the right column. Six timesteps are shown altogether, from top to bottom. The two phases of dynamical evolution shown in Fig. 4.3 and Fig. 4.4 can be seen in the evolving phase space of this group, as the galaxies move upwards in phase space as the group approaches its pericentric passage, and then from left to right after this.

For clarity, schematics are shown in the bottom-right of each panel, similarly to those used in Fig. 4.4, to show where the group is along its orbit through the cluster. From top to bottom, the six timesteps show the group immediately after its first infall, one snapshot after pericentric passage, shortly before exiting the cluster, shortly after exiting the cluster, at apocentre, and immediately after its second infall.

The right panels show the overall behaviour of groups that we find throughout this work – a relatively compact group of galaxies remains coherent for a short period after entering a cluster, but then becomes heavily disrupted, and the galaxies are separated from each other to large distances (shown in the final panel).

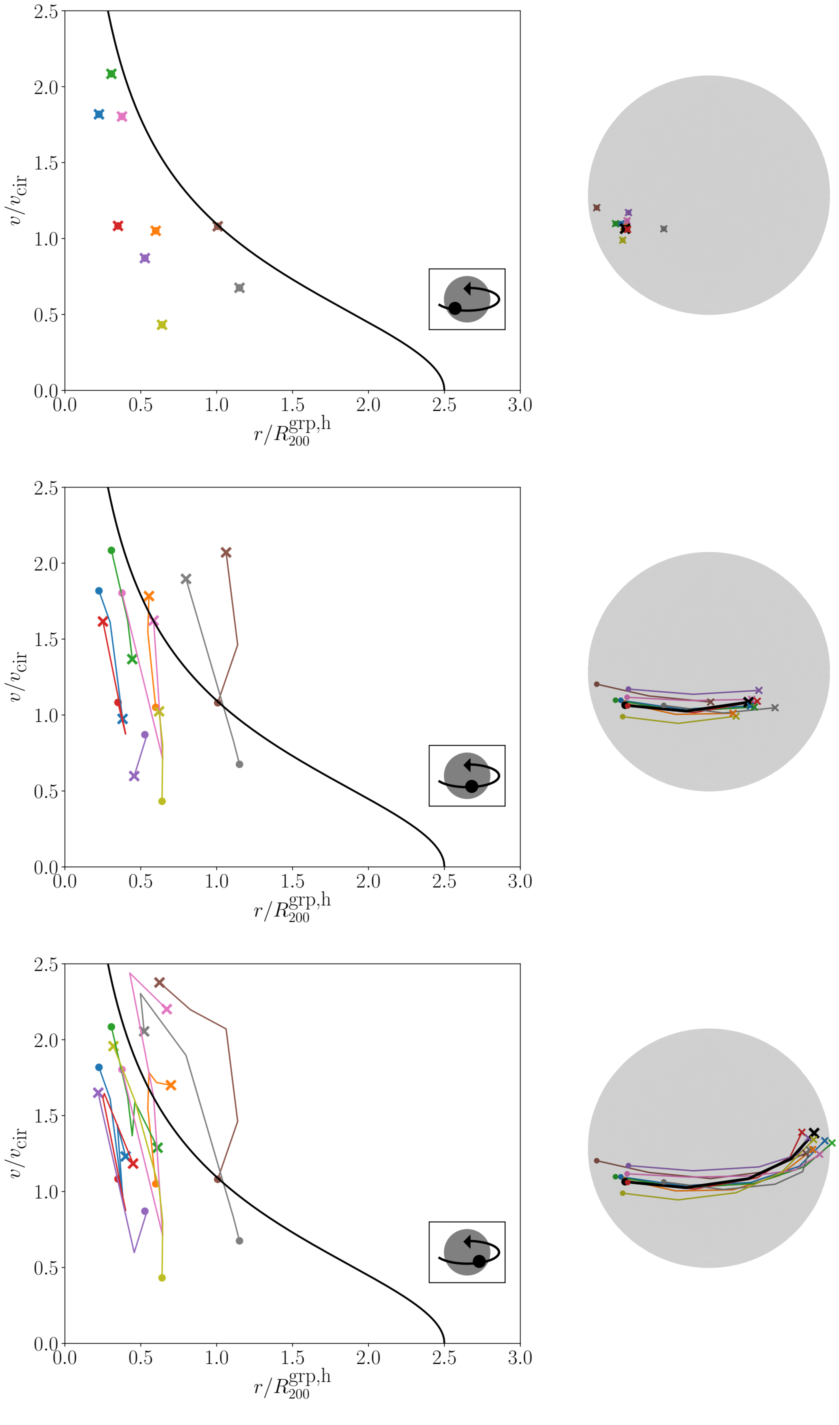

Figure B.1: Phase space evolution of one group as it enters a cluster, passes through, leaves the cluster and then re-enters, showing the initial increase in v, and subsequent increase in r, of the member galaxies. Right panels show the cluster halo (large grey circle), and a 2D projection of the paths taken through it by the group halo (thick black line) and the galaxies bound to this group at infall (thin coloured lines). Left panels show the corresponding paths taken by these galaxies in the phase space diagram used in Section 4.3, from infall to the current snapshot. For clarity, the positions of galaxies at infall and at the present time are represented by dots and crosses respectively *(caption and figure continued on next page)*.

