Probing the Environmental Dependence of Star Formation in Satellite Galaxies using Orbital Kinematics

by

Kyle Oman

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Physical processes regulating star formation in satellite galaxies represent an area of ongoing research, but the projected nature of observed coordinates makes separating different populations of satellites (with different processes at work) difficult. The orbital history of a satellite galaxy leads to its present-day phase space coordiantes; we can also work backwards and use these coordinates to statistically infer information about the orbital history. The estimated orbital histories of a large population of satellites can then be compared to their star formation histories. This is expected to reveal both a trigger time when the star forming properties of satellites begin to be affected by a high density environment, and the timescale on which environmental quenching operates. Finally, this can be related back to the physical process(es) regulating star formation in high density environments.

We use merger trees from the MultiDark Run 1 N-body simulation to compile a catalogue of the orbits of satellites in cluster environments. We parameterize the orbital history by the time since crossing within 2.5 $r_{\rm vir}$ of the cluster centre and use our catalogue to estimate the probability density over a range of this parameter given a set of present-day projected phase space coordinates. We show that different populations of satellite haloes occupy distinct regions of phase space, and semi-distinct regions of projected phase space. We generalize this result by producing a probability distribution function (PDF) of possible infall times at every point in projected phase space. Although our method focusses on time since infall, it is easily generalized to other orbital parameters.

We apply our method to determining the infall time PDFs of a large sample of observed cluster satellite candidates from the Sloan Digital Sky Survey for which spectroscopic redshifts and g- and r-band bulge-disk decompositions are available. We use galaxy colour as a proxy for star formation history (SFH) and model the distribution of satellite galaxy colours as two gaussian populations, i.e. a red peak and a blue peak. We derive a Markov Chain Monte-Carlo (MCMC) method to obtain the colour distribution as a function of the time since infall into the cluster environment. Although our analysis focusses on this single measure of SFH, it can be generalized to other parameterizations as well. Our implementation of this method is still being tested and tuned, but we use a second simpler and faster (but much cruder) method to obtain a preliminary estimate of the evolution of the colour distribution. Our results are suggestive of a quenching process that begins within perhaps ± 1 Gyr of virial radius crossing and which slows after pericentric passage. Although the regions of projected phase space occupied by satellites just before and just after pericentric passage are partially degenerate, the colour evolution we observe – redder disks in the post-pericentre population – is not symmetric about the pericentre crossing. We stress that results obtained with this second method come with important caveats.

Acknowledgements

I would like to thank my M.Sc. supervisor, Prof. Mike Hudson, for innumerable comments, suggestions and discussions throughout the course of my studies. Dr. Peter Behroozi who provided some of the codes, much of the data, help troubleshooting both and valuable feedback on the manuscript of my first paper. The members of my M.Sc. and defence committees (Profs. Michael Balogh, James Taylor, Niayesh Afshordi) for their feedback on the project and this manuscript. The authors whose publicly available codes and data were used in this work. The developers and maintainers of the NASA ADS for effectively taking care of all the details of my reference list in my place.

Dedication

This thesis is dedicated to my family, very soon to include one more member.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

It has long been known that galaxy properties, notably morphology, colour and star formation rate, are dependent on environment – whether a galaxy is isolated in space or part of a larger structure. We aim to gain some insight into the dominant process(es) that drive change when a galaxy moves from a low density to a high density environment. We use a combination of numerical simulation and observational data, presented in Chapter 2, to tackle this problem. Simulation output is used to learn about the statistical properties of cluster satellite orbits and their relationship to their position in projected phase space in Chapter 3. We then apply this knowledge of orbit statistics to a sample of observed galaxies and relate their orbits to their colours in Chapter 4. Galaxy colour is correlated with the star formation history (SFH) of the galaxy, which can in turn be related to processes that quench (or trigger) star formation. We summarise our findings and discuss them in the broader context of the litterature in Chapter 5. We also suggest possible extensions of this work.

First, we provide some context on the formation of large-scale structure in the Universe (§1.1) and some of the fundamentals of cosmological numerical simulations (§1.2). We give an overview of some well-studied mechanisms thought to affect star formation in high density environments in §1.3. In §1.4 we provide some background on the properties of orbits in clusters and how they are expressed in terms of the intrinsic properties of the host system. We also identify which properties are directly observable. We outline the relationship between the SFH and colour of a galaxy in §1.5. Finally, we give our assumptions regarding cosmological parameters in §1.6.

A substantial portion of the results presented in this thesis were published in Oman et al. (2013) – in particular the majority of Chapter 3 (excepting sections concerning "interlop-

ers"). The results in Chapter 4 are new; we hope to include them in a future publication.

1.1 Gravitational Collapse and Cluster Formation

Galaxy clusters are the largest virialized structures in the Universe, and they are valuable cosmological laboratories, allowing the study of not only structures and processes intrinsic to the cluster, but also the measurement of several more fundamental parameters of cosmology (e.g. the universal baryon fraction, σ_8). Cluster formation is an inherently nonlinear process and is therefore difficult to model in its entirety. Nevertheless, there is a rich theory describing the portions of the process which can be linearized, and some of the nonlinear processes under simplifying assumptions or symmetries as well. Background for the remainder of this section is drawn from Binney & Tremaine (2008) except where indicated otherwise. Rather than give an overview of the theory itself, we focus instead on some of the most interesting consequences of the theory and those that are most relevant to the present thesis. We restrict ourselves to the standard Λ-Cold Dark Matter (ΛCDM) theory, ignoring more exotic models which, while potentially correct, have yet to produce a front-runner with an overall match to observations superior to ΛCDM.

At early times, the Universe had a homogeneous and isotropic distribution of matter on large scales with random density fluctuations described by a characteristic power spectrum. The background Friedmann-Lemaître-Robertson-Walker (FLRW) metric is expanding. Initially the energy density of the Universe is radiation dominated, but as expansion progresses transition to matter domination occurs. Density perturbations become unstable and begin to collapse, with overdense regions collapsing to form haloes and underdense regions "collapsing" to form voids. During the era of matter domination, perturbations grow as $t^{\frac{2}{3}}$ while in the linear regime, i.e. until the density of the perturbation is no longer of the same order as the background density. Initially, only perturbations to the dark matter (DM) component grow since baryons are coupled to the radiation field, but once the Universe cools sufficiently for recombination of protons and electrons this coupling ceases and the baryons begin to collapse into the (already partially formed) DM structure. Without DM (i.e. if all matter coupled to the radiation field), structures smaller than ~ 10 Mpc could not form.

As the perturbations continue to grow, they leave the linear regime. Gravity begins to locally "win out" against the global expansion of space, and shells of matter begin to undergo a net collapse rather than decelerating expansion. This transition is termed the

turnaround of the shell. It occurs at a characteristic time:

$$t_{\max} = 1.095 \frac{t_i}{\delta_i^{\frac{3}{2}}} \tag{1.1}$$

where t_i is the initial time (set around $z = 3100$, when matter domination begins) and δ_i is the overdensity enclosed by the shell at that time. A series of collapsing shells comprise a single structure which, depending on their total mass, will form a galaxy, group or cluster halo. The largest overdensities and smallest structures – i.e. the most concentrated – are the first to collapse. Subsequently, collapsed structures fall together to form ever larger organized structures; the largest today are the filaments, walls and nodes of the cosmic web, which is the natural structure formed by the expansion of underdense voids. Concisely, structure formation proceeds "bottom-up". The outermost shell bound to a structure in an Einstein-de Sitter cosmology is located at a critical radius given by Gunn & Gott (1972):

$$R_{\text{crit}} = \left(\frac{9GM_\infty t^2}{2} \right)^{\frac{1}{3}} \tag{1.2}$$

where M_∞ is the mass of the structure as $t \to \infty$. They also give the radius of the zero-velocity surface, which is that of the shell just turning around at time t:

$$R_{v=0} = \left(\frac{8GM(2t)t^2}{\pi^2} \right)^{\frac{1}{3}} \tag{1.3}$$

where $M(t)$ is the mass of the density perturbation at time t (which can be derived straightforwardly from the initial collapse time of the perturbation, its initial mass, the magnitude of the initial overdensity, the background density of the Universe and the critical density of the Universe at the time of initial collapse). The corresponding expressions in a general FLRW metric are somewhat more complicated; see Mo et al. (2010) for a discussion.

As collapse continues, other nonlinear effects – specifically violent relaxation and phase mixing – intervene and cause the virialization of the collapsing halo at about $t = 2t_{\max}$. Once this process is complete the halo is expected to at least approximately satisfy the virial theorem. At this time the mean density inside the virialized region of the halo turns out to be about 200 times the critical density, leading to the characteristic radius R_{200} and the mass enclosed within it M_{200}.

The mass distribution of DM haloes is quantified using the Press-Schechter theory (Press & Schechter, 1974), which was extended to take subhaloes into consideration by Bond et al. (1991) and to predict the merger rate of haloes as a function of mass by Lacey & Cole (1993).

The mass function is given by:

$$\frac{dn}{dM}(M,t)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_b}{M^2}\frac{\delta_c(t)}{\sigma(M)}\left|\frac{d\log\sigma}{d\log M}\right|\exp\left[-\frac{\delta_c(t)^2}{2\sigma^2(M)}\right]dM \tag{1.4}$$

where $\sigma^2(M)$ is the variance of the overdensity field in a spherical region containing mass M. The mass distribution of haloes obeys an approximate power law with negative slope at low masses, and is cut off exponentially at masses greater than a characteristic mass M_c, which at present turns out to be $\sim 7.4 \times 10^{14}\,\mathrm{M}_\odot$. This is also the present characteristic mass for clustering of DM structure.

The merger rate, stated as the probability that a halo of mass M_1 will merge with another halo of mass $\Delta M = M_2 - M_1$ at time t, is:

$$\begin{aligned}\frac{d^2p}{d\log\Delta M dt}(M_1 \to M_2|t) = \sqrt{\frac{2}{\pi}}\frac{1}{t}\left|\frac{d\log\delta_c}{d\log t}\right|\left(\frac{\Delta M}{M_2}\right)\left|\frac{d\log\sigma_2}{d\log M_2}\right|\frac{\delta_c(t)}{\sigma_2} \\ \times\frac{1}{\left(1-(\sigma_2/\sigma_1)^2\right)^{\frac{3}{2}}}\exp\left[-\frac{\delta_c(t)^2}{2}\left(\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\right)\right]\end{aligned} \tag{1.5}$$

This merger rate is averaged only over accretions onto primary haloes; it does not account for mergers between subhaloes. The current merger rate for equal-mass mergers is predicted to be in sharp decline over the past few Gyr for haloes $\lesssim 10^{14}\ \mathrm{M}_\odot$.

The extended Press-Schechter theory is limited in that it does not predict the spatial and velocity (phase space) distribution of haloes, and does not describe subhaloes after they join their host.

Bertschinger (1985) (see also Fillmore & Goldreich, 1984) complements the Press-Schechter theory nicely with some analytic work on the dynamics of matter shells falling into a structure after it has already formed (secondary infall). He finds that these shells re-expand out to consecutively smaller radii over the course of a few orbits (orbit here meaning collapse to a point then re-expansion), then settle down to a periodic orbit out to a maximum radius of $0.81 r_{\mathrm{turnaround}}$. His fig. 6 shows the idealized phase space trajectory of a secondary infall shell. These results assume spherical symmetry, so all orbits are fully radial and ignore effects such as dynamical friction. Still, the backsplash radius of $0.81 r_{\mathrm{turnaround}}$ can be considered an upper limit, and simulations show that the trajectory through phase space is at least a qualitatively correct description of the movement of subhaloes.

1.2 N-Body & Hydrodynamic Simulations

Once analytic treatment of nonlinear behavior in structure formation becomes impossible, the natural tool to turn to is simulation. Broadly speaking, the problem to solve has three components. The first is the expanding background FLRW metric, governed by the Friedmann equations. These have an analytic solution and are trivially encoded in simulations. The second and third are the collisionless and collisional fluid components. The evolution of these components must be solved in two stages; the distribution of material must be discretized, then the equations for the force(s) acting on each discrete element must be solved and used to advance the distribution in time.

1.2.1 N-Body Simulations

The simplest cosmological simulations include only DM, which is modelled as a collisionless fluid. The matter distribution is discretized as an ensemble of particles in all mainstream modern codes, with each particle representing a (large) collection of physical fundamental DM particles. The particles are assumed collisionless, so the force between them is "softened" so that two particles that pass through each other behave more like two clouds of mass (a close analogy would be a pair of star clusters) passing through each other, rather than a pair of point masses experiencing a close approach. The most intuitive method of solving for the gravitational force felt by each particle is simply to sum up the contribution of each other particle as:

$$\vec{F_i} = -\sum_{j \neq i} G \frac{m_i m_j (\vec{r_i} - \vec{r_j})}{(\epsilon^2 + |\vec{r_i} - \vec{r_j}|^2)^{\frac{3}{2}}} \tag{1.6}$$

where ϵ is the softening term. This solution yields a result as accurate as the discretization allows, but is computationally very slow. In practice it is only used to calculate the contribution from very nearby particles, while the contribution from more distant particles is approximated using faster methods. A tree method divides the distribution of particles into a hierarchy of bins. For instance, the simulation box may be cut into octants, then each octant which contains multiple particles is divided itself into octants, proceeding recursively until each cell contains a single particle. This hierarchical organization can then be used to rapidly compute the centre of mass of all cells (including parent cells that contain subdivisions). The force on each particle is then calculated using contributions from other cells, with nearby contributions calculated from finer cells for accuracy while contributions from collections of more distant particles are combined into a single coarser cell for speed.

Perhaps the most well-known tree algorithm is that of Barnes & Hut (1986). But again in practice a tree code is somewhat unwieldly, with a very direct tradeoff between accuracy and speed. Furthermore, periodic boundaries – not referring to the simple contribution of a particle near one edge of the box being considered nearby to a particle near the opposite edge, but the very long range forces over multiple box lengths – are awkward to handle. Tree methods are therefore usually reserved for the calculation of intermediate range contributions to the gravitational force.

Particle-mesh (PM) solvers are the approach usually used on large scales. The simulation box is overlaid with a grid or mesh, and the particle distribution is projected onto the mesh, giving a sampling of the density field at each mesh point. Poisson's equation:

$$\nabla^2\phi = 4\pi G\rho \tag{1.7}$$

can then be solved for the gravitational potential ϕ extremely efficiently using Fourier methods, which account naturally for a periodic boundary condition. These methods struggle to accurately resolve forces in regions where the mean interparticle distance is much less than the mesh spacing, and using an extremely fine mesh becomes very computationally intensive. In many ways PM methods are complementary to tree methods, and the two are usually combined into "treePM" methods in state-of-the-art codes (Xu, 1995, is the first example). Until recently so-called adaptive mesh refinement (AMR) codes were a popular alternative to treePM methods, using Fourier solutions on nested meshes to increase resolution where required, but these have fallen somewhat out of favour on grounds of computational efficiency (O'Shea et al., 2005).

Given a particular method to solve for the gravitational force on each particle tuned to a desired accuracy, there are essentially two remaining parameters which will control the computational demands of a simulation. The first is the number of particles which has grown rapidly from humble simulations of a few tens of particles a few decades ago to simulations of approximately 10^{10} particles today. The second parameter is the size of the box to be simulated. The largest boxes are on the order of a few Gpc on a side. For a fixed number of particles, a larger box is easier to simulate than a smaller box (Boylan-Kolchin et al., 2009). A reduction in box size corresponds to an increase in mass resolution, and the lowest mass structures are the first in which nonlinear effects become important. Furthermore, there are more particles per (nonlinearly-behaving) structure, making calculations more involved, and the time resolution also needs to be increased to accurately trace the orbits of particle near the centres of haloes. A final key point relating to resolution is that the time between the accretion of a subhalo onto a host and the eventual "complete" (as far as the simulation is concerned) disruption of the subhalo increases with the resolution of the simulation – this is clearly an unphysical effect.

1.2.2 Hydrodynamical Simulations

Different methods are required to include baryonic gas physics in a simulation. There are, broadly speaking, two classes of methods – the following brief overview is based on the summary of Springel (2010), and references therein. The first is smoothed-particle hydrodynamics (SPH), which adopts a Lagrangian point of view. The required fields (e.g. density, pressure, temperature, velocity, etc.) are computed at the location of each particle by "smoothing" over the properties of nearby particles. The relevant equations of fluid dynamics and any other desired baryonic physics are then solved on the smoothed fields. This method's strength is a natural adaptive scale, where high resolution can be achieved in high-density regions for accuracy alongside low resolution in low-density regions for speed. The primary weakness of SPH codes is difficulty in resolving shocks, and more generally in dealing with any sharp boundary or discontinuity, since these will forcibly be smoothed out by the method. The second class of methods use an Eulerian point of view; these mesh codes are operationally similar to the PM methods described in the context of gravity solvers, but solve the equations of hydrodynamics (and possibly also radiative transfer, etc.) instead of gravity. AMR is frequently employed to increase resolution where it is most needed. AMR codes are excellent at resolving shocks and discontinuities. Their primary drawback is that they are "non galilean-invariant", in the sense that it is difficult to deal with bulk flows (especially supersonic flows) and keep the refined mesh spatially matched to the region which requires high resolution, at least in a way that is computationally efficient. Some of the latest methods use a "moving mesh" which purports to have most of the advantages of both SPH and AMR with few of the drawbacks. These methods are relatively new, the first presented by Springel (2010), and have yet to see extensive use, but present exciting prospects.

1.3 Quenching

We know that galaxies in clusters are typically more 'red and dead' than their counterparts in the field (Hogg et al., 2004; Balogh et al., 2004), as well as dominantly elliptical rather than spiral (the so-called morphology-density relation, see Dressler, 1980). This is thought to be due to a mechanism or combination of mechanisms that halt the collapse of cold gas into stars in a satellite galaxy as it orbits within a deep potential well – either by heating or removing the gas or by preventing the cooling or accretion of additional gas and consuming the existing supply. There is a long list of such mechanisms; we describe some of the most studied ones – ram pressure stripping; tidal stripping; mergers, harassment

and cannibalism; strangulation – in §§1.3.1-1.3.4 below. Numerous arguments have been presented in favour of and against each, but in the end it seems reasonable to assume that several or all are at work (Lewis et al., 2002; Gill et al., 2005; van den Bosch et al., 2008); the objective then becomes to determine which is dominant in driving the observed star formation/colour and morphological trends (which may each be driven by separate mechanisms). There are also mechanisms entirely internal to a galaxy that can contribute to its quenching, a few of these are briefly described in §1.3.5. While these cannot explain trends that are a function of environment, they contribute to galaxy evolution and must be taken into account in a complete model (see, for instance, Peng et al., 2010). Some of these may even couple to external quenching mechanisms, for instance by making part of the galactic gas supply more susceptible to stripping by an external mechanism.

In addition to the trends observed with environment, an evolution of the population of galaxies within clusters is observed with redshift. The classic example is the Butcher-Oemler effect of Butcher & Oemler (1978), which shows that in similar clusters at low and high redshift, those at high redshift contain proportionally more blue (presumed spiral) galaxies than those at low redshift, which instead have a surplus of S0s (lenticulars). The cause of the Butcher-Oemler effect is still a topic of active discussion (especially its extension to group environments, see McGee et al., 2009; Li et al., 2012), and the effect has been extended to include an evolution in the mid-infrared properties of cluster galaxies (Haines et al., 2009).

It has also become clear that environmental effects on galaxy evolution are not limited to cluster environments; the group environment is also hostile to star formation. While the distinction between a "poor cluster" and a "rich group" can be somewhat nebulous, Balogh & McGee (2010) argue that groups with masses of $10^{13}\,M_{\odot}$ impact star formation, and potentially even lower mass groups as well. It is widely agreed (Balogh et al., 2000; De Lucia et al., 2012; Mahajan et al., 2012; Wetzel et al., 2013a) that "group preprocessing" will be an important aspect of a complete model of environmental effects on star formation.

One way to attempt to determine which mechanism is dominant is to attempt to determine (i) when quenching starts and (ii) the quenching timescale. While knowing these two quantities would not definitively identify a dominant mechanism, they can be used to constrain the mechanism or combination of mechanisms at work. The measurement of these two quantities has been attempted by several authors (e.g. Balogh et al., 2000; Kauffmann et al., 2004; van den Bosch et al., 2008; Weinmann et al., 2009; Peng et al., 2010; McGee et al., 2011; Taranu et al., 2012; Wetzel et al., 2013a; Haines et al., 2013; Wetzel et al., 2013b), but it is difficult to make and no consensus on their values has yet been reached.

It has so far been found difficult to produce a semi-analytic model for quenching that both reproduces the observed star formation rate (SFR) distribution in clusters (Wetzel et al. (2013a), but see Weinmann et al. (2010) for some recent improvements). However, all the environmental quenching mechanisms listed above have at least one common characteristic : each is sensitive to the path taken through the cluster. Tidal stripping is strongest near the cluster centre. Ram pressure stripping varies with the local density of hot cluster gas and the relative velocity of the satellite, both of which vary radially in the cluster. Mergers are more probable in the outskirts of clusters. Harassment is most effective during high speed encounters, which occur near the cluster core. Strangulation is triggered by the removal of the halo gas of the satellite, ostensibly via one of the aforementioned mechanisms. An environmental quenching model can therefore reasonably be expected to depend on the previous positions and velocities, or more concisely the orbital history, of satellite galaxies.

1.3.1 Ram Pressure Stripping

Gunn & Gott (1972) give a simple theoretical argument for the removal of material from a gas disk moving through the intracluster medium (ICM). The ram pressure exerted by the ICM on a galaxy moving through it is:

$$P_r \approx \rho_{\rm ICM}(r_c)|v|^2 \tag{1.8}$$

where $\rho_{\rm ICM}(r_c)$ is the density of the ICM at the cluster-centric radius of the galaxy and $|v|$ is the speed of the galaxy relative to the cluster. Stripping will occur if the ram pressure exceeds the force per unit area binding gas to the galaxy. For a typical disk/spiral, this force per unit area is (using the more modern notation of Jáchym et al., 2007):

$$F_b = \left.\frac{\partial\Phi(r_g,z_g)}{\partial z}\right|_{\rm max} \Sigma_{\rm ISM} \tag{1.9}$$

where $\Phi(r_g, z_g)$ is the total gravitational potential in galacto-centric cylindrical coordinates and $\Sigma_{\rm ISM}$ is the surface density of the interstellar medium (ISM). The condition for ram pressure stripping is therefore:

$$\rho_{\rm ICM}(r_c)|v|^2 \geq \left.\frac{\partial\Phi(r_g,z_g)}{\partial z}\right|_{\rm max} \Sigma_{\rm ISM} \tag{1.10}$$

The more recent work of Abadi et al. (1999) and Jáchym et al. (2007) uses numerical simulations to explore the limitations of the simple model of Gunn & Gott (1972). They

raise several important follow-up points to this simple model. In particular, ram pressure stripping is most efficient if the disk orientation is face-on to the wind direction. Stripping can be enhanced via other mechanisms, e.g. harassment (see §1.3.3) can puff up the disk and reduce its binding energy, or star formation activity can disrupt GMCs and allow efficient coupling with the wind, which is otherwise impossible for the cold gas phase. Even if full-blown ram pressure stripping is not in effect, thermal evaporation of the ISM through contact with the hot ICM or viscous stripping of the disk as the ICM wind flows by can remove gas, though at a lower rate. Stripping can even be enhanced by the debris left behind by earlier stripping events, which leave higher-density pockets in the ICM, thus increasing the ram pressure. On the other hand, the approximation of Gunn & Gott (1972) can also overestimate the amount of stripping by neglecting the impulse-like nature of the ram pressure interaction which only acts for a very short time as the galaxy passes pericenter, and also by neglecting the possible re-accretion of material back onto the ISM.

There is also direct observational support for ram pressure stripping, which produces a characteristic bending of the disk and formation of a tail behind the galaxy. Smith et al. (2010b) is one example of a recent analysis which identifies 13 ram pressure stripping event candidates in the Coma cluster. They point out that the majority of tails are oriented away from the cluster centre, suggesting that ram pressure stripping acts rapidly on the first approach to the cluster core.

1.3.2 Tidal Stripping

The differential force from a host cluster on a satellite galaxy can also remove mass from the satellite. The satellite DM halo is the first component affected as it is the most diffuse, followed by the hot gas halo and, in the case of an especially strong tidal field, possibly the gas and stellar disks.

In a very simplified picture where the host and satellite are idealized as well separated, it is easy to derive an approximate measure of the tidal stripping effect. Along an axis connecting the centre of the host and satellite (which are separated by R), the tidal acceleration felt by a test particle a distance r from the centre of the satellite is, to first order:

$$a_{\rm tidal} = \frac{2rGM_{\rm host}(< R)}{R^3} \tag{1.11}$$

Naïvely, if the magnitude of this acceleration exceeds the gravitational acceleration due to

the satellite, we expect the test particle to be stripped, giving the condition:

$$\frac{GM_{\rm sat}(< r)}{r^3} < \frac{2rGM_{\rm host}(< R)}{R^3} \tag{1.12}$$

Or, in terms of the average densities of the host and satellite[1]:

$$\bar{\rho}_{\rm sat}(< r) < 2\bar{\rho}_{\rm host}(< R) \tag{1.13}$$

Thus, very crudely, the satellite will be stripped until its mean density is on the order of the mean density of its host inside its orbit. A much more detailed analytical treatment is given by Merritt (1983).

Simulations are a powerful tool to study the tidal stripping of DM haloes since the simulation requries only gravitational physics. For instance, using simulations Gao et al. (2004) find that 92 per cent of the mass in subhaloes of clusters at $z = 1$ has been stripped by $z = 0$ (i.e. over a period of about 8 Gyr). Warnick et al. (2008) find that a typical subhalo loses about 30 per cent of its mass per orbit to tidal stripping, while a subhalo on a very eccentric orbit may lose up to 80 per cent of its mass (both these quantities do not take into account satellites that are completely disrupted, i.e. lose 100 per cent of their mass in an orbit). A more complete analysis requires the inclusion of baryonic physics, both to examine the effect of tides on the galaxy hosted by a halo, and to account for the additional gravitational force provided by the galaxy. The first work studying tidal stripping specifically in a full-fledged hydrodynamical simulation is by Limousin et al. (2009). Their results suggest that the inclusion of baryonic physics makes the stripping of DM haloes more efficient in the cores of clusters, and that the stellar component of galaxies is much less sensitive to tidal stripping than the DM halo.

Even if the tidal field is not strong enough to actually strip matter from the satellite, it may trigger other effects. Tidal heating may inject energy into the gas, making it easier to remove by the action of ram pressure. Conversely, a bar instability can be triggered, which acts to funnel gas toward the center of the galaxy and bind it more tightly (Mayer et al., 2006). Tidally induced collisions of GMCs can trigger star formation, and more broadly tides are capable of driving the morphological evolution of a galaxy (Byrd & Valtonen, 1990).

[1] A slightly more detailed calculation by Binney & Tremaine (2008) gives nearly the same result: $\bar{\rho}_{\rm sat}(< r) < 3\bar{\rho}_{\rm host}(< R)$

1.3.3 Mergers, Harassment and Cannibalism

These three types of interaction describe interactions between galaxies in the cluster. Major mergers are dramatic events capable of completely changing the morphology and star forming properties of the participating galaxies, but are relatively rare in the cluster environment due to the high relative speeds of galaxies (Tormen et al., 1998). This makes major mergers an unlikely candidate to drive the evolution toward a population of quescient galaxies. Nevertheless, when a major merger does occur, a starburst is triggered and consumes up to 80 per cent of the initial mass in gas to stars, leaving little for continued star forming activity (Cox et al., 2006).

Minor mergers are also rare for the same reason, but close high speed encounters between satellites are common. Moore et al. (1996) showed that a typical satellite galaxy experiences a close encounter with an L_* (or larger) galaxy about once every Gyr. Even a single such interaction can be sufficient to trigger a bar instability and pull out long streams of material which are subsequently easily stripped, and may trigger starbursts which rapidly consume gas as well. At first this high interaction rate seems contradictory with the observation that very few cluster satellites are seen to be interacting (Dressler et al., 1994), but this is explained by the very short duration of the interactions. The collective effect of these repeated encounters is termed "harassment". A handful of harassing encounters are sufficient to drive evolution from a strong spiral morphology to an S0 (lenticular).

