DISSERTATION

SUBMITTED TO THE

COMBINED FACULTIES OF THE NATURAL SCIENCES AND MATHEMATICS

OF THE RUPERTO-CAROLA-UNIVERSITY OF HEIDELBERG, GERMANY

FOR THE DEGREE OF

DOCTOR OF NATURAL SCIENCES

DISSERTATION

PUT FORWARD BY

SURHUD SHRIKANT MORE

BORN IN: ALIBAG, INDIA

ORAL EXAMINATION: JULY 27^{th}, 2009

GALAXY-DARK MATTER CONNECTION

INSIGHTS FROM SATELLITE KINEMATICS

REFEREES: PROF. DR. HANS-WALTER RIX
PROF. DR. FRANK C. VAN DEN BOSCH

असतो मा सद्गमय तमसो मा ज्योतिर्गमय।

From ignorance, lead me to the truth, from darkness lead me to light

माझी आजी

कै. सुलोचना सखाराम सावंत

हिला समर्पित

Dedicated to the fond memory of my late grandmother

Kai. Sulochana Sakharam Sawant

Abstract

A thorough knowledge of the connection between the mass of dark matter haloes and the properties of their central galaxies is crucial to understand the physics of galaxy formation. The kinematics of satellite galaxies is an excellent technique to measure the dark matter halo masses. However, the kinematics can be measured with high signal-to-noise only by stacking the signal around central galaxies with similar properties, which results in various systematic biases and complicates the interpretation of the signal. This thesis presents an analytical framework that accounts for systematic biases and selection effects and aids in the interpretation of the kinematics of satellite galaxies. A new method is established to obtain the average scaling relations between halo mass and central galaxy properties, and the scatter in these relations *simultaneously*. After a thorough testing of this method using a realistic mock galaxy catalogue, it is applied to the Sloan Digital Sky Survey to extract the halo mass-luminosity and halo mass-stellar mass relationship of central galaxies and their scatter. Comparisons with other probes of these scaling relations, such as galaxy−galaxy lensing, show good agreement which implies that these scaling relations are well established and supported by various astrophysical probes. Physical insights about these scaling relations, in particular their scatter, gained by the analysis of a semi-analytical model for galaxy formation are also presented. Finally, the inferred scaling relations crucially depend on the transparency of the Universe. By performing a test of the "Etherington relation" between the distances measured by standard rulers and by standard candles, a quantitative measure of the cosmic transparency, which is relatively free from astrophysical assumptions, is obtained.

Zusammenfassung

Um die physikalischen Mechanismen der Galaxienentwicklung zu verstehen, müssen die Zusammenhänge zwischen der Masse eines Dunkle-Materie-Halos und den Eigenschaften seiner Zentralgalaxie bekannt sein. Die Bestimmung der Kinematik von Satellitengalaxien ist eine etablierte Methode zur Messung von Halomassen. Allerdings sind kinematische Messungen mit hoher Signifikanz nur dadurch möglich, dass die Signale über viele Zentralgalaxien mit ähnlichen Eigenschaften gemittelt werden. Dies hat systematische Fehler zur Folge und erschwert die Interpretation kinematischer Beobachtungen. In dieser Dissertation wird ein analytisches Modell vorgestellt, mit dessen Hilfe systematische Verzerrungen und Auswahleffekte korrigiert werden können, was die Auswertung kinematischer Daten von Satellitengalaxien erleichtert. Hieraus folgt ein Verfahren zur Bestimmung der Skalierungsvorschrift zwischen Halomasse und Eigenschaften der Zentralgalaxie, die *gleichzeitig* die Streuung über diese Relation

vorhersagt. Diese Gültigkeit dieses Verfahrens wird anhand eines künstlichen Galaxienkatalogs mit bekannten, aber realistischen Eigenschaften verifiziert. Aus der Anwendung auf die Beobachtungen der Sloan Digital Sky Survey werden die Skalierungsvorschriften zwischen Halomasse einerseits und Leuchtkraft sowie stellarer Masse andererseits für Zentralgalaxien mitsamt dazugehöriger Streuung bestimmt. In Vergleichen zeigt sich eine gute übereinstimmung mit anderen Messverfahren für diese Vorschriften, wie z.B. der Beobachtung von Gravitationslinseneffekten der Zentralgalaxien auf andere Galaxien. Dies bestätigt die Zuverlässigkeit der hier vorgestellten Skalierungen. Darüber hinaus ergeben sich aus der Analyse eines semiempirischen Modells für Galaxienentstehung neue Erkenntnisse über die physikalischen Ursachen der vorgestellten Skalierungen und insbesonderere ihrer Streuung. Schlielich hängen die aus unserem Modell erhaltenen Skalierungen entscheidend von der Transparenz des Universums ab. Vermittels einer überprüfung des "Etherington'schen Gesetzes" für den Zusammenhang zwischen Entfernungen, die aus Standardlängen beziehungsweise Standardkerzen bestimmt werden, wird die kosmische Transparenz in einer Weise bestimmt, die relativ unabhängig von astrophysikalischen Annahmen ist.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Galaxy Formation in a Dark Universe

The vast ocean of space is full of starry islands called galaxies, such as our own Milky Way. Galaxies act as lighthouses in this vast ocean, serving as an interface with which we can explore and understand our Universe. The majesty and the variety of galaxies has often boggled the human mind. It is more curious however that, in our current understanding, galaxies form a very small portion of the energy content of the Universe. Most of the energy content of the Universe today is "dark" – the two dominant components "dark energy" and "dark matter" account for nearly 95 percent of the energy density of the Universe. The rest is ordinary matter primarily present in the form of gas in the intergalactic medium and around galaxies (referred to as baryons). How do galaxies come into existence in this dark Universe and how do they evolve? What is the relation of galaxies to the dark components in the Universe? What shapes the properties of different galaxies? How are different properties of galaxies correlated with each other and what is the physics that drives these correlations? These questions, among others, currently drive the research field of galaxy formation and evolution.

The origin and the nature of both the dark components of the Universe is still a mystery. Although the presence of dark matter can be motivated theoretically from (currently untested) ideas in particle physics that are based upon supersymmetry (Preskill et al. 1983; Ellis et al. 1984), the presence of dark energy and its ubiquitous nature has little theoretical motivation (see e.g., Dolgov 2004). The evidence for the presence of both the components is purely astrophysical. Dark energy manifests itself through the recently-discovered accelerated expansion of the Universe (Riess et al. 1998; Perlmutter et al. 1999; Kowalski et al. 2008) while dark matter makes its presence felt only through its gravitational effects (Zwicky 1933; Rubin et al. 1982). All attempts to detect the elusive dark matter through a wide range of non-gravitational experiments have largely been inconclusive (see e.g., Benoit et al. 2002; Akerib et al. 2003, 2004; Sanglard et al. 2005). These negative results have led some to favour the radical approach of modifying Newton's theory of gravity in the weak field limit (Milgrom 1983a,b,c) and its relativistic version (Bekenstein 2004). Whether this approach can explain all the observed as-

trophysical phenomena which suggest the presence of dark matter – the jury is still out (see e.g., McGaugh & de Blok 1998; Sanders & McGaugh 2002; Sanders 2003; Clowe et al. 2004; Klypin & Prada 2009).

In the simplest picture, that conforms to a wide range of astrophysical observations, dark energy is assumed to be an all-pervading, non-clumpy form of energy and is generally attributed to the vacuum. Dark matter, on the other hand, is supposed to be dynamically cold and collisionless, but clumpy due to the effects of gravity. These two dark components form the backbone of the ΛCDM theory (Λ stands for dark energy, CDM for cold dark matter). According to this theory, the early Universe started off as a dense hot soup of elementary particles and underwent a rapid inflationary phase where the tiny fluctuations of a (hypothesized scalar) quantum field were stretched to cosmologically large scales. These fluctuations were imprinted onto the initial density field of the particles. Dark matter, the most abundant gravitationally unstable component in the Universe, was then responsible for the formation of structure in the Universe. The tiny initial fluctuations in the density field grew over time by the action of gravity and formed bound structures (haloes). Baryons were trapped within the gravitational potential of these dark matter haloes and they underwent a series of complex physical processes to form the galaxies that we observe today.

The great thing about the ΛCDM theory is its ability to make testable predictions. Given the power spectrum of the initial density fluctuations and the energy density parameters of the various components of the Universe, the statistical properties of the dark matter distribution can be accurately predicted (see e.g., Eisenstein & Hu 1999). As the dark matter distribution is not directly observable, establishing the link between the observable galaxies to their dark matter haloes is central to test this prediction (see e.g., Tegmark et al. 2004). The theory also predicts that the properties of dark matter haloes, in particular the mass, should shape the properties of galaxies that form within them. Precise measurements of the scaling relations between different galaxy properties with the mass of a dark matter halo can also provide key insights into the galaxy-dark matter connection predicted by this theory. This is precisely the aim of this thesis. We wish to investigate the connection between different galaxy properties and their dark matter haloes.

There are various approaches to study the galaxy-dark matter connection. One approach is to study such a connection via direct numerical simulations (see e.g., Katz et al. 1992; Evrard et al. 1994; Frenk et al. 1996, 1999; Katz et al. 1996; Navarro et al. 1997; Pearce et al. 1999; Kravtsov 1999). This involves following the evolution of the density field and its fluctuations, and the various astrophysical processes that transform the baryons into luminous galaxies. In this approach, the gravitational and hydrodynamical equations need to be solved in full generality. The resultant population of galaxies can then be compared with the observed population of galaxies in the Universe. One drawback of this approach is the tremendous computational expense of simulating a cosmologically meaningful volume with sufficient resolution and within a reasonable amount of time.

The second approach, called semi-analytical modelling, improves upon the former by separating the evolution of the dark matter component and the baryonic component (see e.g., White & Rees 1978; White & Frenk 1991). The evolution of the dark matter component is followed numerically (or by using Monte-Carlo techniques) while the evolution of the baryonic component in the distribution of dark matter is followed by using simple analytical recipes. This approach of modelling the formation of galaxies can be used to compute the properties of a large population of galaxies and establish their link to the underlying dark matter distribution (e.g., White & Frenk 1991; Kauffmann & White 1993; Cole et al. 1994; Kauffmann 1996; Kauffmann et al. 1997; Baugh et al. 1998; Somerville & Primack 1999; Cole et al. 2000; Benson et al. 2002; Springel et al. 2005; Croton et al. 2006; De Lucia & Blaizot 2007). The advantage of this method is its flexibility. It is relatively easy to test the effects of the various assumptions and parameters involved in the modelling on the final properties of the modelled galaxies (see e.g., Cole et al. 2000). The first approach is complementary to the semi-analytical approach, because the simple analytical recipes often have to be calibrated against high resolution hydrodynamical simulations which focus on a small volume. One drawback of the semi-analytical approach is that any back reaction of the baryons on the dark matter haloes are either neglected or are included a posteriori.

The third approach to investigate the link between galaxies and their dark matter haloes is statistical in its nature. In this approach the connection between galaxies and dark matter haloes is specified by a halo occupation model (for an excellent review, see Cooray & Sheth 2002). The model uses a few parameters to specify the distribution of various properties of galaxies as a function of the mass of the halo in which they reside. Given the properties of the dark matter haloes, such as their abundance, their clustering strength and their density profiles (usually obtained from dark matter only numerical simulations), these models can be easily used to make analytical predictions for various observational properties of galaxies. The observed properties of the real-world galaxies can then be used to constrain the parameters of the halo occupation model and thus establish the link between galaxies and their dark matter haloes (Bullock et al. 2002; Berlind & Weinberg 2002; Berlind et al. 2003; Wang et al. 2004; Abazajian et al. 2005; Zheng et al. 2005; van den Bosch et al. 2007; Zheng et al. 2007; Cacciato et al. 2009). In this thesis, we will focus on this third approach and infer the halo occupation distribution of galaxies by probing the dark matter haloes of galaxies.

Various observational probes can be used to measure the masses of dark matter haloes and subsequently connect it to the properties of the galaxies. These include various methods that focus on measuring the kinematics of a tracer population in the halo around individual galaxies. For example, kinematic measurements of the gas around spiral galaxies observed in the optical and the radio wavelengths have been traditionally used as a strong evidence for the presence of dark matter haloes (e.g., Shostak 1973; Roberts & Whitehurst 1975; Bosma 1978; Rubin et al. 1978, 1982; Sofue & Rubin 2001). Unfortunately, these probes do not trace the entire extent of the dark matter halo and hence can at best be used to measure the masses of only the inner

regions of the haloes. The X-ray emission from hot gas in clusters can also be used to measure the dark matter halo masses in these systems under simplifying assumptions of hydrostatic equilibrium and spherical symmetry (e.g., Mushotzky et al. 1978). Strong gravitational lensing, manifested by the presence of multiple images or highly magnified arcs of background objects, is yet another important probe of the halo masses in individual systems (see e.g., Schneider et al. 1992). Being a purely gravitational effect, lensing has the advantage of being able to probe mass without the need for simplifying assumptions about the relaxedness of the system under consideration. However, degeneracies in the modelling of lens systems and projection effects can cause some trouble in the interpretation of the lensing observations.

With the advent of large scale galaxy redshift surveys in the last decade, such as the the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Two degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001), new methods have been developed to investigate the galaxy-dark matter connection. These methods do not focus on individual systems but rather examine the statistical properties of the galaxy distribution to infer the halo mass properties on average. The galaxy redshift surveys can be used to reliably determine the abundance of galaxies as a function of their properties such as luminosity or stellar mass (e.g., Norberg et al. 2002b; Bell et al. 2003; Blanton et al. 2003b; Panter et al. 2004). Another statistical property is the clustering of the galaxy distribution, given by the two-point correlation function measured as a function of galaxy properties (e.g., Zehavi et al. 2002; Norberg et al. 2002a; Madgwick et al. 2003; Zehavi et al. 2004, 2005; Wang et al. 2007). Since the abundance and clustering of dark matter haloes is a function of the mass of the halo, the abundance and clustering of galaxies can also be used to constrain the halo occupation distribution of galaxies (e.g., Jing & Suto 1998; Peacock & Smith 2000; Bullock et al. 2002; Berlind & Weinberg 2002; Wang et al. 2004; Abazajian et al. 2005; van den Bosch et al. 2007; Zheng et al. 2007; Cacciato et al. 2009).

The dark matter haloes around galaxies can cause weak tangential distortions in the shapes of background galaxies. This effect known as galaxy-galaxy lensing is yet another statistical way to probe the dark matter haloes around galaxies (e.g., Tyson 1987; Brainerd et al. 1996; dell'Antonio & Tyson 1996; Hudson et al. 1998; Hirata et al. 2004; Mandelbaum et al. 2006b). Measurements of the galaxy−galaxy lensing signal can be used to constrain the properties of dark matter haloes around galaxies (e.g., Schneider & Rix 1997; Wilson et al. 2001; Guzik & Seljak 2002; Parker et al. 2007; Cacciato et al. 2009). Since this effect is very weak, a stacking procedure has to be adopted in which the signal around galaxies of similar properties is added to improve the signal-to-noise ratio. Such a measurement then probes the average dark matter halo of galaxies as a function of the property used to stack the galaxies.

In this thesis, we scrutinize another powerful method that involves measuring the kinematics of satellite galaxies that orbit the dark matter haloes of central galaxies to measure the masses of these haloes. Satellite galaxies trace the dark matter halo in its entirety and hence are useful to probe the dark matter haloes around galaxies. This method is of historical significance as its application to the Coma cluster of galaxies had led to the discovery of dark matter (Zwicky

1933). Precise measurement of the kinematics of satellite galaxies is only possible in systems, such as clusters, that have a large number of satellites (e.g., Carlberg et al. 1996, 1997). The number of satellites in low mass systems is too small to provide a reliable measure of the kinematics. However, stacking methods were soon pioneered that enabled the measurement of the kinematics of satellites of central galaxies in low mass haloes stacked by their luminosities (Erickson et al. 1987; Zaritsky et al. 1993; Zaritsky & White 1994; Zaritsky et al. 1997). These studies involved a modest number of satellite galaxies ($\lesssim 100$), but were nevertheless succesful in establishing the presence of extended dark matter haloes around spiral galaxies.

The sample of satellite galaxies used to measure the kinematics received an order of magnitude boost in number after data from large scale redshift surveys became available (McKay et al. 2002). This has led to a number of interesting studies that have measured the scaling relations between galaxy properties and their dark matter haloes (Brainerd & Specian 2003; Prada et al. 2003; van den Bosch et al. 2004; Becker et al. 2007; Conroy et al. 2007; Norberg et al. 2008). Qualitatively all studies agree on the fact that the velocity dispersion of satellites correlates positively with the property (luminosity/stellar mass) used to stack central galaxies which in turn implies that the mass of dark matter haloes increases as a function of the stacking property. However, there are quantitative disagreements about the exact scaling relations that are inferred by these studies. The contrast between the results of different studies was recently highlighted in Norberg et al. (2008). The differences between various studies were attributed to the different criteria used to select the samples of centrals and satellites by the previous studies.

The research work presented in this thesis aims to understand how selection effects and systematic biases affect the kinematics of satellite galaxies and establish a new method based on satellite kinematics that can be used to measure the scaling relations between dark matter haloes and central galaxies in an unbiased manner. We also discuss the implications of these results on the physics of galaxy formation.

1.2 Thesis Overview

Chapter 2 introduces the reader to the theoretical framework that can be used to analyse the kinematics of satellites around central galaxies stacked according to their properties. In this chapter, we present a degeneracy problem that has been hitherto ignored in the analysis of satellite kinematics. We show that the kinematics of satellite galaxies cannot be used to infer a unique relation between halo masses and the property used to stack central galaxies if the scatter in this relation is unknown. In this chapter, we also present a novel method that has the potential to break this degeneracy and measure both the average scaling relation and its scatter from the kinematics of satellite galaxies.

In Chapter 3, we test the feasibility of the application of the method presented in Chapter 2 to realistic galaxy surveys. For this purpose, we first create an artificial universe by populating galaxies in a dark matter only simulation. By mimicking the flux limited nature of galaxy

observations in redshift surveys, we create a mock galaxy catalogue from this artificial universe. This mock catalogue is then used to test how various selection effects affect the measurement of the kinematics of satellite galaxies. We also show that the new method we proposed can reliably recover both the average scaling relation between halo masses and luminosity and its scatter originally present in the mock catalogue.