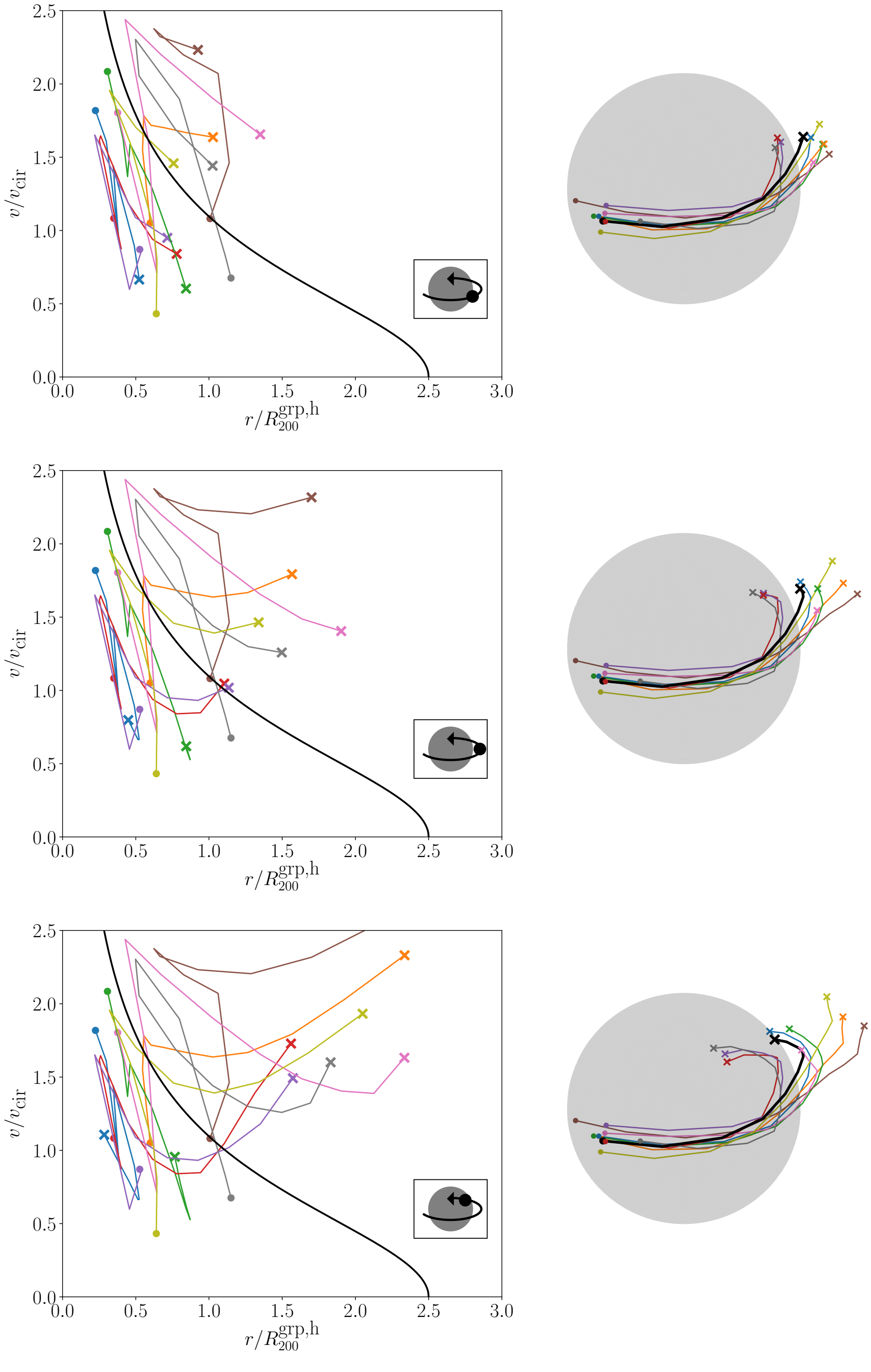

Figure B.1: *(continued from previous page)* Six snapshots are shown altogether, with time increasing from top to bottom. Schematics in the bottom-right of the left panels show where the group is on its cluster orbit – for instance, the fifth timestep shows the group at apocentre. The first and last panels are separated by 10 snapshots, covering ~ 3.2 Gyr between $z = 0.25$ and $z = 0$.

Appendix C

The fates of group galaxies as a function of time

In Section 4.4.3, we showed how the state of galaxies changes over the course of a passage through a cluster, for galaxies that were group members at the moment of infall. Specifically, we showed the fraction of galaxies that remained bound to the group, became unbound, were 'disrupted', or merged, from the first infall of the group to pericentre, apocentre, and second infall. Furthermore, we showed this for all galaxies in these groups (Fig. 4.5), slow-moving galaxies in the group centres (Fig. 4.6), and galaxies in the group outskirts (Fig. 4.7).

Much of the group evolution occurs between cluster infall and the first pericentric passage. As we also showed in Section 4.4.3, the time between these two epochs is also short, compared to the whole journey of the group through a cluster. This means that the changes in the galaxies' states are actually even more sudden than they appear in Fig. 4.5, Fig. 4.6 and Fig. 4.7.

To demonstrate this, we reproduce those three figures here. Instead of spacing the points in the groups' orbits equally on the horizontal axis, we instead plot the average time after first infall at which the pericentre, apocentre, and second infall occur. This demonstrates how quickly the groups reach pericentre and lose the majority of their member galaxies, compared to the relatively long time they spend receding from the group, existing in the group outskirts, and undergoing a second infall. A more detailed description of this style of plot can be found in Section 4.4.3.