More recent work by Smith et al. (2010a) – which focusses on dwarfs, but contains some results that are more generally applicable – points out that harassment is highly dependent on the orbit of a galaxy, with most encounters occuring close to the centre of the cluster. This suggests that the effects of harassment should not become apparent until a galaxy approaches its pericenter. Harassment effects are particularly strong for galaxies that have small apocentric distances (and so continue to be harassed on a continuous basis).

Galactic cannibalism refers to a large central galaxy consuming smaller galaxies as they pass, leading to the formation of a cD-type "supergiant elliptical" (Ostriker & Hausman, 1977). While this process certainly stops star formation in the accreted galaxies (if it had not halted already), it also completely destroys them, so cannibalism cannot contribute to the abundance of red galaxies in clusters. Furthermore, cannibalism is inefficient in the cluster environment due to the high velocity dispersion of the satellites; it occurs primarily before cluster virialization (Tovmassian & Andernach, 2012).

1.3.4 Strangulation

Strangulation (sometimes also called starvation) is a mechanism – or perhaps more accurately class of mechanisms – in which the gas supply of a galaxy is cut off so that star formation eventually halts after the gas supply in place is depleted. The gas supply can be cut off in a number of ways, e.g. removal of a hot gas halo (McCarthy et al., 2008; Font et al., 2008), removal of HI (Boissier et al., 2012), direct removal of molecular gas (Koopmann & Kenney, 2004), etc. The original argument by Larson et al. (1980) was that the gas consumption timescale in a typical spiral galaxy is much shorter than the Hubble time, so the gas must be replenished for star-forming galaxies to exist at all. Clusters could provide a truncation mechanism, but to explain the Butcher-Oemler effect the truncation must not occur abruptly upon accretion onto the cluster. They also note that a typical depleted spiral can internally replenish its gas supply via stellar winds, so active stripping must be ongoing for a galaxy to remain quenched.

A number of authors have since provided strong indirect observational evidence for strangulation models (e.g. Balogh et al., 2000; van den Bosch et al., 2008; Haines et al., 2013), and more broadly quenching timescales are typically found to be on the order of a few Gyr, which is consistent with a strangulation model (though see Wetzel et al., 2013a, which finds a much shorter timescale for quenching, but also a long delay before quenching begins, which might also be associated with strangulation).

1.3.5 Internal Mechanisms

There are also mechanisms with no direct environmental dependence that can act to quench star formation. Star formation feedback (often called supernova (SN) feedback) occurs when a new stellar population is formed. The hottest, brightest, most massive stars burn rapidly and explode as SNe after a relatively short delay. This injects large amounts of energy into the surrounding gas, potentially enough to heat it and halt star formation or even eject material from the galaxy entirely (Dekel & Silk, 1986; Veilleux et al., 2005).

Active galactic nucleus (AGN) heating occurs when powerful radiation and outflows driven by accretion onto an AGN heat or expel galactic gas (McNamara et al., 2006; Croton et al., 2006; Gabor et al., 2010). The end result is similar to SN feedback; heating or expulsion of gas prevents star formation.

The same mechanism that maintains the hot gas halo in a galaxy cluster can also occur in galaxies. Under the right conditions (for instance a halo mass of at least $10^{12}\,M_{\odot}$, see Dekel & Birnboim, 2006), gas is heated as it falls into the DM halo and forms a standing

shock front, which then shock heats additional infalling gas (Birnboim & Dekel, 2003). This prevents direct 'cold-mode' accretion from feeding the galactic disk which supplies the fuel for star formation.

While internal mechanisms cannot directly account for the observed environmental dependence of morphology, colour and SFR, these mechanisms can couple to environmental effects – e.g. by heating material, making it more susceptible to environmental stripping mechanisms – and must be taken into account in a model that seeks to capture a complete picture of galaxy formation and evolution.

1.4 Coordinates, Observables and Orbits

In order to compare the results of different authors, it is necessary to introduce several scale radii. These are analogous in definition to the R_{200} mentioned above except that depending on context different overdensities may be used (100 rather than 200, for instance) and the reference density may be the critical density or the background density ($\rho_{\rm c} = \Omega_{\rm m}^{-1}\rho_{\rm b}$). Subscripts are used to denote both the overdensity and the reference density, so $R_{100{\rm c}}$ refers to the radius enclosing a mean density 100 times the critical density while $R_{360{\rm b}}$ is the radius enclosing 360 times the background density. This last scale radius is the one used throughout this thesis; we frequently call it $R_{\rm vir}$. Any other scale radius will be explicitly specified.

Satellite orbits in a cluster are not randomly distributed. Wetzel (2011) finds, at $z = 0$, that their mean circularity is $\sqrt{1-e^2} = 0.52$ (i.e. eccentricity of 0.85), and the median pericentre is $0.21R_{200{\rm c}} \sim 0.15R_{360{\rm b}}$. Both quantities have a standard deviation of about 0.2. The median pericentre and mean circularity both decrease in higher mass host haloes, but no trend with satellite mass is noted. Gill et al. (2004) find somewhat different values, with a mean eccentricity of 0.61 and a mean pericentre of $0.35R_{340{\rm b}} \sim 0.35R_{360{\rm b}}$, again with a scatter of 0.1–0.2. Mamon et al. (2004) gives the maximum apocentric, or "backsplash", distance for satellites at $2.5R_{100{\rm c}} \sim 2.5R_{360{\rm b}}$, with typical backsplash distances anywhere from $1.0R_{360{\rm b}}$ to $2.5R_{360{\rm b}}$. Wetzel et al. (2013b) give a somewhat lower maximum backsplash distance of about $2.5R_{200{\rm b}} \sim 1.4R_{360{\rm b}}$ We would expect higher mass satellites to have lower backsplash distances since the magnitude of dynamical friction is proportional to the mass of the satellite.

The orbital history of a galaxy is not directly observable; if we hope to compare orbit-based environmental quenching models to observational data, we need some way to characterize it. We might parametrize the orbital history of a galaxy in any of a number of ways

– for instance by the time since infall into the cluster potential, the distance of closest approach to the cluster centre or the number of orbits completed since infall – then attempt to find correlations between these parameters and observables using N-body simulation data. The observables we consider available are the distance between the satellite and the cluster centre projected on to the plane of the sky and the component of the velocity of the satellite relative to the cluster centre along the line of sight (LoS). In some nearby clusters it may also be possible to obtain distances to the cluster centre and the satellite via direct distance measurements (e.g. Tully-Fisher relation, surface brightness fluctuations, SNIa luminosities, etc.), but in general these will be much less precise than the other two coordinates – we will consider this information inaccessible (for an interesting analysis of satellite orbits similar to ours but focussing on the satellites of the Milky Way for which distances and sometimes proper motions are available, see Rocha et al., 2012).

Several authors (Balogh et al., 2000; Gao et al., 2004; Wang et al., 2011; De Lucia et al., 2012; Taranu et al., 2012) have shown that the present radial distance from the cluster centre is negatively correlated to the time since infall into the cluster and other orbital parameters. Gill et al. (2005) were amongst the first to compare the velocity distributions of different satellite populations (see also Wang et al., 2005), and more recently Mahajan et al. (2011) and Noble et al. (2013) have used methods for seperating satellite populations based on, amongst other parameters, their LoS velocities.

1.5 Star Formation and Galaxy Colour

Very roughly speaking, the relationship between galaxy colour and SFH is simple. A galaxy that has had active star formation recently will include bright, blue (e.g. O- and B-type) stars as well as dimmer, redder stars. The brightest, bluest stars are the shortest lived, so as they burn out only the longer lived redder stellar populations remain and the overall colour of the galaxy reddens with increasing time since active star formation. In more detail a stellar synthesis model must be adopted. The model of Maraston (2005, see references therein for some examples of the many other models that have been proposed) was used to guide some of the work in the present thesis. Bruzual & Charlot (2003) provide a concise example of how the SFH and metallicity of a galaxy can be constrained from its colour using a stellar population synthesis model.

Observationally, two populations of galaxies are identified by their colours. Strateva et al. (2001) identified a red population and a blue population in the Sloan Digital Sky Survey (SDSS), and a correlation with morphology, with red galaxies being predominantly early-type (E/S0) and blue galaxies late-type (spiral). They note that, of all possible colours

defined using SDSS filters, $(u-r)$ gives the best separation between the two populations. Blanton et al. (2003) also find a significant bimodality in $(g-r)$ colour, and note a strong luminosity dependence; more luminous galaxies are redder. The population of red galaxies is essentially identical in low and high density environments (Hogg et al., 2004). There are some weak trends, with luminosity, size and velocity dispersion increasing with the density of the local environment and surface brightness decreasing (Bernardi et al., 2003). The main evolution with environmental density is the relative abundance of red and blue galaxies (Hogg et al., 2004; Balogh et al., 2004).

Baldry et al. (2004) showed that the $(u-r)$ colour distribution at low redshift is well-modelled by a pair of gaussian functions, one for each colour peak. They also note that the colour distribution is a strong function of absolute magnitude M, with an evolution that is approximately linear with respect to M, but with a correction proportional to $\tanh M$. Given the similar (but less pronounced) bimodality in $(g-r)$ colour, we predict that a pair of gaussians will provide a good fit to this distribution as well.

1.6 Cosmology

We assume the same cosmology used in the Bolshoi and MultiDark Run 1 (MDR1) simulations with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.0469$, $n_s = 0.95$, $h_0 = 0.70$, $\sigma_8 = 0.82$ (Prada et al., 2012).

Chapter 2

Data

Our study of the colour evolution of cluster satellites requires large samples of both simulated and observational data, presented below in §2.1 and §2.2 respectively.

2.1 Simulations

To obtain a large sample of satellite orbits, we use the output of the MDR1 DM only simulation (we also consider the output of the Bolshoi simulation as a means to extend the mass range of our analysis in Appendix A) with 1 Gpch^{-1} box side length, 2048^3 particles, current cosmology (Wilkinson Microwave Anisotropy Probe (WMAP)5/WMAP7), $8.7 \times 10^9 h^{-1}\,\mathrm{M}_\odot$ mass resolution and 7 kpc force resolution. The simulation was run from $z = 65$ to the present; the majority of snapshots are output at even intervals in scale factor a with some irregular intervals at small a. The resolution in scale factor is of 0.0304 before $a \sim 0.7$ ($z \sim 0.43$) and doubles to 0.0152 afterwards (corresponding to a time resolution of about 0.210 Gyr at $z = 0$). This is taken into account in our discussion below. Full details of the simulation parameters are described in Prada et al. (2012). The simulation snapshots were processed by the ROCKSTAR halo finder (Behroozi et al., 2013b) and the merger tree code presented in Behroozi et al. (2013c). To produce the data used in this paper, the merger tree code was slightly modified so that a halo may contain satellite haloes at distances of up to 2.5 times its virial radius ($r_{\rm vir} = r_{\rm 360b}$) from its centre; the default code finds satellites within only 1.0 virial radii. This allows us to track satellites out to the largest apocentric distances which Mamon et al. (2004) (see also Gill et al., 2005; Balogh et al., 2000; Ludlow et al., 2009) show to be about $\sim 2.5 r_{\rm 200c} (\sim 2.2 r_{\rm 360b})$ by using host-satellite linking as a proxy for cluster membership.

In the following discussion, we will denote full 6D cluster-centric coordinates (r, v). The position and velocity centre of a halo is determined by averaging the positions and velocities of the subset of halo particles which minimizes the expected Poisson error in the coordinates, i.e. the particles occupying the region of highest local density (Behroozi et al., 2013b, §3.5.1 provides full details on how coordinates are determined by the halo finder).

Projected coordinates will be denoted (R, V). Projection is done along the arbitrarily chosen third (z-)axis of the simulation box and includes a correction to the velocities to account for the Hubble flow. With this correction, the simulated velocity differences can be directly compared to observation data, where velocity differences will be measured using a redshift difference. The projected distance between two points is $R_{12} = \sqrt{(r_{2,x} - r_{1,x})^2 + (r_{2,y} - r_{1,y})^2}$ and the relative velocity of point 2 with respect to point 1 is $V_{12} = |(v_{2,z} - v_{1,z}) + H(r_{2,z} - r_{1,z})|$. Note that $V \geq 0$; this encodes our assumption that the distances to the two points are not known precisely, so only the magnitude of their relative velocity can be determined.

For ease of comparison between satellites of different hosts, all spatial coordinates are normalized to the virial radius $r_{\rm vir}$ of the host halo, which is defined using the formula of Bryan & Norman (1998): the radius enclosing a region with an overdensity of 360 times the background density at $z = 0$. For readers more accustomed to normalization by $r_{\rm 200c}$, an approximate conversion at $z = 0$ is $\frac{r_{\rm 200c}}{r_{\rm vir}} \sim 0.73$. All velocity coordinates are normalized to the rms velocity dispersion σ of the host halo, measured in 3D.

2.1.1 Interloper Sample

Due to the lack of precise distances to clusters and their satellites in observation, it is impossible to impose a cut such as "all galaxies within a given 3D radius of the cluster center". Instead, a cut is imposed on the projected radius, with a second cut imposed on the LoS velocity to remove galaxies that are far removed from the cluster in velocity space (and therefore most likely far removed in LoS distance as well). This gives a cylindrical volume in projected position-LoS velocity space (strictly speaking the region has the shape of a cone with its top sliced off, but is well approximated by a cylinder for distant clusters). Provided the velocity cut is chosen sufficiently large, this cylinder contains the sphere centered around the cluster containing the galaxies of interest, and a number of galaxies outside this sphere, which we term "interlopers". Some of these interlopers will likely someday fall into the cluster while others may eventually move off into a neighbouring structure.

Our sample of interlopers is constructed by selecting all haloes with $\frac{R}{r_{\rm vir}} < 2.5$, $\frac{V}{\sigma_{\rm 3D}} < 2.0$

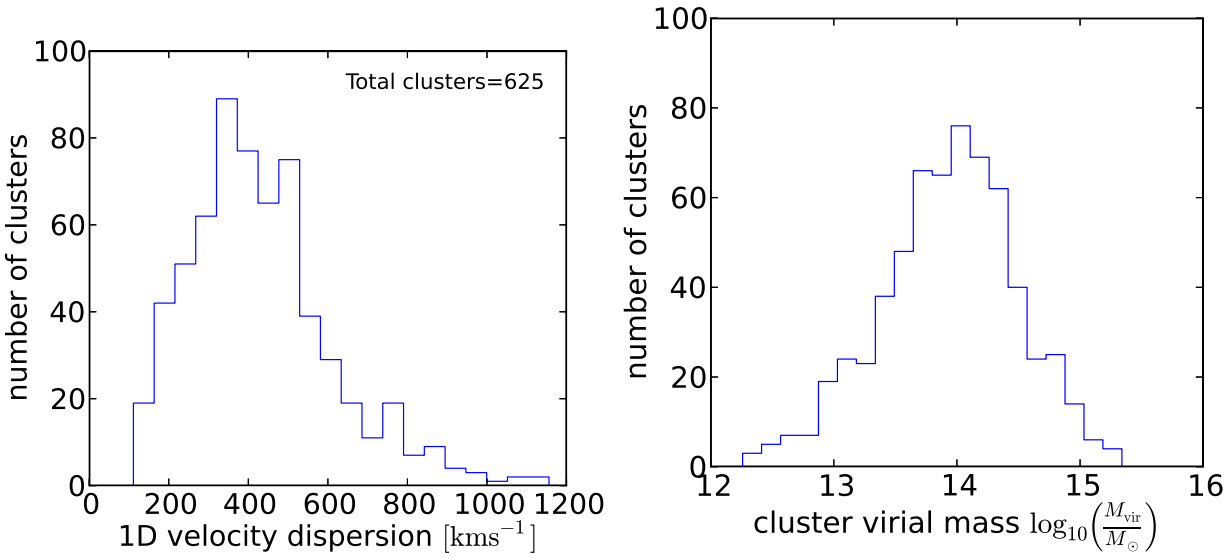

Figure 2.1: *Left panel:* Velocity dispersion distribution of the clusters in the catalogue of von der Linden et al. (2007). *Right panel:* Distribution of halo masses determined using an empirical relation (see Fig. 2.2) for the same sample of clusters.

and $\frac{r}{r_{\rm vir}} > 2.5$, i.e. all haloes within the cylinder described above, but outside a sphere centered around the cluster. See §3.1.1 and §3.3 for details on the interloper sample and discussion, respectively.

2.2 Observations

To obtain a large sample of clusters and their satellites we use the cluster catalogue of von der Linden et al. (2007). This provides the right ascension (RA), declination (Dec), redshift and velocity dispersion of 625 clusters. The velocity dispersion and halo mass distributions of the cluster sample are shown in Fig. 2.1. Halo masses were estimated using an empirical relation derived from our simulation dataset which provides both the mass and velocity dispersion of each halo; see Fig. 2.2. The redshift distribution of the clusters is shown in Fig. 2.3.

Our sample of satellites is drawn from the SDSS DR7 (Abazajian et al., 2009) with the bulge + disk decompositions of Simard et al. (2011). The decomposition is performed using Galaxy IMage 2D (GIM2D) (Simard, 1998) and fits a pure exponential disk + a

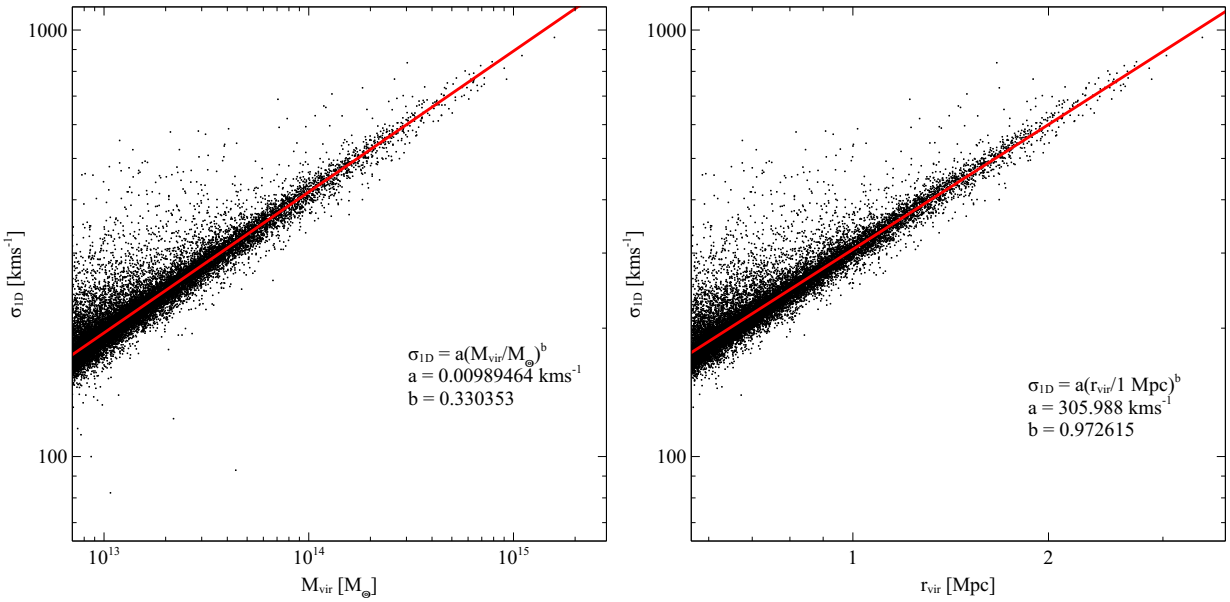

Figure 2.2: *Left panel:* Velocity dispersion as a function of virial mass for haloes in the MDR1 simulation. The data are fit using a power law which is used to estimate the virial mass of our observed sample of clusters. *Right panel:* Velocity dispersion as a function of virial radius for haloes in the MDR1 simulation. Again, a power law is fit to the data, which is used to estimate the virial radius of our observed sample of clusters. In both panels, only $\frac{1}{20}$ of the points are shown for clarity.

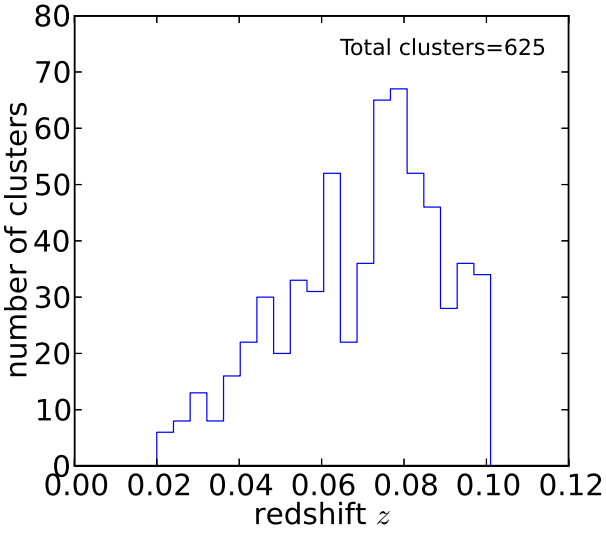

Figure 2.3: Redshift distribution of the clusters in the catalogue of von der Linden et al. (2007).

de Vaucouleurs bulge with Sérsic index $n_b = 4$. The fit is performed simultaneously in two bands (g and r) by matching the bulge radius, ellipticity and position angle, and the disk scale length, inclination and position angle in both bands. This criterion decreases the error in the fits, but also makes them insensitive to colour gradients in the bulge/disk components. The resulting catalogue provides total magnitudes in the g and r bands, decompositions in the same bands as well as redshifts for 1,123,718 galaxies. However, we use only galaxies with spectroscopic data – and thus more reliable redshifts – of which there are 674,701. These data were supplemented with RA and Dec coordinates from the SDSS DR7 archive and stellar masses obtained from the latest release[1] of the data originally presented in Kauffmann et al. (2003).

Figs. 2.4–2.6 illustrate, respectively, the total colour, disk colour and bulge colour distributions for a subset[2] of the SDSS DR7 galaxies as a function of absolute r-band magnitude M_r and a measure of local environmental density defined in terms of the distance to the

[1] `http://home.strw.leidenuniv.nl/~jarle/SDSS/index.html`

[2] d_N values (see Eq. 2.1) were readily available for the SDSS DR6 (Adelman-McCarthy et al., 2008), which includes 215,099 galaxies from our DR7 spectroscopic dataset.

$N^{\rm th}$ nearest neighbour of the galaxy d_N (Baldry et al., 2006):

$$\Sigma_N = \frac{N}{\pi d_N^2} \tag{2.1}$$

Galaxies with $\Sigma_5 < 0.2$ are completely isolated field galaxies, while $\Sigma_5 > 6.5$ is typical for a cluster core. Figs. 2.4–2.6 are intended to be easily compared to fig. 1 of Balogh et al. (2004), using the same M_r and Σ_5 binning, but $(g-r)$ instead of $(u-r)$ colour.

In total colour (Fig. 2.4), we note a (sometimes weakly) bimodal distribution with peaks near $(g-r) = 0.4$ and $(g-r) = 0.7$. The distribution is a strong function of environment for faint galaxies with higher density environments hosting a higher relative abundance of red galaxies (as compared to blue galaxies). Luminous galaxies are predominantly red regardless of environment.

The disk colour distribution (Fig. 2.5) are qualitatively similar to the total colours, but the two colour peaks are slightly broader and closer together, weakening the bimodality. Edge-on disks are subject to dust-reddening; we verified that this does not significantly affect the colour distributions by comparing the distributions for low inclination disks ($i < 50°$) and all disks. Bulge colours (Fig. 2.6) again show the same qualitative trends with environment and luminosity, but more weakly.

We construct our sample of observed satellites as similarly as possible to our sample of simulated satellites. The virial radius of each cluster is determined from its virial mass (which is estimated from the velocity dispersion, see Fig. 2.2):

$$M_{\rm vir} = \frac{4}{3}\pi r_{\rm vir}^3 \left(360\Omega_m \rho_c\right) \tag{2.2}$$

A galaxy is flagged as a satellite candidate of a cluster if it is within 2.5 $r_{\rm vir}$ of the cluster centre and its LoS velocity offset $|\Delta v_{\rm LoS}|$ from the cluster is less than 2.0 $\sigma_{\rm 3D}$. The brightest cluster galaxys (BCGs) are included in the satellite population in our analysis. We note that by construction, the BCGs have coordinates $(R, V) = (0, 0)$ in their respective clusters since von der Linden et al. (2007) define the cluster centres in their catalogue as the location of the BCG. This assumes that the BCG is hosted by the cluster halo (rather than a satellite halo), and no corresponding orbits would appear in our simulated orbit catalogues. We note that the BCGs account for less than 2 per cent of our satellite sample and so we do not expect their presence to impact our conclusions.

The projected radius is determined by first finding the angular separation between the galaxy and the cluster:

$$\Delta\theta = \left|\arccos\left(\sin\left({\rm Dec}_g\right)\sin\left({\rm Dec}_c\right) + \cos\left({\rm Dec}_g\right)\cos\left({\rm Dec}_c\right)\cos\left({\rm RA}_g - {\rm RA}_c\right)\right)\right| \tag{2.3}$$

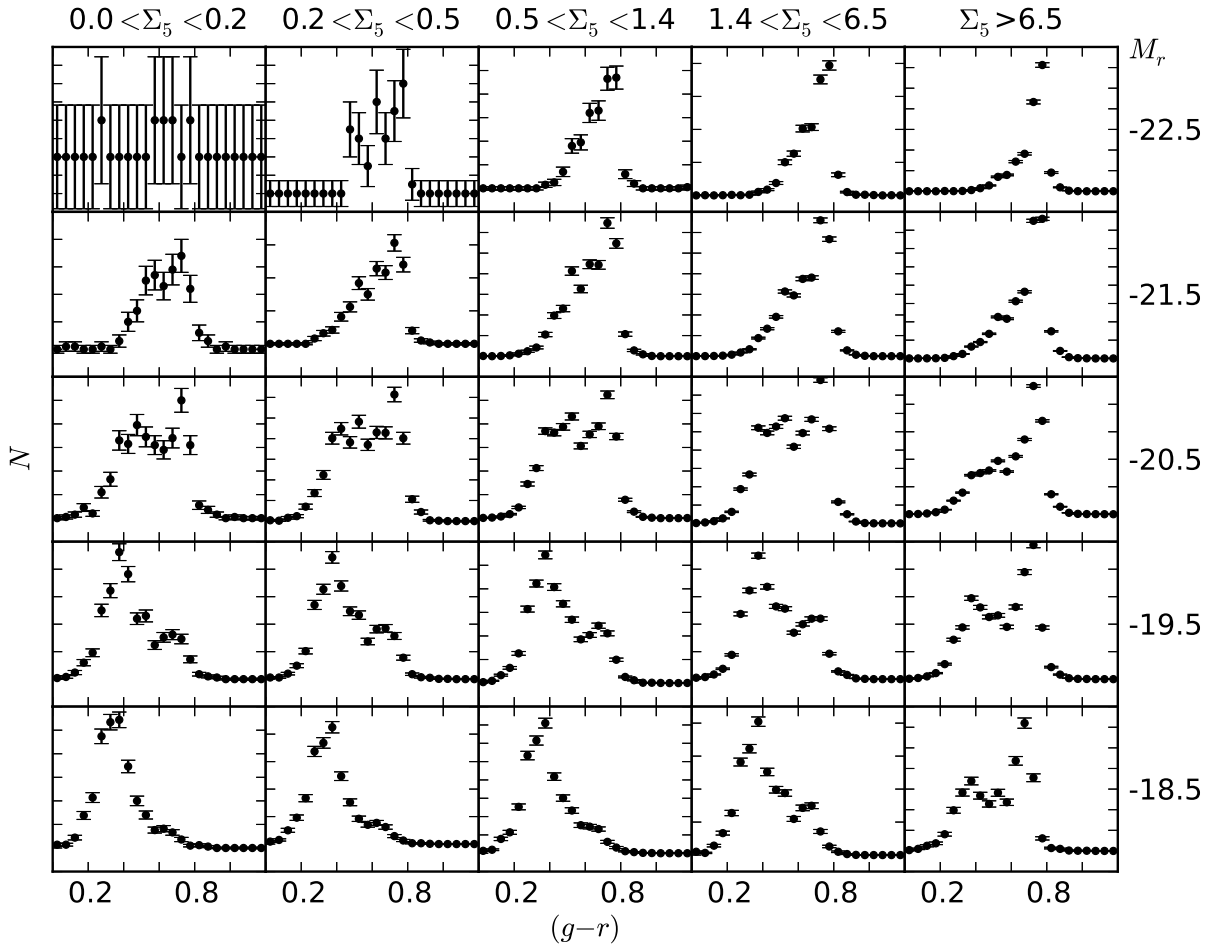

Figure 2.4: Total (bulge + disk) $(g-r)$ colour distribution as a function of absolute magnitude M_r (1.0 mag bins) and local density Σ_5 (see Eq. 2.1) for SDSS DR6 galaxies. The y axis displays absolute number counts; each panel is scaled to fill the available space. The error bars have length $\sqrt{N+2}$. Galaxies with r-band bulge to total light ratios $(B/T)_r < 0.1$ and $(B/T)_r > 0.9$ have been omitted. The same Σ_5 and M_r binning as in fig. 1 of Balogh et al. (2004) is used. We note two peaks in colour, one blue near $(g-r) = 0.4$ and one red near $(g-r) = 0.7$. The relative abundance of red galaxies as compared to blue galaxies increases with increasing local density and galaxy luminosity. Luminous galaxies are preferentially red, regardless of environment.