The next two chapters deal with the application of these methods to actual data from the SDSS. Chapter 4 focusses on the halo mass−luminosity relationship of central galaxies and infer this relation from the kinematics of satellite galaxies. The focus of Chapter 5 is the halo mass-stellar mass relationship of central galaxies. We show that the average scaling relations derived by our method are in excellent agreement with several other probes of these relations. We find that both the relations demonstrate an appreciable scatter and present quantitative measurements of the same. The scatter in these relations is a result of the stochasticity in galaxy formation.

In Chapter 6, we attempt to gain physical insights on the origin of the stochasticity in galaxy formation that we constrained using satellite kinematics. For this purpose, we use a semi-analytical model of galaxy formation. We analyse the scatter in the merger histories of haloes of similar masses and explore its effects on the properties of the galaxies that form at its center. We quantify the merger histories of haloes by considering various definitions for the formation times of haloes and show that haloes that form earlier on average host central galaxies that have a larger stellar mass.

The properties of galaxies that we observe such as their luminosity are often based upon the assumption that the Universe is transparent. In Chapter 7, we obtain a quantitative measure of the transparency of the local Universe in the optical bands. We perform a test of the "Etherington relation" by checking for the consistency of the luminosity distances obtained by supernova Ia experiments and the angular diameter distances obtained by experiments that detect the baryon accoustic feature in the power spectrum of galaxies. Note that, such measures of the transparency of the universe are important given the fact that the only evidence of the presence of dark energy is the dimming of distant supernovae in the Universe and this could be mimicked by the presence of opacity in the Universe.

Finally, in Chapter 8, a summary of the results obtained in this thesis is presented with a short discussion on the possibilities of future work in this field.

Chapter 2

Satellite Kinematics: The Analytical Formalism

The contents of this chapter are based upon the article More et al. (2009b) published in the Monthly Notices of the Royal Astronomical Society. The reference is
More, S., van den Bosch, F. C., & Cacciato, M. 2009, MNRAS, 392, 917.
The introduction from the article published above has been slightly modified to avoid repetition of material from the introductory chapter in the thesis.

2.1 Introduction

According to the current paradigm, the mass of a dark matter halo is believed to strongly influence the process of galaxy formation and thus shape the properties of the galaxies that form and reside at their centres (hereafter referred to as central galaxies). Hence, a reliable determination of scaling relations between halo mass and properties of their central galaxies can provide important constraints on the physics of galaxy formation. Determination of such scaling relations require precise measurements of the halo mass. Numerous methods are available to probe the masses of dark matter haloes (see Chapter 1). In this thesis, we scrutinize in detail the method which uses the kinematics of satellite galaxies that orbit within the halo of their central galaxies, to measure the masses of dark matter haloes.

The kinematics of satellite galaxies in individual cluster-sized haloes can be reliably measured as they host a large number of satellite galaxies which properly sample the line-of-sight (hereafter los) velocity distribution of their haloes. The extension of this analysis to group-scale and galaxy-scale haloes necessitates the use of stacking methods (Erickson et al. 1987; Zaritsky et al. 1993; Zaritsky & White 1994; Zaritsky et al. 1997). Under the assumption that galaxies with similar properties (e.g. luminosities) reside in haloes of similar mass, these methods combine the velocity information of satellite galaxies that revolve around such central galaxies. The kinematics of such a stacked system is then used to infer the average halo mass of the

stacked central galaxies. More recent studies (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; van den Bosch et al. 2004; Conroy et al. 2007; Becker et al. 2007) apply similar stacking procedures to central galaxies selected from the large homogeneous galaxy redshift surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Two degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001). All these studies find that the los velocity dispersion of satellite galaxies, $\sigma_{\rm sat}$, increases with the luminosity of the host (central) galaxy, $L_{\rm c}$. This is in agreement with the expectation that more massive haloes host more luminous centrals. In a recent study, Norberg et al. (2008) have shown that there exist quantitative discrepancies between these previous studies and these discrepancies arise mainly due to the differences in the criteria used to select central hosts and their satellites. This underscores the necessity for a careful treatment of selection effects in order to extract reliable mass estimates from satellite kinematics.

Except for van den Bosch et al. (2004), all previous studies have been extremely conservative in their selection of hosts and satellites. Consequently, despite the fact that the redshift surveys used contain well in excess of 100,000 galaxies, the final samples only contained about $2000 - 3000$ satellite galaxies. This severely limits the statistical accuracy of the velocity dispersion measurements as well as the dynamic range in luminosity of the central galaxies for which halo masses can be inferred. The main motivation for using strict selection criteria is to select only 'isolated' systems, with satellites that can be treated as tracer particles (i.e., their mass does not cause significant perturbations in the gravitational potential of their host galaxy). Let $P(M|L_{\rm c})$ denote the conditional probability distribution that a central galaxy of luminosity $L_{\rm c}$ resides in a halo of mass M. If the scatter in $P(M|L_{\rm c})$ is sufficiently small, preferentially selecting 'isolated' systems should yield an unbiased estimate of $\langle M\rangle(L_{\rm c})$, which is the first moment of $P(M|L_{\rm c})$. However, very little is known about the actual amount of scatter in $P(M|L_{\rm c})$ and different semi-analytical models for galaxy formation make significantly different predictions (see discussion in Norberg et al. 2008). If appreciable, the scatter will severely complicate the interpretation of satellite kinematics, and may even cause a systematic bias (van den Bosch et al. 2004; More et al. 2009c). Furthermore, even if the scatter is small, in practice, satellites of central galaxies stacked in finite bins of luminosity are used to measure the kinematics. If the satellite sample is small, one has to resort to relatively large bins in order to have sufficient signal-to-noise. Therefore, even if the distribution $P(M|L_{\rm c})$ is relatively narrow, this still implies mixing the kinematics of haloes spanning a relatively large range in halo masses.

In this chapter, we demonstrate that whenever the scatter in $P(M|L_{\rm c})$ is non-negligible, the $\sigma_{\rm sat}(L_{\rm c})$ inferred from the data has to be interpreted with great care. In particular, we demonstrate that there is a degeneracy between the first and second moments of $P(M|L_{\rm c})$, in that two distributions with different $\langle M\rangle(L_{\rm c})$ and different scatter can give rise to the same $\sigma_{\rm sat}(L_{\rm c})$. Therefore, a unique $\langle M\rangle(L_{\rm c})$ cannot be inferred from satellite kinematics without a prior knowledge of the second moment of $P(M|L_{\rm c})$. However, not all hope is lost. In fact, we demonstrate that by using two different methods to measure $\sigma_{\rm sat}(L_{\rm c})$, one can actually break

this degeneracy and thus constrain both the mean and the scatter of $P(M|L_{\rm c})$. In this chapter we introduce the methodology, and present the analytical framework required to interpret the data, taking account of the selection criteria used to identify the central host galaxies and their satellites. In Chapter 4, we apply this method to the SDSS to infer both the mean and the scatter of $P(M|L_{\rm c})$, which we show to be in good agreement with the results obtained from clustering and galaxy−galaxy lensing analyses. In addition, in Chapter 4 we demonstrate that (i) the scatter in $P(M|L_{\rm c})$ can not be neglected, especially not at the bright end, and (ii) the strict isolation criteria generally used to select centrals and satellites result in a systematic underestimate of the actual $\langle M\rangle(L_{\rm c})$.

This chapter is organized as follows. In Section 2.2, we present two different schemes to measure the velocity dispersion, the satellite-weighting scheme and the host-weighting scheme. In Section 2.3, we present a toy model which serves as a basis for understanding the dependence of velocity dispersion estimates on the different parameters of interest. In Sections 2.4 and 2.5 we refine our toy model by including selection effects and by using a realistic halo occupation distribution (HOD) model for the central galaxies. We use these more realistic models to investigate how changes in the halo occupation statistics of central galaxies affect the velocity dispersion of satellite galaxies, and we demonstrate how the combination of satellite-weighting and host-weighting can be used to infer both the mean and the scatter of the mass−luminosity relation. We summarize our findings in Section 2.6. Throughout this chapter, M denotes the halo mass in units of $h^{-1}{\rm M}_{\odot}$.

2.2 Weighting Schemes

In order to estimate dynamical halo masses from satellite kinematics one generally proceeds as follows. Using a sample of satellite galaxies, one determines the distribution $P(\Delta V)$, where ΔV is the difference in the line-of-sight velocity of a satellite galaxy and its corresponding central host galaxy. The scatter in the distribution $P(\Delta V)$ (hereafter the velocity dispersion), is then considered to be an estimator of the depth of the potential well in which the satellites orbit, and hence of the halo mass associated with the central. In order to measure the velocity dispersion as a function of central galaxy luminosity, $\sigma_{\rm sat}(L_{\rm c})$, with sufficient signal-to-noise, one has to combine the los velocity information of satellites which belong to centrals of the same luminosity, $L_{\rm c}$. This procedure is influenced by two effects, namely *mass-mixing* and *satellite-weighting*, which we now discuss in turn.

Mass-mixing refers to combining the kinematics of satellites within haloes of different masses. The mass−luminosity relation (hereafter MLR) of central galaxies can have an appreciable scatter, i.e., the conditional probability distribution $P(M|L_{\rm c})$ is not guaranteed to be narrow. In this case, the satellites used to measure $\sigma_{\rm sat}(L_{\rm c})$ reside in halo masses drawn from this distribution, and $\sigma_{\rm sat}(L_{\rm c})$ has to be interpreted as an average over $P(M|L_{\rm c})$.

In most studies to date, the technique used to measure $\sigma_{\rm sat}(L_{\rm c})$ implies satellite weighting.

This can be elucidated as follows. Let us assume that one stacks $N_{\rm c}$ central galaxies, and that the $j^{\rm th}$ central has N_j satellites. The total number of satellites $N_{\rm sat}$ is given by $\sum_{j=1}^{N_{\rm c}} N_j$. Let ΔV_{ij} denote the los velocity difference between the $i^{\rm th}$ satellite and its central galaxy j. The average velocity dispersion of the stacked system, $\sigma_{\rm sw}$, is such that

$$\sigma^2_{\rm sw} = \frac{\sum_{j=1}^{N_{\rm c}} \sum_{i=1}^{N_j} (\Delta V_{ij})^2}{\sum_{j=1}^{N_{\rm c}} N_j} = \frac{1}{N_{\rm sat}} \sum_{j=1}^{N_{\rm c}} N_j \sigma_j^2 \,. \tag{2.1}$$

Here σ_j is the velocity dispersion in the halo of the $j^{\rm th}$ central galaxy. The velocity dispersion measured in this way is clearly a satellite-weighted average of the velocity dispersion σ_j around each central galaxy [1]. Although not necessarily directly using Eq. (2.1), most previous studies have adopted this satellite-weighting scheme (McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; Norberg et al. 2008).

In principle, the satellite-weighting can be undone by introducing a weight $w_{ij} = 1/N_j$ for each satellite−central pair in the los velocity distribution (van den Bosch et al. 2004; Conroy et al. 2007). The resulting *host-weighted* average velocity dispersion, $\sigma_{\rm hw}$, is such that

$$\sigma^2_{\rm hw} = \frac{\sum_{j=1}^{N_{\rm c}} \sum_{i=1}^{N_j} w_{ij} (\Delta V_{ij})^2}{\sum_{j=1}^{N_{\rm c}} w_{ij} N_j} = \frac{1}{N_{\rm c}} \sum_{j=1}^{N_{\rm c}} \sigma_j^2 \,, \tag{2.2}$$

and it gives each halo an equal weight.

Consider a sample of central and satellite galaxies with luminosities $L > L_{\rm min}$. The velocity dispersions in the satellite-weighting and host-weighting schemes can be analytically expressed (see also van den Bosch et al. 2004) as follows:

$$\sigma^2_{\rm sw}(L_{\rm c}) = \frac{\int_0^\infty P(M|L_{\rm c}) \, \langle N_{\rm sat} \rangle_M \, \langle \sigma^2_{\rm sat} \rangle_M \, {\rm d}M}{\int_0^\infty P(M|L_{\rm c}) \, \langle N_{\rm sat} \rangle_M \, {\rm d}M} \,, \tag{2.3}$$

$$\sigma^2_{\rm hw}(L_{\rm c}) = \frac{\int_0^\infty P(M|L_{\rm c}) \, \langle \sigma^2_{\rm sat} \rangle_M \, {\rm d}M}{\int_0^\infty P(M|L_{\rm c}) \, {\rm d}M} \,. \tag{2.4}$$

Here $\langle N_{\rm sat} \rangle_M$ denotes the average number of satellites with $L > L_{\rm min}$ in a halo of mass M, and $\langle \sigma^2_{\rm sat} \rangle_M$ is the square of the los velocity dispersion of satellites averaged over the entire halo.

Consider a MLR of central galaxies that has no scatter, i.e. $P(M|L_{\rm c}) = \delta(M - M_0)$, where M_0 is the halo mass for a galaxy with luminosity $L_{\rm c}$. In this case both schemes give an equal measure of the velocity dispersion, i.e., $\sigma^2_{\rm sw} = \sigma^2_{\rm hw} = \langle \sigma^2_{\rm sat} \rangle_{M_0}$. Most studies to date have assumed the scatter in $P(M|L_{\rm c})$ to be negligible, and simply inferred an average MLR, $M_0(L_{\rm c})$ using $\sigma^2_{\rm sw}(L_{\rm c}) = \langle \sigma^2_{\rm sat} \rangle_{M_0}$ (McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; Norberg et al. 2008). However, as shown in van den Bosch et al. (2004), and as evident from the above equations (2.3) and (2.4), whenever the scatter in $P(M|L_{\rm c})$ is non-negligible,

[1] Note that the velocity dispersion should always be averaged in quadrature as is evident from Eq. (2.1).

$\sigma^2_{\rm sw}(L_{\rm c})$ and $\sigma^2_{\rm hw}(L_{\rm c})$ can differ significantly[2] (see also Chapter 3).

In this chapter, we show that ignoring the scatter in the MLR of central galaxies can result in appreciable errors in the inferred mean relation between mass and luminosity. We show, though, that these problems can be avoided by simultaneously modeling $\sigma^2_{\rm sw}(L_{\rm c})$ and $\sigma^2_{\rm hw}(L_{\rm c})$. In particular, we demonstrate that the ratio of these two quantities can be used to determine the actual scatter in the MLR of central galaxies.

2.3 Toy Model

In the previous section, we have shown that both $\sigma^2_{\rm sw}(L_{\rm c})$ and $\sigma^2_{\rm hw}(L_{\rm c})$ can be analytically expressed in terms of the probability function, $P(M|L_{\rm c})$, the satellite occupation, $\langle N_{\rm sat}\rangle_M$, and the kinematics of the satellite galaxies within a halo of mass M specified by $\langle\sigma^2_{\rm sat}\rangle_M$. In fact, the inversion of equations (2.3) and (2.4) presents an opportunity to constrain $P(M|L_{\rm c})$ using the observable $\sigma^2_{\rm sw}$ and $\sigma^2_{\rm hw}$. In this section we use a simple toy model to demonstrate that the combination of $\sigma^2_{\rm sw}$ and $\sigma^2_{\rm hw}$ can be used to constrain the first two moments (i.e., the mean and the scatter) of $P(M|L_{\rm c})$.

For convenience, let us assume that $P(M|L_{\rm c})$ is a lognormal distribution

$$P(M|L_{\rm c})\,{\rm d}M = \frac{1}{\sqrt{2\pi\sigma^2_{\ln M}}}\exp\left[-\left(\frac{\ln(M/M_0)}{\sqrt{2\,\sigma^2_{\ln M}}}\right)^2\right]\frac{{\rm d}M}{M}\,. \tag{2.5}$$

Here M_0 is a characteristic mass scale which obeys

$$\ln M_0 = \int_0^\infty P(M|L_{\rm c})\ln M{\rm d}M = \langle\ln M\rangle\,, \tag{2.6}$$

and $\sigma^2_{\ln M}$ reflects the scatter in halo mass at a fixed central luminosity and is given by

$$\sigma^2_{\ln M} = \int_0^\infty P(M|L_{\rm c})(\ln M - \ln M_0)^2{\rm d}M. \tag{2.7}$$

In addition, let us assume that both $\langle\sigma^2_{\rm sat}\rangle_M$ and $\langle N_{\rm sat}\rangle_M$ are simple power laws,

$$\langle N_{\rm sat}\rangle_M = \tilde{N}\left(\frac{M}{10^{12}}\right)^\alpha\,, \tag{2.8}$$

$$\langle\sigma^2_{\rm sat}\rangle_M = \tilde{S}^2\left(\frac{M}{10^{12}}\right)^\beta\,. \tag{2.9}$$

with α and β two constants, $\tilde{N}$ the average number of satellites in a halo of mass $10^{12}\,h^{-1}{\rm M}_\odot$, and $\tilde{S}$ the corresponding los velocity dispersion.

[2]Note that $\sigma^2_{\rm sw}\neq\sigma^2_{\rm hw}$ is a sufficient but not a necessary condition to indicate the presence of scatter in $P(M|L_{\rm c})$; after all, if $\langle N_{\rm sat}\rangle_M$ does not depend on mass then $\sigma^2_{\rm sw}=\sigma^2_{\rm hw}$ independent of the amount of scatter.