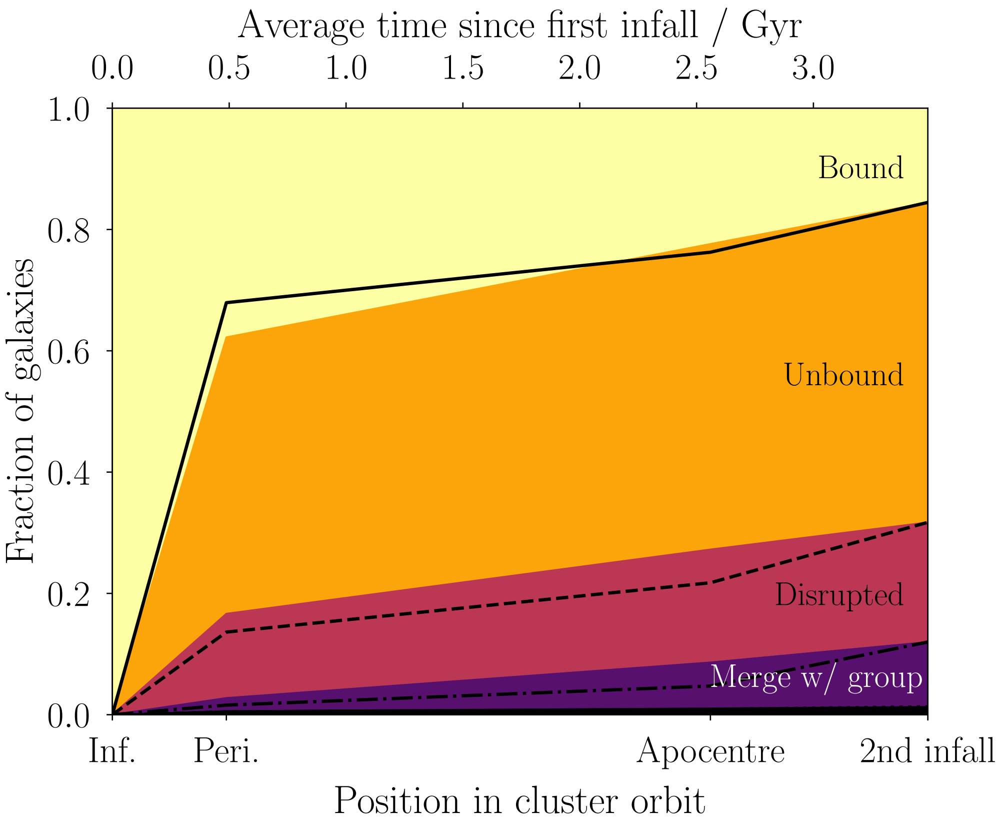

Figure C.1: Same as Fig. 4.5, showing status of galaxies as their host group enters and passes through a cluster, for groups that leave the cluster and experience a second infall. Differently to in Fig. 4.5, the horizontal axis is scaled to show the time elapsed between infall, pericentric passage, apocentric passage, and second infall.