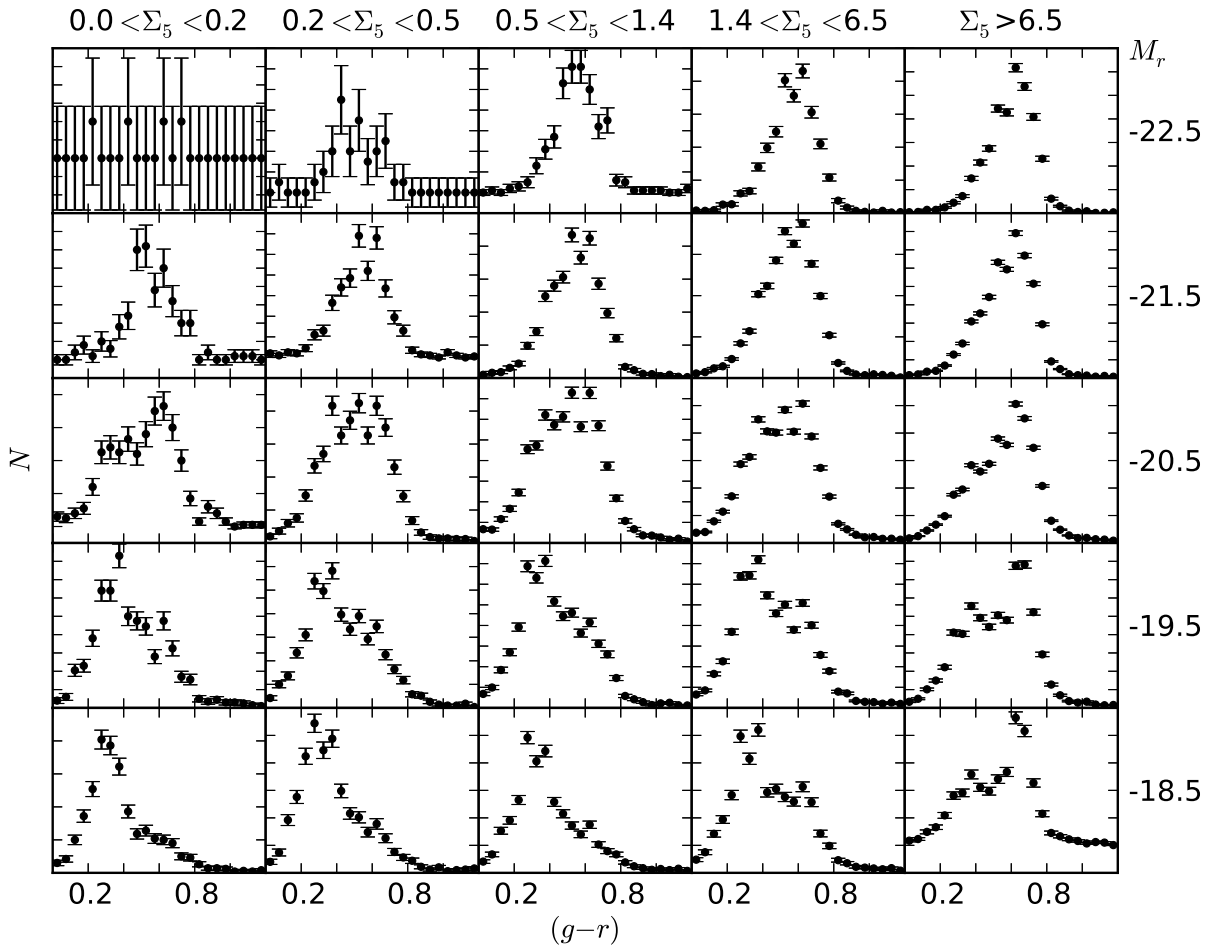

Figure 2.5: As in Fig. 2.4, but for disk colours. There are again two peaks, one red near $(g-r) = 0.65$ and one blue, whose position appears to vary as a function of Σ_5. The peaks are typically slightly broader and closer together than in the total colour distribution, weakening the bimodal character of the distributions.

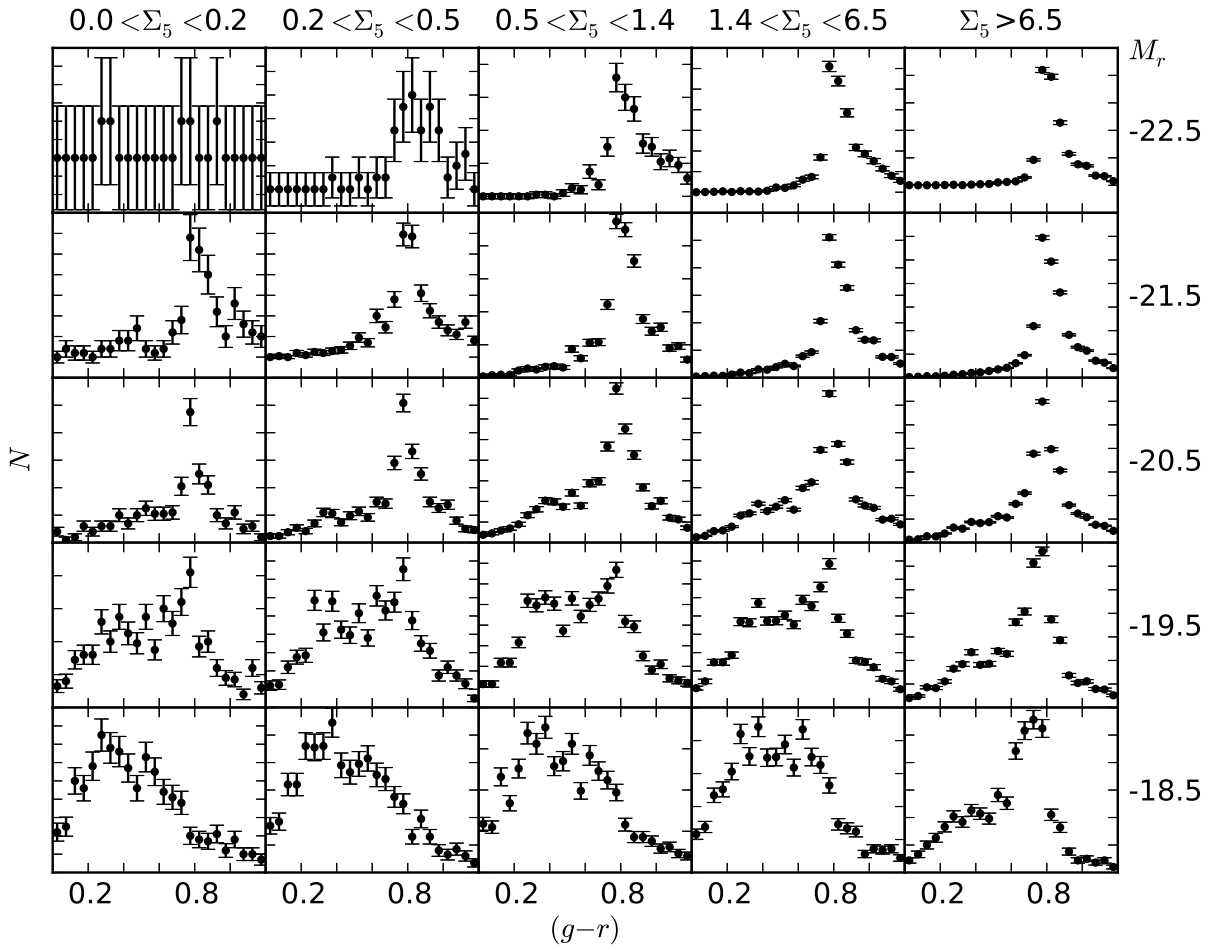

Figure 2.6: As in Fig. 2.4, but for bulge colours. Except for the lowest luminosity galaxies, the bulge colours are dominated by a red peak near $(g-r) = 0.8$ regardless of environment.

where the subscripts g and c indicate the RA and Dec of the galaxy and cluster respectively. Then, given the angular diameter distance d_A at the redshift of the cluster assuming our cosmological parameters (see §1.6), the projected radius of the galaxy is simply:

$$\frac{R}{r_{\rm vir}} = \frac{d_A \Delta\theta}{r_{\rm vir}} \tag{2.4}$$

The velocity offset is calculated as:

$$\frac{|\Delta v_{\rm LoS}|}{\sigma_{\rm 3D}} = \frac{c|z_g - z_c|}{(1+z_c)\sqrt{3}\sigma_{\rm 1D}} \tag{2.5}$$

where z_g is the redshift of the galaxy, z_c is the redshift of the cluster and $\sigma_{\rm 1D}$ is the velocity dispersion of the cluster measured along the LoS. The velocity dispersion is assumed to be isotropic, giving the simple relation:

$$\sigma_{\rm 3D} = \sqrt{3}\sigma_{\rm 1D} \tag{2.6}$$

This assumption is verified using our MDR1 data in Fig. 2.7. In most cases, Eq. 2.6 gives $\sigma_{\rm 3D}$ accurate to withing 15 per cent or better.

This process of associating galaxies to clusters yields a sample of 49,107 satellite candidates, of which we expect about half to be interlopers (see §3.1 and §3.1.1). The distribution of the number of satellites in each cluster is shown in Fig. 2.8, and the distributions of satellite candidate stellar masses and halo masses are shown in Fig. 2.9. Halo masses are estimated using the results of Velander et al. (2013). We use their relation between stellar mass and halo mass (M_{200}, conversion from M_{200} to $M_{\rm vir}$ was based on the work of Bullock et al. (2001)) for 'blue lenses' shown in their fig. 10. We make this choice since we believe the properties of our satellite candidates correspond better to these than the 'red lenses', and we expect that most satellites best modelled as 'red lenses' fall in the mass range where the two relations cross over. In any case, our objective here is only to obtain a rough estimate of the halo mass of the satellites. The exact relation used is:

$$M_{\rm vir} = 10^{0.84\log_{10} M_* + 3.69} \tag{2.7}$$

The fit parameter uncertainties of Velander et al. (2013) are indicative of systematic errors in the conversion from M_* to $M_{\rm 200c}$ of about 15 per cent. There is a further contribution to the systematic error in the conversion from $M_{\rm 200c}$ to $M_{\rm vir}$. This conversion is

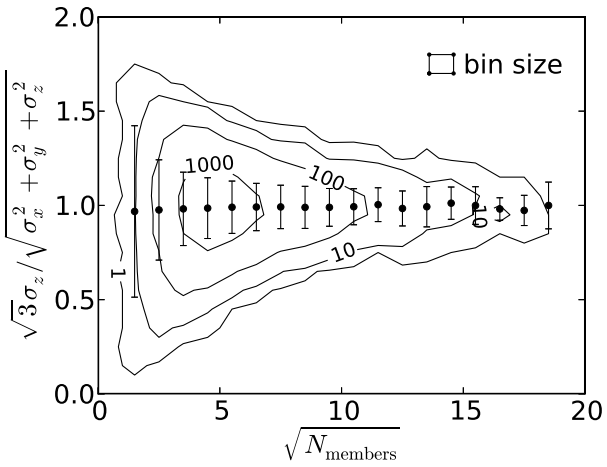

Figure 2.7: The velocity dispersion of each cluster in MDR1 was "measured" along the three orthogonal axes of the simulation box by computing the standard deviation of the satellite halo velocities. Contours indicate the ratio of $\sqrt{3}\sigma_z$ (an approximation of the velocity dispersion if only one component can be measured, this assumes an isotropic velocity dispersion) and $\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$ (exactly the velocity dispersion provided σ_x, σ_y and σ_z are measured accurately) as a function of the square root of the number of cluster members used to measure the velocity dispersion components $\sqrt{N_{\rm members}}$. The bin size used to determine contour heights is shown. The points and error bars show the mean and 1σ scatter of the distribution in bins of $\sqrt{N_{\rm members}}$. For clusters with a small number of members, the scatter is dominated by the shot noise from measuring σ_x, σ_y and σ_z from a small number of members. As $N_{\rm members} \to \infty$, the remaining scatter is due to the spread in cluster velocity dispersion anisotropies.

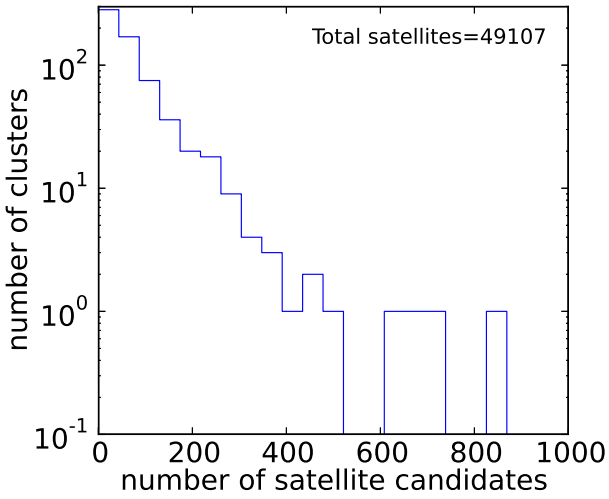

Figure 2.8: Histogram of the number of satellite galaxy candidates drawn from the catalogue of Simard et al. (2011) associated to each cluster from the catalogue of von der Linden et al. (2007). Full details of the process used to associate satellite galaxy candidates to clusters is given in §2.2.

exact except for the choice of a halo concentration parameter c. The dependence on c is relatively weak; we expect that this error is much smaller than the error in converting from stellar to halo mass. There is also a random error in the stellar to halo mass conversion. Moster et al. (2010) suggests a scatter of 0.15 dex in the conversion from a halo mass to a stellar mass, constant with halo mass. We estimate the scatter in the reverse conversion by numerically inverting the functional form of Moster et al. (2010) (using their fit values as well). The scatter in the M_* to $M_{\rm vir}$ conversion is mass dependent but is bounded between 0.08 dex (at low $M_{\rm vir}$) and 0.4 dex (at high $M_{\rm vir}$).

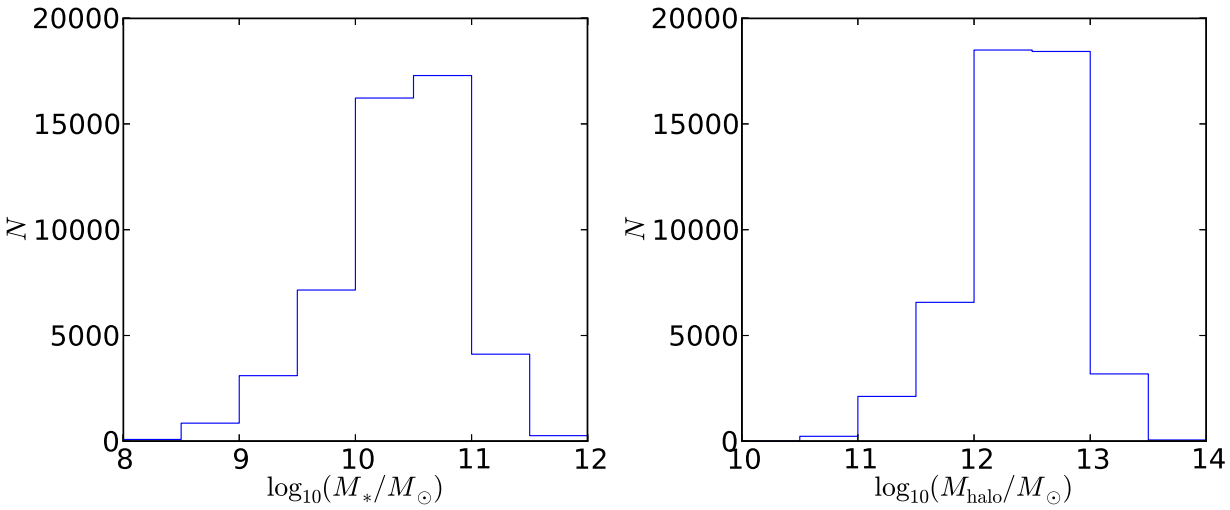

Figure 2.9: *Left panel:* Distribution of satellite galaxy candidate stellar masses. Stellar masses were obtained from the latest release of the data originally presented in Kauffmann et al. (2003). *Right panel:* Distribution of satellite galaxy candidate halo mass estimates. Halo mass was estimated using the relation to stellar mass outlined in Velander et al. (2013) (see §2.2 for details of the estimation).

Chapter 3

Results: Orbit Statistics and Tracking

From the MDR1 simulation dataset presented in §2.1 we create a collection of probability distribution functions (PDFs) of time since infall for cluster satellites as a function of their projected phase space coordinates. The process to achieve this is described in §3.1. There are several systematic effects inherent in this method that need to be understood before using the PDFs. We will discuss the impact of projecting the data in both the radial and velocity coordinates in §3.2. In §3.3 we discuss the impact of the interloper sample on our analysis. In §3.4 we will present the PDFs and discuss some of their features. In §3.5 we discuss the effects of both host and satellite mass on the distribution of satellites in phase space. In §3.6 we discuss the impact of numerical resolution effects on our results (see also Appendix A).

3.1 Method: Orbit Catalogues and Infall Time PDFs

First, the orbital history of a satellite and its host at $z = 0$ are determined. Then the infall time of the satellite into that host halo is defined as the earliest time at which the satellite's progenitor identifies the $z = 0$ host's progenitor as its host. A host-satellite pair may be identified as soon as the satellite is within 2.5 times the virial radius of the host; though other criteria must still be met, as outlined in Behroozi et al. (2013c), in practice these are usually met immediately when the satellite crosses within $2.5r_{\rm vir}$. Only the orbital history with respect to the final ($z = 0$) host is considered, and in cases where a satellite has a

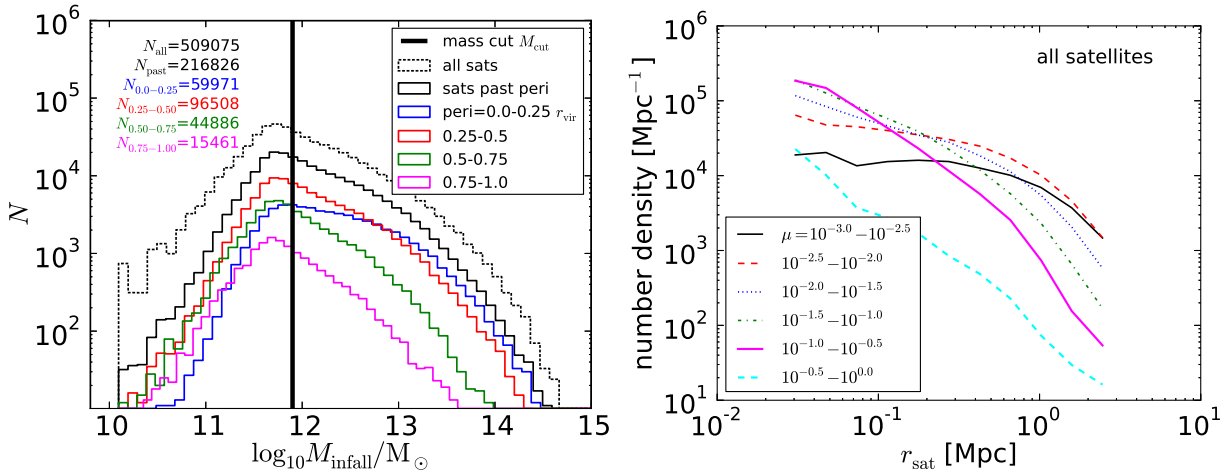

Figure 3.1: These two panels provide a measure of the importance of resolution effects in our analysis; see §3.1 and §3.6 for discussion. *Left panel:* The mass function of all satellite haloes in our initial sample (dotted black line) and the mass function of satellite haloes that have experienced at least one pericentric passage (solid black line), where in both cases the masses are measured at the time of infall on to their host. The coloured lines separate the satellites that have experienced a pericentric passage into bins of pericentric distance, as labelled. The total number of satellites contributing to each mass function are labelled N. We impose a halo mass cut as indicated by the thick vertical line ($M_{\rm cut} = 10^{11.9}$ M$_\odot$), yielding a sample which is as complete as reasonably possible above $M_{\rm cut}$. The relative deficit of haloes with masses $\lesssim 10^{13}$ M$_\odot$amongst those that experience the closest pericentric passages (blue curve) is due to artificial disruption of haloes in the simulation. *Right panel:* Number density as a function of satellite radial position for various mass ratios μ, where the mass of the satellite is measured at the time of infall. The host haloes have masses $\geq 10^{14}$ M$_\odot$. Satellites with low mass ratios are underabundant toward the centre of the cluster environment due to artificial disruption of haloes in the simulation.

hierarchy of hosts, intermediate hosts are ignored and the satellite identifies the largest as its host; in other words, the present work ignores "group preprocessing". Hosts are selected to be "cluster-sized", which we define as a halo mass $> 10^{14}$ $M_\odot$. With the MDR1 dataset, this yields a catalogue of $1,053,807$ satellite orbits belonging to $35,838$ different hosts.

To ensure a sample complete in satellite mass, and that we are minimally sensitive to artificial disruption of satellite haloes, we impose a mass cut at $M_{\rm cut} = 10^{11.9}$ $M_\odot$, where the mass is measured *at the time of infall*. Using the stellar-to-halo mass ratios outlined in Behroozi et al. (2013a), this corresponds to a cut in stellar mass at $\sim 10^{10.3}$ $M_\odot$. The left panel of Fig. 3.1 shows the mass function (Eq. 1.4) for all satellites, and for satellites that have experienced at least one pericentric passage in bins of proximity of pericentric passage. Our mass cut is safely above the mass resolution limit of the simulation. As in essentially all cosmological simulations, MDR1 haloes experience artifical disruption in high density environments (Kitzbichler & White, 2008; Klypin et al., 1999). This is the reason for the underabundance of lower mass satellites that have experienced a close approach to a cluster centre (blue curve in Fig. 3.1, left panel). Comparing the integrals of these mass functions, we estimate that less than 20 per cent of haloes with masses above our mass cut and pericentric distances in the range $0.0 - 0.25r_{\rm vir}$ have been artificially disrupted. Most of the artificially disrupted satellites are in the mass range $10^{11.9} - 10^{13.1}$ $M_\odot$. We estimate that these missing satellites account for less than 4 per cent of the total halo population above our mass cut. The right panel of Fig. 3.1 shows the number density of satellite haloes as a function of radial position for bins of satellite mass relative to the host halo. Observations, along with the assumption that the luminosity of a galaxy hosted by a halo is correlated to the maximum mass at any time of the halo, constrain the slope of this power law relation to be between -1.7 and -1.5, regardless of mass ratio (Behroozi et al., 2013b; Tinker et al., 2012). The slopes shown in Fig 3.1 are somewhat steeper at about -1.9 for mass ratio $\mu = 10^{-0.5} - 10^{0}$. The reason for this steeper slope is not precisely known, but the higher resolution Bolshoi simulation exhibits the same slope of -1.9 so we surmise that it is not due to a resolution effect. The key feature that we wish to highlight is that the slope (and shape) of this relationship is mass ratio *dependent* in our dataset, with satellites that are smaller relative to their host being less abundant (due to resolution effects) closer to the centre of the host. The underabundances of satellites highlighted by each panel of Fig. 3.1 are due to the same population of satellites; those which have orbited to within $\lesssim 0.25r_{\rm vir}$ of their host and have a mass $\lesssim 10^{13.1}$ $M_\odot$. See §3.6 for a discussion of the impact of these missing satellites on our results.

We impose one final cut, removing satellites that existed for less than 3 simulation snapshots before falling into a cluster. This prevents haloes near the mass resolution limit of the simulation from appearing suddenly inside a cluster and being assigned meaningless

infall times. The remaining 466,556 satellites were binned in 100 projected position bins in $0.0 \leq R/r_{\rm vir} \leq 2.5$ and 100 projected velocity bins in $0.0 \leq V/\sigma_{\rm 3D} \leq 2.0$. The set of infall times in each bin was used to create a PDF of infall times for each bin. Separate collections of PDFs were created for different bins of satellite mass, we elaborate on these bins in §4.3.

3.1.1 Interlopers

Our interloper sample, initially containing 1,094,552 haloes, is cut with the same lower mass limit as our satellite sample, leaving 444,182 haloes. These are binned in the same way as the satellites and added to the PDFs, tagged with a nonsensical infall time to distinguish their contribution from that of satellite haloes.

3.2 Projection

3.2.1 Radial Projection

While a simulation provides accurate values for all six phase space coordinates of an object, a typical astronomical observation can only measure three. The RA and Dec give two spatial coordinates (the distance is unknown), while comparing the spectra of two objects can give the difference in velocity between them, but only for motion in a direction along the LoS. Since we ultimately wish to infer properties of observed objects from their coordinates, we must restrict our knowledge of simulated objects to these same coordinates. This can be achieved by ignoring one of the spatial coordinates of the simulation box – in our case, the third – and considering only the velocity coordinate corresponding to the ignored spatial coordinate. Additionally, we include a correction to the projected velocity to account for the Hubble flow. This transformation applied to the spatial coordinates of a point $\mathbf{x}$ and its velocity $\mathbf{u}$ relative to a reference point $\mathbf{y}$ and its velocity $\mathbf{w}$ can be expressed:

$$\begin{aligned} \mathbf{r} = (\mathbf{x} - \mathbf{y}) &= (x_1 - y_1, x_2 - y_2, x_3 - y_3) \\ \Rightarrow \mathbf{R} &= (x_1 - y_1, x_2 - y_2) \end{aligned} \tag{3.1}$$

$$\begin{aligned} \mathbf{v} = (\mathbf{u} - \mathbf{w}) &= (u_1 - w_1, u_2 - w_2, u_3 - w_3) \\ \Rightarrow \mathbf{V} &= (|(u_3 - w_3) + H(x_3 - y_3)|) \end{aligned} \tag{3.2}$$

We predict the effect of projection on the radial coordinate by considering a random uniform distribution of points on a spherical shell. For this distribution the relationship between the actual (r) and projected (R) radial coordinates is characterized by

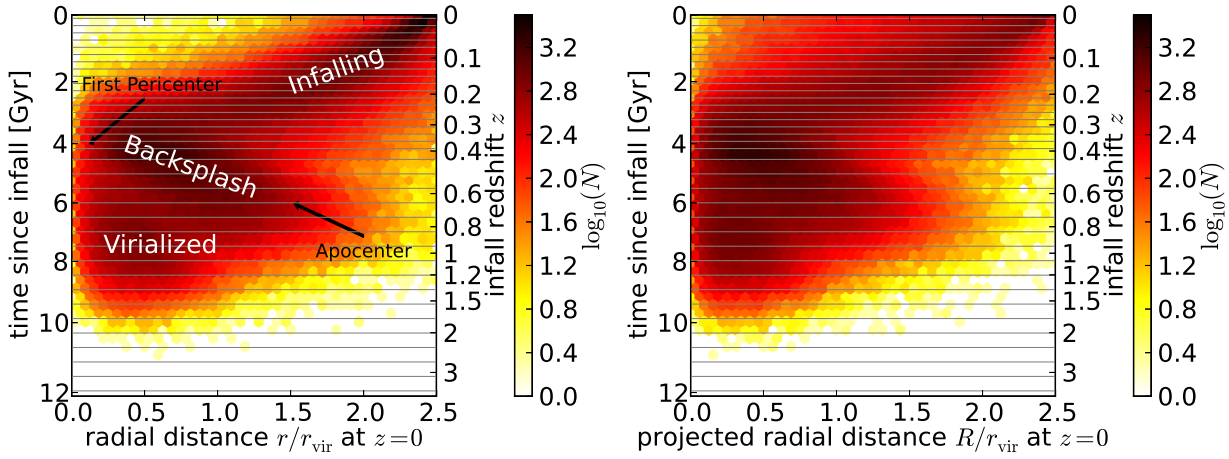

Figure 3.2: *Left panel:* Abundance of satellites in bins of cluster infall time, defined as first inward crossing of 2.5 $r_{\rm vir}$, and radial distance from cluster centre at $t = 0$ Gyr. In the interval $t = 0$ to $t \sim 3.5$ Gyr the 'infalling' population is visible, composed of satellites that have not yet experienced a pericentre passage. From $t \sim 3.5$ Gyr to $t \sim 6$ Gyr the 'backsplash' population is visible, composed of satellites that have passed pericentre once and are approaching apocentre ($t \sim 6$ Gyr). The 'virialized' population ($t > 6$ Gyr) are on a second or subsequent orbit. Gray horizontal lines illustrate the times of simulation snapshots. *Right panel:* same as left panel, but using radial distance from cluster centre projected along one axis of the simulation box (simulating coordinates accessible in observations). This widens the distribution of radii of satellites at a given infall time, but the same populations as in the left panel are visible.