Substituting Eqs. (2.5), (2.8) and (2.9) in Eqs. (2.3) and (2.4) yields

$$\sigma^2_{\rm sw}(L_{\rm c}) = \tilde{S}^2 \left(\frac{M_0}{10^{12}}\right)^{\beta} \exp\left[\frac{\sigma^2_{\ln M}\beta^2}{2}\left(1+2\frac{\alpha}{\beta}\right)\right] , \tag{2.10}$$

$$\sigma^2_{\rm hw}(L_{\rm c}) = \tilde{S}^2 \left(\frac{M_0}{10^{12}}\right)^{\beta} \exp\left[\frac{\sigma^2_{\ln M}\beta^2}{2}\right] . \tag{2.11}$$

The velocity dispersions $\sigma_{\rm sw}(L_{\rm c})$ and $\sigma_{\rm hw}(L_{\rm c})$ depend on both M_0 and $\sigma_{\ln M}$, elucidating the degeneracy between the mean mass $M_0(L_{\rm c})$ and the scatter $\sigma_{\ln M}(L_{\rm c})$ of the distribution $P(M|L_{\rm c})$. In particular, if only $\sigma_{\rm sw}(L_{\rm c})$ or $\sigma_{\rm hw}(L_{\rm c})$ is measured, one cannot deduce $M_0(L_{\rm c})$ without having an independent knowledge of the scatter $\sigma_{\ln M}(L_{\rm c})$. However, the latter can be inferred from the *ratio* of the satellite-weighted to the host-weighted velocity dispersion. In particular, in the case of our toy model,

$$\sigma^2_{\ln M} = \frac{1}{\alpha\beta} \ln\left(\frac{\sigma^2_{\rm sw}}{\sigma^2_{\rm hw}}\right) \tag{2.12}$$

Thus, by measuring *both* $\sigma_{\rm sw}(L_{\rm c})$ and $\sigma_{\rm hw}(L_{\rm c})$ one can determine both $M_0(L_{\rm c})$ and its scatter $\sigma_{\ln M}(L_{\rm c})$, provided that the constants α and β are known. Since virialized dark matter haloes all have the same average density within their virial radii, $\beta = 2/3$ (e.g. Klypin et al. 1999; van den Bosch et al. 2004). Previous studies have obtained constraints on α that cover the range $0.7 \lesssim \alpha \lesssim 1.1$ (e.g. Yang et al. 2005a; van den Bosch et al. 2007; Tinker et al. 2007; Yang et al. 2007). Since $\sigma^2_{\ln M} \propto \alpha^{-1}$, this uncertainty directly translates into an uncertainty of the inferred scatter. Therefore, in Chapter 4, we do not use the constraints on α available in the literature to infer the mean and scatter of the MLR from real data. Instead, we treat α as a free parameter and use the average number of observed satellites as a function of the luminosity of central as an additional constraint. Note that the relation between $\sigma_{\ln M}$ and the ratio of $\sigma_{\rm sw}$ to $\sigma_{\rm hw}$ specified by Eq. (2.12) is model-dependent, i.e. we have assumed particular functional forms for the halo occupation statistics of centrals and satellites to arrive at Eq. (2.12). Furthermore, we have not accounted for any selection effects. In what follows, we present a careful treatment of selection effects and more realistic halo occupation models.

2.4 Selection Effects

The toy model presented in the previous section illustrates that measurements of the satellite-weighted and host-weighted kinematics of satellite galaxies can be used to infer the mean and scatter of the MLR of central galaxies, $P(M|L_{\rm c})$. However, in practice one first needs a method to select central galaxies and satellites from a galaxy redshift survey. In general, central galaxies are selected to be the brightest galaxy in some cylindrical volume in redshift space, and satellite galaxies are defined as those galaxies that are fainter than the central by a certain amount and located within a cylindrical volume centered on the central. In this section we show how these

selection criteria impact on $\sigma^2_{\rm sw}$ and $\sigma^2_{\rm hw}$, and how this can be accounted for in the analysis.

No selection criterion is perfect, and some galaxies will be selected as centrals, while in reality they are satellites (hereafter 'false centrals'). In addition, some galaxies will be selected as satellites of a certain central, while in reality they do not reside in the same halo as the central (hereafter 'interlopers'). The selection criteria have to be tuned in order to minimize the impact of these false centrals and interlopers. Here we make the assumption that interlopers can be corrected for, and that the impact of false centrals is negligible. Using mock galaxy redshift surveys, van den Bosch et al. (2004) have shown that one can devise adaptive, iterative selection criteria that justify these assumptions (see also Chapter 3). Here we focus on the impact of these iterative selection criteria on the satellite kinematics in the absence of interlopers and false centrals. Our analytical treatment for selection effects follows the one presented in van den Bosch et al. (2004) except for the averaging of velocity dispersions in quadrature and the inclusion of an extra selection effect. We state and quantify these differences in Section 2.A. For completeness, we outline our treatment below.

In general, satellite galaxies are selected to lie within a cylindrical volume centered on its central galaxy, and specified by $R_{\rm p} < R_{\rm s}$ and $|\Delta V| < (\Delta V)_{\rm s}$. Here $R_{\rm p}$ is the physical separation from the central galaxy projected on the sky and ΔV is the los velocity difference between a satellite and its central. Usually, $(\Delta V)_{\rm s}$ is chosen sufficiently large, so that it does not exclude true satellites from being selected. However, in the adaptive, iterative selection criteria of van den Bosch et al. (2004), which we will use in the subsequent chapters, the aperture radius is tuned so that $R_{\rm s} \simeq 0.375\, r_{\rm vir}$, where $r_{\rm vir}$ is the virial radius of the dark matter halo hosting the central−satellite pair. This means that $\langle N_{\rm sat}\rangle_M$ and $\langle \sigma^2_{\rm sat}\rangle_M$ in Eqs. (2.3) and (2.4) need to be replaced by $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ and $\langle \sigma^2_{\rm sat}\rangle_{{\rm ap},M}$, respectively. Here $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ is the average number of satellites in a halo of mass M that lie within the aperture, and $\langle \sigma^2_{\rm sat}\rangle_{{\rm ap},M}$ is the square of the los velocity dispersion of satellite galaxies averaged over the aperture.

The number of satellites present within the aperture, $\langle N_{\rm sat}\rangle_{{\rm ap},M}$, is related to the number of satellites given by the halo occupation statistics, $\langle N_{\rm sat}\rangle_M$, via

$$\langle N_{\rm sat}\rangle_{{\rm ap},M} = \begin{cases} f_{\rm cut}\,\langle N_{\rm sat}\rangle_M & \text{if } R_{\rm s} < r_{\rm vir} \\ \langle N_{\rm sat}\rangle_M & \text{if } R_{\rm s} \geq r_{\rm vir} \end{cases} \tag{2.13}$$

with

$$f_{\rm cut} = \frac{4\pi}{\langle N_{\rm sat}\rangle_M} \int_0^{R_{\rm s}} R\,{\rm d}R \int_R^{r_{\rm vir}} n_{\rm sat}(r|M)\,\frac{r\,{\rm d}r}{\sqrt{r^2-R^2}}\,. \tag{2.14}$$

Here $n_{\rm sat}(r|M)$ is the number density distribution of satellites within a halo of mass M, which is normalized so that

$$\langle N_{\rm sat}\rangle_M = 4\pi \int_0^{r_{\rm vir}} n_{\rm sat}(r|M)\, r^2\,{\rm d}r\,. \tag{2.15}$$

Under the assumption that the satellites are in virial equilibrium within the dark matter halo, and that the velocity dispersion of satellite galaxies within a given halo is isotropic, the

los velocity dispersion of satellites within the cylindrical aperture of radius $R_{\rm s}$ is given by

$$\begin{aligned}\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M} &= \frac{4\pi}{\langle N_{\rm sat}\rangle_{{\rm ap},M}}\int_0^{R_{\rm s}} {\rm d}R\, R \\ &\quad \int_R^{r_{\rm vir}} n_{\rm sat}(r|M)\,\sigma^2_{\rm sat}(r|M)\,\frac{r\,{\rm d}r}{\sqrt{r^2-R^2}}\,. \end{aligned} \tag{2.16}$$

Here $\sigma_{\rm sat}(r|M)$ is the local, one-dimensional velocity dispersion which is related to the potential Ψ of the dark matter halo via the Jeans equation

$$\sigma^2_{\rm sat}(r|M) = \frac{1}{n_{\rm sat}(r|M)}\int_r^\infty n_{\rm sat}(r'|M)\,\frac{\partial\Psi}{\partial r'}(r'|M)\,{\rm d}r'. \tag{2.17}$$

The radial derivative of the potential Ψ represents the radial force and is given by

$$\frac{\partial\Psi}{\partial r}(r|M) = \frac{4\pi G}{r^2}\int_0^r \rho(r'|M)\,r'^2\,{\rm d}r'\,, \tag{2.18}$$

with $\rho(r|M)$ the density distribution of a dark matter halo of mass M. The assumptions of virial equilibrium and orbital isotropy are supported by results from numerical simulations which show that dark matter subhaloes (and hence satellite galaxies) are in a steady state equilibrium within the halo and that their orbits are nearly isotropic at least in the central regions (Diemand et al. 2004). Furthermore, van den Bosch et al. (2004) have demonstrated that anisotropy has a negligible impact on the average velocity dispersion within the selection aperture.

Finally, there is one other effect of the selection criteria to be accounted for which has not been considered in van den Bosch et al. (2004) or in Eq. (2.4). When selecting central−satellite pairs, only those centrals are selected with at least one satellite inside the search aperture. This has an impact on the host-weighted velocity dispersions that needs to be accounted for. The probability that a halo of mass M, which on average hosts $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ satellites within the aperture R_s, has $N_{\rm sat}\geq 1$ satellites within the aperture, is given by

$$\begin{aligned} P(N_{\rm sat}\geq 1) &= 1-P(N_{\rm sat}=0) \\ &= 1-\exp\left[-\langle N_{\rm sat}\rangle_{{\rm ap},M}\right] \\ &\equiv \mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M}). \end{aligned} \tag{2.19}$$

Here, for the second equality, we have assumed Poisson statistics for the satellite occupation numbers. Note that, in the satellite-weighting scheme, haloes that have zero satellites, by definition, get zero weight. Therefore only the host-weighted velocity dispersions need to be corrected for this effect.

Thus, in light of the selection effects, Eqs. (2.3) and (2.4) become

$$\sigma^2_{\rm sw}(L_{\rm c}) = \frac{\int_0^\infty P(M|L_{\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|L_{\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}, \tag{2.20}$$

and

$$\sigma^2_{\rm hw}(L_{\rm c}) = \frac{\int_0^\infty P(M|L_{\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|L_{\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,{\rm d}M}. \tag{2.21}$$

Note that $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M}) \simeq \langle N_{\rm sat}\rangle_{{\rm ap},M}$ when $\langle N_{\rm sat}\rangle_{{\rm ap},M} \rightarrow 0$. This implies that $|\sigma_{\rm sw}-\sigma_{\rm hw}| \rightarrow 0$ for faint centrals (i.e. when $L_{\rm c}$ becomes comparable to $L_{\rm min}$, the minimum luminosity adopted to select the satellites). Therefore, the ability to detect the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ depends on how bright $L_{\rm c}$ is compared to $L_{\rm min}$. In principle, this can be overcome by decreasing $L_{\rm min}$ (detecting faint satellite galaxies), such that $\langle N_{\rm sat}\rangle_{{\rm ap},M} \gg 0$ and $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M}) \rightarrow 1$. However, faint satellite galaxies can only be detected out to a very small distance due to the flux limit of a survey. The number of galaxies in a volume-limited sample with low $L_{\rm min}$ is small. This in turn makes the detection of the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ difficult due to small number statistics. Therefore, there is a trade-off involved in the choice of $L_{\rm min}$, which limits the significance with which one can detect the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$. Since this selection effect was not taken into account in Section 2.3, Eq. (2.12), which relates the ratio $\sigma_{\rm sw}/\sigma_{\rm hw}$ to the scatter in halo masses, $\sigma_{\ln M}$, does not reveal this dependence on $L_{\rm min}$.

2.5 More Realistic Models

Using the methodology described above, we now illustrate how satellite kinematics can be used to constrain the mean and the scatter of the MLR of central galaxies, $P(M|L_{\rm c})$. We improve upon the toy model described in Section 2.3 by considering a realistic model for the halo occupation statistics and take the impact of selection criteria into account.

As is evident from the discussion in the previous section, calculating $\sigma^2_{\rm sw}(L_{\rm c})$ and $\sigma^2_{\rm hw}(L_{\rm c})$ requires the following input:

- the density distributions of dark matter haloes, $\rho(r|M)$
- the number density distribution of satellites, $n_{\rm sat}(r|M)$
- the halo occupation statistics of centrals, $P(M|L_{\rm c})$.

We assume that dark matter haloes follow the NFW (Navarro et al. 1997) density distribution

$$\rho(r|M) = \frac{M}{4\pi r_s^3 \mu(c)} \left(\frac{r}{r_{\rm s}}\right)^{-1} \left(1+\frac{r}{r_{\rm s}}\right)^{-2}. \tag{2.22}$$

Here, $r_{\rm s}$ is a characteristic scale radius, $c = r_{\rm vir}/r_{\rm s}$ is the halo's concentration parameter, and

$$\mu(x) \equiv \ln(1+x) - \frac{x}{1+x}\,. \tag{2.23}$$

Throughout we use the relation between c and M given by Macciò et al. (2007).

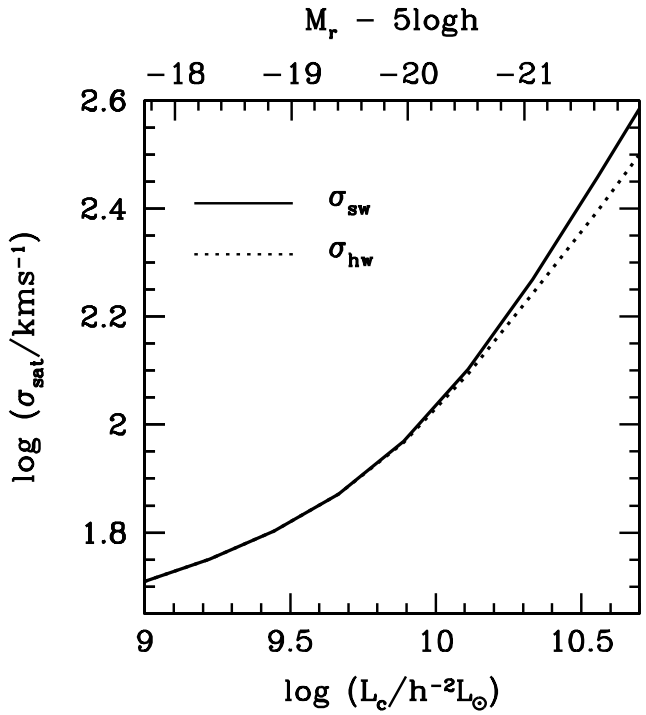

Figure 2.1: The satellite-weighted ($\sigma_{\rm sw}$) and host-weighted ($\sigma_{\rm hw}$) velocity dispersions of satellite galaxies for model G1. Note that $\sigma_{\rm sw}(L_{\rm c}) > \sigma_{\rm hw}(L_{\rm c})$ at the bright end, indicating that the MLR of central galaxies, $P(M|L_{\rm c})$, has a non-negligible amount of scatter.

Table 2.1: Different models for the HOD of centrals

Model	$\sigma_{\log L}$	γ_1	γ_2	L_0	M_1
G1	0.14	3.27	0.25	9.94	11.07
G2	0.25	3.27	0.25	9.94	11.07
G3	0.14	1.80	0.40	9.80	11.46

Three different models describing the MLR of centrals used to predict $\sigma_{\rm sw}(L_{\rm c})$ and $\sigma_{\rm hw}(L_{\rm c})$.

We assume that the number density distribution of satellite galaxies is given by the generalised NFW profile,

$$n_{\rm sat}(r|M) \propto \left(\frac{r}{\mathcal{R}r_s}\right)^{-\gamma} \left(1+\frac{r}{\mathcal{R}r_s}\right)^{\gamma-3}, \tag{2.24}$$

where γ represents the slope of the number density distribution of satellites as $r \to 0$ and $\mathcal{R}$ is a free parameter. In this chapter, we assume $\gamma = 1$ and $\mathcal{R} = 1$, i.e. the number density distribution of satellite galaxies is spatially unbiased with respect to the distribution of dark matter particles. Note that this is a fairly simplistic assumption. We address the issue of potential spatial antibias of satellite galaxies in Chapter 4.

Substituting $\rho(r|M)$ and $n_{\rm sat}(r|M)$ in Eqs. (2.18) and (2.17) gives

$$\sigma^2_{\rm sat}(r|M) = \frac{c\,V^2_{\rm vir}}{\mathcal{R}^2\mu(c)} \left(\frac{r}{\mathcal{R}r_s}\right)^{\gamma} \left(1+\frac{r}{\mathcal{R}r_s}\right)^{3-\gamma} \int_{r/r_s}^{\infty} \frac{\mu(x){\rm d}x}{(x/\mathcal{R})^{\gamma+2}(1+x/\mathcal{R})^{3-\gamma}},$$

where $V_{\rm vir} = (GM/r_{\rm vir})^{1/2}$ is the circular velocity at $r_{\rm vir}$.