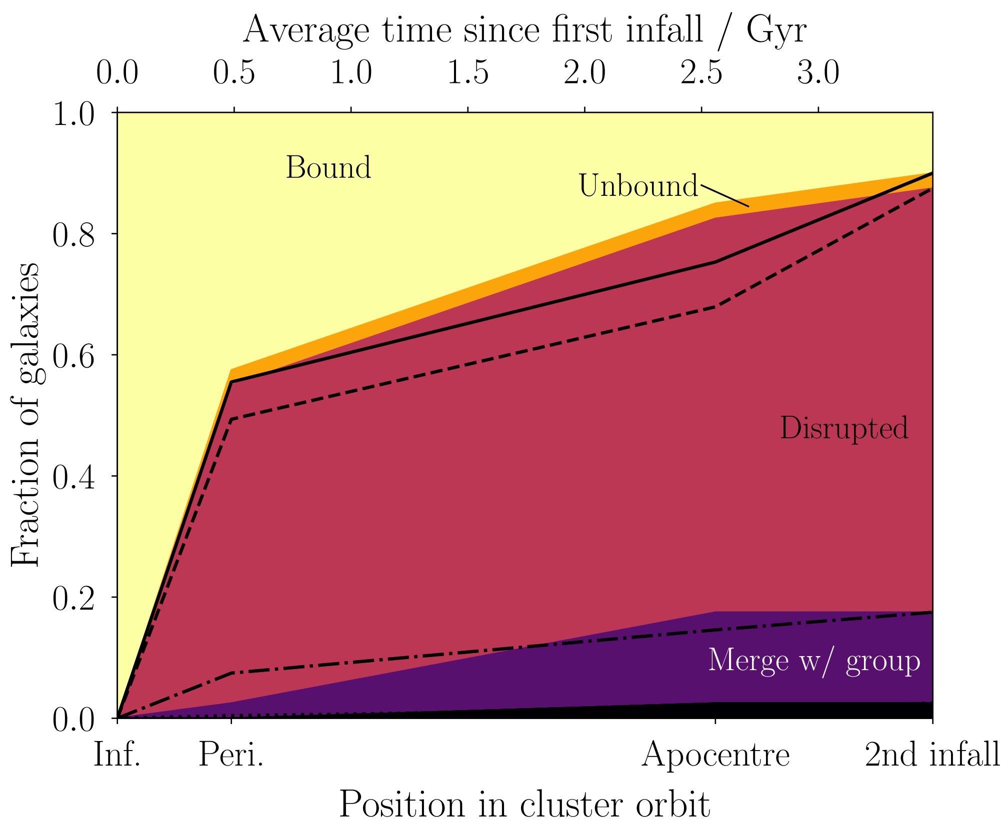

Figure C.2: Same as Fig. 4.6, showing status of galaxies as their host group passes through a cluster. Here, only slow-moving galaxies in the group centres are included ($r < 0.5R_{200}^{\mathrm{grp,h}}$, $v < 0.5v_{\mathrm{cir}}$). As in Fig. C.1, the horizontal axis is scaled to show the time elapsed between infall, pericentric passage, apocentric passage, and second infall.