$\langle \frac{R}{r} \rangle = \frac{\pi}{4} \pm \sqrt{\frac{2}{3} - \frac{\pi^2}{16}} \sim 0.79 \pm 0.22$ (1σ scatter). In our sample of satellite galaxies, we find that $\langle \frac{R}{r} \rangle$ agrees with this prediction to within 1 per cent, both in the mean and in the scatter. The projected radial coordinate tracks the 3D radial coordinate more closely and more consistently at larger projected radii; at $R \sim r_{\rm vir}$, $\langle \frac{R}{r} \rangle \sim 0.83 \pm 0.17$ while at $R \sim 0.1 r_{\rm vir}$, $\langle \frac{R}{r} \rangle \sim 0.51 \pm 0.31$. The observed velocity coordinate does not have as straightforward a relationship with the actual quantity of interest – the radial component of the velocity difference between two points – but the information about two components is lost instead of one, so we expect a much larger typical difference between projected velocity and true velocity. We also lose the sign of the one remaining component of velocity since, without precise knowledge of the distances to the two points, we cannot know whether the distance between them is increasing or decreasing.

The left panel of Fig. 3.2 shows the distribution of satellite infall times[1] as a function of radial distance from the host centre *measured at* $z = 0$ before projection. Satellites that have recent infall times (near the top of the diagram) are necessarily concentrated near the edge of our definition of the cluster at 2.5 $r_{\rm vir}$; they have not had enough time to move anywhere else. A typical crossing time for our sample of clusters is about 6-8 Gyr, so most satellites that fell in 3-4 Gyr ago are near the centre of the host. Note that this seemingly long crossing time is due to our definition of the edge of the cluster at 2.5 $r_{\rm vir}$; the typical time to cross from $r_{\rm vir}$ to pericentre and back to $r_{\rm vir}$ is about 2 Gyr. The satellite also spends a significant amount of time outside the virial radius after this initial crossing; it takes a further $\sim$1 Gyr for a typical satellite to reach apocentre after making its first outbound crossing of $r_{\rm vir}$.

There is a spread in the time taken to reach pericentre and the radii of the pericentres caused by the variety of possible orbits and details of the host potentials. The first apocentre after infall typically occurs after about 6 Gyr; the population of objects between first pericentre and apocentre is termed 'backsplash'. Satellites with infall times earlier than ~ 7 Gyr have less distinct features in their distribution due to the increasing impact of variations in orbital history, but the majority are confined within ~ 1 $r_{\rm vir}$; we call this population 'virialized'.

The right panel of Fig. 3.2 shows the effect of projection in the radial coordinates on the same distribution as in the left panel. Features are shifted somewhat to lower radii (consistently with our expectation of $\langle \frac{R}{r} \rangle = \frac{\pi}{4}$) and broadened by the scatter about this mean deformation. All the populations and features discussed above are still identifiable. One notable omission from this diagram is any foreground/background objects, or interlopers, which are common in observational samples, that could be confused with the satellite populations. These objects are further discussed in §3.3.

There is one other important effect to consider when interpreting Fig. 3.2 (and others involving infall times). As a host halo accretes mass, its virial radius grows (slowly, except in the case of major mergers). Because of this, an orbiting satellite may appear to move further in coordinates scaled to $r_{\rm vir}$ than it otherwise would as the coordinates grow around it. This does not have a large impact on the positions, but contributes some of the scatter in the radial coordinate. One might consider choosing some radius that is constant with time to scale each halo, but other choices, such as the virial radius at the last snapshot, also introduce similar effects.

[1]The infall times produced by the method of §3.1 occur at discrete times – those of the simulation snapshots. For the purposes of visualization only, some uniform random scatter was added to the times, redistributing each discrete time t_d over the interval $\left[\frac{1}{2}(t_{d-1} + t_d), \frac{1}{2}(t_d + t_{d+1})\right]$.

3.2.2 Velocity Projection

Next, we consider the effect of projection on the velocity coordinates. The upper left panel of Fig. 3.3 shows the distribution of satellite haloes in phase space at $z = 0$. A typical halo would, given enough time, progress from large radii and low velocities down to low radii and high negative velocities, then switch to high positive velocity as it passes pericentre. From there it follows a series of progressively shrinking concentric semicircles or chevrons, switching from negative to positive velocity at each pericentric passage (for a more in depth theoretical background, see Bertschinger, 1985, especially fig. 6 therein). This normal progression is well represented by the distribution of our halo sample, however the individual orbital 'shells' are not visible since they overlap, and we only expect 1-2 shells given the orbital timescales and ages of these systems. Also, close encounters redistribute some haloes off this idealized track.

Based on this expected movement through phase space in time, satellite haloes with different infall times should occupy different regions of phase space. This is shown in Fig. 3.3, where the phase space distribution of haloes is plotted for a variety of bins in infall time. Fig. 3.4 shows the same binned distributions, but in projected coordinates, simultaneously showing the effects of both radial and velocity projections, but emphasizing the latter. Much of the structure visible in Fig. 3.3 is lost in projection, but haloes with different infall times still occupy different regions of phase space.

3.3 Interloper Projected Phase Space Distribution

In addition to the actual cluster satellite haloes and the galaxies they host, an observed sample of objects near a cluster will inevitably include interlopers. The selection of our simulated interloper sample is described in §§2.1.1, 3.1.1. The number of interlopers increases with projected radius, primarily because there is more volume at larger projected radii. It also increases with decreasing LoS velocity; haloes that will eventually fall into the cluster contribute to this distribution at low LoS velocities. The left panel of Fig. 3.5 shows the distribution of our interloper sample in projected phase space. The right panel shows the combination of the satellite and interloper samples.

Both the interloper and the satellite distributions occupy essentially the entire projected phase space plane, but peak in distinct regions. Fig. 3.6 shows the interloper fraction as a function of projected phase space position. The upper right region of the phase space plane is dominated by interlopers while the lower left region contains mostly cluster satellites.

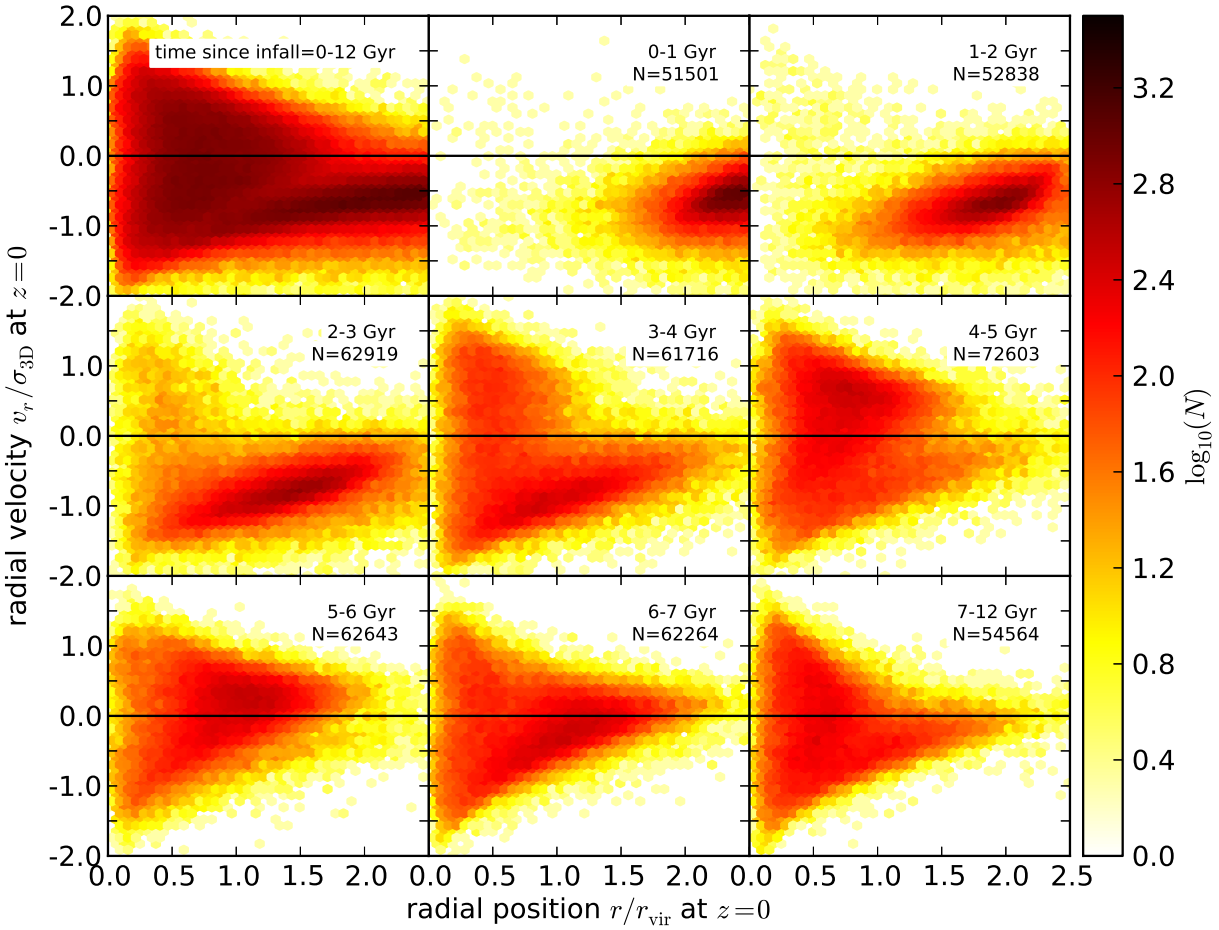

Figure 3.3: The upper left panel shows the phase space distribution of satellite haloes (no projection). The infalling population of haloes is particularly distinct, forming the long dark bar with $v < 0$. The backsplash population forms the upper and lower edges of the rest of the main distribution, while the virialized population fills in the centre. The other panels correspond to bins by satellite infall time, as labelled (most bins are 1 Gyr). Each panel shows the $z = 0$ distribution of satellites in phase space. Different satellite populations occupy distinct regions of phase space; compare for instance the upper right panel showing mostly infalling satellites, the centre right panel showing mostly backsplash satellites and the lower right panel showing mostly virialized satellites. Each bin is also labelled with the number of satellites N contained in the bin.

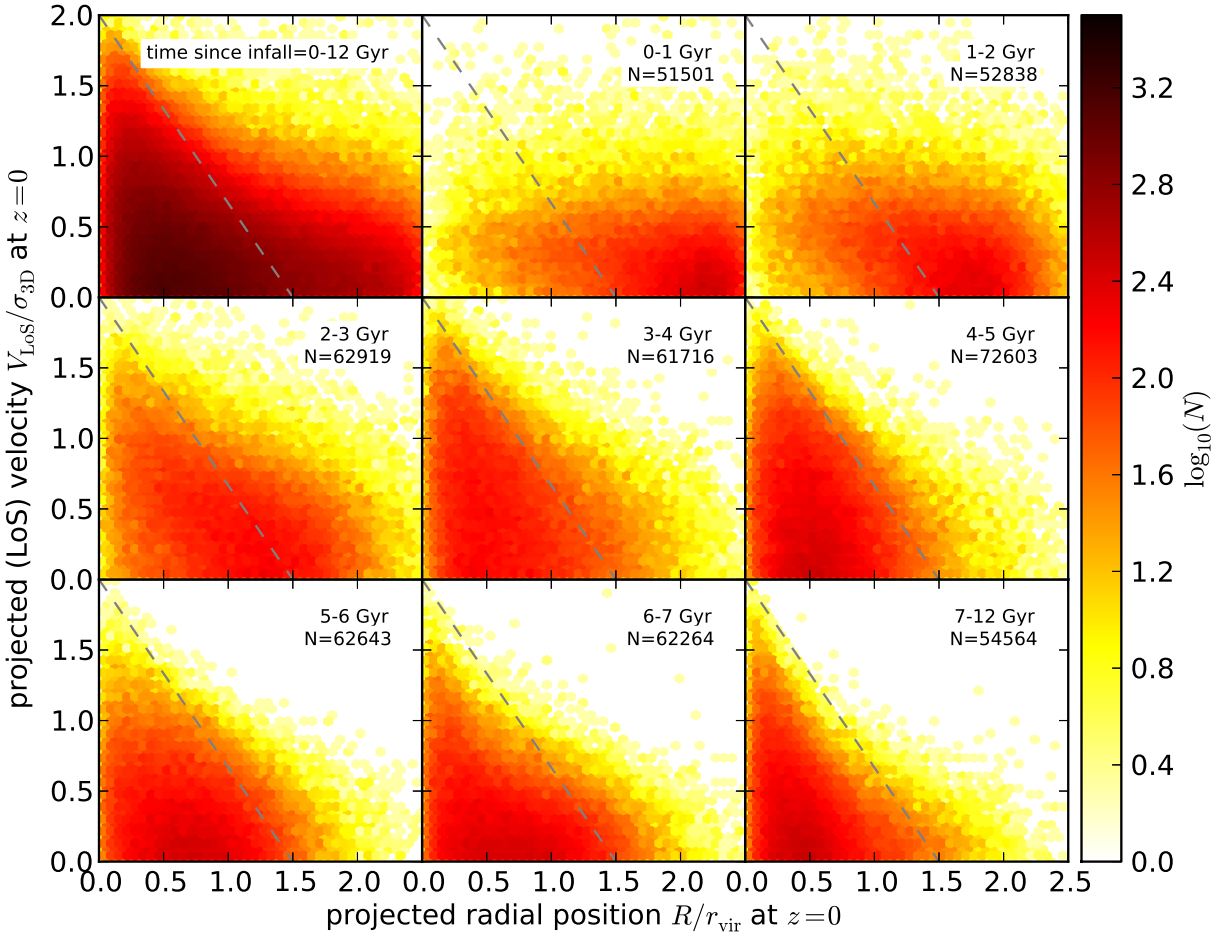

Figure 3.4: As in Fig. 3.3, but using radial distance from cluster centre projected along one axis of the simulation box and velocity projected along the same axis (again simulating coordinates of an observed system). Much of the structure visible in Fig. 3.3 is lost in projection, but the use of LoS velocity data allows better separation of different satellite populations than using the radial coordinate alone. The dashed gray line marks $\frac{V}{\sigma_{\rm 3D}} = -\frac{4}{3}\frac{R}{r_{\rm vir}} + 2$; see §3.4.

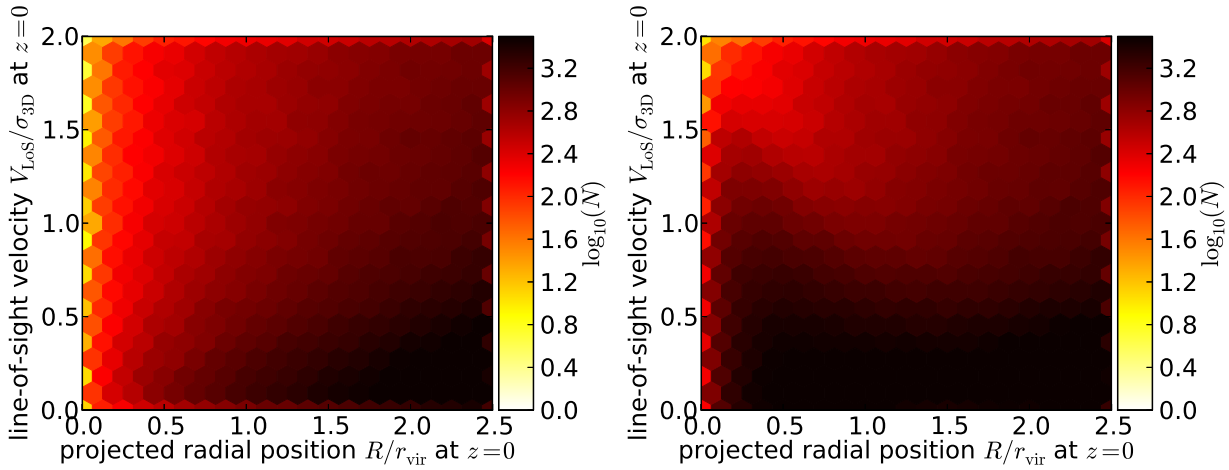

Figure 3.5: *Left panel:* Phase space distribution of haloes outside a sphere of radius $2.5r_{\rm vir}$ but with projected coordinates $R < 2.5r_{\rm vir}$ and $V < 2.0\sigma_{\rm 3D}$. These "interloper" haloes contaminate the cluster satellite halo population. *Right panel:* Combined phase space distribution of cluster satellite haloes and interloper haloes.

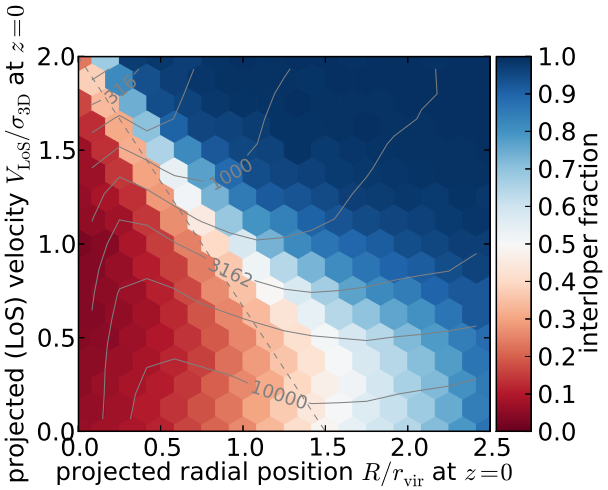

Figure 3.6: Interloper fraction as a function of projected phase space position. Blue regions are dominated by interlopers, red regions by satellites, and white regions contain an even mixture of both populations. Contours show the number of haloes in each bin. The gray dashed line identifies the same region as in Fig. 3.4. The population of haloes with large R and large V is dominated by interlopers, while the populations with low R and low V is dominated by satellites.

3.4 Infall Time PDFs

Since our ultimate goal is to model quenching, perhaps the most directly useful question we can ask of our dataset is, given a position in projected phase space, what is the distribution of possible infall times (or another parameter of interest), and what is the likelihood of each? This question is easily answered in a statistical manner by sampling all satellites in a small region of phase space and binning them by infall time. The result is shown in Fig. 3.7 for a selection of points in phase space, varying both the radial coordinate (left-right across the panels) and the velocity coordinate (up-down). We focus here on the trends with the velocity coordinate. The first trend we note is that at a given radius, satellites with higher $v_{\rm LoS}$ typically have slightly more recent infall times than their low $v_{\rm LoS}$ counterparts. This is because a satellite with high $v_{\rm LoS}$ is more likely to have a high total speed and can penetrate deeper into the host potential in a given amount of time than a satellite with low speed.

In some cases, the velocity coordinate allows us to discriminate between different satellite populations. Consider the rightmost panels ($R = 1.5$) of Fig. 3.7. One peak in the distribution of infall times is clearly visible at all values of $v_{\rm LoS}$, at $t \approx 2$ Gyr; it is made up of infalling satellites. A second peak at $t \approx 6$ Gyr is made up of backsplash satellites, and is only present for low values of $v_{\rm LoS}$. The backsplash galaxies have lower kinetic energy relative to the host potential than recent infalls due to a combination of mass accretion by the host (increasing σ, therefore causing an apparent slowing of haloes) and dynamical friction.

Another visualization of the PDFs is shown in Fig. 3.8. The velocity distribution of satellites with various infall times around a few fixed radii are plotted. Populations with more recent infall times have a broader distribution of velocities. They are also generally dominant over populations with earlier infall times at larger projected radii.

Given a sufficiently large sample of observed galaxies, the PDFs we have created will allow us to statistically assign an infall time to each one using all the available dynamical information and a meaningful uncertainty in our assignment.

As a test of the reliability of our PDFs, we use them to estimate the infall time of each simulated satellite halo in our sample. This estimate is then compared to the actual infall time of the satellite, which we know from tracking its orbit. This is roughly equivalent to measuring the width of the PDF corresponding to each region of phase space. Interlopers are not included in this analysis. The distribution of $\Delta t = t_{\rm infall}({\rm actual}) - t_{\rm infall}({\rm estimated})$ is plotted in Fig. 3.9. A Δt of 0 represents a correct guess, so a stronger peak about $\Delta t = 0$ represents a higher rate of success. As a quantitative measure of the strength of

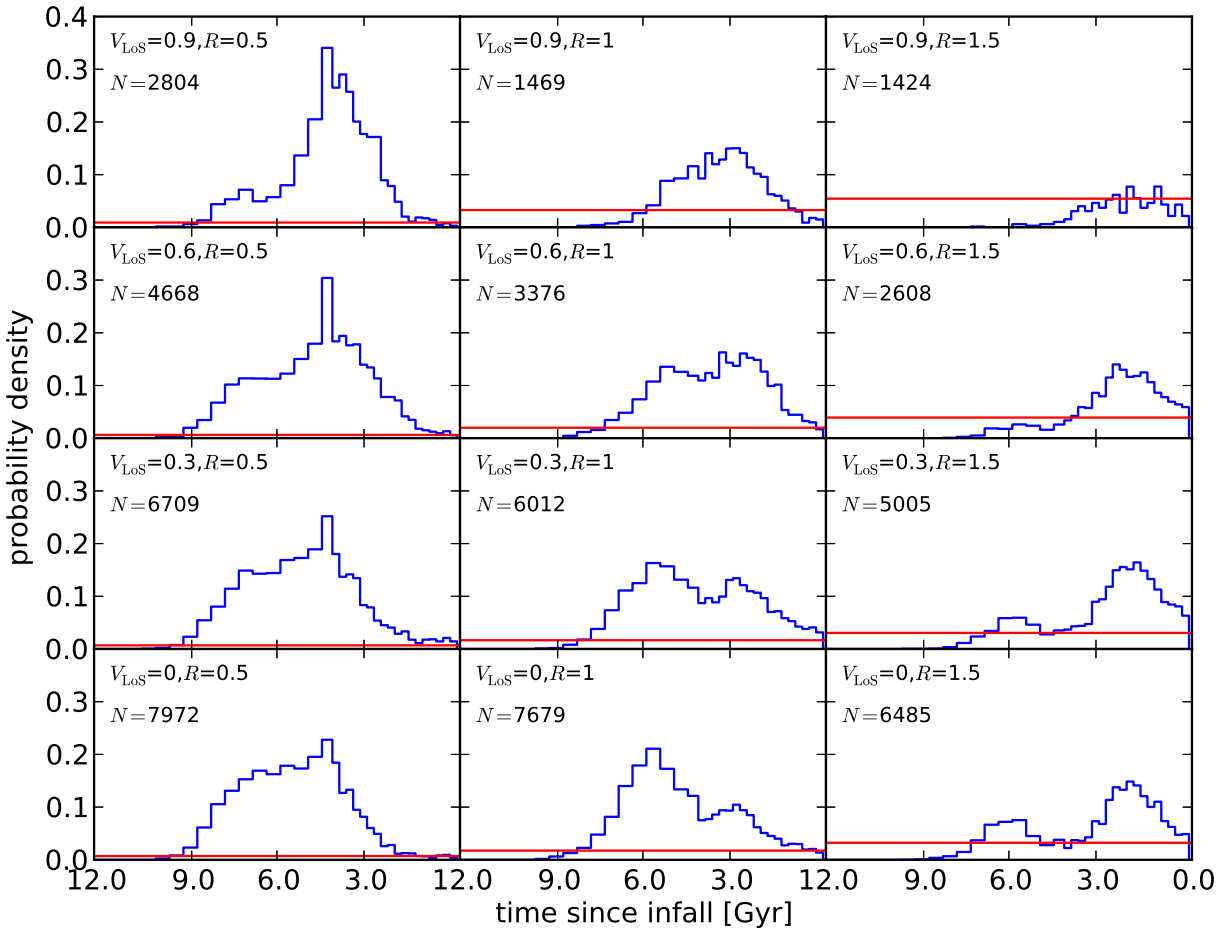

Figure 3.7: PDFs of infall times for a selection of regions $(R/r_{\rm vir}, V/\sigma_{\rm 3D})$ in projected phase space. Each region is sampled in a small area around the labelled point. The number of satellites in each region is labelled N, giving a measure of the statistics of each PDF. The horizontal red line provides a measure of the number of interlopers in each region of phase space. The axes have no meaning relating to the interlopers, rather the height of the line is fixed so that the ratio of the integrals of the red and blue curves is proportional to the ratio of interlopers to satellites in each region.