The final ingredient is a realistic model for the halo occupation statistics of centrals and satellites. To that extent, we use the conditional luminosity function (CLF) presented in Cacciato et al. (2009). The CLF, denoted by $\Phi(L|M){\rm d}L$, specifies the average number of galaxies with luminosities in the range $L \pm {\rm d}L/2$ that reside in a halo of mass M, and is explicitly written as the sum of the contributions due to central and satellite galaxies, i.e. $\Phi(L|M) = \Phi_{\rm c}(L|M) + \Phi_{\rm s}(L|M)$. From this CLF, the probability distribution $P(M|L_{\rm c})$ follows from Bayes' theorem according to

$$P(M|L_{\rm c}) = \frac{\Phi_{\rm c}(L_{\rm c}|M)\,n(M)}{\int_0^\infty \Phi_{\rm c}(L_{\rm c}|M)\,n(M)\,{\rm d}M}, \tag{2.25}$$

with $n(M)$ the halo mass function, while the average number of satellites with $L \geq L_{\rm min}$ in a halo of mass M is given by

$$\langle N_{\rm sat}\rangle_M = \int_{L_{\rm min}}^{\infty} \Phi_{\rm s}(L|M)\,{\rm d}L\,. \tag{2.26}$$

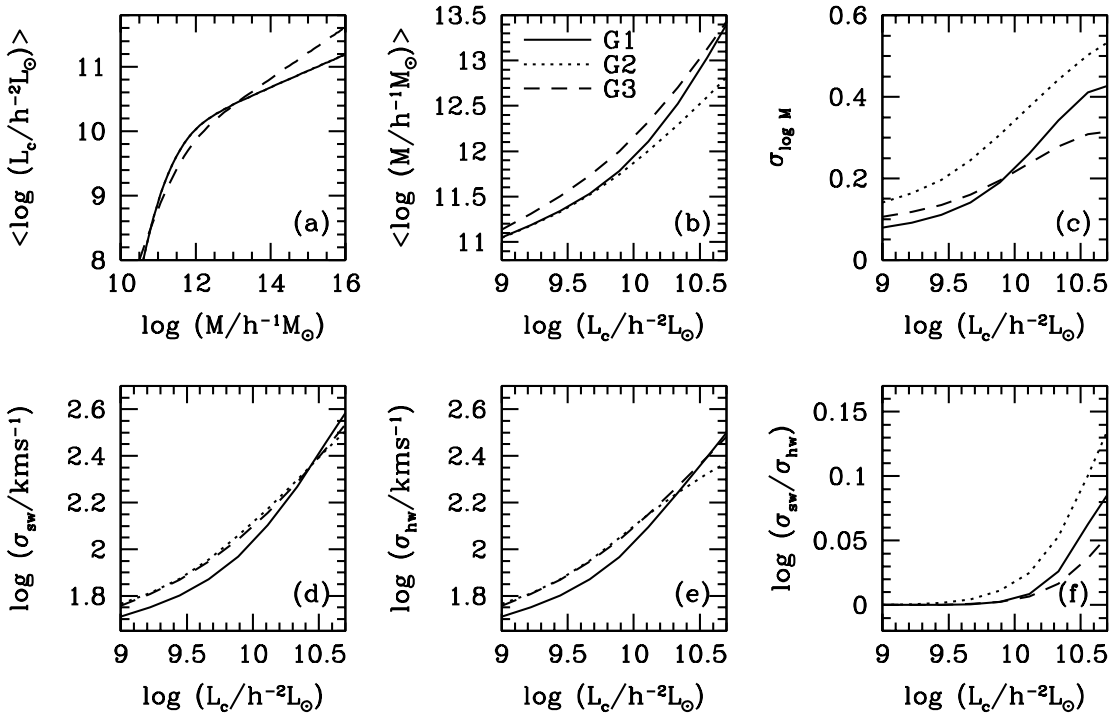

Figure 2.2: Comparison of three models with different HODs for the central galaxies. In all panels the solid line corresponds to model G1, the dotted line to model G2 and the dashed line to model G3 (see Table 1 for the parameters). Panels (a), (b) and (c) show $\langle \log L_{\rm c} \rangle (M)$, $\langle \log M \rangle (L_{\rm c})$ and $\sigma_{\log M}(L_{\rm c})$, respectively. Panels (d) and (e) show the predicted satellite-weighted and host-weighted velocity dispersions as function of luminosity, and panel (f) shows the logarithm of the ratio between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$. See text for a detailed discussion.

The parametric forms for $\Phi_{\rm c}(L|M)$ and $\Phi_{\rm s}(L|M)$ are motivated by the results of Yang et al. (2008, hereafter YMB08), who determined the CLF from the SDSS group catalogue of Yang et al. (2007). In particular, $\Phi_{\rm c}(L|M)$ is assumed to follow a log-normal distribution

$$\Phi_{\rm c}(L|M){\rm d}L = \frac{\log {\rm e}}{\sqrt{2\pi}\,\sigma_{\log L}} \exp\left(-\left[\frac{\log(L/L_c^*)}{\sqrt{2}\sigma_{\log L}}\right]^2\right)\frac{{\rm d}L}{L}\,, \tag{2.27}$$

with $\sigma_{\log L}$ a free parameter that we take to be independent of halo mass, and

$$L_{\rm c}^*(M) = L_0 \frac{(M/M_1)^{\gamma_1}}{[1+(M/M_1)]^{\gamma_1-\gamma_2}} \tag{2.28}$$

which has four additional free parameters: two slopes, γ_1 and γ_2, a characteristic halo mass, M_1, and a normalization, L_0. Note that, $L_{\rm c}^* \propto M^{\gamma_1}$ for $M \ll M_1$ and $L_{\rm c}^* \propto M^{\gamma_2}$ for $M \gg M_1$. Cacciato et al. (2009) constrained the free parameters, $\sigma_{\log L}$, γ_1, γ_2, M_1 and L_0, by fitting the SDSS luminosity function of Blanton et al. (2003b) and the galaxy$-$galaxy correlation lengths as a function of luminosity from Wang et al. (2007). The resulting best-fit parameters are listed in the first row of Table 1, and constitute our fiducial model G1. We also consider two alternative models for $\Phi_{\rm c}(L|M)$, called G2 and G3, the parameters of which are also listed in Table 1. For $\Phi_{\rm s}(L|M)$ we adopt the model of Cacciato et al. (2009) throughout, without any modifications: i.e. models G1, G2, and G3 only differ in $P(M|L_{\rm c})$ and have the same $n_{\rm sat}(r|M)$.

Having specified all necessary ingredients, we now compute the satellite weighted and host-weighted satellite kinematics for our fiducial model G1 using Eqs. (2.20) and (2.21). The results are shown as solid and dotted lines in Fig. 2.1, where we have adopted a minimum satellite luminosity of $L_{\rm min} = 10^9 h^{-2}\ {\rm L}_\odot$. At the faint-end, the velocity dispersions $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ are equal, this simply reflects the fact that $\langle N_{\rm sat}\rangle_M \to 0$ if $L_{\rm c} \to L_{\rm min}$. At the bright end, though, the non-zero scatter in $P(M|L_{\rm c})$ causes the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ to increase systematically with increasing $L_{\rm c}$. This is a generic trend for any realistic halo occupation model (see also van den Bosch et al. 2004). It is important to note here that the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ depends upon the central galaxy luminosity and how bright this luminosity is compared to $L_{\rm min}$ (see discussion at the end of Section 2.4). In Chapter 3, we show that the difference between the velocity dispersions in the two schemes is detectable from current datasets. Previous studies (McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; Conroy et al. 2007; Norberg et al. 2008) did not have sufficient number statistics to detect the difference between the two schemes given the measurement errorbars.

The upper panels of Fig. 2.2 show the mean and scatter of the MLR of central galaxies in models G1 (solid lines), G2 (dotted lines) and G3 (dashed lines). Panel (a) plots $\langle \log L_{\rm c}\rangle(M) = \log(L_{\rm c}^*)$, which reveals the double power-law behavior of Eq. (2.28), panel (b) shows the inverse relation,

$$\langle \log M\rangle(L_{\rm c}) = \int_0^\infty P(M|L_{\rm c}) \log M\, {\rm d}M\,. \tag{2.29}$$

and panel (c) shows the scatter in the MLR, $\sigma_{\log M}(L_{\rm c})$, deduced by using

$$\sigma^2_{\log M} = \int_0^\infty P(M|L_{\rm c}) \left[\log M - \langle \log M \rangle (L_{\rm c})\right]^2 \, {\rm d}M \, . \tag{2.30}$$

Note that $\sigma_{\log M}(L_{\rm c})$ increases with increasing $L_{\rm c}$, even though the scatter $\sigma_{\log L}$ is constant with halo mass. This simply owes to the fact that the slope of $\langle \log L_{\rm c} \rangle (M)$ becomes shallower with increasing $L_{\rm c}$, as illustrated in Fig. 2.3.

The comparison between models G1 and G2 illustrates the effect of changing the scatter $\sigma_{\log L}$ in $\Phi_{\rm c}(L|M)$. Both models have exactly the same $\langle \log L_{\rm c} \rangle (M)$ (the solid line overlaps the dotted line in panel a). However, because the scatter $\sigma_{\log L}$ in G2 is larger than in G1 (see Table 1), the $\langle \log M \rangle (L_{\rm c})$ of G2 is significantly lower than that of G1 at the bright end ($\sim$ 0.5 dex at the bright end). This is due to the shape of the halo mass function. Increasing the scatter adds both low mass and high mass haloes to the distribution $P(M|L_{\rm c})$ (cf. Eqs. [2.25] and [2.29]), and the overall change in the average halo mass depends on the slope of the halo mass function. Brighter galaxies live on average in more massive haloes where the halo mass function is steeper. In particular, when the halo mass range sampled by $P(M|L_{\rm c})$ lies in the exponential tail of the halo mass function, an increase in the scatter adds many more low mass haloes than massive haloes, causing a shift in the average halo mass towards lower values. On the other hand, fainter galaxies live in less massive haloes, where the slope of the halo mass function is much shallower. Consequently, a change in the scatter does not cause an appreciable change in the average mass. Finally, as expected, the scatter in the MLR, $\sigma_{\log M}(L_{\rm c})$, in G2 is higher than for G1 at all luminosities (see panel c).

Panels (d) and (e) of Fig. 2.2 show the analytical predictions for $\sigma_{\rm sw}(L_c)$ and $\sigma_{\rm hw}(L_c)$, respectively. Note that models G1 and G2 predict satellite kinematics that are significantly different (which can be distinguished given the typical measurement errors in Chapter 4), even though both have exactly the same $\langle \log L_{\rm c} \rangle (M)$. In particular, model G2 predicts larger $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ at the faint end, but lower $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ at the bright end. The trend at the faint is due to the fact that the scatter $\sigma_{\log M}(L_{\rm c})$ is higher in G2 than in G1. Quantitatively, this is evident from Eqs. (2.10) and (2.11), which demonstrate that both the satellite weighted and host weighted satellite kinematics increase with increasing scatter. At the bright end, however, the drastic decrease in $\langle \log M \rangle (L_{\rm c})$ for G2 with respect to G1 overwhelms this boost and causes $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ to be lower in G2.

Now consider model G3. This model has the same amount of scatter as model G1, but we have tuned its parameters $(\gamma_1, \gamma_2, M_1, L_0)$ that describe $\langle \log L_{\rm c} \rangle (M)$ such that its $\sigma_{\rm sw}(L_{\rm c})$ closely matches that of model G2 (the dotted and dashed curves in panel (d) are almost overlapping). As is evident from panels (a)$-$(c), though, the MLR of G3 is very different from that of G2. Note that the higher values of $\langle \log M \rangle (L_{\rm c})$ for G3 are compensated by its lower values of $\sigma_{\log M}(L_{\rm c})$, such that the satellite-weighted kinematics are virtually identical. This clearly illustrates the degeneracy between the mean and the scatter of the MLR: One can decrease the

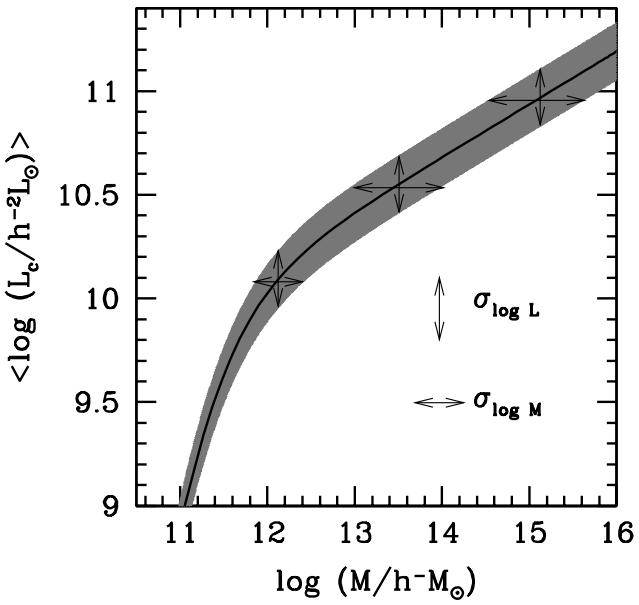

Figure 2.3: Illustration of the MLR of central galaxies. The solid black line indicates the mean of the $L_{\rm c}$-M relation, while the gray scale region reflects the scatter. In this particular case the scatter in $P(L_{\rm c}|M)$ (indicated by vertical arrows) is taken to be constant with halo mass. Note, though, that the scatter in $P(M|L_{\rm c})$ (indicated by horizontal arrows) increases with increasing $L_{\rm c}$; this simply is due to the fact that the slope of the mean $L_{\rm c}$-M relation becomes shallower with increasing halo mass.

mean of the MLR and yet achieve the same $\sigma_{\rm sw}$ by increasing the scatter of the MLR. It also shows that $\sigma_{\rm sw}$ alone does not yield sufficient information to uniquely constrain the MLR.

Note, though, that although $\sigma_{\rm sw}$ is the same for models G2 and G3, their host-weighted satellite kinematics, $\sigma_{\rm hw}(L_{\rm c})$, are different at the bright end. In fact, the ratios $\sigma_{\rm sw}/\sigma_{\rm hw}$ for models G2 and G3 are clearly different. The logarithm of this ratio, shown in panel (f), follows the same trend as $\sigma_{\log M}(L_{\rm c})$, i.e. it is higher for model G2 than for G3. This is in agreement with our toy model, according to which the ratio $\sigma_{\rm sw}/\sigma_{\rm hw}$ increases with the scatter $\sigma_{\log M}(L_{\rm c})$ (cf. Eq. [2.12]). This illustrates once again that the *combination* of $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ allows one to constrain both the mean and the scatter of the MLR simultaneously (see also Chapter 4).

2.6 Summary

The kinematics of satellite galaxies is a powerful probe of the masses of the dark matter haloes surrounding central galaxies. With the advent of large, homogeneous redshift surveys, it has become possible to probe the mass−luminosity relation (MLR) of central galaxies spanning a significant range in luminosities. Unfortunately, since most centrals only host a few satellite galaxies with luminosities above the flux limit of the redshift survey, one generally needs to stack a large number of central galaxies within a given luminosity bin and combine the velocity information of their satellites. Because of the finite bin-width, and because the MLR has intrinsic scatter, this stacking results in combining the kinematics of satellite galaxies in haloes of different masses, which complicates the interpretation of the data. Unfortunately, most previous studies have ignored this issue, and made the oversimplified assumption that the scatter is negligible.

Using realistic models for the halo occupation statistics, and taking account of selection effects, we have demonstrated a degeneracy between the mean and the scatter of the MLR: one can change the mean relation between halo mass, M, and central galaxy luminosity, $L_{\rm c}$, and simultaneously change the scatter around that mean relation, such that the observed satellite kinematics, $\langle\sigma_{\rm sat}\rangle(L_{\rm c})$, are unaffected.

We have also presented a new technique to break this degeneracy, based on measuring the satellite kinematics using two different weighting schemes: host-weighting (each central galaxy gets the same weight) and satellite weighting (each central galaxy gets a weight proportional to its number of satellites). In general, for central galaxies close to the magnitude limit of the survey, the average number of satellites per host is close to zero, and the satellite-weighted velocity dispersion, $\sigma_{\rm sw}$, is equal to the host-weighted velocity dispersion, $\sigma_{\rm hw}$. This is because only those centrals with at least one satellite are used to measure the satellite kinematics. For brighter centrals, however, $\sigma_{\rm sw} > \sigma_{\rm hw}$ and the actual ratio of these two values is larger for MLRs with more scatter (see Eq. [2.12] and panels c and f of Fig. 2.2). Hence, the combination of $\sigma_{\rm sw}(L_{\rm c})$ and $\sigma_{\rm hw}(L_{\rm c})$ contains sufficient information to constrain both the mean and the scatter of the MLR of central galaxies. In Chapter 3, we apply this method to a mock catalogue,

and show that the difference between $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$ can be detected with sufficient significance to constrain both the mean and the scatter of the MLR of central galaxies. In Chapter 3, we also address the issues of measurement errors, sampling effects and interlopers. In Chapter 4, we apply this method to data from the SDSS and show that the MLR and its scatter inferred from the data are in excellent agreement with other, independent constraints.

In a recent study, Becker et al. (2007) analyzed the kinematics of MaxBCG clusters (Koester et al. 2007) and inferred the mean and the scatter of the mass−richness relation (here richness is a measure for the number of galaxies that reside in the cluster). Becker et al. (2007) combined the kinematics of satellite galaxies in finite bins of cluster richness and measured the second and fourth moments of the host-weighted velocity distribution. They used these two moments simultaneously to determine the mean and the scatter of the mass−richness relation. This method is complementary to that presented here, and it will be interesting to compare both methods and investigate their relative strengths and weaknesses. We intend to address this in a future study.

Finally we emphasize that the scatter in the conditional probability function $P(M|L_{\rm c})$ is expected to increase with increasing $L_{\rm c}$. This is due to the fact that the slope of $\langle L_{\rm c}\rangle(M)$, which is the mean of $P(L_{\rm c}|M)$, becomes shallower with increasing halo mass. Hence, when stacking haloes according to the luminosity of the central galaxy, one cannot ignore the scatter in M, even when the scatter in $P(L_{\rm c}|M)$ is small. This has important implications for any technique that relies on stacking, such as satellite kinematics and galaxy−galaxy lensing (see e.g. Tasitsiomi et al. 2004; Cacciato et al. 2009)

Appendix

2.A Comparison with van den Bosch et al.

The analytical treatment of the selection effects presented in Section 2.4 closely follows van den Bosch et al. (2004, hereafter vdB04) except for two subtle differences. First of all, vdB04 incorrectly averaged the velocity dispersion directly rather than in quadrature as done here (Eqs. 2.16, 2.20 and 2.21). Secondly, vdB04 failed to account for the factor $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})$ (hereafter $\mathcal{P}$ for brevity) that corrects for the centrals that do not host any satellite and hence do not contribute to the host-weighted velocity dispersion of satellites (Eq. 2.21).

Fig. 2.4 quantifies the error in the estimates of vdB04 due to these differences. The dotted line shows the relative error on $\sigma_{\rm hw}$ caused due to direct averaging instead of averaging in quadrature. Direct averaging leads to an underestimate of the velocity dispersion which is negligible at the faint end, but grows to $\sim 5\%$ at the bright end. The satellite-weighted velocity dispersion is also underestimated by a similar amount. In their paper, vdB04 only compared the analytical estimate of the *satellite-weighted* velocity dispersion to real data. Fortunately, the small error in the estimate of the satellite-weighted velocity dispersion due to direct averaging does not change any of their conclusions.