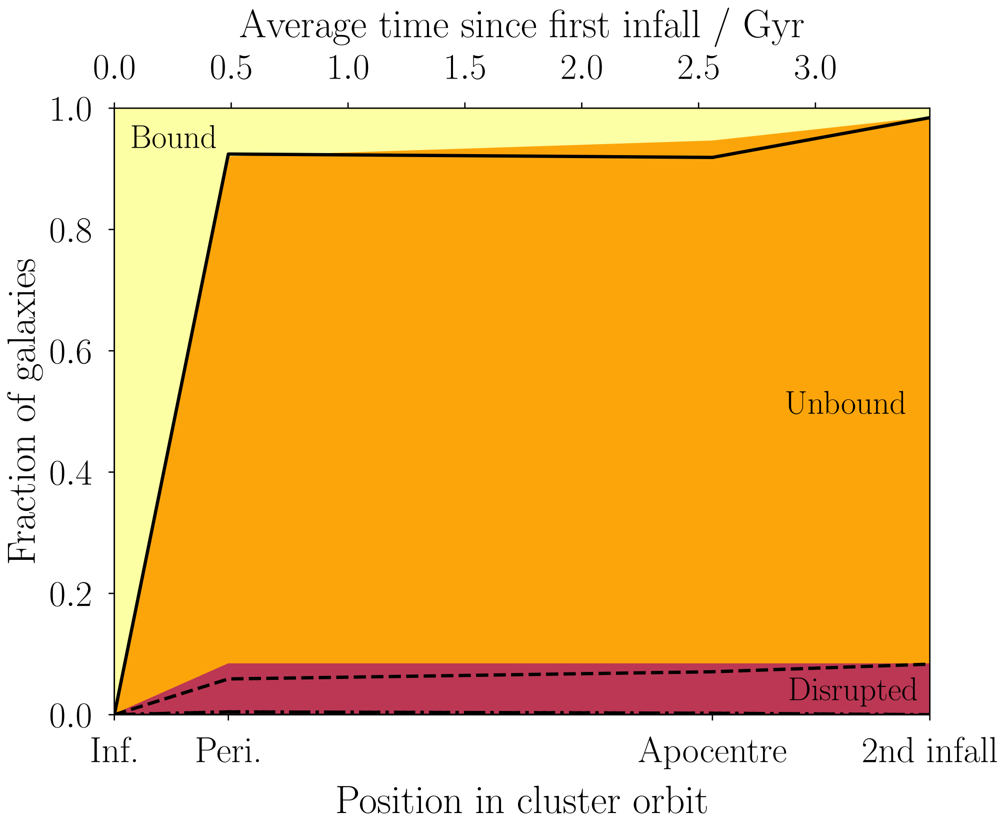

Figure C.3: Same as Fig. 4.7, showing status of galaxies as their host group passes through a cluster. Here, only galaxies in the outskirts of groups are included ($r > 0.8R_{200}^{\mathrm{grp,h}}$). As in Fig. C.1, the horizontal axis is scaled to show the time elapsed between infall, pericentric passage, apocentric passage, and second infall.