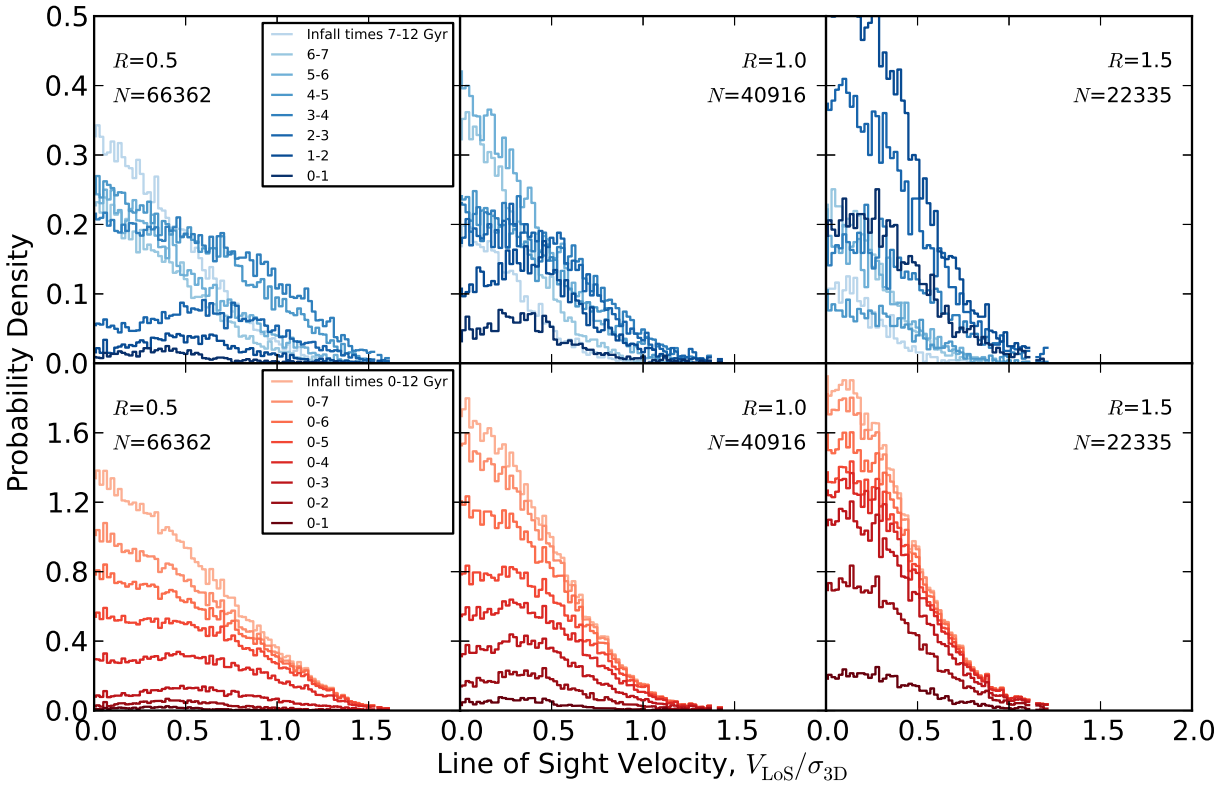

Figure 3.8: PDFs of LoS velocities for 3 radii (each radius is sampled in a small region around the labelled point). The number of satellites used to construct each panel is labelled N. The first row shows PDFs for consecutive bins in infall time (lookback time to infall), while the second row shows cumulative bins in infall time.

the peak, we calculate the standard deviation of the distribution. By using our PDFs based on knowledge of the (R, V) coordinates of each satellite, we assign the infall time correctly to within ± 2.58 Gyr in 68 per cent of cases. For comparison, we also plot the distribution of Δt obtained if only the position coordinate R of the haloes is known, accurate to within ± 2.64 Gyr in 68 per cent of cases, and if none of the coordinates are known; in this case, we simply draw values from the distribution of infall times of our entire sample at random, and find we are accurate within ± 3.10 Gyr in 68 per cent of cases. We also compare to the case where 3D radius r and the radial velocity v_r of the satellite are known (these are typically not observable quantities in extragalactic systems). In this case the prediction is understandably better than when projected coordinates are used, with an accuracy of ± 1.94 Gyr in 68 per cent of cases. Finally, we plot a curve showing the distribution of Δt for a particular subsample of satellites chosen by eye which have $\frac{V}{\sigma_{3\mathrm{D}}} > -\frac{4}{3}\frac{R}{r_{\mathrm{vir}}} + 2$. This subsample is designed to retain primarily infalling satellites (see Fig. 3.4), and in this case the assignment of infall times is even more reliable. This demonstrates that if a particular population of satellites is of interest, it is often possible to choose a region of phase space that maximises the likelihood of correctly assigning t_{infall}. For this example, we correctly assign t_{infall} within ± 2.48 Gyr in 68 per cent of cases. In practice, this makes it possible to increase the purity of an observational sample of galaxies at the expense of completeness of the sample by limiting the region of phase space from which the sample is drawn.

Fig. 3.10 shows the accuracy of the assignment of t_{infall} at the 68 per cent confidence level for bins in (R, V) space. The PDFs give the most reliable results in those regions with both good accuracy and a large population of satellites. In the outskirts of clusters, an accuracy Δt of about 3 Gyr is sufficient to robustly separate infalling and backsplash satellites.

We conclude that using all observable coordinates of a satellite (R, V) offers an improvement in assigning t_{infall} correctly over using only the position coordinate R comparable to, but somewhat less than, the improvement seen when using the position coordinate R rather than no knowledge of satellite coordinates at all. Careful selection of the region of phase space to be sampled can, in many cases, further increase the reliability in estimating t_{infall}.

The above estimations of infall time assume that the cluster satellite population can be separated from the interloper population before any predictions are made. In practice, in regions dominated by interlopers, it is essentially impossible to pick out satellites based on their phase space coordinates alone (though a selection on some other physical property of the galaxy hosted by the satellite halo may be of use). In regions with a more even mixture of satellites and interlopers (suppose there is an interloper fraction f_i), an infall time prediction has the caveat that either the infall time of the halo is drawn from the distribution defined by the PDF at that location in phase space with probability $(1 - f_i)$, or

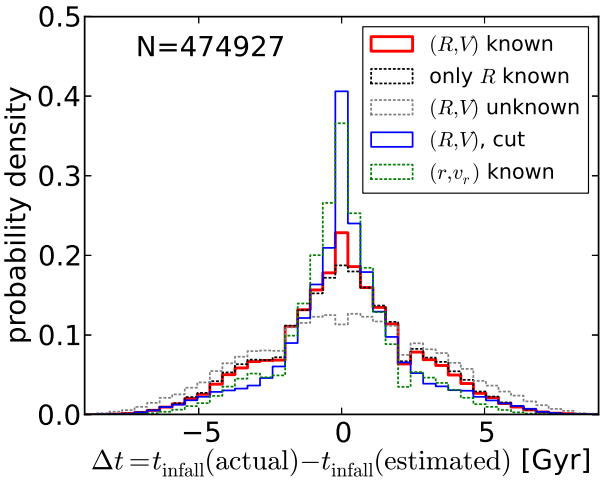

Figure 3.9: The distribution of Δt, a measure of the rate of success of using our PDFs to estimate the infall time $t_{\rm infall}$ of satellite haloes in our sample, is plotted with a thick red solid line. $\Delta t = 0$ represents a successful estimation, while increasing $|\Delta t|$ represents increasing an increasing discrepancy with the correct value. A sharp peak at $\Delta t = 0$ therefore represents a high rate of success. The black dotted line shows the distribution if we assume no knowledge of the LoS velocity of the satellites, while the gray dotted line shows the distribution if we assume no knowledge whatsoever of the coordinates of the satellites. Each additional coordinate used increases the reliability of the predictions. For comparison, we also show the distribution for a scenario where both the 3D radius r and the radial velocity v_r are known (green dotted line). The thin solid blue curve shows the distribution of Δt for a subsample of satellites selected to study a particular population of satellites (see text for full details). This shows an example of restricting the region of phase space considered to increase the reliability in estimating $t_{\rm infall}$. We quantify the degree of success in estimating $t_{\rm infall}$ by the sharpness of the peak in the distribution of Δt, measured by its standard deviation. The standard deviations for the five curves are 2.58 Gyr, 2.64 Gyr, 3.10 Gyr, 2.48 Gyr, 1.94 Gyr (in the same order as labels in the legend). The number satellites used to test the PDFs is labelled N.

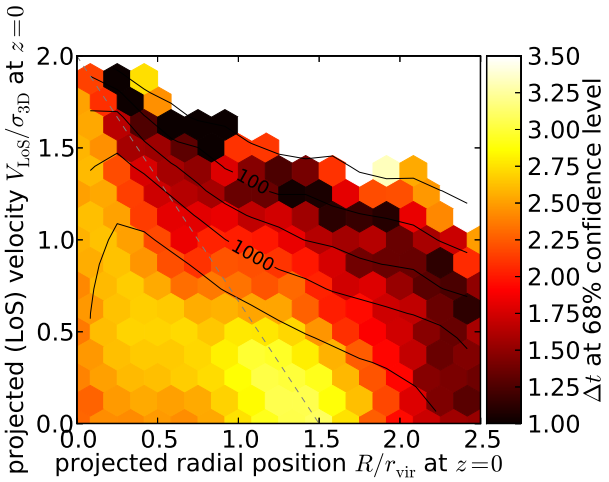

Figure 3.10: The distribution of Δt at the 68 per cent confidence level in the (R, V) plane; our PDFs will be most accurate when applied to satellites that occupy the reddest (darkest) regions of the space. Use of the PDFs also requires a reasonable sample size in the region of interest; contours (0.5 dex intervals) illustrate the number of satellites per cell. We find that the PDFs become too sparse for reliable use in cells containing less than a few hundred satellites. The gray dashed line identifies the same region as in Fig. 3.4.

it is an interloper with probability f_i. This information is still useful, however, as we expect the interloper population to have properties generally similar to recent infalls, and it is no surprise that these two populations overlap in their projected phase space distributions as we have only separated them by a line drawn at an arbitrary radius. The interlopers are also not especially problematic for the virialized population since these satellites preferentially lie in the region of projected phase space with the lowest interloper density. The backsplash population, however, overlaps substantially either with the virialized population or with the interlopers depending on which region of projected phase space is considered. This makes it the most difficult to separate out which is unfortunate since it is one of the most interesting populations in the context of star formation quenching by the cluster environment.

3.5 Mass Trends

While haloes of different masses are expected to be self-similar in most ways, we still find some trends with both host halo mass and satellite halo mass. We study these trends by separating the haloes of interest – either hosts or satellites – crudely into a 'high mass' and a 'low mass' bin, then compare the distribution of satellites in phase space and infall time for the two bins.

Fig. 3.11 shows a comparison of satellites around high- and low-mass hosts in infall time vs. radial position space (left panel), and in phase space (right panel). Low mass hosts are defined to have 10^{14} $M_\odot < M < 10^{14.5}$ $M_\odot$while high mass hosts are in the range $10^{14.5}$ $M_\odot < M < 10^{15}$ $M_\odot$. The population of satellites around each type of host is normalized, then the two are subtracted in each infall time - radial distance bin and compared to the fiducial abundance of satellites in that bin. Cells with a reddish colouration are preferentially occupied by satellites around high mass hosts, while cells with a bluish colouration are preferentially occupied by satellites around low mass hosts.

In the infalling population of satellites ($t < 4$ Gyr in Fig. 3.11, left panel) the satellites of high mass hosts appear to fall into their host somewhat faster than the satellites of low mass hosts; at a given time since infall, a satellite of a low mass host has a larger typical radial distance than a satellite of a low mass host. We propose two possible explanations for this effect. First, satellites of high mass hosts tend to have more radial orbits than satellites of low mass hosts (Wetzel, 2011), causing them to move radially inward more rapidly on average. Alternately, because the hosts are continually accreting mass, their virial radius gradually increases. Hosts with higher masses have higher present-day accretion rates; this is visible in Fig. 3.11 (right panel) where the region of phase space occupied by infalling

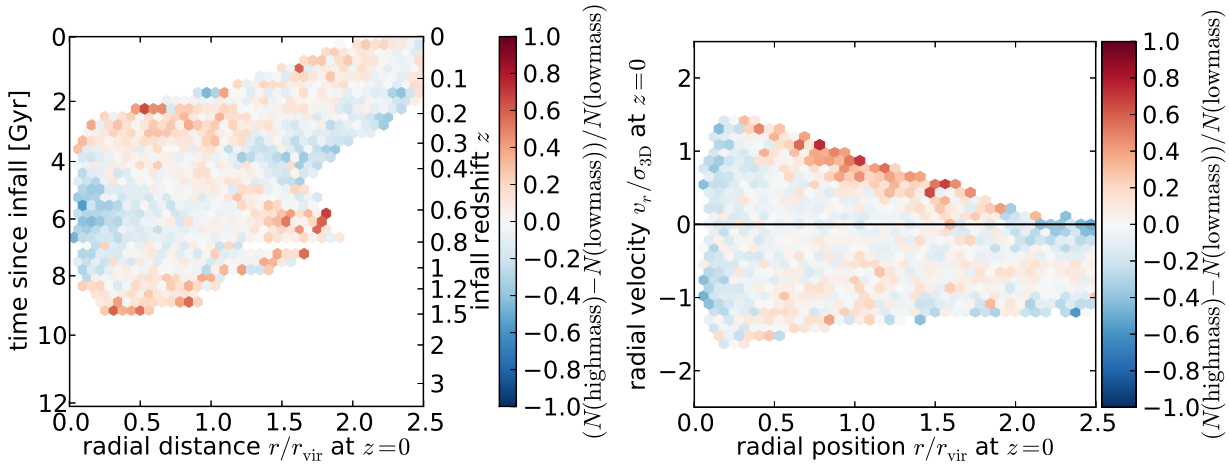

Figure 3.11: Relative abundance of satellites in the space of infall times vs. radial positions (left panel) and phase space (right panel), comparing satellites around low mass ($10^{14} - 10^{14.5}$ $\rm M_\odot$) hosts and high mass ($10^{14.5} - 10^{15}$ $\rm M_\odot$) hosts. The two populations are normalized, subtracted in each bin, then compared to the fiducial abundance of satellites in that bin (arbitrarily chosen to be the population with low mass hosts). Only bins containing at least 100 satellites are shown.

satellites shows an excess of satellites around higher mass haloes (see also Wechsler et al., 2002, fig. 3 therein especially). Their satellites therefore appear to move inward slightly faster than those of low mass hosts in this coordinate system. These two explanations could be considered an argument for a different normalization of the radial coordinate, however other choices (e.g. the virial radius of the host at the infall time of each satellite) still introduce similar effects, and the present virial radius of the host has the advantage of being more closely related to observable quantities. A more practical approach is to compute PDFs for narrow host mass bins (~ 0.5 dex) and use whichever is applicable to an observed system of interest; our sample contains enough satellites to make this feasible.

We also consider trends with satellite mass. Because satellite mass typically decreases with increasing time spent in a cluster environment, we use the satellite mass at time of infall (first crossing of 2.5 $r_{\rm vir}$) as a characteristic mass for this analysis. Fig. 3.12 shows that lower mass satellites experience backsplashes to larger radii than high mass satellites. This is because the slowing due to dynamical friction is proportional to the mass of the satellite, so that higher mass satellites lose a larger fraction of their kinetic energy during a passage through a cluster (see, for instance, Tormen et al., 1998). To account for this trend with satellite mass, separate PDFs can be produced for relatively narrow bins in

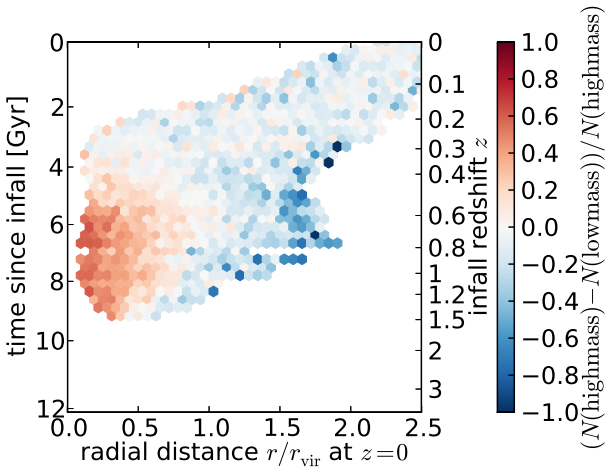

Figure 3.12: Relative abundance of satellites in the space of infall times vs. radial positions comparing low mass ($10^{11.9} - 10^{12.2}$ $M_\odot$) and high mass ($10^{12.2} - 10^{12.8} h^{-1}$ $M_\odot$) satellites. The two populations are normalized, subtracted in each bin, then compared to the fiducial abundance of satellites in that bin (arbitrarily chosen to be the high mass population). Only bins containing at least 100 satellites are shown. High mass satellites are slowed more than low mass satellites by dynamical friction and mass accretion by their host, and therefore have lower typical backsplash distances than their low mass counterparts.

satellite mass.

Care must be taken when using our method to interpret the distribution of galaxies in a cluster. Baryons (especially the stellar component) are typically more tightly bound than their associated DM halo, and should therefore outlive the halo in the cluster environment. This restricts the satellites, as traced by galaxies, that can be studied using our method to those which we expect to have a surviving DM halo. With the MDR1 dataset, this corresponds to galaxies with associated DM halo masses of $\sim 10^{11.9}$ $M_\odot$or greater. Smaller satellites may still be studied by using higher resolution simulations (e.g. the Bolshoi simulation of Klypin et al. (2011), which has identical parameters to MDR1 but with higher resolution and a smaller box size, see Appendix A). Low mass satellites are much more abundant, so the smaller volume of a higher resolution simulation should not impact our ability to obtain a large sample, except perhaps for low mass satellites around very large hosts (of which there may only be a few in a simulation such as Bolshoi).

3.6 Numerical Resolution Effects

As shown in section 3.1, some haloes which experience a close approach to the centre of a larger halo experience artificial disruption. We estimate that any given bin in our PDFs is less than 20 per cent incomplete, and that our sample of satellites as a whole is less than 4 per cent incomplete down to $10^{11.9}$ $M_\odot$. The effect on the predictive power of our method was investigated by generating a test catalogue of orbits with extra entries added to compensate for artificially disrupted haloes. The number of artificially disrupted satellites as a function of satellite mass is evident from the mass functions of satellites binned by pericentric distance (Fig. 3.1 left panel). Randomly selected undisrupted satellites with masses and trajectories matching those expected for the artificially disrupted population were duplicated to create the test catalogue. This revealed that the effect of artificial disruption on the distributions of Δt (such as those shown in Fig. 3.9) is very small, causing changes on the order of a few tenths of a per cent.

Our initial interest for applications of our method is to large satellites (halo masses $\gtrsim 10^{12}$ $M_\odot$) around hosts with halo masses $\gtrsim 10^{14}$ $M_\odot$, which motivated our choice of the MDR1 simulation as a starting point. Applying our method to smaller satellites of smaller hosts is simply a matter of choosing an appropriate simulation; our method is easily applied to any simulation that can be processed by the ROCKSTAR code. The Bolshoi simulation is of particular interest since it is identical to MDR1 in all respects save resolution and could therefore be used in conjunction with MDR1 to extend the range of our method to lower mass ranges while maintaining good statistics at the high mass end. At present we have only preliminary results from the Bolshoi output; these are presented in Appendix A.

Chapter 4

Results: Modelling Colour Distributions

We are now armed with knowledge of the orbital history of satellites, parameterized by the time since their infall into their host cluster, in the form of PDFs. Next, we use this information to extract the time evolution of the colour distribution of galaxies in clusters in a large observed sample. We hope to eventually relate this colour evolution to the SFH of the galaxies, and then back to the mechanism(s) which governed the SFH.

First, we ensure that the projected phase space distribution of our observed sample corresponds to that of our simulation dataset, i.e. that we are comparing "apples to apples", in §4.1. §4.2 presents the results of the analysis of a very simple toy model that was used to guide the formulation of the method presented in §4.3. This method simultaneously fits a model for the colour distribution to the observed data in a sequence of bins of infall time, effectively deconvolving the colour distribution and the infall time distribution of the satellites. This method is still being tested and tuned, but in §4.2, we present a second, greatly simplified (at the cost of accuracy) method to explore the evolution of the colour distribution, and the results of this rough estimation.

4.1 Compatibility of Observed and Simulated Data

Before applying our infall time PDFs to the analysis of our observed sample of cluster satellites (recall §2.2 describes the dataset), it is important to verify that our PDFs are in fact applicable to this dataset. Our assumptions that there is a one-to-one correspondence

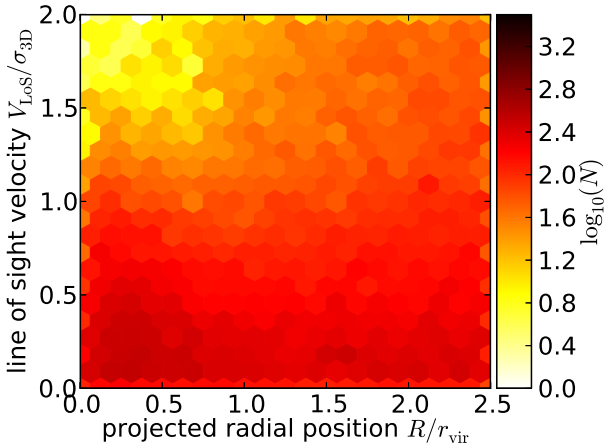

Figure 4.1: Phase space distribution of our sample of observed satellite galaxies, expected to be comparable to the right panel of Fig. 3.5.

between haloes and satellite galaxies and that galaxies follow exactly the same trajectories as their associated haloes could lead to a mismatch between the projected phase space distribution of satellite galaxies and haloes if they are invalid (for instance, if there are "dark haloes" which host no galaxy). Fig 4.1 shows the projected phase space distribution of our observed sample of cluster satellite galaxy candidates (including an interloper population). Encouragingly it appears qualitatively similar to the projected phase space distribution of the sum of our simulated satellite and interloper populations (Fig. 3.5, right panel) with a different normalization. We compare the two distributions more carefully in Fig. 4.2. The two distributions are matched in normalization and compared using the simulated distribution as the fiducial value. In most of the projected phase space plane the two distributions match to within 5 per cent or less. The exception is at low projected radius and low LoS velocity where there is an excess in the observed galaxy sample. The magnitude of the excess is consistent with the amount of artificial halo destruction expected in the central regions of simulated clusters. There may also be some galaxies missing in the observed sample in this region due to fibre collisions during the observing process. Finally, this region includes the observed BCGs, for which we do not expect a corresponding halo in our simulated orbit catalogue since BCGs are assumed to be hosted by the cluster halo. Overall, we deem the match sufficiently close to justify the application of our infall time PDFs to our observed sample of galaxies.

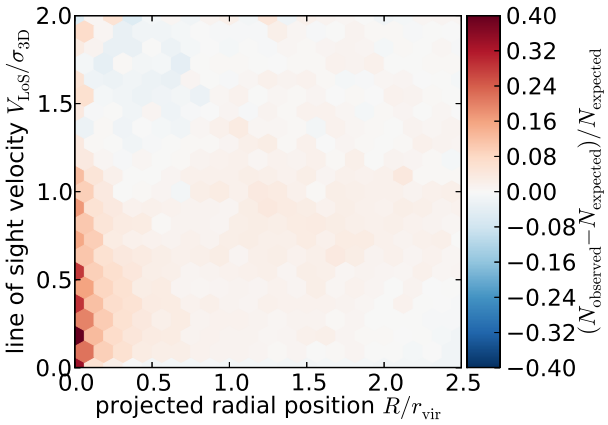

Figure 4.2: Fractional excess of observed galaxies relative to amount expected from simulation halo counts as a function of phase space position. $N_{\rm expected}$ is computed by renormalizing the phase space distribution of simulation haloes so that it matches the normalization of the observed sample. Most regions are limited to variations of a few per cent; the excess of observed satellites at low R and low V is consistent with the amount of artificial disruption of simulation haloes near the centre of clusters discussed in §3.1.

4.2 Comparison of Pre- and Post-Pericentre Satellites

To begin examining the star forming properties of satellite galaxies, we adopt a very crude model where some change is triggered in a satellite galaxy when it passes pericentre and is reflected in the colour of the galaxy. This choice is only vaguely physically motivated in the sense that pericentre is when satellites are exposed to the highest local densities, relative velocities and tidal fields. It will suffice for some exploratory analysis.

Before examining the distributions of galaxy colours as a function of pericentric passage, we familiarize ourselves with the colour distribution as a function of position in the (R, V) plane. Fig. 4.3 shows the total colour distributions for a selection of broad regions in (R, V) space, and Fig. 4.4 shows the same for disk colours. Galaxies with small projected radii and low LoS velocities, which we expect to have the oldest infall times, have more red galaxies and less blue galaxies than the cluster average, while galaxies which we expect to be infalling or interlopers have an excess of blue galaxies and a deficit of red galaxies.

We consider a quantity easily derived from our infall time PDFs: the probability p_{pp} that a satellite (or an interloper, though these have by definition a probability of 0) has passed pericentre. Here, we obtain this by integrating the PDFs described above from an infall time of ~ 4 Gyr (the typical time taken to reach pericentre after infall) to ~ 14 Gyr. The

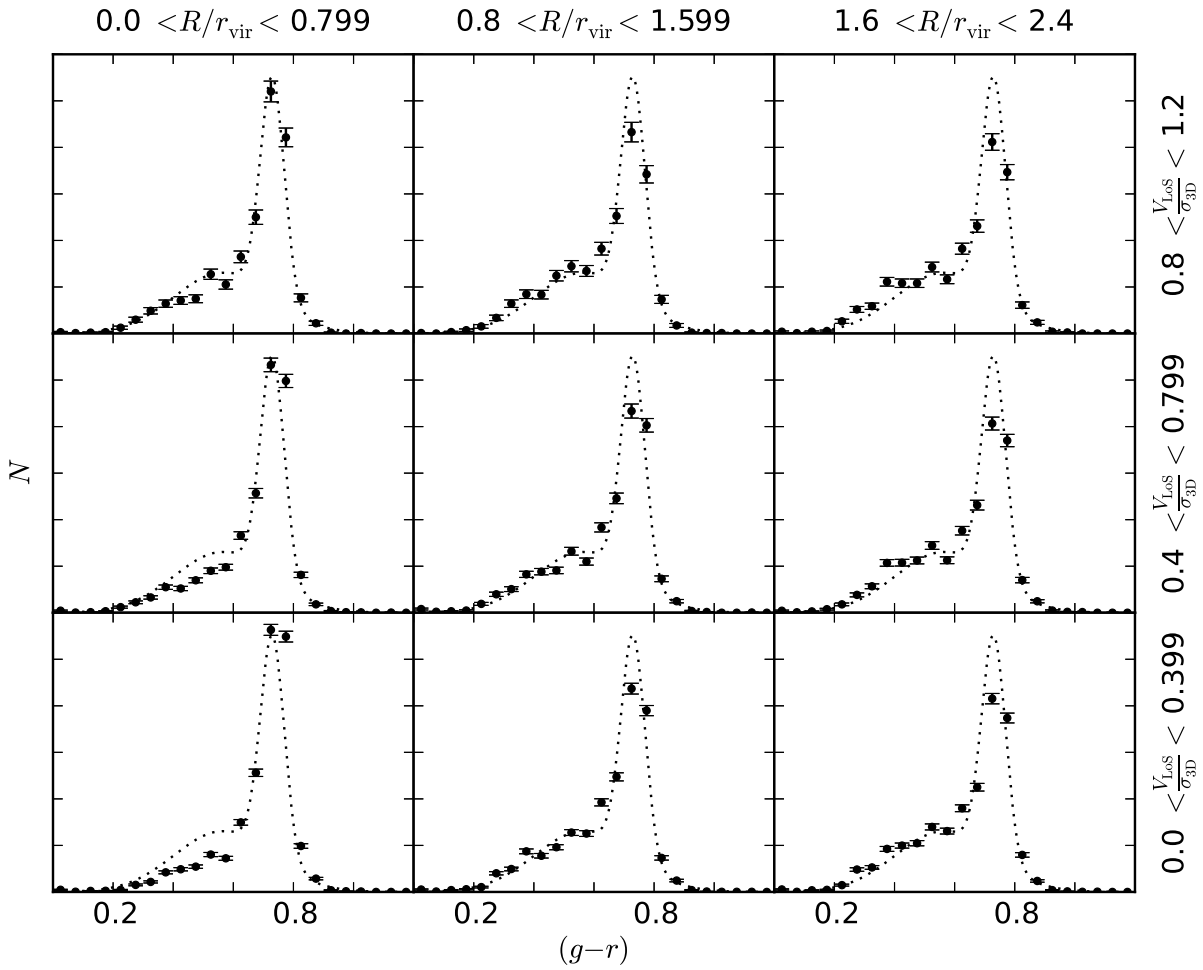

Figure 4.3: Galaxy $(g-r)$ colour distribution for nine regions in projected phase space. The y axis shows absolute number counts and is scaled to fill the available space in each panel. The error bars have length $\sqrt{N}$. The dotted line is meant to serve as a reference curve and shows the best-fit double gaussian function to the colour distribution for all galaxies in $0 < R/r_{\rm vir} < 2.5$ and $0 < V_{\rm LoS}/\sigma_{\rm 3D} < 2$. Galaxies with low R and V have an excess of red galaxies and a deficit of blue galaxies as compared to the cluster mean while the reverse is true for regions we expect to be dominated by infalling or interloper galaxies.