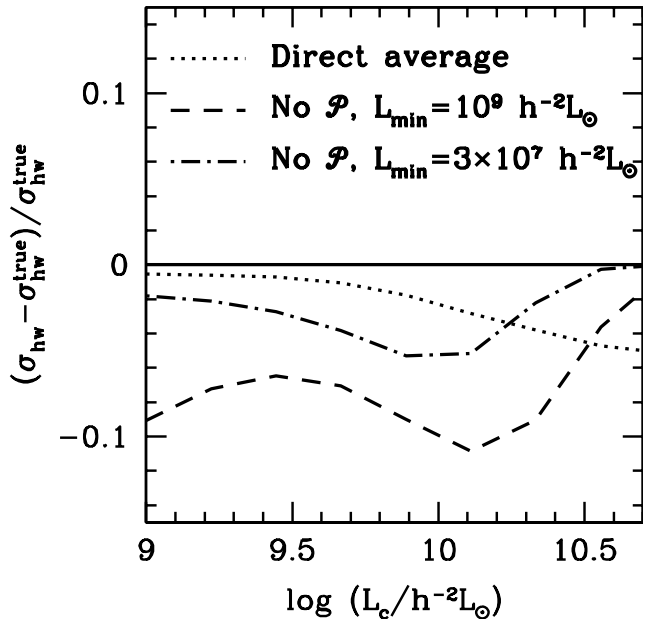

Figure 2.4: The relative error in the estimates of the host-weighted velocity dispersion by vdB04. The dotted line shows the error caused due to direct averaging of the velocity dispersion. The dashed and the dot-dashed lines show the error when $\mathcal{P}$ is ignored for $L_{\rm min} = 10^9\ h^{-2}{\rm L}_\odot$ and $L_{\rm min} = 3 \times 10^7\ h^{-2}{\rm L}_\odot$ respectively.

The dashed line in Fig. 2.4 shows that $\sigma_{\rm hw}$ is underestimated by $\sim 10\%$ if the factor $\mathcal{P}$ in Eq. (2.21) is ignored. The factor $\mathcal{P}$ depends on the minimum luminosity, $L_{\rm min}$, adopted to select the satellites: lower values of $L_{\rm min}$ result in larger number of satellites, which imply $\mathcal{P} \rightarrow 1$. For the dashed line, $L_{\rm min} = 10^9\ h^{-2}{\rm L}_\odot$. However, vdB04 adopted $L_{\rm min} = 3 \times 10^7\ h^{-2}{\rm L}_\odot$ in their analysis. In this case (shown with the dot-dashed line), the relative error decreases to $< 5\%$. Note that the factor $\mathcal{P}$ affects only the host-weighted velocity dispersion (see Section 2.4). Since vdB04 only compared their estimates of the satellite-weighted velocity dispersion to data, their results are not influenced by the fact that they failed to account for $\mathcal{P}$ in their equations for host-weighting.

Chapter 3

Satellite Kinematics: Tests on a Mock Catalogue

The contents of this chapter are based upon the article More et al. (2009c) published in the Monthly Notices of the Royal Astronomical Society. The reference is
More, S., van den Bosch, F. C., Cacciato, M., et al. 2009b, MNRAS, 392, 801.
The original article also contains the results of the analysis with SDSS data. To maintain a coherent flow in this thesis these results are presented in the next chapter. In addition, unnecessary repetition of certain equations is avoided by referring to the previous chapter.

3.1 Introduction

According to the standard picture of galaxy formation, galaxies form in dark matter haloes. The complex astrophysics of galaxy formation and evolution is primarily believed to be governed by the mass of the dark matter halo in which it occurs. Quantifying scaling relations between central galaxy properties and their dark matter halo masses is, hence, an important stepping stone towards understanding galaxy formation. The kinematics of satellite galaxies is a powerful probe of the halo masses of central galaxies and can be used to determine the scaling relations between central galaxy properties and their halo masses.

In the past, several studies have used the kinematics of satellite galaxies to determine the halo mass−luminosity relation (MLR) of central galaxies (McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; van den Bosch et al. 2004; Conroy et al. 2005, 2007) and to study the density profiles of dark matter haloes (Prada et al. 2003; Klypin & Prada 2009). Although the results obtained by these studies appear consistent with each other, Norberg et al. (2008, hereafter N08) have demonstrated a quantitative disagreement in the kinematice obtained by these studies and showed that this disagreement is largely due to subtle differences in the selection criteria used to identify central and satellite galaxies. Therefore, it is crucial to understand how selection effects bias the MLR of central galaxies inferred from satellite kinematics and

to test the methods used to quantify the kinematics of satellites in order to identify potential systematic biases that can affect the measurements.

The present chapter is aimed at understanding the various selection biases that affect the analysis of the kinematics of satellite galaxies and the subsequent determination of the halo mass−luminosity relationship (MLR). We construct a realistic mock galaxy catalogue for this purpose. In this chapter, we create an analysis pipeline, which takes a galaxy redshift catalogue as input, performs the analysis of the kinematics of satellite galaxies and outputs both the mean and the scatter of the MLR of central galaxies based upon the the novel method presented in Chapter 2. This analysis pipeline is rigorously tested using the mock galaxy catalogue. In particular, we show that our central−satellite selection criteria and the method to measure the kinematics reliably recover the true kinematics present in the mock catalogue. We also show that the mean and the scatter of the MLR inferred from the kinematics match the corresponding true relations in the mock catalogue. In the subsequent chapters, this analysis pipeline will be used on data from the Sloan Digital Sky Survey (SDSS).

This chapter is organized as follows. In Section 3.2 we describe the construction of the mock catalogue that is used to test our method of analysis of the kinematics of satellites. In Section 3.3 we briefly outline the iterative selection criteria used to select centrals and satellites. In Section 3.4 we describe and test the method used to measure the kinematics of satellites as a function of the central luminosity. The inference of the MLR from the kinematics of satellite galaxies requires the knowledge of the number density distribution of satellites within a halo. In Section 3.5 we show that this distribution can also be inferred from the selected satellites. In Section 3.6 we describe our model to interpret the measured velocity dispersions and show that this model is able to recover the true mean and scatter of the MLR of central galaxies from the mock catalogue. We summarize our results in Section 3.7.

3.2 Mock Catalogue Construction

It is important to carefully identify central and satellite galaxies from a redshift survey in order to study the kinematics of satellite galaxies. Furthermore, it is also important to reliably quantify the kinematics of the selected satellites as a function of central luminosity which in turn can yield the MLR of central galaxies. We monitor the performance of our method of analysis for each of these tasks using a realistic mock galaxy catalogue (MGC) which serves as a control dataset. The halo occupation of galaxies in the MGC is known a priori, thereby allowing an accurate assessment of the level of contamination of the selected sample of centrals and satellites due to false identifications and also a comparison between the kinematics recovered from the selected satellites and the actual kinematics present in the MGC.

The two essential steps to construct a MGC are to obtain a distribution of dark matter haloes and to use a recipe to populate the dark matter haloes with galaxies. For the former purpose, we use a numerical simulation of dark matter particles in a cosmological setup. For the latter, we

use the conditional luminosity function (CLF) which describes the average number of galaxies with luminosities in the range $L \pm \mathrm{d}L/2$ that reside in a halo of mass M.

A distribution of dark matter haloes is obtained from a $N-$body simulation for a ΛCDM cosmology with the following parameters, matter density $\Omega_{\rm m} = 0.238$, energy density in the cosmological constant $\Omega_\Lambda = 0.762$, the linearly extrapolated root mean square variance of the density fluctuations on scales of $8\,h^{-1}{\rm Mpc}$ $\sigma_8 = 0.75$, the spectral index of the initial density fluctuations $n_{\rm s} = 0.95$ and the Hubble parameter $h = H_0/100\,{\rm km\,s^{-1}\,Mpc^{-1}} = 0.73$. The simulation consists of $N = 512^3$ particles within a cube of side $L_{\rm box} = 300\,h^{-1}{\rm Mpc}$ with periodic boundary conditions. The particle mass is 1.33 x $10^{10}\,h^{-1}{\rm M}_\odot$. Dark matter haloes are identified using the friends$-$of$-$friends algorithm (Davis et al. 1985) with a linking length of 0.2 times the mean inter$-$particle separation. Haloes obtained with this linking length have a mean overdensity of 180 (Porciani et al. 2002). We consider only those haloes which have at least 20 particles or more.

To populate the dark matter haloes with galaxies, we need to know the number and the luminosities of galaxies to be assigned to each halo. Furthermore, we also need to assign phase space coordinates to each of these galaxies. We use the CLF described in Cacciato et al. (2009) for the first purpose. The CLF is a priori split into a contribution from centrals and satellites, i.e. $\Phi(L|M) = \Phi_{\rm c}(L|M) + \Phi_{\rm s}(L|M)$. Here, $\Phi_{\rm c}(L|M){\rm d}L$ denotes the conditional probability that a halo of mass M harbours a central galaxy of luminosity between L and $L + \mathrm{d}L$, and $\Phi_{\rm s}(L|M){\rm d}L$ denotes the average number of satellites of luminosity between L and $L + \mathrm{d}L$. The parameters that describe the CLF are constrained using the luminosity function (Blanton et al. 2005) and the luminosity dependence of the correlation length of galaxies (Wang et al. 2007) in SDSS.

Let us consider a halo of mass M. The luminosity of the central galaxy within this halo is sampled from the distribution $\Phi_{\rm c}(L|M)$. The average number of satellites that have a luminosity greater than $L_{\rm min} = 10^9\,h^{-2}{\rm L}_\odot$ and reside within haloes of mass M is given by

$$\langle N_{\rm sat}\rangle(M) = \int_{L_{\rm min}}^{\infty} \Phi_{\rm s}(L|M){\rm d}L\,. \tag{3.1}$$

We assume Poisson statistics for the occupation number of satellites (Kravtsov et al. 2004; Yang et al. 2005a, 2008) and assign $N_{\rm sat}$ galaxies to the halo where $N_{\rm sat}$ is drawn from

$$P(N_{\rm sat}|M) = \exp(-\mu)\frac{\mu^{N_{\rm sat}}}{N_{\rm sat}!}\,, \tag{3.2}$$

with $\mu = \langle N_{\rm sat}\rangle(M)$. The luminosities of these satellite galaxies are drawn from the distribution $\Phi_{\rm s}(L|M)$.

Phase space coordinates are assigned to the galaxies in the following manner. The central galaxy is assumed to reside at rest at the centre of the halo. Therefore, it has the same phase space coordinates as the parent dark matter halo. As in Chapter 2, we assume that the halo

is spherical and that the dark matter density distribution, $\rho(r|M)$, follows the universal NFW profile (Navarro et al. 1997) given by Eq. (2.22). Similarly the number density distribution of satellite galaxies, $n_{\rm sat}(r|M)$, is assumed to follow the profile given by Eq. (2.24) which has two additional parameters γ and $\mathcal{R}$ which allow the satellite galaxies to be spatially biased with respect to dark matter particles. For populating the MGC, we adopt $\gamma = \mathcal{R} = 1$ which implies that the satellites trace the dark matter density distribution in an unbiased manner. The distribution, $n_{\rm sat}(r|M)$, is normalized such that

$$\langle N_{\rm sat}\rangle(M) = 4\pi \int_0^{r_{\rm vir}} n_{\rm sat}(r|M)\, r^2\, {\rm d}r\,. \tag{3.3}$$

The radial coordinates of the satellite galaxies with respect to the center of the halo are sampled from the distribution $n_{\rm sat}(r|M)$. The satellite distribution around centrals is assumed to be spherically symmetric and random angular coordinates are assigned to the satellite galaxies. At the assigned position for every satellite galaxy, velocities along each of the three axes are drawn from a Gaussian,

$$f(v_j) = \frac{1}{\sqrt{2\,\pi}\sigma_{\rm sat}(r|M)} \exp\left[-\frac{v_j^2}{2\sigma^2_{\rm sat}(r|M)}\right], \tag{3.4}$$

where v_j denotes the relative velocity of the satellite with respect to the central along axis j and $\sigma^2_{\rm sat}(r|M)$ denotes the radial velocity dispersion at a distance r from the centre of the halo. Here isotropy of orbits is assumed, i.e. the velocity dispersion along the $j^{\rm th}$ axis, σ_j^2, equals $\sigma^2_{\rm sat}(r|M)$. The radial velocity dispersion is related to $\rho(r|M)$ and $n_{\rm sat}(r|M)$ via the Jeans equation. We use Eq. (2.25) from Chapter 2 to determine the radial velocity dispersion, $\sigma^2_{\rm sat}(r|M)$, within the halo. The radial velocity dispersion is used in the distribution given by Eq. (3.4) to assign velocities to satellites. The entire procedure of assigning central and satellite galaxies is repeated for all the dark matter haloes within the simulation.

Our aim is to construct a mock redshift survey that mimics the SDSS. Therefore, 2x2x2 identical galaxy−populated simulation boxes (which have periodic boundary conditions) are stacked together. A $({\rm RA}, {\rm DEC})$ coordinate frame is defined with respect to a virtual observer at one of the corners of the stack. The apparent magnitude of each galaxy is computed according to its luminosity and distance from the observer. The line-of-sight (los) velocity of the galaxy is calculated by adding its peculiar velocity to the velocity of the cosmological flow. A random velocity drawn from a Gaussian distribution with a dispersion of $35\ \ {\rm km\,s^{-1}}$ is further added along the los to account for the spectroscopic redshift errors present in the SDSS. The redshift as seen by the virtual observer is then computed using the total velocity. We only consider galaxies with an observed redshift $z < 0.15$ and an apparent magnitude brighter than 17.77. This flux limited catalogue is denoted henceforth by MOCKF and has 289,500 galaxies above an absolute luminosity of $10^9\,h^{-2}{\rm L}_\odot$. MOCKF is used in Appendix 3.A to investigate potential selection biases associated with the selection of central galaxies. In addition to MOCKF, we construct a volume limited sample, MOCKV, of galaxies that lie in the redshift range $0.02 \leq z \leq 0.072$ and

have luminosities greater than $10^{9.5}\,h^{-2}\mathrm{L}_{\odot}$. It consists of 69,512 galaxies. In what follows, we use the volume limited sample MOCKV to validate our method for quantifying the kinematics (Section 3.4), to validate the method to infer the number density distribution of satellites (Section 3.5) and finally to confirm that the mean and scatter of the MLR can be reliably recovered using the kinematics of satellites (Section 3.6).

Note that we have made the simplifying assumption that the satellites are unbiased tracers of the dark matter for the construction of our mock catalog. The effects of a bias in the satellite number density distribution were investigated in detail by van den Bosch et al. (2004). In particular, they have shown that if the satellites are spatially antibiased with respect to the dark matter, then the velocity dispersion of satellites is systematically higher than the dark matter velocity dispersion. In Chapter 4, when we use data from SDSS to constrain the MLR, we do take into account the fact that satellite galaxies may be spatially antibiased with respect to the dark matter. We have also assumed that the angular distribution of satellites is uniform which is not realistic. Note that this is not a concern as the random stacking of haloes to infer the kinematics of satellite galaxies will wash away any non-uniformities.

3.3 Selection Criteria to identify Centrals and Satellites

Large-scale galaxy redshift surveys such as the SDSS allow the selection of a statistically significant sample of satellites. Since the observed galaxies cannot be a priori classified as centrals and satellites, it is important to use selection criteria that can correctly identify central galaxies and the satellites which orbit around them. In this section, we describe the selection criteria that we use to identify the central and satellite galaxies.

A galaxy is identified as a central if it is at least $f_{\rm h}$ times brighter than every other galaxy within a cylindrical volume specified by $R < R_{\rm h}$ and $|\Delta V| < (\Delta V)_{\rm h}$ (see Fig. 3.1). Here, R is the physical separation from the candidate central galaxy projected on the sky and ΔV is the los velocity difference. Around each of the identified centrals, satellites are those galaxies that are at least $f_{\rm s}$ times fainter than their central galaxy and lie within a cylindrical volume specified by $R < R_{\rm s}$ and $|\Delta V| < (\Delta V)_{\rm s}$. The identification of the central galaxies depends on the parameters $R_{\rm h}, (\Delta V)_{\rm h}$ and $f_{\rm h}$, while the selection of satellites depends on the parameters $R_{\rm s}, (\Delta V)_{\rm s}$ and $f_{\rm s}$. The values of these parameters also determine the level of contamination of the sample due to falsely identified centrals and falsely identified satellites (hereafter interlopers). The false identification of centrals can be minimized by choosing large values of $R_{\rm h}, (\Delta V)_{\rm h}$ and $f_{\rm h}$ so that the selected central is the dominant galaxy in a large volume. On the other hand, minimizing the interlopers requires small values of $R_{\rm s}$ and $(\Delta V)_{\rm s}$. A large value of $f_{\rm s}$ further guarantees that the selected satellites are small and do not dominate the kinematics of the halo (i.e. can safely be considered as test particles). Although stricter restrictions yield cleaner samples, they also reduce the sample size significantly. This makes the velocity dispersion measurements noisy. Thus, there is a tradeoff between the contamination level and the

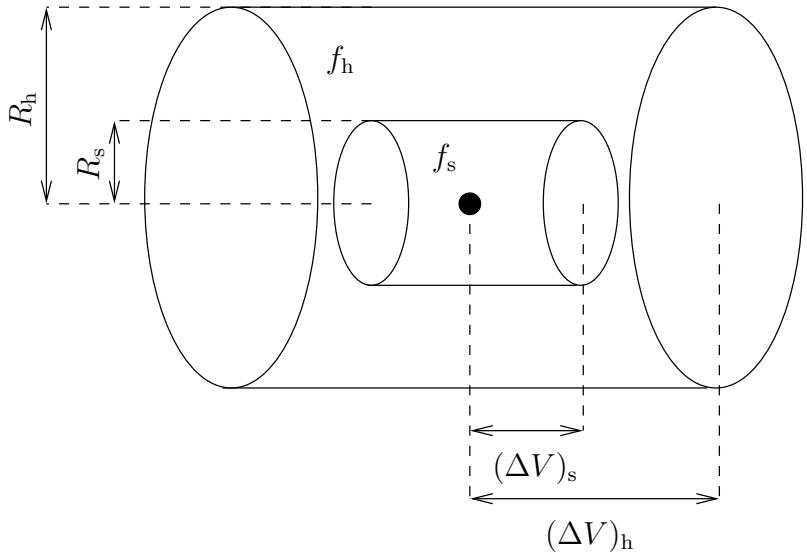

Figure 3.1: Schematic diagram of a selection criterion. Two coaxial cylinders are defined around each galaxy (represented by a solid dot). The axis is along the los while the face of each cylinder is parallel to the plane of the sky.

sample size.

Most authors have chosen fixed values for the selection criteria parameters, independent of the luminosity of the galaxy under consideration (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; Norberg et al. 2008). Since brighter centrals on average reside in more extended haloes, van den Bosch et al. (2004) advocated an aperture which scales with the virial radius of the halo around the galaxy. They used iterative criteria which scale the cylindrical aperture based upon the estimate of the velocity dispersion around the central after every iteration. In this thesis, we also use these iterative criteria to select centrals and satellites. In Appendix 3.A, we compare the performance of our iterative criteria with the restrictive selec-

Table 3.1: Selection criteria

SC	$R_{\rm h}$ Mpc/h	$(\Delta V)_{\rm h}$ km/s	$f_{\rm h}$	$R_{\rm s}$ Mpc/h	$(\Delta V)_{\rm s}$ km/s	$f_{\rm s}$
ITER	2.0	4000	1.0	0.5	4000	1.0
	$0.8\sigma_{200}$	$1000\sigma_{200}$	1.0	$0.15\sigma_{200}$	4000	1.0
N08	1.0	2400	2.0	0.4	1200	8.0

The parameters used to specify the inner and the outer cylinders around a galaxy for the selection criteria used in this chapter (ITER) and the selection criteria used in N08. The first row for ITER denotes the parameters used in the first iteration, while the second row denotes the parameters used in subsequent iterations. The velocity dispersion, $\sigma_{\rm sat}$ in units of $200\ {\rm km\,s^{-1}}$ is denoted by σ_{200} and is used to scale the cylinders in every iteration.

tion criteria used by N08 in there analysis. The parameter set $\{R_{\rm h}, (\Delta V)_{\rm h}, f_{\rm h}, R_{\rm s}, (\Delta V)_{\rm s}, f_{\rm s}\}$ that defines the inner and outer cylinders for the iterative criteria (ITER) is listed in Table 3.1. The first row lists the parameters for the first iteration while the next row lists the scaling of these parameters in the subsequent iterations. In short, we proceed as follows:

1. Use fixed values of the aperture size to select centrals and satellites in the first iteration.
2. Fit the velocity dispersion of the selected satellites as a function of the central galaxy luminosity, $\sigma_{\rm sat}(L_{\rm c})$, with a simple functional form (see Section 3.4.1).
3. Select new centrals and satellites by scaling the inner and the outer cylinders based on the estimate of the velocity dispersion.
4. Repeat 2 and 3 until $\sigma_{\rm sat}(L_{\rm c})$ has converged to the required accuracy.

For step 3, we adopt the aperture scalings used in van den Bosch et al. (2004). These aperture scalings were optimised to yield a large number of centrals and satellites, but at the same time reduce the interloper contamination. The values chosen for $R_{\rm h}$ and $R_{\rm s}$ approximately correspond to 2 and 0.375 times the virial radius, $r_{\rm vir}$.

3.4 Satellite Kinematics

In this section, we describe how to measure and model the velocity dispersion−luminosity relation, $\sigma_{\rm sat}(L_{\rm c})$, using the satellites identified by the selection criteria. The relation $\sigma_{\rm sat}(L_{\rm c})$ can be measured either by binning the satellites by central galaxy luminosity or by using an unbinned estimator. We use the unbinned estimate after every iteration of the selection criteria to scale the selection aperture. However, to quantify the kinematics of the final sample of satellites, we use the binned estimator, for reasons which we describe further in the text. In the following subsections, we describe the unbinned and the binned estimators for $\sigma_{\rm sat}(L_{\rm c})$ and finally an analytical model for the same.

3.4.1 Unbinned Estimates

We use a maximum likelihood method to estimate the relation $\sigma_{\rm sat}(L_{\rm c})$ from the velocity information of the selected satellites after every iteration of the selection criteria. Let σ_{200} denote $\sigma_{\rm sat}(L_{\rm c})$ in units of 200 $\rm km\,s^{-1}$ and L_{10} denote the luminosity of the central galaxy in units of $10^{10}\,h^{-2}{\rm L}_{\odot}$. Following van den Bosch et al. (2004), we parametrize σ_{200} as,

$$\sigma_{200}(\log L_{10}) = a + b\,(\log L_{10}) + c\,(\log L_{10})^2. \tag{3.5}$$

Let $f_{\rm int}$ denote the interloper fraction and assume that this fraction is independent of the luminosity of the central galaxy and ΔV. The probability for a selected satellite to have a los

Table 3.2: Selection criteria parameters

Sample	a	b	c
MOCKV	2.06	0.45	0.25
MOCKF	2.05	0.50	0.23

The parameters used in Eq. (3.5) to define σ_{200} as a function of the luminosity of a galaxy in the final iteration for samples MOCKV and MOCKF.

velocity difference of $\Delta V\ \mathrm{km\,s^{-1}}$ with respect to the central is then given by

$$P(\Delta V) = \frac{f_{\rm int}}{2(\Delta V)_{\rm s}} + \frac{1-f_{\rm int}}{\bar{\omega}} \exp\left[-\frac{(\Delta V)^2}{2\sigma_{\rm eff}^2}\right], \tag{3.6}$$

where, $\sigma_{\rm eff} = [\sigma_{\rm sat}^2 + \sigma_{\rm err}^2]^{1/2}$ is the effective velocity dispersion in the presence of the redshift errors and the factor

$$\bar{\omega} = \sqrt{2\pi}\sigma_{\rm eff}\,{\rm erf}\left[\frac{(\Delta V)_{\rm s}}{\sqrt{2}\sigma_{\rm eff}}\right], \tag{3.7}$$

is such that the $P(\Delta V)$ is properly normalized to unity. In our attempt to mimic SDSS, we have added a Gaussian error of $35\ \mathrm{km\,s^{-1}}$ to the velocity of each galaxy in the mock catalog. Therefore, the error on the velocity difference, ΔV, of the central and satellite galaxies is $\sigma_{\rm err} = \sqrt{2}\ \times\ 35\ \mathrm{km\,s^{-1}}$ which adds in quadrature to $\sigma_{\rm sat}$ to yield $\sigma_{\rm eff}$.

We use Powell's direction set method to determine the parameters ($a,\ b,\ c,\ f_{\rm int}$) that maximize the likelihood $\mathcal{L} = \sum_i \ln[P(\Delta V)]_i$, where the summation is over all the selected satellites. This yields a continuous estimate of $\sigma_{\rm sat}(L_{\rm c})$ without the need to bin the los velocity information of satellites according to the luminosity of the central galaxy. The parameter set ($a,\ b,\ c$) fitted in the last but one iteration determines the size of the apertures used to select the final sample of satellites. The values of these parameters for the samples investigated in this chapter are listed in Table 3.2.

3.4.2 Binned Estimates

We use a binned estimator to quantify the kinematics of the final sample of satellites. The binned estimator allows us to relax the simplistic assumption of $f_{\rm int}$ being independent of $L_{\rm c}$. More importantly, the binned estimator allows us, in a straightforward manner, to measure $\sigma_{\rm sat}(L_{\rm c})$ using two different weighting schemes $-$ satellite-weighting and host-weighting. Most studies in the literature have used one of these two weighting schemes to infer the mean of the MLR. However, as demonstrated in Chapter 2, the mean of the MLR inferred from the velocity dispersion in any one of these two schemes is degenerate with the scatter in the MLR. This degeneracy can be broken by modelling the velocity dispersions in both schemes *simultaneously*. In what follows, we briefly explain these two weighting schemes in turn and then verify that the velocity dispersions in both schemes can be accurately recovered from the MGC.

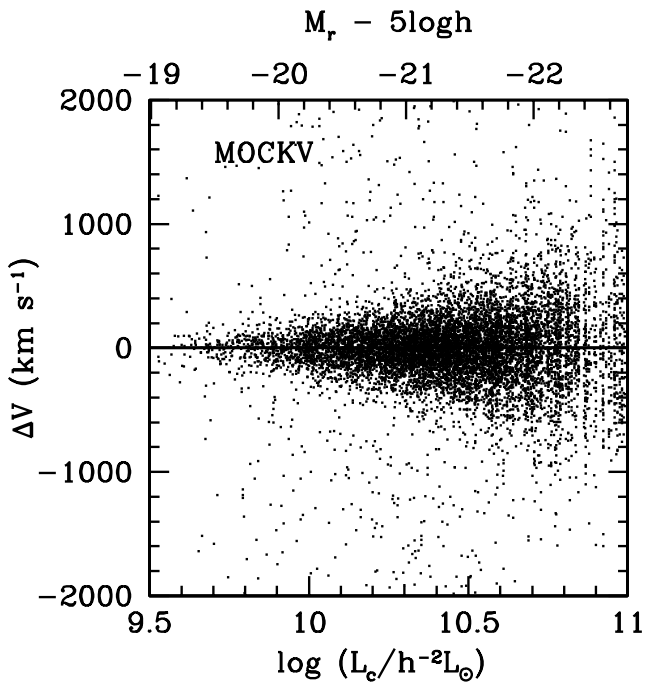

Figure 3.2: Scatter plot of the velocity difference, ΔV, between the satellites and their centrals as a function of the central galaxy luminosity. The satellites were obtained by applying the iterative selection criteria to MOCKV.

To measure the velocity dispersion of satellites in the satellite-weighting scheme, we obtain the distribution of velocities of the satellites, $P(\Delta V)$, with respect to their centrals for several bins of central galaxy luminosity. Each bin has a width $\Delta \log[L_{\rm c}] = 0.15$. In this scheme, the centrals that have a larger number of satellites clearly contribute more to the $P(\Delta V)$ distribution than those which have a smaller number of satellites. Therefore, the resulting scatter in $P(\Delta V)$ is a satellite-weighted average of the velocity dispersions around the stacked centrals (see Chapter 2 for a detailed discussion). The dispersion obtained using this scheme is denoted henceforth by $\sigma_{\rm sw}$.

One has to undo the satellite-weighting described above in order to measure the host-weighted velocity dispersion. This can be accomplished by introducing a weight $w = N^{-1}$ for each central−satellite pair while constructing the $P(\Delta V)$ distribution (van den Bosch et al. 2004; Becker et al. 2007; Conroy et al. 2007). Here, N denotes the number of satellites selected around the central under consideration. Therefore, in this scheme each central receives a total weight of unity irrespective of the number of satellites it hosts. The scatter in this weighted $P(\Delta V)$ distribution is the host-weighted velocity dispersion and is denoted henceforth by $\sigma_{\rm hw}$.

The procedure to obtain the scatter in the $P(\Delta V)$ distributions is the same for both the satellite-weighted and the host-weighted case. This procedure must account for the interlopers and the redshift errors present in MOCKV. In what follows, we illustrate this procedure only for the satellite-weighted case.

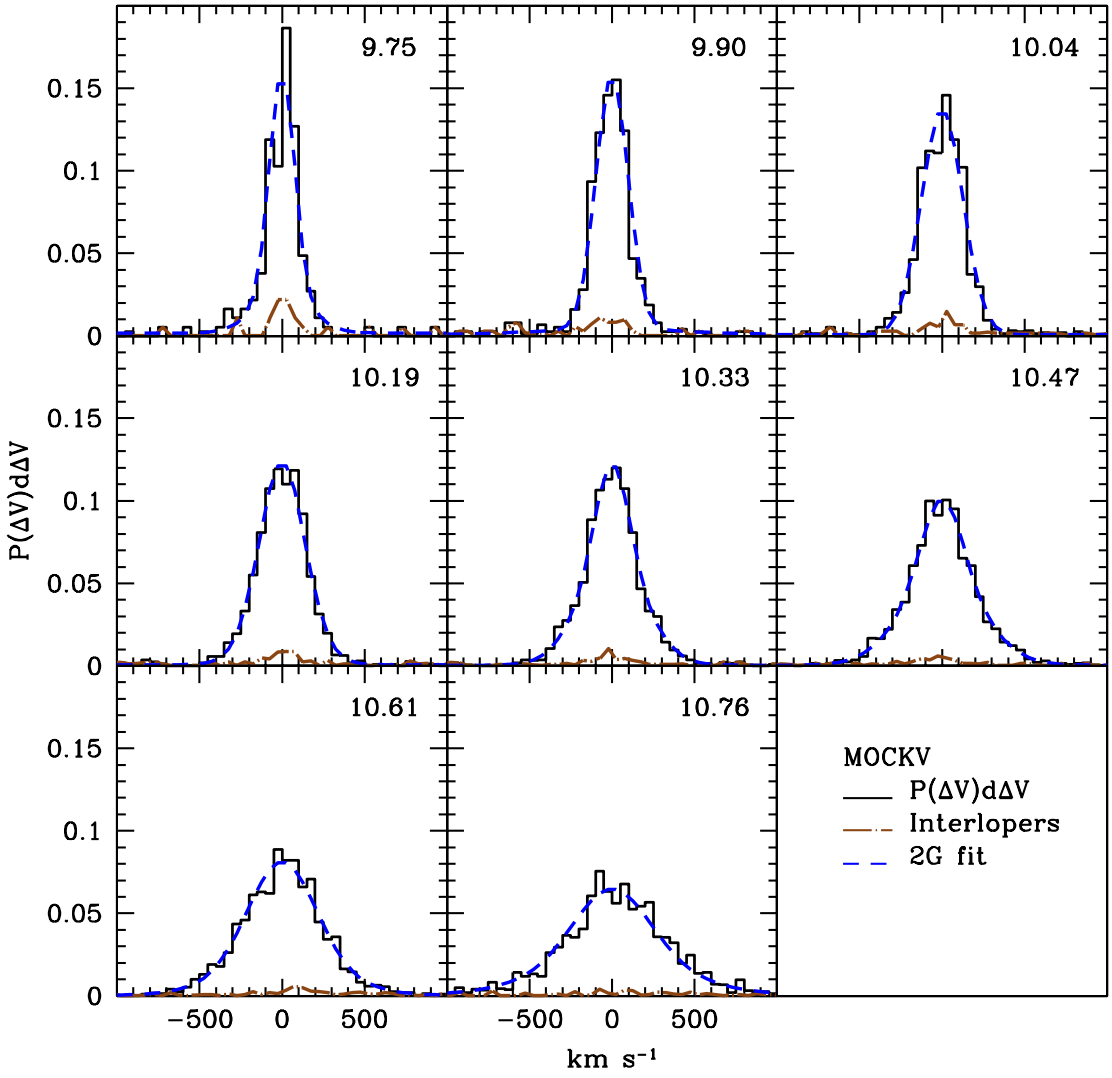

Figure 3.3: The satellite-weighted $P(\Delta V)$ distributions of satellites around centrals selected in several luminosity bins from MOCKV. The average $\log(L_{\rm c}/\ h^{-2}{\rm L}_{\odot})$ for each bin is indicate at the upper right corner of every panel. The (brown) dot-dashed line at the bottom of each distribution shows the contamination of the $P(\Delta V)$ distributions due to the interlopers. The (blue) dashed lines indicate the double-Gaussian fits.

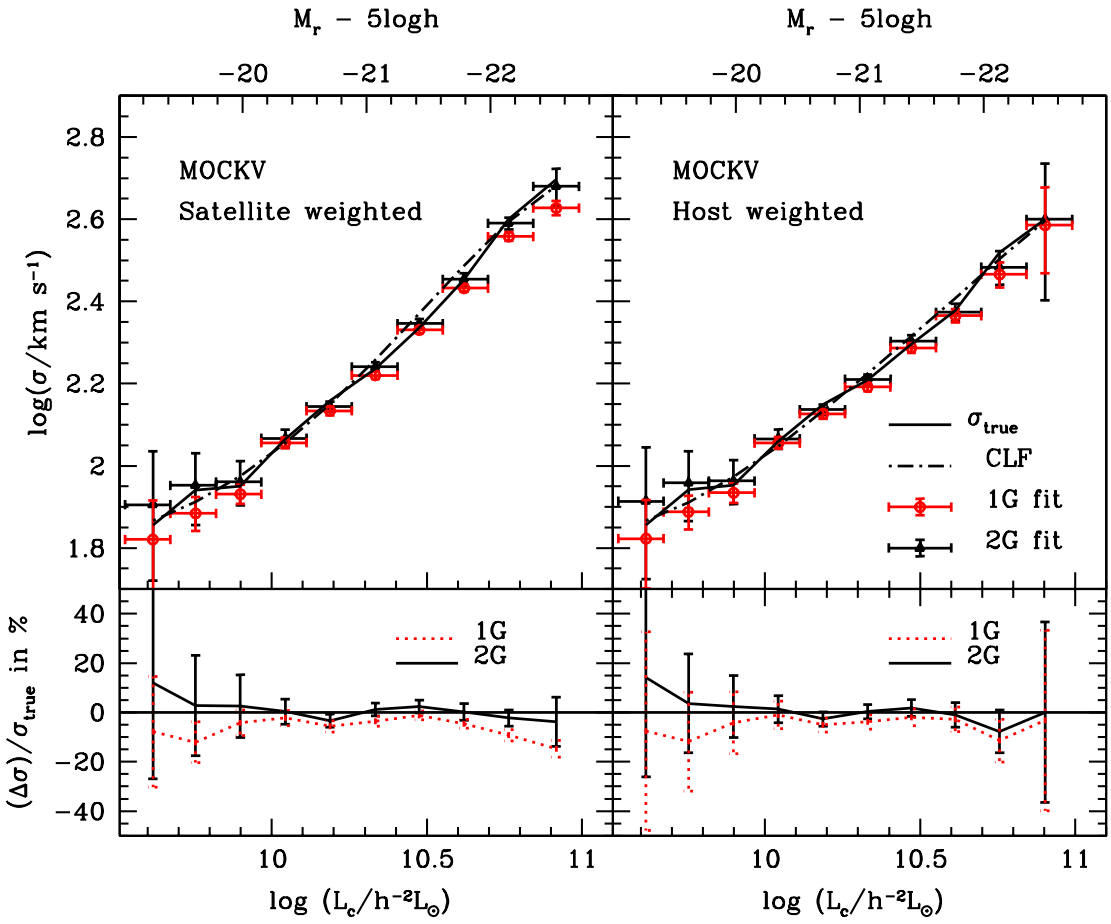

Figure 3.4: Upper panels show the satellite-weighted and the host-weighted velocity dispersions recovered from MOCKV. The (red) circles show values recovered from a single Gaussian fit while the (black) triangles show those from the double Gaussian fit. The solid line shows the variance of the true satellites and the dot-dashed line shows the analytical prediction using the halo occupation statistics of centrals from the MGC. The bottom panels show the percentage deviation of the single and double Gaussian fits from the variance of the true satellites.

Fig. 3.2 shows the scatter plot of velocity difference ΔV of the selected satellites and the centrals as a function of the luminosity of the centrals. The satellite-weighted $P(\Delta V)$ distributions of the satellites selected from MOCKV for several central luminosity bins are shown in Fig. 3.3. Dot-dashed lines show the contamination of the $P(\Delta V)$ distributions due to interlopers and are barely visible at the bottom of each distribution. This confirms the claims in van den Bosch et al. (2004) that the iterative criteria yield a small fraction of interlopers with a weak dependence on L_c and that the interlopers can be modelled as a constant contribution to the velocity distribution independent of ΔV.