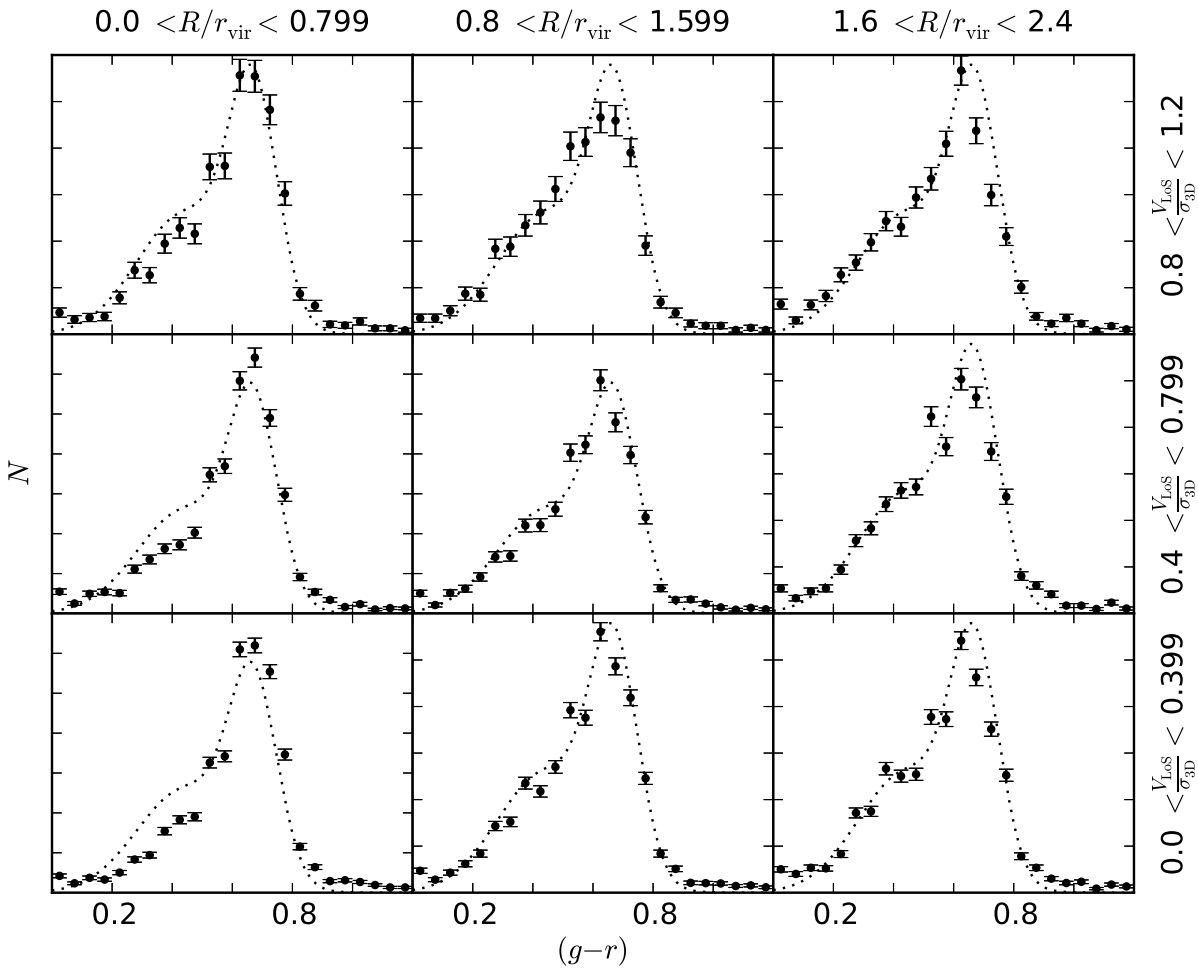

Figure 4.4: As in Fig. 4.3, but for disk colours.

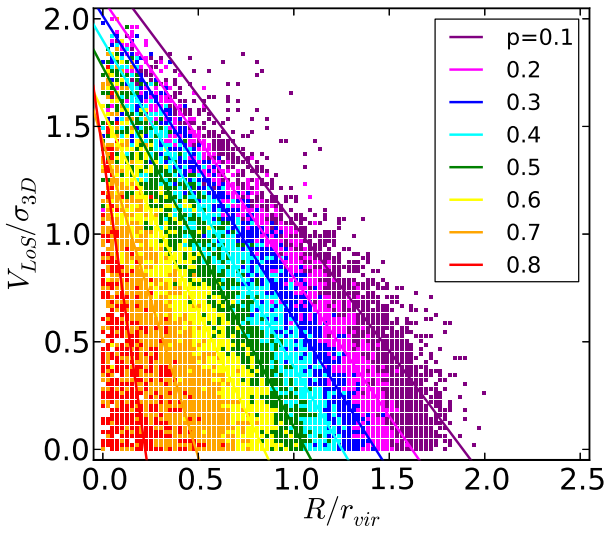

Figure 4.5: Probability that a satellite halo has experienced a pericentric passage as a function of phase space position. Points with very low probabilities ($p < 0.05$, i.e. the interloper population in the upper right region of the phase space plane) are omitted. Each probability bin is fit by minimizing the square of the perpendicular offsets of points from the fit line. These fits allow crude selections; for instance, choosing all objects with phase space coordinates left of the green line (p=0.5) will give a sample of objects which have experienced a pericentric passage with at least 50 per cent certainty. Table 4.1 gives the fit parameters for each line.

probabilities could be calculated more precisely by constructing new PDFs that explicitly encode the time of pericentric passage, but again we are focussing on a rough first pass over the data for the moment. p_{pp} as a function of projected phase space position is plotted in Fig. 4.5. The strongest trend is the increase of p_{pp} with decreasing R, but satellites with higher $V_{\rm LoS}$ at a fixed radius also have a lower p_{pp}. The probabilities are sorted into bins of width 0.1. We perform a linear fit on each bin (fit parameters are given in Table 4.1), allowing a crude selection of, for instance, "all satellites with a $p_{pp} > 0.7$" by choosing only satellites below the appropriate line.

As our trigger time for a change in galaxy colour in this simple model is pericentre, we construct two samples of galaxies. One is comprised of all galaxies with a low probability of having experienced a pericentric passage ($p_{pp} < 0.2$ while the other is comprised of galaxies which have most likely experienced a pericentric passage ($p_{pp} > 0.8$). Fig. 4.6 shows the colour distribution for all galaxies (black curve) and for our sample of high and low p_{pp}

Table 4.1: Fit parameters for the lines shown in Fig. 4.5, and r^2 value for each fit. The function fit is of the form $\left(\frac{V_{\rm LoS}}{\sigma_{\rm 3D}}\right) = a\left(\frac{R}{r_{\rm vir}}\right) + b$

p	a	b	r^2
0.1	-1.18	2.23	0.969
0.2	-1.29	2.09	0.982
0.3	-1.41	2.01	0.983
0.4	-1.51	1.90	0.981
0.5	-1.67	1.77	0.977
0.6	-1.86	1.57	0.965
0.7	-2.94	1.43	0.905
0.8	-6.21	1.38	0.916

galaxies. The distribution is reasonably well fit by two gaussians in all cases. The two gaussians have equal normalizations, and the total normalization is fixed by the size of the sample. The distribution has a sharp peak at the red end which appears to increase in amplitude with pericentric passage but is nonetheless prominent in the pre-pericentre sample. We speculate that this peak is made up of a combination of galaxies quenched by the cluster environment, "group pre-processed" galaxies which have already been reddened by a group environment and "mass quenched" galaxies where star formation has been cut off by an internal mechanism. There is also a much broader blue peak (or perhaps "blue hump" is a better term). Our fits suggest that the blue hump reddens (shift in the mean) and appears to decrease in amplitude as one goes from galaxies which have not passed pericentre to those that have. For this plot (and Figs. 4.7-4.8) we have made the simple choice that the two peaks are constrained to have equal numbers of galaxies (there was no *a priori* reason for this choice, but it seems to provide reasonable fits). The amplitudes of the peaks are fixed by the other parameters. For instance, the increase in amplitude in the red peak is caused by a slight narrowing of the peak and an increasing contribution from the tail of the blue hump as it reddens.

Fig. 4.7 is similar to Fig. 4.6, but shows the colour distribution of the bulge component of the galaxies. This time we find only a strong red peak; the second gaussian has nearly the same mean as the first but is much broader, effectively just increasing the amplitude of the tails of the distribution. There is perhaps a slight reddening after pericentric passage. Qualitatively, this is as we might expect. Galactic bulges are typically not expected to host active star formation, leading to predominantly red colours; they are already quenched.

Fig. 4.8 is again similar to the previous two, but shows the colour distribution of the

disk component. Similarly to the total colour distribution, we see a red peak and a blue hump. According to our gaussian fit parameters, the red peak is again approximately fixed in mean and width (though with a slight narrowing with higher p_{pp}); much of the apparent change is again due to the changing contribution from the tail of the blue hump. The blue hump appears to redden with increasing p_{pp} and changes only very slightly in width. The red and blue distributions are overlapping substantially. It is likely that they would be better separated in $(u-r)$ colour, but the bulge-disk decomposition data of Simard et al. (2011) is limited to $(g-r)$ colour. We speculate that this is because the u-band is noisier than the g-band, making reliable decomposition more difficult.

Fig. 4.9 compares the r-band bulge-to-total light ratio of high and low p_{pp} galaxies. We do not note any obvious trend with p_{pp}, except that the number of bulgeless galaxies might decrease after pericentric passage. We are uncertain as to whether this is a real effect or whether these bulgeless galaxies are artifacts of the decomposition process. There is perhaps a very slight trend toward higher $(B/T)_r$ after pericentric passage (in the bottom panel, after ignoring the first bin, the red curve has a very slight positive slope and the blue curve has a very slight negative slope), but the magnitude of the effect is on the same order as the scatter in the distribution.

4.3 Method: Colour Distribution Evolution Fitting

Our observed sample is first cut to match the domain of applicability of our PDFs. We discard any galaxies with an estimated halo mass of less than $10^{11.9}\,\mathrm{M}_\odot$, and any with $(g-r)$ colours outside the main distribution of colours, specifically those with colours outside the interval $[0.0, 1.5]$. We also split the sample into three consecutive bins of halo mass with edges at $10^{11.9}$, $10^{12.3}$, $10^{12.6}$ and $10^{15.0}\,\mathrm{M}_\odot$ so that we may investigate a possible mass dependence of the colour distribution evolution. The bins were selected to have a similar number of galaxies in each bin. These are the same mass bins used when creating the PDFs described in §3.1. We focus on the colour of the disk, and so we impose one final cut, removing galaxies with an r-band bulge to total light ratio $\left(\frac{B}{T}\right)_r$ of greater than 0.9 (i.e. systems with little or no disk component). This gives us a sample of 8,143 disks in our lowest mass bin, 10,134 in our middle bin and 12,461 in our highest mass bin.

The $(g-r)$ colour distribution of our sample of cluster satellites is bimodal, both in total colour and disk colour, as expected. We model the distribution using two gaussian curves. Each gaussian has two parameters, a mean and a standard deviation, which we denote μ_b, σ_b, μ_r and σ_r with subscripts denoting the 'red' and 'blue' peaks. We also allow the relative normalization of the gaussians to vary using a parameter for the 'blue

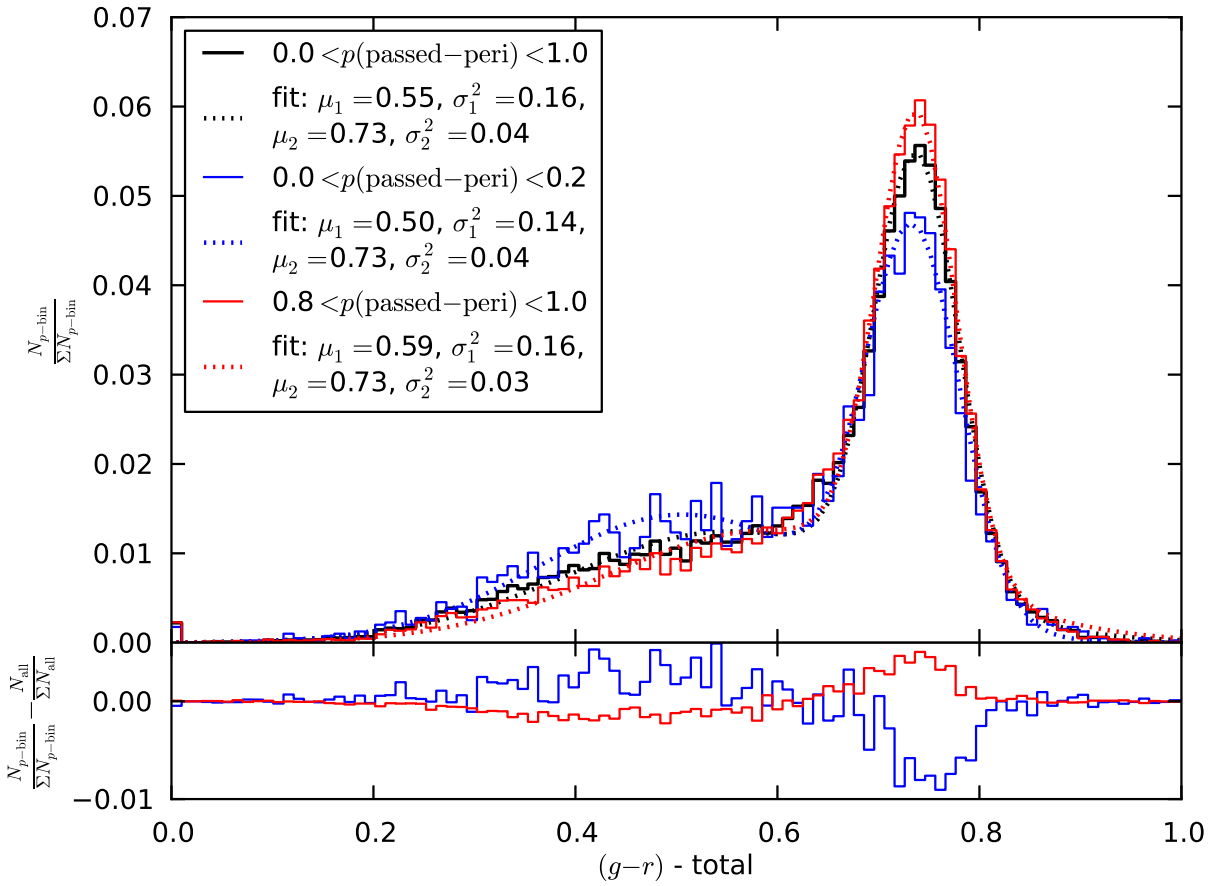

Figure 4.6: Total (bulge + disk) $(g-r)$ colour distribution of cluster satellites. *Top panel:* The absolute distribution for all satellites (black histogram), and the distribution for satellites with a probability of having experienced a pericentric passage of 20 per cent or less (blue histogram) and 80 per cent or more (red histogram). The distributions are normalized so that they all have the same integral for easier comparison. The histograms were fit with two gaussians (dotted lines), with fit parameters given in the legend. *Bottom panel:* Distribution of 20 per cent (blue) and 80 per cent (red) chance of pericentric passage relative to the distribution for all satellites.

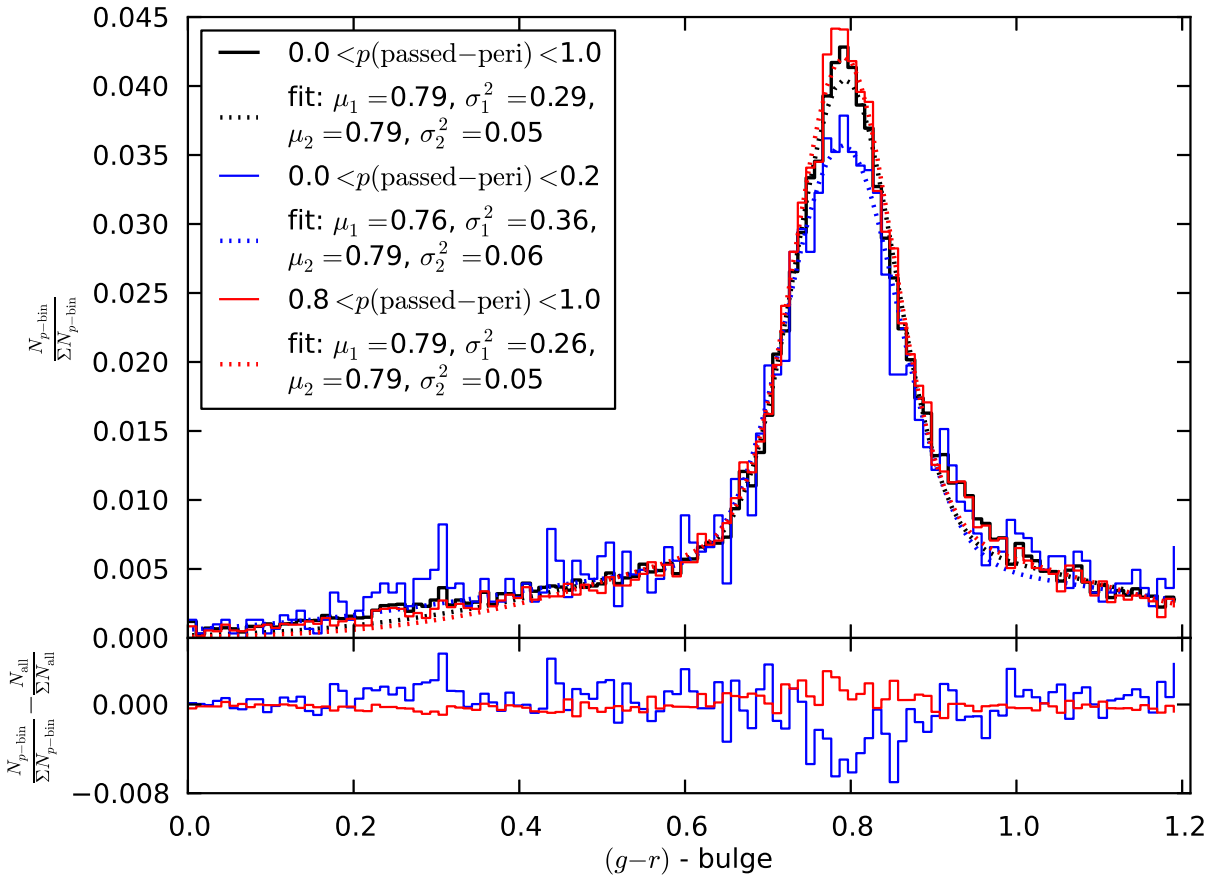

Figure 4.7: As in Fig. 4.6, but for the bulge colour.

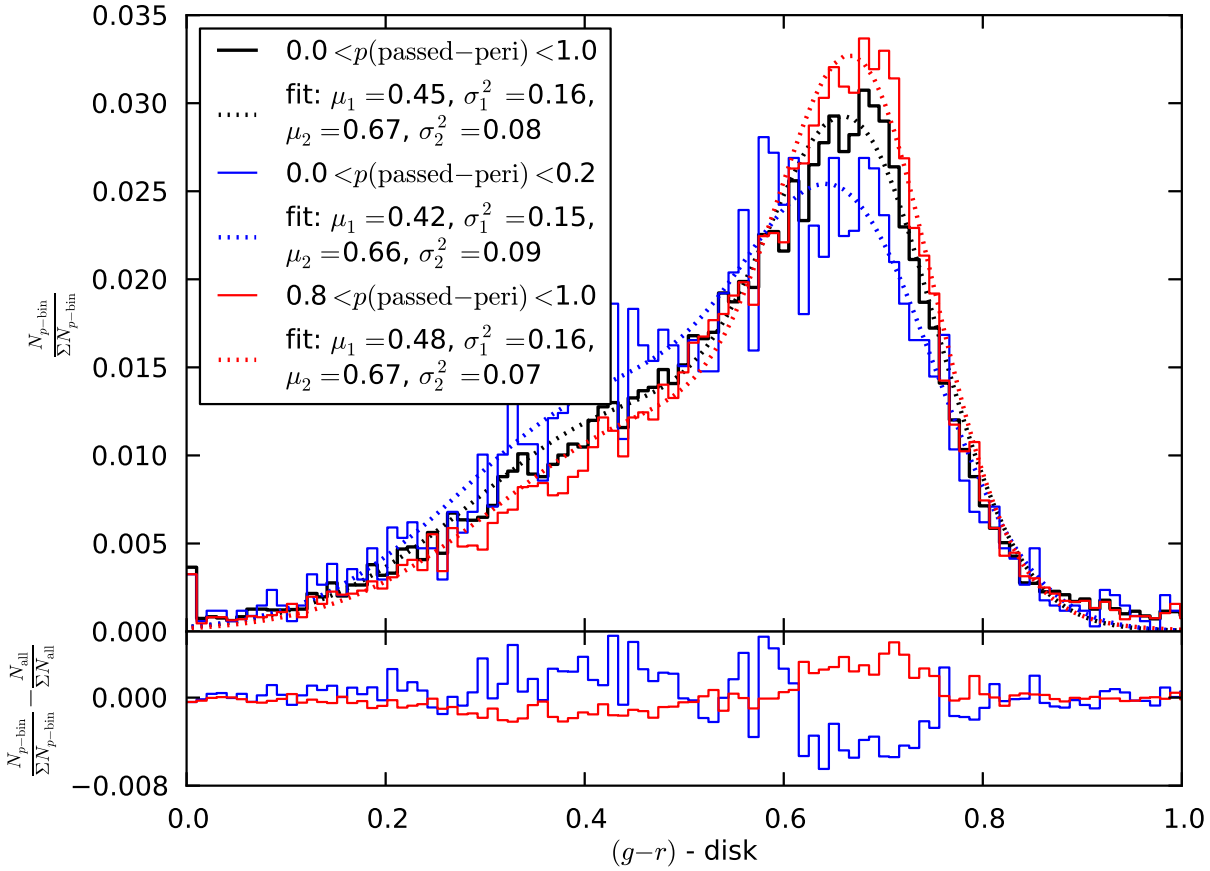

Figure 4.8: As in Fig. 4.6, but for the disk colour.

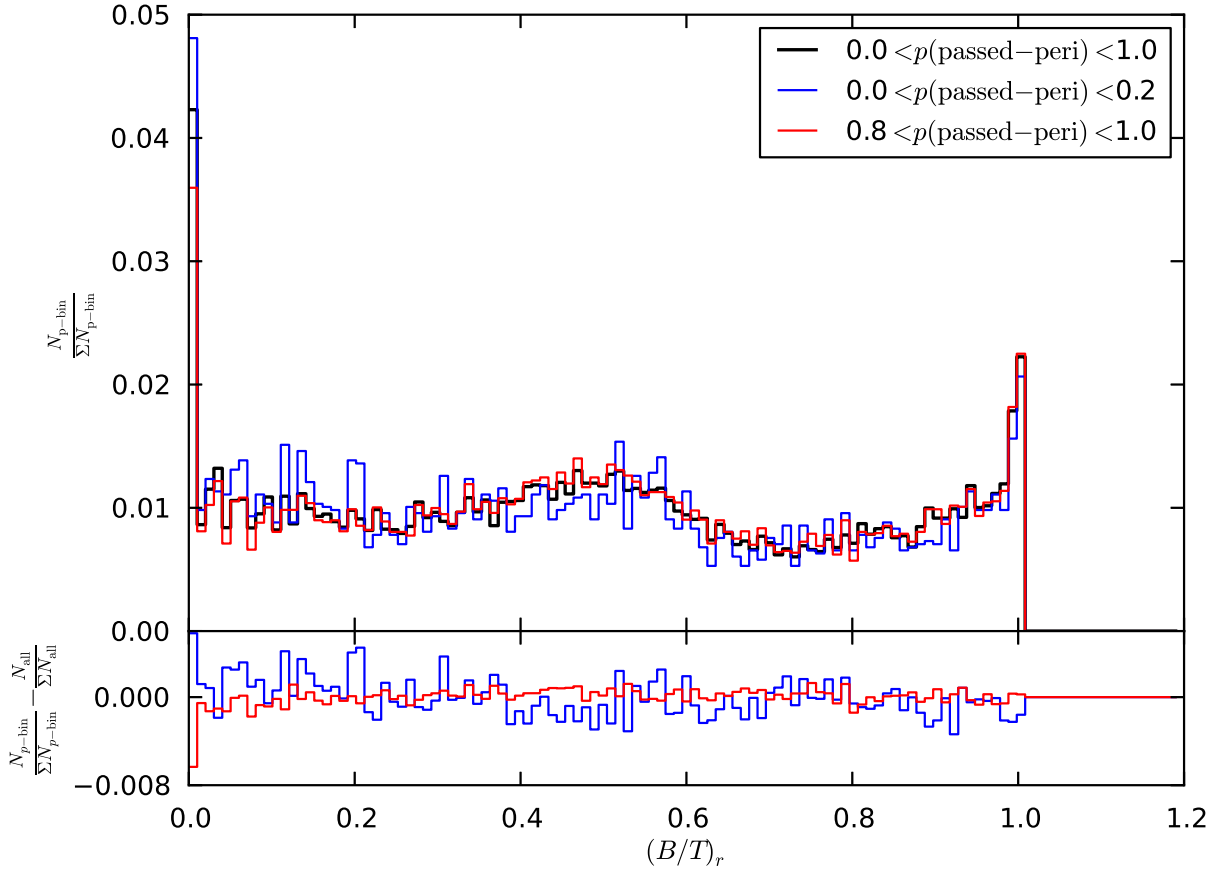

Figure 4.9: Bulge-to-total ratio in the r-band distribution of cluster satellites. *Top panel:* The absolute distribution for all satellites (black histogram), and the distribution for satellites with a probability of having experienced a pericentric passage of 20 per cent or less (blue histogram) and 80 per cent or more (red histogram). The distributions are normalized so that they all have the same integral for easier comparison. *Bottom panel:* Distribution of 20 per cent (blue) and 80 per cent (red) chance of pericentric passage relative to the distribution for all satellites.

fraction' f_b so that the blue peak is weighted by f_b and the red peak by $(1-f_b)$. The total normalization is fixed by our sample size. A preliminary examination of the data (discussed in detail in §4.2, see especially Figs. 4.6, 4.7, 4.8) revealed that μ_r, σ_r and σ_b do not change significantly between satellites that have most likely not experienced a pericentric passage and satellites that have most likely experienced a pericentric passage. In the following analysis we fix these parameters to their best-fit values for the disk sample (see Fig. 4.8 and Table 4.2 for an example of parameter choices, though these can be refined e.g. with the results in §4.4.2). This leaves two parameters (μ_b and f_b) to be fit.