A simple way to estimate the scatter of a $P(\Delta V)$ distribution is to fit a Gaussian plus a

constant model given by

$$P(\Delta V) = a_0 + a_1 \exp\left[\frac{-(\Delta V)^2}{2\sigma_{\rm eff}^2}\right]. \tag{3.8}$$

Here, a_0 denotes the constant (with respect to ΔV) interloper background, a_1 is the normalization of the Gaussian and $\sigma_{\rm eff}$ is the effective dispersion in the presence of the redshift errors.

The velocity dispersion obtained using a single Gaussian plus constant model fit can be systematically affected if the $P(\Delta V)$ distribution is intrinsically non-Gaussian. Diaferio & Geller (1996) demonstrated that the velocity distribution can be non-Gaussian partly due to mass mixing (which is the result of stacking haloes of different mass) and partly due to the unrelaxed state of a halo. The second moment of such a non-Gaussian distribution can be estimated with a double Gaussian plus a constant model (Becker et al. 2007) given by

$$P(\Delta V) = a_0 + a_1 \exp\left[\frac{-(\Delta V)^2}{2\sigma_1^2}\right] + a_2 \exp\left[\frac{-(\Delta V)^2}{2\sigma_2^2}\right]. \tag{3.9}$$

The scatter, $\sigma_{\rm eff}$, in this case is such that

$$\sigma_{\rm eff}^2 = \frac{a_1\sigma_1^3 + a_2\sigma_2^3}{a_1\sigma_1 + a_2\sigma_2} = \sigma_{\rm sw}^2 + \sigma_{\rm err}^2. \tag{3.10}$$

The dashed lines in Fig. 3.3 show the double-Gaussian fits to the $P(\Delta V)$ distributions, respectively.

Fig. 3.4 shows the velocity dispersions obtained from the satellite-weighted $P(\Delta V)$ distributions in the upper left panel and those obtained from the host-weighted $P(\Delta V)$ distributions in the upper right panel. The (red) circles and the (black) triangles indicate the single and the double Gaussian fits respectively. Since the true satellites of centrals selected from MOCKV are known, they can be used to judge the goodness of the fits. The satellite-weighted and the host-weighted velocity dispersions of the true satellites (among the satellites selected using the iterative criteria) are obtained using

$$\sigma_{\rm true}^2 = \frac{\sum_{j=1}^{N_{\rm c}} \sum_{i=1}^{N_j} w_{ij} (\Delta V)_{ij}^2}{\sum_{j=1}^{N_{\rm c}} \sum_{i=1}^{N_j} w_{ij}} - \sigma_{\rm err}^2. \tag{3.11}$$

Here, $N_{\rm c}$ denotes the number of true centrals, N_j denotes the number of true satellites of the $j^{\rm th}$ central and $(\Delta V)_{ij}$ denotes the los velocity difference of the $j^{\rm th}$ central with respect to its $i^{\rm th}$ satellite. The weight $w_{ij} = 1$ for the satellite-weighted case and $w_{ij} = N_j^{-1}$ for the host-weighted case. The true velocity dispersions thus obtained are shown as solid curves in Fig. 3.4.

The bottom panels of Fig. 3.4 show the percentage deviation of both the single and the double Gaussian fits from the velocity dispersions of the true satellites. The single Gaussian fit (the dotted line) underestimates the dispersions systematically by about $5 - 10\%$. The dou-

ble Gaussian fit (the solid line) on the other hand gives an unbiased estimate of both velocity dispersions. Therefore, in what follows, we use the double Gaussian fit for measuring both the satellite-weighted and the host-weighted velocity dispersions (cf. Becker et al. 2007).

3.4.3 Analytical Estimates

We now compare the velocity dispersions obtained from the satellite-weighted and the host-weighted schemes to their analytical expectation values. As detailed in Chapter 2, the satellite-weighted and the host-weighted velocity dispersions depend on the distribution of halo masses of central galaxies specified by $P(M|L_{\rm c})$. The analytical expressions describing the velocity dispersion in these two weighting schemes are

$$\sigma^2_{\rm sw}(L_c) = \frac{\int_0^\infty P(M|L_{\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|L_{\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}{\rm d}M}\,, \tag{3.12}$$

$$\sigma^2_{\rm hw}(L_c) = \frac{\int_0^\infty P(M|L_{\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|L_{\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,{\rm d}M}\,. \tag{3.13}$$

Here, the average number of satellites and the average velocity dispersion of satellites, within the aperture $R_{\rm s}$ in a halo of mass M, are denoted by $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ and $\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}$, respectively. We use results from Chapter 2 to describe these quantities. In particular, we use Eqs. (2.13) and (2.14) to describe $\langle N_{\rm sat}\rangle_{{\rm ap},M}$, and Eq. 2.16 to describe $\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}$.

Note that, when measuring the host-weighted velocity dispersions only satellites of those centrals that have at least one satellite within the search aperture are used. The fraction of such centrals is denoted by $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})$ and is given by the probability that a halo of mass M, which on average hosts $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ satellites within the aperture $R_{\rm s}$, has $N_{\rm sat} \geq 1$ within the aperture. We assume that the satellite occupation numbers (cf. Eq. 3.2) follow Poisson statistics, which is supported by numerical simulations (Kravtsov et al. 2004) and by results from group catalogs based on SDSS (Yang et al. 2005a, 2008) and use the expression for $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})$ given by Eq. 2.19 from Chapter 2. The factor $\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})$ is not considered in the analytical estimate in the satellite-weighting scheme as haloes with zero satellites, by definition, contribute zero weight.

From the analytical description presented in Chapter 2, it is clear that the analytical estimates for the velocity dispersions require the knowledge of

- the density distribution of dark matter haloes, $\rho(r|M)$
- the number density distribution of satellites, $n_{\rm sat}(r|M)$
- the halo occupation statistics of centrals, $P(M|L_{\rm c})$, and the halo occupation number of satellites, $\langle N_{\rm sat}\rangle(M)$.

We assume that the density distribution of dark matter haloes is given by Eq. (2.22) and that the number density distribution of satellites is given by Eq. (2.24) with $\gamma = \mathcal{R} = 1$. The halo

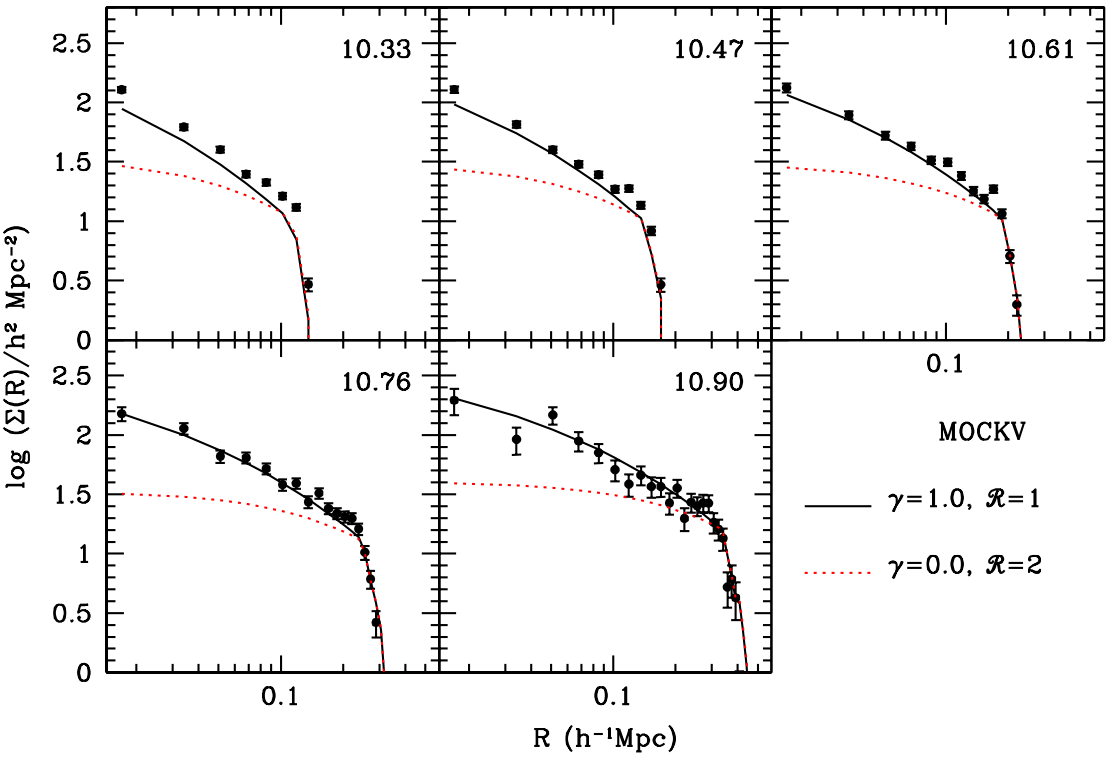

Figure 3.5: The projected number density distributions of the satellites selected from MOCKV as a function of the projected radius in the brightest central luminosity bins. The average luminosity of the bin is indicated at the top right corner in each panel. The errorbars assume Poisson statistics for the number of satellites in each bin. The (black) solid lines indicate the analytical predictions and assume that the satellite number density distribution follows the dark matter distribution in an unbiased manner, i.e. $\mathcal{R} = 1$ and $\gamma = 1$ in eq. (2.24). For comparison, the (red) dotted lines show the analytical predictions that assume $\mathcal{R} = 2$ and $\gamma = 0$.

occupation statistics of centrals, $P(M|L_{\rm c})$ is given by

$$P(M|L_{\rm c}) = \frac{\Phi_{\rm c}(L_{\rm c}|M)n(M)}{\int \Phi_{\rm c}(L_{\rm c}|M)n(M){\rm d}M} \tag{3.14}$$

where $\Phi_{\rm c}(L_{\rm c}|M)$ is the conditional luminosity function of central galaxies and $n(M)$ is the halo mass function. The number of satellite galaxies in a halo of mass M is given by Eq. (3.1). We adopt the $\Phi_{\rm c}(L_{\rm c}|M)$ and $\Phi_{\rm s}(L|M)$ that were used in Section 3.2 to populate the MGC. With this input, we compute the analytical estimates for the velocity dispersions of satellites as a function of luminosity using Eqs. (3.12) and (3.13). The results thus obtained are shown as dot-dashed curves in the corresponding panels of Fig. 3.4. Overall the agreement with the velocity dispersions obtained from the satellites in the MGC is very good, except at intermediate luminosities where the analytical estimates are ~ 5 percent higher than $\sigma_{\rm true}$. This indicates that the central galaxies selected from the MGC do not properly sample the full $P(M|L_{\rm c})$. This can be due to two reasons: (i) a systematic problem with the criteria used to select central galaxies, or (ii) cosmic variance due to the finite volume probed by MOCKV. As we demonstrate in Appendix 3.A our iterative criteria accurately sample the true $P(M|L_{\rm c})$, except for the fact that it misses the haloes of those centrals which have zero satellites. However, this sampling effect is accounted for in our analytical model via Eq. (2.19). In fact detailed tests show that the discrepancies between $\sigma_{\rm true}$ and our analytical estimates are entirely due to cosmic variance in the MGC.

In Appendix 3.A, we also show that the strict selection criteria, that have been abundantly used in the literature, lead to a sample of central galaxies that is biased to reside in relatively low mass haloes. Consequently, the resulting MLR of central galaxies is similarly biased, and has to be interpreted with great care.

3.5 Number Density Distribution of Satellites

As described above, the number density distribution of satellite galaxies, $n_{\rm sat}(r|M)$, is a necessary input to analytically compute the velocity dispersions. The projected number density distribution of satellites, $\Sigma(R|L_{\rm c})$, around centrals of a given luminosity, directly reflects the functional form of $n_{\rm sat}(r|M)$. The distribution $\Sigma(R|L_{\rm c})$ can be directly measured by combining the satellites around centrals of a given luminosity, $L_{\rm c}$, chosen by the selection criteria. However, it is necessary to first assess the impact of the interloper contamination on the measurement of $\Sigma(R|L_{\rm c})$, for which we again make use of the satellite sample selected from MOCKV.

Fig. 3.5 shows, for the five brightest luminosity bins, the azimuthally averaged projected number density distributions of the satellites selected from MOCKV. The errorbars reflect the Poisson noise on the number of satellites in each radial bin. The abrupt cutoff at large R is an artefact due to the parameter $R_{\rm s}$ in the selection criteria which describe the maximum projected

radius within which satellites get selected. Note that, since $R_{\rm s}$ depends upon the luminosity of central galaxies under consideration, this cutoff shifts to larger R with increasing central galaxy luminosity.

The projected number density distribution of satellites around centrals stacked according to luminosity, $\Sigma(R|L_{\rm c})$, can be analytically expressed as,

$$\Sigma(R|L_{\rm c}) = \frac{\int P(M|L_{\rm c})\,\Sigma(R|M)\,{\rm d}M}{\int P(M|L_{\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle(M)_{{\rm ap},M})\,{\rm d}M}. \tag{3.15}$$

Here, $\Sigma(R|M)$ is the projection of $n_{\rm sat}(r|M)$ along the line-of-sight and is given by

$$\Sigma(R|M) = \int_R^{r_{\rm vir}} \frac{n_{\rm sat}(r|M)\,2r\,{\rm d}r}{\sqrt{r^2-R^2}}, \tag{3.16}$$

Using $n_{\rm sat}(r|M)$ given by Eq. (2.24) with $\mathcal{R} = \gamma = 1$ and the true $P(M|L_{\rm c})$ present in the MGC, we analytically compute the expected number density distribution of satellites around centrals of a given luminosity. The solid lines in Fig. 3.5 show the results of this analytical expectation. The small differences between the measured and the analytically obtained distributions are due to the interlopers in the sample. However, the differences become negligible in the brighter luminosity bins. For comparison, the (red) dotted lines show the expected $\Sigma(R|L_{\rm c})$ for $\mathcal{R} = 2$ and $\gamma = 0$. This shows that the parameters $\mathcal{R}$ and γ, that characterize the number density distribution of satellites, can be inferred from the projected number density distributions of the selected satellites.

3.6 Mass−Luminosity Relationship

In the previous sections, using a mock catalog, we have demonstrated that the satellite-weighted velocity dispersions, the host-weighted velocity dispersions and the projected number density distributions of satellites around centrals of a given luminosity can be reliably measured starting from a volume limited redshift catalogue of galaxies. Next, we attempt to infer the MLR of central galaxies from the velocity dispersions measured from MOCKV. The aim is to invert Eqs. (3.12) and (3.13) which describe the dependence of the velocity dispersions on the MLR of central galaxies. In addition to the velocity dispersions, we also measure the average number of satellites per central of a given luminosity, $\langle N_{\rm sat}\rangle(L_{\rm c})$, and use this as a constraint. The dependence of $\langle N_{\rm sat}\rangle(L_{\rm c})$ on the MLR of central galaxies is given by

$$\langle N_{\rm sat}\rangle(L_c) = \frac{\int_0^\infty P(M|L_{\rm c})\langle N_{\rm sat}\rangle_{{\rm ap},M}{\rm d}M}{\int_0^\infty P(M|L_{\rm c})\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M}){\rm d}M}. \tag{3.17}$$

In this section, we first describe the model we use to infer the mean and the scatter of the MLR from the observables $\sigma_{\rm sw}$, $\sigma_{\rm hw}$ and $\langle N_{\rm sat}\rangle$. Next, we use this model to infer the mean and the scatter of the MLR in the MGC and compare it to the true relations present in the MGC.

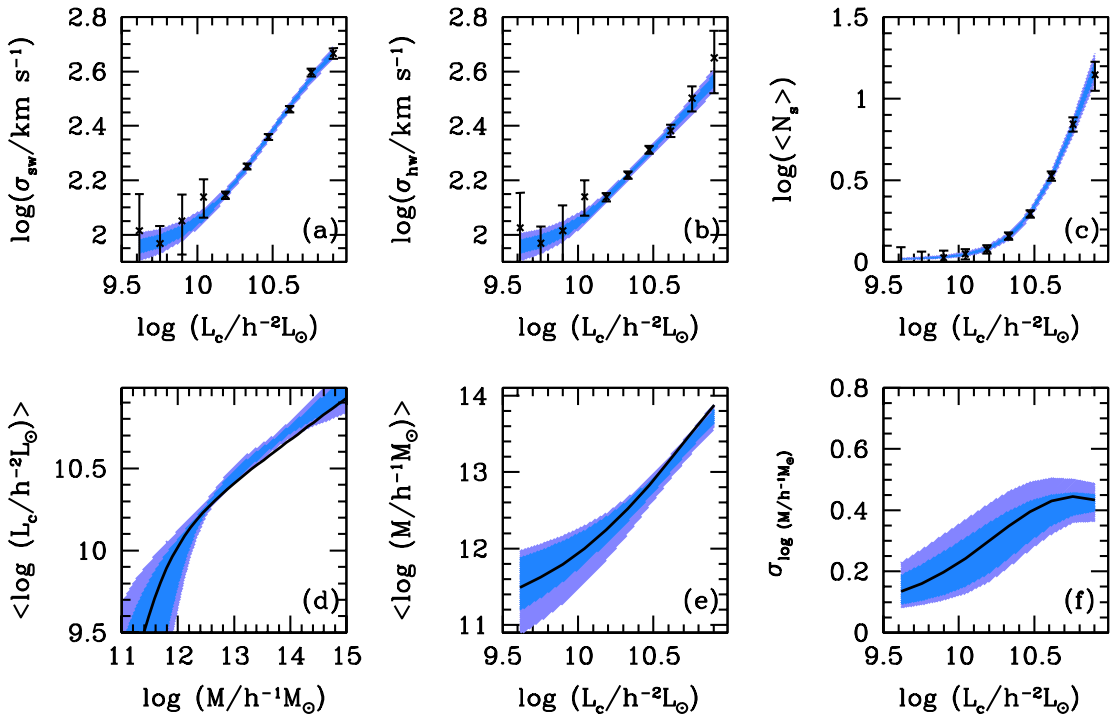

Figure 3.6: The results of the MCMC analysis of the velocity dispersions obtained from MOCKV. Crosses with errorbars in the upper panels denote the data used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the average number of satellites per central in panel (c). The relations recovered from the MCMC analysis are shown in the bottom panels; $\langle \log L_{\rm c} \rangle (M)$ in panel (d), $\langle \log M \rangle (L_{\rm c})$ in panel (e) and $\sigma_{\log M}(L_{\rm c})$ in panel (f). In each panel, the blue and purple colours denote the 68% and the 95% confidence levels. The solid lines in the lower panels denote the true relations present in MOCKV.