We perform a bayesian statistical analysis (additional background and detail is given in Appendix B) to simultaneously fit this model to our disk colour data (we choose the disk component because it is expected to be most sensitive to environmental effects, see Hudson et al., 2010) in several bins of time since infall. We use our infall time PDFs to weight the contribution of each data point to each time bin. There is a pair of parameters (μ_b, f_b) for each time bin; we further subscript them with an index $\{j \in \mathbb{N} | 0 \leq j < N_t\}$ where N_t is the number of time bins. Our data are subscripted with an index $\{i \in \mathbb{N} | 0 \leq i < N_d\}$ where N_d is the number of data points. Thus the phase space coordinates and $(g-r)$ colour of a single data point are denoted $(R,V)_i$ and $(g-r)_i$, respectively. Formally, from Bayes' theorem:

$$P(A|B) \propto P(B|A)P(A) \tag{4.1}$$

$$\begin{aligned} P_i(\{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i | (g-r)_i) \propto \\ P_i((g-r)_i | \{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i) P_i(\{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i) \end{aligned} \tag{4.2}$$

We define the likelihood $P(B|A)$ as:

$$P_i((g-r)_i | \{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i) = \sum_{j=0}^{N_t} p_j((R,V)_i) P_i((g-r)_i | \mu_{bj}, f_{bj}) \tag{4.3}$$

where p_j expresses the probability that datapoint i belongs to time bin j. This is easily computed from our PDFs by integrating the PDF over the time bin (the PDF corresponding to the phase space coordinates of the data point is used):

$$p_j((R,V)_i) = \int_{t_j}^{t_{j+1}} \mathrm{PDF}((R,V)_i) dt \tag{4.4}$$

Finally, $P_i((g-r)_i | \mu_{bj}, f_{bj})$ is simply our double gaussian model evaluated at the data

point:

$$P_i((g-r)_i|\mu_{bj}, f_{bj}) = \frac{f_{bj}}{\sigma_b\sqrt{2\pi}} \exp\left[-\frac{((g-r)_i - \mu_{bj})^2}{2\sigma_b^2}\right] + \frac{(1-f_{bj})}{\sigma_r\sqrt{2\pi}} \exp\left[-\frac{((g-r)_i - \mu_r)^2}{2\sigma_r^2}\right] \tag{4.5}$$

We adopt a flat prior for all our parameters f_{bj} and μ_{bj}. An example of possible parameter ranges are shown in Table 4.2. We use an implementation of our model in the Bayesian Analysis Toolkit (BAT) software package[1] (Caldwell et al., 2009) to compute a Markov Chain Monte-Carlo (MCMC) estimate of the posterior distribution of our parameters by maximizing:

$$\log \prod_i P_i((g-r)_i|\{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i) \iff \sum_i \log P_i((g-r)_i|\{\mu_{bj}, f_{bj}\}_{j=0\ldots N_t}, (R,V)_i) \tag{4.6}$$

This method is implemented and being tested, debugged and tuned. This must be done carefully since a full-scale run on our dataset can take on the order of several days to a few weeks. Because of this, no conclusive results are yet available. However, we have constructed a simpler, faster framework to obtain a rough picture of how the colour distribution of cluster satellites evolves as a function of their time spent in the cluster. We present this method and its results in §4.4.

4.4 Preliminary: Colour Evolution of Cluster Satellites

In an effort to obtain some preliminary results on the colour distribution evolution of cluster satellites, we devised a second fitting method. It is conceptually simpler than that in §4.3, and is also many orders of magnitude faster computationally, but comes with several important caveats. First, the errors on the results are uncertain; we provide only an estimate. Second, the colour distributions in the various time bins are decoupled and fitted individually, so the amount of information constraining each fit is greatly reduced. Third, and most importantly, the infall time distributions for satellites in each time bin

[1] https://www.mppmu.mpg.de/bat/

Table 4.2: Example of parameter choices, both to fix and to fit, for our MCMC analysis of the total colour distribution evolution of galaxy cluster satellites. The edges of the time bins correspond to simulation snapshot times. The time bins use half-open intervals, e.g. [a,b).

Parameter(s)	Type	Description	Value/Range	Units
f_{bj}, $\{j \in \mathbb{N} \vert 0 \leq j < N_t\}$	fit	blue fraction	$0.2 \leq f_b \leq 0.6$	–
μ_{bj}, $\{j \in \mathbb{N} \vert 0 \leq j < N_t\}$	fit	mean of blue peak	$0.25 \leq \mu_b \leq 0.60$	$(g-r)$ colour
σ_b	fixed	std. dev. of blue peak	0.16	$(g-r)$ colour
μ_r	fixed	mean of red peak	0.67	$(g-r)$ colour
σ_r	fixed	std. dev. of red peak	0.08	$(g-r)$ colour
N_t	fixed	number of time bins	9	–
t_j, $\{j \in \mathbb{N} \vert 0 \leq j < N_t\}$	fixed	time bin lower edges	$\{-\infty^*, 1.96, 2.87, 3.82, 4.78, 6.26, 7.25, 8.24\}$	Gyr since infall
t_{j+1}, $\{j \in \mathbb{N} \vert 0 \leq j < N_t\}$	fixed	time bin upper edges	$\{1.96, 2.87, 3.82, 4.78, 6.26, 7.25, 8.24, 13.11\}$	Gyr since infall

* Denotes that interlopers are included in this bin.

(from the PDFs at the location of each in projected phase space) are convolved with the colour distribution of the satellites. Stated concisely, this smooths out the colour evolution. Each of these caveats is discussed in more detail in §4.4.3. First, though, we present the method in §4.4.1 and its results in §4.4.2.

An overview of the approach is as follows. First, the colour distribution (a histogram) for an interval in infall time is constructed from the probability that each satellite has an infall time in this interval. This distribution is then fit with with a double-gaussian function independently in each time bin.

4.4.1 Fitting Procedure

In this subsection we present the details of this fitting procedure. The notation is kept as consistent as possible with §4.3. We split the range of possible infall times into bins which we label with subscripts $\{j \in \mathbb{N} | 0 \leq j < N_t\}$. We denote the lower edge of bin j by t_j, so the bin's upper edge is at t_{j+1}. For each time bin, we construct a histogram of the colour distribution. Each bin of these histograms is labelled with a subscript $\{i \in \mathbb{N} | 0 \leq i < N_i\}$; each bin has a height $\mu_{j,i}$ with associated error $\sigma^2_{j,i}$ and centre $(g-r)_i$. The half-width $\Delta(g-r)$ is the same for every bin. Finally, we subscript quantities that are defined for every galaxy (data point) with a subscript $\{d \in \mathbb{N} | 0 \leq d < N_d\}$. For instance, the colour of a galaxy is $(g-r)_d$, its projected phase space coordinates are $(R,V)_d$ and the probability that it has an infall time that falls in time bin j is:

$$w_{j,d} = \int_{t_j}^{t_{j+1}} PDF((R,V)_d) dt \tag{4.7}$$

$\mu_{j,i}$ and $\sigma^2_{j,i}$ are constructed from these $w_{j,d}$ by considering them as the mean and standard deviation of a sum of independent random variables. To make the equations more readable,

$(g-r)$ has been replaced with the shorthand C, for colour. The histogram is normalized.

$$\mu_{j,i} = \frac{\sum_{d=0}^{N_d} [w_{j,d}(C_i - \Delta C < C_d \leq C_i + \Delta C)]}{\sum_{d=0}^{N_d} w_d} \tag{4.8}$$

$$\sigma^2_{j,i} = \frac{\sum_{d=0}^{N_d} [w_{j,d}(1-w_{j,d})(C_i - \Delta C < C_d \leq C_i + \Delta C)]}{\left[\left(\sum_{d=0}^{N_d}(C_i - \Delta C < C_d \leq C_i + \Delta C)\right)\left(\sum_{d=0}^{N_d} w_{j,d}\right)\right]^2} \tag{4.9}$$

The inequalities are to be evaluated to 1 or 0 depending on whether they are true or false, respectively. As the number of data points contributing to a bin increases, the distribution described by a $(\mu_{j,i}, \sigma^2_{j,i})$ pair approaches a gaussian distribution by the central limit theorem.

The ensemble of $(\mu_{j,i}, \sigma^2_{j,i})$ pairs describes a series of histograms, one for each time bin, that we now desire to fit with our double gaussian model. First, however, there is a practical difficulty with the $\sigma^2_{j,i}$s. These encode only the statistical uncertainty in the mean of each bin, but the scatter in the $\mu_{j,i}$s is much larger than what would be expected if this was the dominant source of uncertainty; presumably this scatter is fixed by some property of the observations or the bulge-disk decomposition; for simplicity we fix all the $\sigma^2_{j,i}$s to be of the same order as the observed scatter with $\sigma_{j,i} = 0.05$ in all colour bins and in all time bins. We retain the assumption that the errors are gaussian.

The problem of fitting a function (our model function will be denoted F_j; it is defined by a set of parameters for every time bin j) to a data series with gaussian errors is a standard problem in Bayesian analysis. We use an MCMC analysis with the well-known likelihood function:

$$P_j = \prod_{i=0}^{N_i} \frac{1}{\sqrt{2\pi\sigma^2_{j,i}}} \exp\left[-\frac{(\mu_{j,i} - F_j((g-r)_i)^2}{2\sigma^2_{j,i}}\right] \tag{4.10}$$

The large increase in speed in this method comes from the much reduced number of terms contributing to the product in the likelihood function. Because this method is so fast, we can afford to fit all parameters of a double gaussian model, i.e. a mean for each peak $\mu_{j,b}$ and $\mu_{j,r}$ (red and blue peak, respectively), a standard deviation for each peak $\sigma_{j,b}$ and $\sigma_{j,r}$ and the relative normalization of the peaks $f_{j,b}$, $f_{j,r} = (1 - f_{j,b})$. In practice we constrain

the blue peak to be bluer than the red peak by fitting $\Delta\mu_{j,r} = \mu_{j,r} - \mu_{j,b}$ with the constraint $\Delta\mu_{j,r} > 0$. The function to fit is then:

$$F_j((g-r)_i) = \frac{f_{j,b}}{\sqrt{2\pi}\sigma_{j,b}} \exp\left[-\frac{((g-r)_i - \mu_{j,b})^2}{2\sigma_{j,b}^2}\right] + \frac{(1-f_{j,b})}{\sqrt{2\pi}\sigma_{j,r}} \exp\left[-\frac{((g-r)_i - (\mu_{j,b} + \Delta\mu_{j,r}))^2}{2\sigma_{j,r}^2}\right] \tag{4.11}$$

For all fits we adopt the (quite broad) parameter ranges $0.2 < \mu_{j,b} < 0.75$, $0.01 < \sigma_{j,b} < 1.0$, $0.0 < \Delta\mu_{j,r} < 0.6$, $0.0 < f_{j,b} < 1.0$ with flat priors on all parameters.

4.4.2 Colour Evolution Results

The results of the fitting process described in §4.4.1 applied to our total galaxy colour data are summarized in Table 4.3 and applied to our disk colour data in Table 4.4. Of the parameters fit, three appear to show no significant change over the 13 Gyr spread in infall times. The width of both peaks remains unchanged, except in the time bin that includes interlopers, where the blue peak is somewhat wider. The position of the red peak also shows no significant evolution. There may be a weak trend of reddening of red disks with increasing time since infall, but it is consistent with no evolution within our 1σ estimated errors. The position of both peaks and the blue fraction as a function of time since infall are shown in Fig. 4.10 for total colours and in Fig. 4.11 for disk colours. The position of the blue peak moves toward the red with increasing time since infall, and simultaneously the ratio of blue to red galaxies decreases. The evolution is qualitatively similar in both the total colour and disk colour, with faster reddening until about 4 Gyr since infall (i.e. until the mean pericentre time), then slower reddening thereafter. We note that the region of projected phase space occupied by satellites just before pericentre (slightly less than 4 Gyr since infall) and just after pericentre (slightly more than 4 Gyr since infall) is partially degenerate, which might naïvely lead us to the prediction that the colour evolution would be locally symmetric near $t = 4$ Gyr. However, the clear asymmetry in the evolution about this point leads us to believe that this is a time marking a real transition in the colour evolution.

It is not clear whether the colour evolution of individual galaxies is rapid or slow. The gradual drift toward the red of $\mu_{j,i}$ is suggestive of a slow process, but the evolution of $f_{j,b}$ could also lead to the conclusion that individual galaxies "jump" rapidly from one

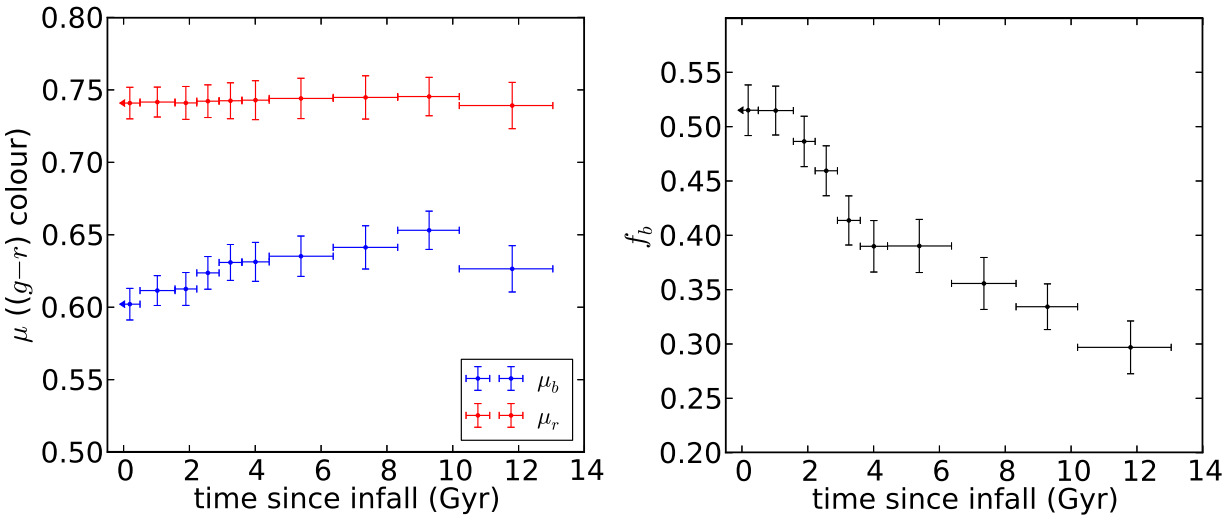

Figure 4.10: *Left panel:* Evolution of the peak of the blue (red) distribution is shown with blue (red) symbols as a function of time for total colours fit by the method of §4.4.1. The horizontal error bars indicate the extent of the time bins, while the vertical error bars are a very rough estimate of the error in the parameters (see §4.4.3). The arrows indicate the inclusion of interlopers in the first time bin. *Right panel:* As *left*, but showing the evolution of the relative normalization of the red and blue distributions.

distribution to the other. It may be that both interpretations are correct and there are processes driving slow and rapid colour evolution operating in tandem.

Care must be taken in interpreting these results for the reasons alluded to above and discussed in §4.4.3 below, but at the very least these results seem suggestive of there being an environmental effect on galaxy colour as early as virial radius crossing, and potentially even before this. The decrease in reddening rate after pericentre passage is also suggestive of this being a key time in the history of a satellite galaxy.

4.4.3 Caveats and Discussion

There are three major caveats relating to the interpretation of the results of §4.4.2: the reliability of the errors quoted on the fits, the effect of decoupling the time bins and the convolution of the infall time distributions with the colour distributions.

The assumption that the error in the colour histogram means $\sigma^2_{j,i}$ is not dominated by the statistical error (equation 4.9) but rather by some other source, motivating us to set

Table 4.3: Fit parameters for the total colour distribution analysis described in §4.4.2. t_j and t_{j+1} give the edges of each time bin, while $\mu_{j,b}$, $\sigma_{j,b}$, $\Delta\mu_{j,r}$ and $f_{j,b}$ are the means of the marginalized parameter distribution for each parameter. The same quantities prepended by δ are the standard deviations of the marginalized parameter distributions about the means. Note that all marginalized distributions are sharply and nearly symmetrically peaked so that mean, median and mode are nearly identical. The most interesting parameters, $\mu_{j,b}$, $\mu_{j,r}$ and $f_{j,b}$ are plotted as a function of time in Fig. 4.10.

Parameter	$j=0$	1	2	3	4	5	6	7	8	9
t_j	13.11	10.12	8.24	6.26	4.30	3.58	2.87	2.18	1.51	0.42
t_{j+1}	10.12	8.24	6.26	4.30	3.58	2.87	2.18	1.51	0.42	$-\infty^\dagger$
$\mu_{j,b}$	0.628	0.650	0.645	0.639	0.634	0.633	0.626	0.614	0.612	0.602
$\delta\mu_{j,b}$	0.016	0.013	0.015	0.014	0.013	0.012	0.011	0.011	0.010	0.011
$\sigma_{j,b}$	0.134	0.135	0.134	0.133	0.134	0.136	0.133	0.138	0.140	0.141
$\delta\sigma_{j,b}$	0.010	0.008	0.008	0.007	0.007	0.007	0.006	0.006	0.006	0.006
$\Delta\mu_{j,r}$	0.111	0.0956	0.100	0.105	0.109	0.110	0.116	0.127	0.130	0.139
$\delta\Delta\mu_{j,r}$	0.016	0.013	0.015	0.014	0.013	0.012	0.011	0.011	0.010	0.011
$\sigma_{j,r}$	0.0439	0.0367	0.0373	0.0377	0.0378	0.0371	0.0375	0.0380	0.038	0.0390
$\delta\sigma_{j,r}$	0.0010	0.0009	0.0010	0.0011	0.0011	0.0011	0.0012	0.0013	0.0013	0.0014
$f_{j,b}$	0.300	0.332	0.361	0.397	0.39	0.418	0.463	0.490	0.516	0.517
$\delta f_{j,b}$	0.024	0.021	0.024	0.024	0.024	0.023	0.023	0.023	0.023	0.023

† Denotes that interlopers are included in this bin.

Table 4.4: As in Table 4.3, but for the disk colour distribution evolution. There is no $j = 0$ time bin because there was insufficient disk colour data in this time interval. The most interesting parameters, $\mu_{j,b}$, $\mu_{j,r}$ and $f_{j,b}$ are plotted as a function of time in Fig. 4.11.

Parameter	$j = 1$	2	3	4	5	6	7	8	9
t_j	10.12	8.24	6.26	4.30	3.58	2.87	2.18	1.509	0.42
t_{j+1}	8.24	6.26	4.30	3.58	2.87	2.18	1.51	0.42	$-\infty^{\ddagger}$
$\mu_{j,b}$	0.588	0.587	0.579	0.578	0.571	0.562	0.554	0.552	0.549
$\delta\mu_{j,b}$	0.017	0.015	0.015	0.017	0.017	0.017	0.017	0.019	0.020
$\sigma_{j,b}$	0.166	0.167	0.165	0.163	0.166	0.166	0.166	0.168	0.167
$\delta\sigma_{j,b}$	0.009	0.008	0.008	0.008	0.008	0.008	0.007	0.008	0.008
$\Delta\mu_{j,r}$	0.105	0.101	0.109	0.109	0.114	0.121	0.126	0.128	0.129
$\delta\Delta\mu_{j,r}$	0.017	0.015	0.015	0.017	0.017	0.017	0.017	0.019	0.019
$\sigma_{j,r}$	0.0611	0.0601	0.0612	0.0612	0.0620	0.0630	0.0640	0.0661	0.0663
$\delta\sigma_{j,r}$	0.0040	0.0042	0.0045	0.0049	0.0049	0.0057	0.0064	0.0073	0.0079
$f_{j,b}$	0.579	0.608	0.626	0.629	0.629	0.659	0.687	0.706	0.718
$\delta f_{j,b}$	0.048	0.045	0.048	0.053	0.052	0.056	0.058	0.065	0.068

‡ Denotes that interlopers are included in this bin.

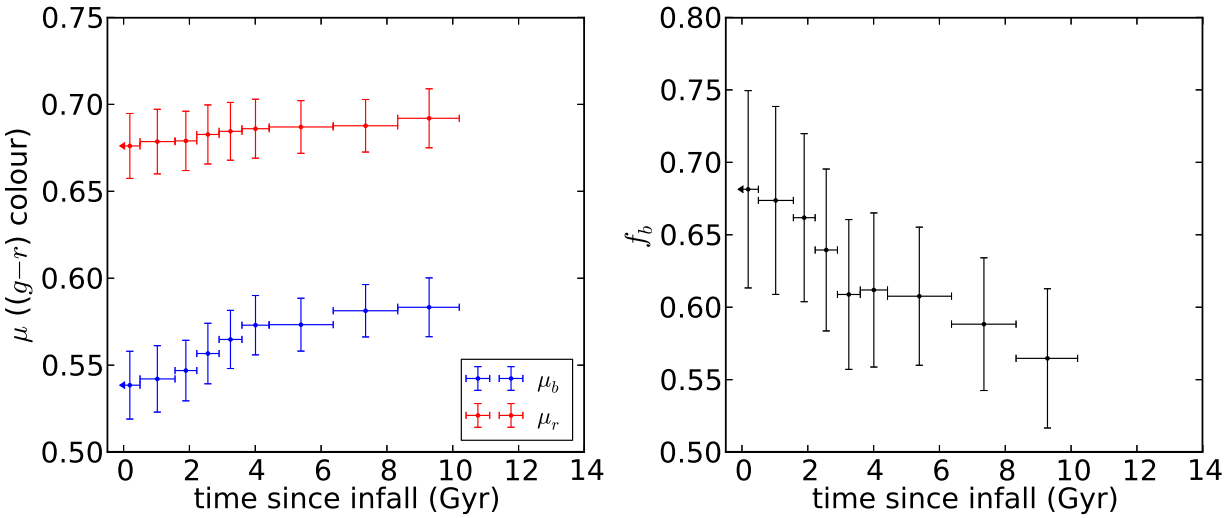

Figure 4.11: As in Fig. 4.10, but for disk colour distribution fits.

$\sigma^2_{i,j} = 0.05$ to approximately match the observed scatter, means that the uncertainty in our fit parameters is only approximate. The vertical error bars in Figs. 4.10-4.11 should therefore be interpreted only as a rough guess at the errors. Our estimated value of $\sigma^2_{i,j} = 0.05$ is likely somewhat generous in some regions, particularly the red peak which appears to be very well modelled by a gaussian curve. In other less occupied regions of colour space (e.g. the tail of the blue distribution), the scatter may be larger than estimated (one likely explanation being shot noise). It is difficult to say what effect improving these estimates might have on the confidences in the fits, so making strong statements about the significance of the observed trends is difficult.

The time bins are decoupled in the sense that the fit for one time interval is independent of the others (the conspicuous correlations visible in Fig. 4.10-4.11 are due to the convolution discussed in the next paragraph). This means that the amount of information constraining each fit is in a sense reduced by a factor of $\frac{1}{N_t}$, and likely contributes to increasing the error in the fits. Performing the fits simultaneously allows for the construction of a likelihood function that attempts to determine the "correct" time bin for a galaxy rather than forcing a contribution to the colour histogram of every time bin from every galaxy as we do here.

The infall time distributions are convolved with the colour distributions. This process has the effect of blurring out the colour distributions[2]. Each histogram still contains some

[2]The effect of convolution of the infall time distributions with the colour distributions can be understood

information since the contributions will be dominated by galaxies from regions with high certainty of being in that particular time bin, but some level of contamination is inevitable, and acts as a sort of smoothing filter on the parameter evolution curves. This effect is most pronounced for intervals of time since infall that do not occupy a relatively distinct region of phase space; for instance the backsplash population overlaps significantly with both infalling and virialized haloes, making the colour histogram for these satellites particularly uncertain. This effect is not accounted for in the error bars in Figs. 4.10-4.11. In any case, it is much more desirable (but also much more computationally demanding) to attempt to *de*-convolve the colour and infall time distributions. One way to do this is to convolve the model function with the infall time distributions then compare to the observed colour data; this is the approach discussed in §4.3 that has not yet produced results.

With these caveats in mind, the two main conclusions that we retain from this analysis are that the effect of environment on colour (and by extension star formation) begins in the infalling phase of the satellite orbit, and that the first pericentric passage marks a transition in the colour evolution.

in more detail by examining the construction of the colour histograms. Consider a toy model where infalling galaxies are blue until they reach pericentre, then immediately become red after passing pericentre (these will be called "virialized" in this example). Suppose we consider two hypothetical regions in (R, V) space, region 1 where we expect 60 per cent infalling and 40 per cent virialized galaxies and region 2 where we expect 40 per cent infalling and 60 per cent virialized galaxies. To construct a colour distribution for the infalling population using convolution, we sum together the colour distribution in region 1 with weight 0.6 and in region 2 with weight 0.4. The colour distribution obtained contains 52 per cent blue galaxies and 48 per cent red galaxies. A similarly constructed colour distribution for the virialized population contains 48 per cent blue galaxies and 52 per cent red galaxies. The true colour distributions (100 per cent blue, 0 per cent red for infalling and 0 per cent blue, 100 per cent red for virialized) have been blurred together.

Chapter 5

Conclusions

We begin by summarizing our results (§5.1). We then place our results in a broader context by comparing them to related work by other authors (§5.2). We finish with some ideas for possible continuations of this work (§5.3).

5.1 Summary

In Chapter 3 we explored the orbital statistics of cluster satellites using cosmological N-body simulation data. We found that the phase space distribution of infalling, backsplash and virialized satellite haloes are different but not everywhere distinct. Different populations of satellite haloes are better separated in phase space than in the radial coordinate alone, but there is no immediately obvious cut that we can impose on the projected phase space coordinates of a satellite galaxy to separate different populations. We examined the entire projected phase space and determined the confidence with which the time since infall can be assigned to a satellite occupying each region. This lets us easily identify regions where the time since infall can be determined accurately enough to separate satellite populations. However, the addition of an interloper population makes matters somewhat worse; the region of maximum discrimination between satellite populations is also the region of maximum overlap between satellites and interlopers.

We have developed a tool that estimates the infall time and a confidence in the estimate for a satellite halo given its projected phase space coordiantes. The possible infall times for a region of projected phase space are weighted according to the frequency of their occurence in simulation, then an infall time is randomly selected from this weighted distribution.

More generally, it also provides a full PDF of the possible infall times of a halo given these same coordinates. This method is easily adapted to parameters other than the time since infall, for instance the closest approach to host centre or the time since first pericentre, or to higher redshift systems.

In Chapter 4 we applied the PDFs derived in Chapter 3 to the study of the colour distributions of cluster satellites as a function of the time since their infall into the cluster. Colour can in principle be related back to a SFH, so tracing the colour evolution can constrain the SFHs of the satellites. This can in turn give some insight into which mechanisms are dominant in driving the observed environmental trends. We present two statistical methods to disentangle the phase space distribution of galaxies and recover the evolution of the satellite colour distribution as a function of time spent in the cluster environment. The first is thought to be superior but is much more expensive computationally and is still being tested and tuned. The second has the advantage in speed but comes with several important caveats. Nevertheless, the results of this second method are suggestive of an initially relatively rapid reddening of blue galaxies upon entering the cluster environment. The reddening may begin around the time that the satellite crosses the virial radius, or potentially even earlier. The reddening appears to slow substantially after a time equal to the time it takes a typical satellite to reach pericentre, suggesting that this is a key time in the life (and death) of a satellite galaxy. It is not clear whether the colour evolution of individual galaxies is rapid or slow. There may be processes driving both rapid and slow colour evolution occurring simultaneously.

5.2 Comparison to Previous Work

Many authors have previously considered trends with cluster-centric radius (colour gradients, for instance), but relatively few have used velocity space information in similar analyses. The velocity distribution for cluster satellites that we find is in good agreement with that of Gill et al. (2005), and like them, we find that satellite populations are better separated in projected phase space than in projected radius alone. The regions of low cross-contamination between satellite populations that we find are a superset of those identified by Mahajan et al. (2011). We build on this result by quantifying the mixture of different satellite populations at every point in phase space. A key point left out by both of these studies is the inclusion of an interloper population in the projected phase space distribution of simulated haloes. This makes separating out the backsplash population substantially more difficult but is crucial to constructing a fair comparison between observation and simulation (unless some independent method to separate out the interlop-

ers can be devised; Knebe et al. (2011) proposes one possible method based on the stellar kinematics and concentrations of the satellites).

Baldry et al. (2004) modelled the $(u-r)$ colour distribution of SDSS galaxies (including satellites, but predominantly field galaxies) with a double gaussian function. We find that this double gaussian fit works equally well in our sample of cluster satellites, but that the bimodality is much less pronounced in $(g-r)$ colour. While it would cost us the ability to trace the colour evolution of the bulge and disk components separately, we feel it may be worth investigating the evolution of $(u-r)$ colour as a function of time since infall; if nothing else, this would give some additional information to be folded into future analyses. Baldry et al. (2004) also notes a strong dependence of the colour distribution on galaxy absolute magnitude (or, approximately, mass). This is a parameter we have yet to explore fully, but it is clearly important.

The key quantities we wish to compare with current measurements are the time at which environment begins to affect a satellite galaxy and the timescale on which quenching processes occur. While our results in this area are still only weakly constrained, we see evidence for (1) environment having an effect on galaxy colour potentially as early as a few Gyr before first pericentre, but certainly before reaching pericentre and (2) colour evolution slowing immediately after the first pericentric passage. Mahajan et al. (2011) argued that star formation quenching is completed in a single passage through the cluster, which seems consistent with our results. In sharp contrast, Wetzel et al. (2012) argues that environment does not even begin to have an effect on star formation in satellites (though the gas supply is of course affected) until several Gyr after crossing R_{200b}, well beyond the first pericentric passage and into the backsplash portion of the orbit. This appears to be in direct contradiction with our results, but we wish to confirm this with the results of our more robust MCMC analysis. Wetzel et al. (2012) also argues that all apparent environmental effects on star formation beyond R_{200b} can be accounted for by backsplash galaxies; in other words, outside this radius, infalling galaxies have yet to be affected by the environment. This is also an interesting claim to verify once results from our more robust MCMC method are available.

5.3 Future Directions

There are several extensions of this work that we believe are worth pursuing. The first (and perhaps most obvious) one is to complete the analysis described in §4.3; our preliminary assessment points to interesting results to come. The logical next step following this is to relate the colour distribution evolution back to its more fundamental counterpart, the

SFH, using population synthesis modelling. During this process, it may be worthwhile to investigate other measures of star formation activity to better constrain the possible star formation histories of individual satellites. Some examples of other indicators that can be used to measure star formation are ultraviolet continuum emission, Hα line emission, far-infrared continuum emission and [OII] line emission (Kennicutt, 1998, gives a review of these and more). The catalogue of Brinchmann et al. (2004) provides SFRs for a large sample ($\sim 10^5$) of SDSS galaxies based on a combination of several such indicators.

So far, our parameterization of orbital history has been very simple, counting the time since infall. It is worth exploring other parameterizations, which may correlate more strongly with SFH. For instance, using the time of first pericentre as a reference time instead of $2.5r_{\rm vir}$ crossing would provide a reference time that is less arbitrarily defined. Our result showing that first pericentre appears to be an important time in the life of a satellite makes this choice even more attractive. However, this means one needs to introduce negative time into the simulation book keeping, e.g. instead of the mean first pericentre occuring at $t = 4$ Gyr, first pericentre occurs at $t = 0$ Gyr and the mean initial crossing of $2.5r_{\rm vir}$ occurs at $t = -4$ Gyr. Furthermore, infalling haloes have not yet had a pericentric passage, so attributing them a time of first pericentric passage is somewhat ill-defined. One way to achieve this in a consistent manner would be to use a simulation that proceeds past $z = 0$, but this is something of a niche application making it difficult to justify the large computational cost of such a simulation.

Another effect that our analysis has thus far completely ignored is "group pre-processing". In principle all the required information to explore the effect of membership in a group environment before accretion into a cluster environment is contained in the merger trees we used to produce our PDFs. One simply needs to replace time of first infall onto a final host above some mass threshold with time of infall onto any host above a given mass. This sustantially complicates the interpretation of the results, motivating our decision to attempt to understand the simpler case first, but it is likely crucial to a complete model of environmental effects on star formation.

We also wish to exploit the superior mass and time resolution of the Bolshoi simulation to extend the dynamic range of our PDFs. Interestingly, Geha et al. (2012) and Teyssier et al. (2012) suggest that for low mass galaxies ($\lesssim 10^9\,{\rm M}_\odot$), the quenching processe is strictly environmentally regulated. This is in contrast to the picture usually invoked for high mass galaxies where both environmental and internal ("mass quenching") mechanisms are thought to be important. This makes the study of such low mass galaxies attractive since the number of processes to disentangle may be much lower. However, observations of large numbers of dwarf galaxies are limited to relatively nearby clusters.

As the volume of high redshift observations continues to grow, applications of our method to these datasets becomes increasingly attractive. At high ($z \sim 1$) redshift, satellites only have time to complete about one orbit in a cluster, which we speculate could substantially simplify the projected phase space distribution of satellites. Star formation rates at high redshift are on average higher than at $z = 0$, so quenching effects may be more pronounced as well.

Finally, we highlight one other possible application of our PDFs. By combining the information on orbital history encoded in the PDFs and gravitational lensing observations, it may be possible to directly probe the effect of environment on the dark matter component of satellite systems.

APPENDICES

Appendix A

Results from the Bolshoi Simulation

The Bolshoi simulation is a higher resolution simulation with the same cosmology, physics and number of particles as the MDR1 simulation used throughout this thesis. It sacrifices a factor of 64 in volume for a factor of 64 improvement in mass resolution and a factor of 7 improvement in force resolution. It also extends to earlier times and has a finer time spacing between its output snapshots; see Table A.1 for a summary of the differences between the two simulations. Because of its close similarity to MDR1 it offers both an excellent way to check for resolution effects in our results and a natural way to extend the applicable mass range of our PDFs.

We process the Bolshoi output in exactly the same way as we did MDR1. We find 592 haloes with a mass $\geq 10^{14}\,\mathrm{M}_{\odot}$, about a factor of 64 less than in MDR1. Before any mass or sanity cuts our catalogue contains 1,053,807 orbits. The mass function for these satellites is shown in Fig. A.1. Note that due to the higher resolution of Bolshoi, we are less

Table A.1: Summary of the differences between the MDR1 simulation and the Bolshoi simulation. Values from Riebe et al. (2011).

Parameter	MDR1	Bolshoi
Box size (h^{-1}Mpc)	1000	250
Number of particles	2048^3	2048^3
Mass resolution ($h^{-1}\,\mathrm{M}_{\odot}$)	8.721×10^9	0.135×10^9
Force resolution (h^{-1}kpc)	7.0	1.0
Initial redshift	65	80
Output snapshots	42	180

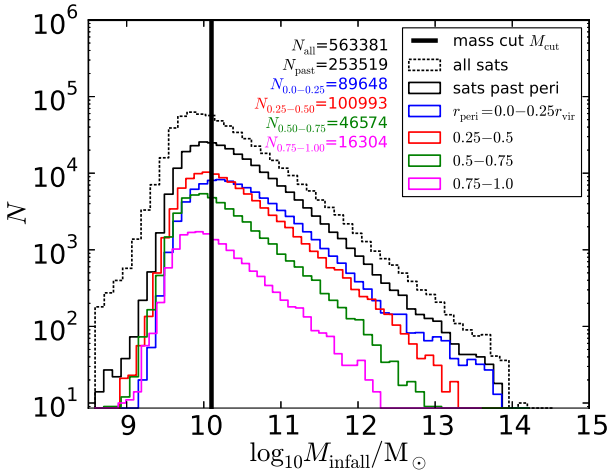

Figure A.1: Mass function of satellites in a sample constructed in the same way as our MDR1 sample (see Fig. 3.1). The mass function of all satellite haloes in our initial sample (dotted black line) and the mass function of satellite haloes that have experienced at least one pericentric passage (solid black line) are shown; in both cases the masses are measured at the time of infall on to their host. The coloured lines separate the satellites that have experienced a pericentric passage into bins of pericentric distance, as labelled. The total number of satellites contributing to each mass function are labelled N. The thick vertical line shows a mass cut 64 times lower than that imposed on the MDR1 satellites ($M_{\rm cut} = 10^{10.1}$ M$_\odot$).

sensitive to the artificial destruction of satellites – there is less flattening of the blue curve. Above the mass cut used in MDR1 ($10^{11.9}\,{\rm M}_\odot$), there is essentially no artificial destruction of satellites, allowing for the possibility of a direct test of the impact of this effect on our results. After a mass cut at $10^{10.1}\,{\rm M}_\odot$ and the same sanity cuts as we used on MDR1, we are left with 481,048 orbits. Our sample of interlopers with masses $\geq 10^{10.1}\,{\rm M}_\odot$ contains 515,000 haloes.

The left panel of Fig. A.2 shows the distribution of satellites in the space of infall time and cluster-centric radius, analogously to Fig. 3.2 which shows the same distribution for the MDR1 simulation. The two distributions are qualitatively very similar, except that in the Bolshoi distribution the oldest virialized satellites have infall times dating back to 1-2 Gyr earlier than those in MDR1. This is because lower mass haloes collapse earlier and therefore have a range of infall times extending to earlier times than high mass haloes, and our mass cut is applied to the infall mass of the haloes. The right panel shows the

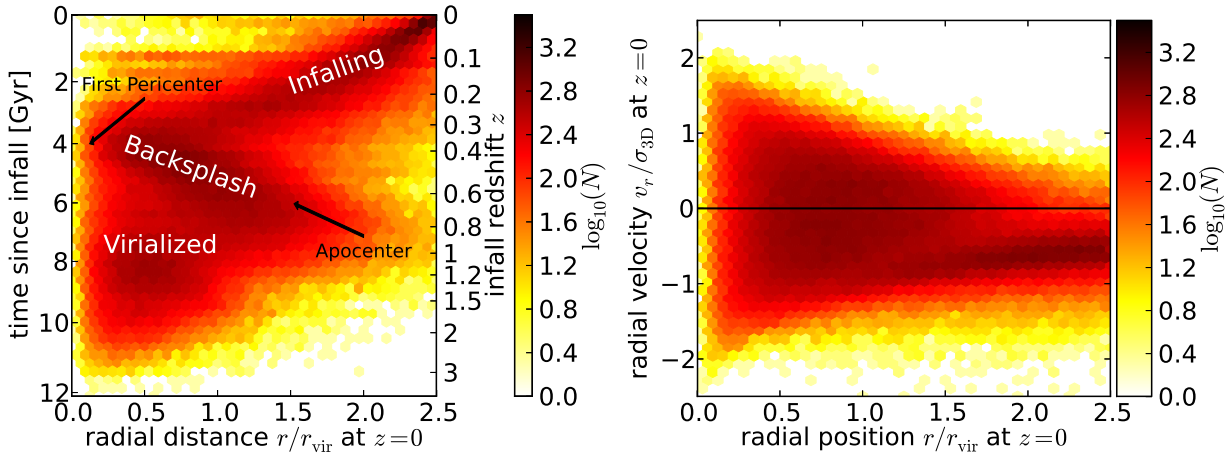

Figure A.2: *Left panel:* Abundance of satellites in bins of cluster infall time, defined as first inward crossing of 2.5 $r_{\rm vir}$, and radial distance from cluster centre at $t = 0$. Compare to MDR1 in left panel of Fig. 3.2. The two distributions are qualitatively very similar, except that in the Bolshoi distribution the oldest virialized satellites have infall times dating back to 1-2 Gyr earlier than those in MDR1. The lines extending horizontally left from the infalling population are artifacts of the binning process and should be ignored. *Right panel:* Phase space distribution of the satellites in Bolshoi. The distribution is qualitatively indistinguishable from that of MDR1 satellites shown in the upper left panel of Fig. 3.3

distribution of satellites in phase space, which is qualitatively indistinguishable from the distribution of MDR1 satellites shown in the upper left panel of Fig. 3.3.

Fig. A.3 shows the projected phase space distribution of Bolshoi satellites in bins of infall time, analogously to Fig. 3.4 for MDR1. Again, the distributions for the two simulations are very similar, except that Bolshoi has a larger relative number of satellites in the bin with the earliest infall times, for the same reasons explained above.

The above discussion serves to show that Bolshoi's higher resolution does not change anything dramatic in the phase space and infall time distributions of the satellites. We would expect to get essentially the same results from Bolshoi with a mass cut at $10^{11.9}\,{\rm M}_\odot$ as from a $(250h^{-1}{\rm Mpc})^3$ sub-volume of MDR1. With a bit of careful calibration, the PDFs we produced from the two simulations can be combined into a single PDF that benefits from both the lower mass resolution limit of Bolshoi and the high cluster count in MDR1.

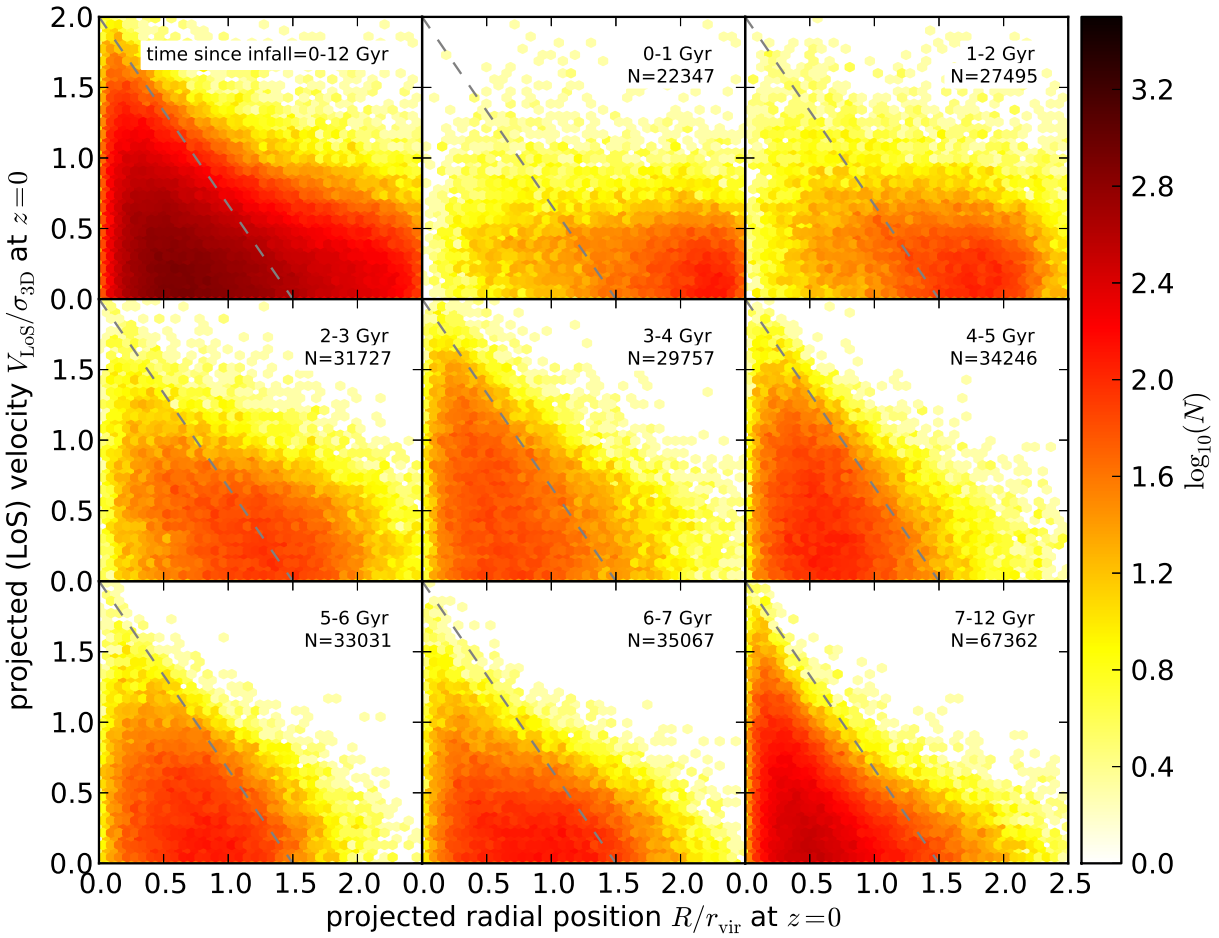

Figure A.3: Projected phase space distribution of Bolshoi satellites in bins of infall time (the top left panel shows the distribution for all times). This distribution is qualitatively identical to the same distribution in MDR1 (Fig. 3.4) except for the bottom right panel which has a higher relative number of satellites in Bolshoi because of the lower mass cut.

Appendix B

Markov Chain Monte-Carlo Analysis

In the framework of Bayesian statistics, the governing principle is Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{B.1}$$

$P(A|B)$ can be read as "the probability of A given B". The left-hand side is called the posterior distribution while on the right hand side is product of the likelihood $P(B|A)$ and the prior distribution $P(A)$, normalized by $P(B)$. Qualitatively, this describes a way in which a new state of knowledge (posterior) can be reached based on a current state of knowledge (prior) and a description of the relative frequency of possible outcomes of a test of that knowledge (likelihood). The normalization is only required if one wishes to discuss absolute probabilities; if relative probabilities are sufficient to the desired analysis (in the context of this thesis, they are), then Bayes theorem expressed as a proportionality between the posterior and the product of the likelihood and the prior is sufficient.

Bayes' theorem has many possibly applications; that of interest here is the estimation of the parameters $\vec{\lambda}$ of a model M based on some data $\vec{D}$. (B.1) is now interpreted as a means to compute the probability that $M(\vec{\lambda})$ is a good description of $\vec{D}$ (posterior) in terms of the probability that $\vec{D}$ would be obtained from a system described by $M(\vec{\lambda})$ (likelihood) and any *a priori* constraints on $M(\vec{\lambda})$ (prior):

$$P(M(\vec{\lambda})|\vec{D}) \propto P(\vec{D}|M(\vec{\lambda}))P(M(\vec{\lambda})) \tag{B.2}$$

If the priors are chosen to give equal weight to all possible parameter values (typically within some interval since the interval $(-\infty, \infty)$ is difficult to deal with computationally) – "flat priors", this describes a maximum-likelihood estimator for the model parameters.

Once the likelihood and priors are suitably defined, all that remains is to evaluate the posterior at every point in parameter space. For small numbers of discrete-valued parameters this may be feasible, but with a large number of continuous parameters, even a coarse sampling across the entire parameter space becomes prohibitively expensive computationally; some other means of estimating the posterior distribution is required. One possible method is the use of Markov chains.

A Markov chain is defined as a sequence of random numbers in which the probability distribution for each element depends only on the value of the previous element in the sequence (and not of a longer history of elements). Applied to a maximum-likelihood estimator, the chain begins at an arbitrary location in parameter space and makes a series of hops to new locations, with each hop being preferentially (but not necessarily) toward positions in parameter space where M fits $\vec{D}$ better than at the current parameter space location. A MCMC maximum likelihood estimator uses multiple Markov chains running in parallel to estimate the variance of the posterior distribution. This can be compared to the variances of the individual chains to determine whether convergence has been achieved.

The MCMC implementation used in the present thesis was constructed using BAT (Caldwell et al., 2009). This toolkit provides all the machinery to perform the analysis, requiring only the definition of a model with a suitably defined likelihood function and list of parameters with their ranges and prior distributions (those we used for our analysis are laid out in §4.3). The algorithm used to compute the chain has two principle elements, a proposal function and an acceptance ratio. The chain is computed in two stages: an initial "pre-run" phase used to calibrate the proposal function and ensure that convergence will be achieved, followed by the actual estimation of the posterior distribution.

The proposal function in BAT is very simple; a proposed state is chosen from a uniform random distribution spanning the entire parameter space – this is the simplest possible proposal function. The proposal function is refined during the pre-run by adjusting the reach of each parameter, so for instance instead of choosing from a uniform random distribution across the full range of a particular parameter, the reach may be adjusted to choose from only half the full range of the parameter, centered around the current parameter value. During the pre-run a new value for each parameter is proposed and accepted or rejected in turn. The ratio of acceptances to proposals for a given parameter defines the "efficiency". The reach of each parameter is adjusted dynamically until all efficiencies are within a desired range (15-50 per cent). A low efficiency is undesireable because it increases the computational expense of producing the chain (evaluating a proposal is expensive; if it is rejected, computation is effectively wasted). A high efficiency is undesireable because it is an indication that the chain may not be sampling the full extent of the posterior distribution – rejections indicate that the chain is exploring the boundaries of the distribution, so

at least a few are desireable.

The acceptance ratio is the expression used to evaluate whether to accept or reject a proposal. It is defined as the ratio of two likelihoods:

$$R = \frac{P(\vec{D}|M(\vec{\lambda}'))}{P(\vec{D}|M(\vec{\lambda}))} \tag{B.3}$$

where $\vec{\lambda}$ represents the current position of the chain in parameter space and $\vec{\lambda}'$ represents the proposed position. In practice the likelihood function is often defined as a product of many terms – one for each data point – which is susceptible to numerical overflow on a computer. To reduce the chance of overflow, it is more practical to work with the log-likelihood, which is simply the logarithm of the likelihood function, and so the product of many terms becomes the sum of the logs of many terms. Once computed, the acceptance ratio is compared to a random number U drawn from a uniform distribution on the inverval $[0, 1]$. The proposal is accepted if $R > U$. The reason for this comparison is worth explaining. If the proposed position is more likely than the current position ($R > 1 \Rightarrow R > U$), the proposal is automatically accepted. If the proposed position is less likely than the current position, it is randomly accepted or rejected, but more likely proposals (R near 1) are accepted more often than less likely proposals (R near 0). This criteria helps ensure that the chain does not get "stuck" indefinitely around a local maximum of the posterior distribution; there is always a chance that a proposal that is unlikely is accepted and takes the chain away to explore another region of parameter space. This particular method is called the Metropolis algorithm.

Once the chain has completed several (in our case 100,000) propose-accept/reject iterations, the density of chain points in parameter space tends toward the probability density of the posterior distribution; an estimate has been constructed. This estimate encodes a wealth of information about the parameters of the model given the data. Not only is a best fit (mode) for each available, but the full probability distribution as well. This allows a straightforward calculation of virtually any desired statistical quantity, for instance confidence intervals for each parameter, a 1σ region about a fit, or correlations between any subset(s) of parameters. The results for each parameter can be evaluated individually or in groups by "marginalizing" over the other parameters if desired. Recalling that the Monte Carlo component of the analysis uses multiple chains in parallel, these allow robust tests of convergence and ergodicity (independence of results from initial conditions).

References

Abadi, M. G., Moore, B., & Bower, R. G. 1999, MNRAS, 308, 947

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, ApJS, 175, 297

Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681

Balogh, M. L., Baldry, I. K., Nichol, R., et al. 2004, ApJ, 615, L101

Balogh, M. L., & McGee, S. L. 2010, MNRAS, 402, L59

Balogh, M. L., Navarro, J. F., & Morris, S. L. 2000, ApJ, 540, 113

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

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013a, ApJ, 770, 57

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013b, ApJ, 762, 109

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., et al. 2013c, ApJ, 763, 18

Bernardi, M., Sheth, R. K., Annis, J., et al. 2003, AJ, 125, 1866

Bertschinger, E. 1985, ApJS, 58, 39

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349

Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, ApJ, 594, 186

Boissier, S., Boselli, A., Duc, P.-A., et al. 2012, A&A, 545, A142

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

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

Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559

Butcher, H., & Oemler, Jr., A. 1978, ApJ, 219, 18

Byrd, G., & Valtonen, M. 1990, ApJ, 350, 89

Caldwell, A., Kollár, D., & Kröninger, K. 2009, Computer Physics Communications, 180, 2197

Cox, T. J., Jonsson, P., Primack, J. R., & Somerville, R. S. 2006, MNRAS, 373, 1013

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11

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

Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2

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

Dressler, A. 1980, ApJ, 236, 351

Dressler, A., Oemler, Jr., A., Sparks, W. B., & Lucas, R. A. 1994, ApJ, 435, L23

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

Font, A. S., Bower, R. G., McCarthy, I. G., et al. 2008, MNRAS, 389, 1619

Gabor, J. M., Davé, R., Finlator, K., & Oppenheimer, B. D. 2010, MNRAS, 407, 749

Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, MNRAS, 355, 819

Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, ApJ, 757, 85

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

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

Gunn, J. E., & Gott, III, J. R. 1972, ApJ, 176, 1

Haines, C. P., Smith, G. P., Egami, E., et al. 2009, ApJ, 704, 126

Haines, C. P., Pereira, M. J., Smith, G. P., et al. 2013, arXiv:1307.1135

Hogg, D. W., Blanton, M. R., Brinchmann, J., et al. 2004, ApJ, 601, L29

Hudson, M. J., Stevenson, J. B., Smith, R. J., et al. 2010, MNRAS, 409, 405

Jáchym, P., Palouš, J., Köppen, J., & Combes, F. 2007, A&A, 472, 5

Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33

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

Kitzbichler, M. G., & White, S. D. M. 2008, MNRAS, 391, 1489

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102

Knebe, A., Libeskind, N. I., Knollmann, S. R., et al. 2011, MNRAS, 412, 529

Koopmann, R. A., & Kenney, J. D. P. 2004, ApJ, 613, 866

Lacey, C., & Cole, S. 1993, MNRAS, 262, 627

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

Lewis, I., Balogh, M., De Propris, R., et al. 2002, MNRAS, 334, 673

Li, I. H., Yee, H. K. C., Hsieh, B. C., & Gladders, M. 2012, ApJ, 749, 150

Limousin, M., Sommer-Larsen, J., Natarajan, P., & Milvang-Jensen, B. 2009, ApJ, 696, 1771

Ludlow, A. D., Navarro, J. F., Springel, V., et al. 2009, ApJ, 692, 931

Mahajan, S., Mamon, G. A., & Raychaudhury, S. 2011, MNRAS, 416, 2882

Mahajan, S., Raychaudhury, S., & Pimbblet, K. A. 2012, MNRAS, 427, 1252

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

Maraston, C. 2005, MNRAS, 362, 799

Mayer, L., Mastropietro, C., Wadsley, J., Stadel, J., & Moore, B. 2006, MNRAS, 369, 1021

McCarthy, I. G., Frenk, C. S., Font, A. S., et al. 2008, MNRAS, 383, 593

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

McGee, S. L., Balogh, M. L., Wilman, D. J., et al. 2011, MNRAS, 413, 996

McNamara, B. R., Rafferty, D. A., Bîrzan, L., et al. 2006, ApJ, 648, 164

Merritt, D. 1983, ApJ, 264, 24

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution (Cambridge University Press)

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

Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903

Noble, A. G., Webb, T. M. A., Muzzin, A., et al. 2013, ApJ, 768, 118

Oman, K. A., Hudson, M. J., & Behroozi, P. S. 2013, MNRAS, 431, 2307

O'Shea, B. W., Nagamine, K., Springel, V., Hernquist, L., & Norman, M. L. 2005, ApJS, 160, 1

Ostriker, J. P., & Hausman, M. A. 1977, ApJ, 217, L125

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E., & Primack, J. 2012, MNRAS, 423, 3018

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

Riebe, K., Partl, A. M., Enke, H., et al. 2011, arXiv:1109.0003

Rocha, M., Peter, A. H. G., & Bullock, J. 2012, MNRAS, 425, 231

Simard, L. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse, 108

Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11

Smith, R., Davies, J. I., & Nelson, A. H. 2010a, MNRAS, 405, 1723

Smith, R. J., Lucey, J. R., Hammer, D., et al. 2010b, MNRAS, 408, 1417

Springel, V. 2010, MNRAS, 401, 791

Strateva, I., Ivezić, Ž., Knapp, G. R., et al. 2001, AJ, 122, 1861

Taranu, D. S., Hudson, M. J., Balogh, M. L., et al. 2012, arXiv:1211.3411

Teyssier, M., Johnston, K. V., & Kuhlen, M. 2012, MNRAS, 426, 1808

Tinker, J. L., Sheldon, E. S., Wechsler, R. H., et al. 2012, ApJ, 745, 16

Tormen, G., Diaferio, A., & Syer, D. 1998, MNRAS, 299, 728

Tovmassian, H. M., & Andernach, H. 2012, MNRAS, 427, 2047

van den Bosch, F. C., Aquino, D., Yang, X., et al. 2008, MNRAS, 387, 79

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

Velander, M., van Uitert, E., Hoekstra, H., et al. 2013, arXiv:1304.4265

von der Linden, A., Best, P. N., Kauffmann, G., & White, S. D. M. 2007, MNRAS, 379, 867

Wang, H., Mo, H. J., Jing, Y. P., Yang, X., & Wang, Y. 2011, MNRAS, 413, 1973

Wang, H. Y., Jing, Y. P., Mao, S., & Kang, X. 2005, MNRAS, 364, 424

Warnick, K., Knebe, A., & Power, C. 2008, MNRAS, 385, 1859

Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, ApJ, 568, 52

Weinmann, S. M., Kauffmann, G., van den Bosch, F. C., et al. 2009, MNRAS, 394, 1213

Weinmann, S. M., Kauffmann, G., von der Linden, A., & De Lucia, G. 2010, MNRAS, 406, 2249

Wetzel, A. R. 2011, MNRAS, 412, 49

Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, MNRAS, 424, 232

Wetzel, A. R., Tinker, J. L., Conroy, C., & van den Bosch, F. C. 2013a, MNRAS, 432, 336

—. 2013b, arXiv:1303.7231

Xu, G. 1995, ApJS, 98, 355