Table 3.3: MOCKV: Parameters recovered from the MCMC

Parameter	Input	16%	50%	84%
$\log(L_0)$	9.93	9.64	10.01	10.32
$\log(M_1)$	11.04	10.48	11.28	11.69
γ_2	0.25	0.18	0.26	0.32
$\sigma_{\log L}$	0.14	0.13	0.15	0.17

The input parameters that describe $P(L_{\rm c}|M)$ are compared to the $16^{\rm th}$, $50^{\rm th}$ and the $84^{\rm th}$ percentiles of the corresponding distributions of parameters obtained from the MCMC.

3.6.1 The Model

As mentioned earlier, the analytical computation of $\sigma_{\rm sw}$, $\sigma_{\rm hw}$ and $\langle N_{\rm sat}\rangle$ requires the knowledge of the density distribution of dark matter haloes, the number density distribution of satellites and the halo occupation statistics of centrals and satellites. We assume that the density distribution of dark matter haloes follows the NFW profile given by Eq. (2.22). For the number density distribution of satellites within a halo of mass M, $n_{\rm sat}(r|M)$, we use Eq. (2.24) with $\mathcal{R} = \gamma = 1$. As shown in Section 3.5, the projected number density distributions of satellites selected from MOCKV is consistent with this analytical expression. Next, we describe our model for the halo occupation statistics of the centrals, specified by $P(M|L_{\rm c})$, and the satellites, specified by $\langle N_{\rm sat}\rangle(M)$.

The distribution $P(M|L_{\rm c})$ is related to the complementary distribution, $P(L_{\rm c}|M)$, by Bayes' theorem

$$P(M|L_{\rm c}) = \frac{n(M)P(L_{\rm c}|M)}{\int n(M)P(L_{\rm c}|M){\rm d}M}, \tag{3.18}$$

where $n(M)$ is the halo mass function. We follow Cacciato et al. (2009) and parametrize the distribution $P(L_{\rm c}|M)$[1] as a lognormal in $L_{\rm c}$,

$$P(L_{\rm c}|M){\rm d}L_{\rm c} = \frac{\log(e)}{\sqrt{2\pi}\sigma_{\log L}} \exp\left[-\frac{(\log[L_{\rm c}/\tilde{L_{\rm c}}])^2}{2\sigma^2_{\log L}}\right]\frac{{\rm d}L_{\rm c}}{L_{\rm c}}. \tag{3.19}$$

Here, $\log \tilde{L_{\rm c}}(M)$ denotes the mean of the lognormal distribution and $\sigma_{\log L}$ is the scatter in this distribution. We use four parameters to specify the relation $\tilde{L_{\rm c}}(M)$: a low mass end slope γ_1, a high mass end slope γ_2, a characteristic mass scale M_1, and a normalisation L_0, such that

$$\tilde{L_{\rm c}} = L_0 \frac{(M/M_1)^{\gamma_1}}{[1+(M/M_1)]^{\gamma_1-\gamma_2}}. \tag{3.20}$$

We assume the scatter $\sigma_{\log L}$ to be independent of mass. We do not explore the faint end slope (γ_1) in our analysis as the velocity dispersions at the faint end are very uncertain due to low

[1]Note that the distribution $P(L_{\rm c}|M)$ is equivalent to the CLF for central galaxies. We model $P(L_{\rm c}|M)$ using a few parameters and use Eq. (3.18) to infer $P(M|L_{\rm c})$.

number statistics. Instead, we keep it fixed at 3.273, which is the value obtained from the analysis of the abundance and clustering of galaxies (see Cacciato et al. 2009). This parametrization is motivated by results of Yang et al. (2008) who measure the conditional luminosity function from the SDSS group catalogue described in Yang et al. (2007).

We model the satellite occupation number, $\langle N_{\rm sat}\rangle(M)$, as a power law distribution, given by

$$\langle N_{\rm sat}\rangle(M) = N_s \left(\frac{M}{10^{12}\,h^{-1}{\rm M}_\odot}\right)^{\alpha}\,, \tag{3.21}$$

which adds two more parameters (N_s, α). Thus, in total, our model has six free parameters ($\sigma_{\log L}$, L_0, M_1, γ_2, N_s, α). Given these parameters and the radial number density distribution of satellites (specified by $\mathcal{R}$ and γ), the velocity dispersions $\sigma_{\rm sw}(L_{\rm c})$ and $\sigma_{\rm hw}(L_{\rm c})$ as well as the number of satellites per central, $\langle N_{\rm sat}\rangle(L_{\rm c})$ in an aperture of a given size can be computed using Eqs. (3.12), (3.13) and (3.17) and compared to the measured values. Crosses with errorbars in panels (a), (b) and (c) of Fig. 3.6 show $\sigma_{\rm sw}$, $\sigma_{\rm hw}$ and $\langle N_{\rm sat}\rangle$ as a function of the luminosity of the central obtained from MOCKV, respectively. We use these measurements to constrain the six free parameters of our model.

3.6.2 Monte-Carlo Markov Chain

To determine the posterior probability distributions of the 6 free parameters in our model, we use the Monte-Carlo Markov Chain (hereafter MCMC) technique. The MCMC is a chain of models, each with 6 parameters. At any point in the chain, a trial model is generated with the 6 free parameters drawn from 6 independent Gaussian distributions which are centred on the current values of the corresponding parameters. The chi-squared statistic, $\chi^2_{\rm try}$, for this trial model, is calculated using

$$\chi^2_{\rm try} = \chi^2_{\rm sw} + \chi^2_{\rm hw} + \chi^2_{\rm ns}\,, \tag{3.22}$$

with

$$\chi^2_{\rm sw} = \sum_{i=1}^{10}\left[\frac{\sigma_{\rm sw}(L_{\rm c}[i]) - \hat{\sigma}_{\rm sw}(L_{\rm c}[i])}{\Delta\hat{\sigma}_{\rm sw}(L_{\rm c}[i])}\right]^2\,, \tag{3.23}$$

$$\chi^2_{\rm hw} = \sum_{i=1}^{10}\left[\frac{\sigma_{\rm hw}(L_{\rm c}[i]) - \hat{\sigma}_{\rm hw}(L_{\rm c}[i])}{\Delta\hat{\sigma}_{\rm hw}(L_{\rm c}[i])}\right]^2\,, \tag{3.24}$$

$$\chi^2_{\rm ns} = \sum_{i=1}^{10}\left[\frac{\langle N_{\rm sat}\rangle(L_{\rm c}[i]) - \hat{N}_{\rm sat}(L_{\rm c}[i])}{\Delta\hat{N}_{\rm sat}(L_{\rm c}[i])}\right]^2\,. \tag{3.25}$$

$$\tag{3.26}$$

Here, $\hat{X}$ denotes the observational constraint X and $\Delta\hat{X}$ its corresponding error. The trial step is accepted with a probability given by

$$P_{\rm accept} = \begin{cases} 1.0, & {\rm if} \chi^2_{\rm try} \leq \chi^2_{\rm cur} \\ \exp[-(\chi^2_{\rm try} - \chi^2_{\rm cur})/2], & {\rm if} \chi^2_{\rm try} > \chi^2_{\rm cur} \end{cases} \tag{3.27}$$

where $\chi^2_{\rm cur}$ denotes the χ^2 for the current model in the chain.

We initialize the chain from a random position in the parameter space and discard the first 10^4 models allowing the chain to sample from a more probable part of the distribution. This is called the burn-in period for the chain. We proceed and construct a chain of models consisting of 10 million models. We thin this chain by a factor of 10^4 to remove the correlations between neighbouring models. This leaves us with a chain of 1000 independent models that sample the posterior distribution. We use this chain of models to estimate the confidence levels on the parameters and relations of interest.

In Table 3.3, we compare the $16^{\rm th}, 50^{\rm th}$ and $84^{\rm th}$ percentiles of the distributions of parameters, which characterize $P(L_{\rm c}|M)$, obtained from the MCMC with the corresponding true values of these parameters present in MOCKV. The true parameter values have been recovered within the 68% confidence intervals. The 68 and 95% confidence levels in panels (a), (b) and (c) of Fig. 3.6 show that the models from the MCMC accurately fit the velocity dispersions, $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$, as well as the average number of satellites per central, $\langle N_{\rm sat}\rangle$ as a function of central galaxy luminosity. The confidence levels for the average luminosity of the centrals as a function of the halo mass, $\tilde{L_{\rm c}}(M)$, are shown in panel (d). The confidence levels on the mean, $\langle \log M\rangle(L_{\rm c})$, and the scatter, $\sigma_{\log M}(L_{\rm c})$, of the distribution $P(M|L_{\rm c})$, i.e. the MLR of central galaxies are shown in panels (e) and (f), respectively. They have been calculated using Eq. (3.18) and

$$\langle \log M\rangle(L_{\rm c}) = \int_0^\infty \log M \, P(M|L_{\rm c}) \, {\rm d}M \, , \tag{3.28}$$

$$\sigma_{\log M}(L_{\rm c}) = \left[\int_0^\infty (\log M - \langle \log M\rangle)^2 \, P(M|L_{\rm c}) \, {\rm d}M\right]^{1/2} . \tag{3.29}$$

The solid lines in the lower panels show the corresponding true relations present in MOCKV. Clearly, our method is able to accurately recover the true MLR.

This completes our tests with the MGC. Employing a variety of tests on a realistic MGC, we have established a proof-of-concept that, starting from a redshift survey of galaxies, one can reliably select central and satellite galaxies, quantify the kinematics of the selected satellites around central galaxies and use this information to infer an unbiased estimate of the mean and the scatter of the MLR of central galaxies.

3.7 Summary

The kinematics of satellite galaxies can be used to statistically relate the mean halo masses of central galaxies to their extensive properties. In this chapter, using a realistic mock catalogue, we showed that it is indeed possible to recover the average and the scatter of the scaling relation between halo mass and a central galaxy property (such as the luminosity) using the kinematics of satellites. We thoroughly tested the analysis method at every step. We first tested the performance of the iterative selection criteria, advocated by van den Bosch et al. (2004), to identify central and satellite galaxies and our method to measure the kinematics of the selected satellites. We showed that the kinematics recovered from the selected satellites are a fair representation of the true kinematics of satellite galaxies present in the mock catalogue. We presented an analytical model that properly accounts for the selection biases and showed that the predictions of this analytical model are in good agreement with the measured kinematics of the selected satellites.

In Chapter 2, we have shown that the velocity dispersion of satellites can be measured using two different weighting schemes: satellite-weighting and host-weighting. We have demonstrated a degeneracy between the mean and the scatter of the MLR obtained from either the satellite-weighted or the host-weighted velocity dispersion alone. However, we have also shown that this degeneracy can be broken by using the velocity dispersions in the two schemes simultaneously. In this chapter, we first tested our method using a mock galaxy catalogue. We fitted the measured satellite-weighted and host-weighted velocity dispersions simultaneously using a parametric model for the halo occupation statistics of central and satellite galaxies, and demonstrated that we can reliably obtain confidence levels on the true mean and scatter of the mass−luminosity relation of central galaxies.

In the next two chapters of this thesis, we apply the method developed in this chapter to data from the Sloan Digital Sky Survey. In Chapter 4, we use satellite kinematics to infer the scaling relation between halo mass and luminosity of central galaxies and quantify the scatter in this scaling relation. In Chapter 5, we use satellite kinematics to learn more about the scaling relation between halo mass and the stellar mass of central galaxies along with the scatter in this relation.

We would like to conclude this chapter by emphasizing that, satellite kinematics need not be restricted to the study of isolated haloes of galaxies as is routinely done in the literature. By using a relaxed criteria to identify central−satellite systems and properly accounting for the selection biases, we have shown that satellite kinematics can be effectively used to probe the halo masses in a wide range of environments.

Appendix

3.A Sampling of Central Galaxies

The ultimate goal of satellite kinematics is to probe the halo mass−luminosity relationship (MLR) of central galaxies. In principle, an unbiased estimate for the MLR requires that the central galaxies identified by the selection criteria are an unbiased (sub-)sample with respect to their corresponding dark matter haloes. In this section, we investigate, using the MGC, how our iterative criteria perform in this respect and compare them with the strict criteria used in the literature.

For reasons that will become clear later, we use the flux-limited sample MOCKF for this test. The solid lines in Fig. 3.7 show the distributions of halo masses, $P(M|L_{\rm c})$, for all central galaxies in MOCKF divided in 5 luminosity bins. The average logarithm of the luminosities of central galaxies in each bin is indicated at the top right corner. The dotted lines show the distributions, $P(M|L_{\rm c})$, for all central galaxies that have at least one satellite in the selection aperture defined by our iterative selection criteria. Finally, the histograms show the distributions, $P(M|L_{\rm c})$, of the centrals selected by our iterative criteria. Clearly, the centrals selected by our iterative criteria sample the distribution of halo masses from the dotted lines (and not the solid lines). However, as discussed in Section 3.4.3, this bias is taken into account while modelling the kinematics (see Eq. [2.19]) and therefore allows us to make an unbiased estimate. As shown in Section 3.6, we indeed recover an unbiased MLR from the kinematics of the selected satellites measured around the centrals selected by our iterative criteria.

For comparison, we now repeat this exercise using the strict criteria employed in previous studies. In particular, we adopt the criteria N08 (see Table 3.1) used in Norberg et al. (2008). These criteria identify a galaxy as a central if it is at least $f_{\rm h} = 2$ times brighter than any other galaxy in a fixed (irrespective of the luminosity of the galaxy) aperture cylinder (see Table 3.1) around itself. Satellites are identified as those galaxies that are at least $f_{\rm s} = 8$ times fainter than the centrals and reside in a smaller aperture cylinder defined around the centrals. The values of $f_{\rm h}$ and $f_{\rm s}$ in the N08 criteria are conservative, as the principle goal of their study was to select isolated central galaxies. Applying the N08 criteria to MOCKV selects only 126 satellites around 96 central galaxies. Therefore, to do a meaningful comparison, we apply the N08 criteria to MOCKF for which it selects 657 satellites around 395 centrals. For comparison, our iterative criteria yields $39,951$ satellites around $21,206$ centrals.

Solid lines in the lower panels of Fig. 3.7 are the same as in the upper panels and show the distributions of halo masses, $P(M|L_{\rm c})$, for all central galaxies in MOCKF divided in 5 central luminosity bins. The dotted lines show the $P(M|L_{\rm c})$ for those centrals that have at least one satellite around them which is $f_{\rm s}(=8)$ times fainter than themselves. There is a negligibly small difference in the dotted lines in the two rows due to different values of $f_{\rm s}$. Finally, the histograms show the $P(M|L_{\rm c})$ distributions of the sample of centrals selected by the N08 criteria. Clearly, these do not sample the distributions shown by the dotted lines and the

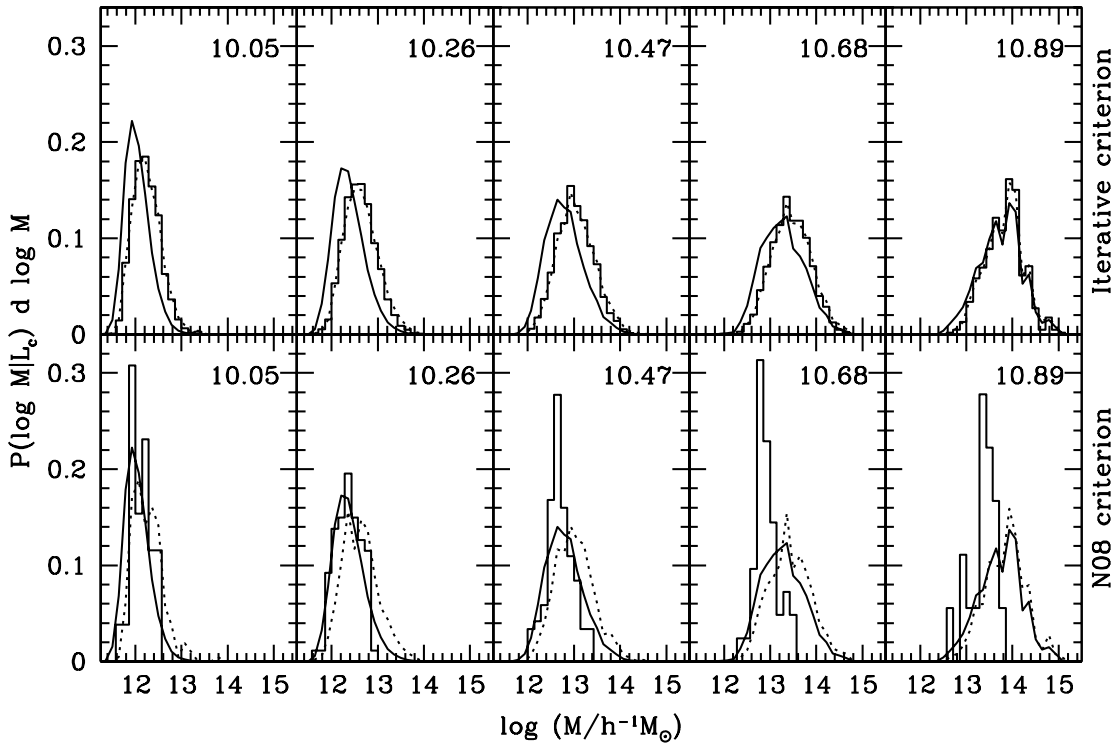

Figure 3.7: Comparison of the sampling of central galaxies using the iterative selection criteria (ITER) used in this chapter and the criteria used in N08. The histograms in the upper (bottom) panel show the distributions of halo masses of central galaxies selected according to ITER (N08). The average $\log(L_{\rm c}/\ h^{-2}{\rm L}_{\odot})$ for central galaxies in each bin is indicated at the top right corner of each panel. The solid lines show the true distributions of halo masses for all the central galaxies and the dotted lines show the distribution of halo masses of those central galaxies that have at least one satellite more than $f_{\rm s}$ times fainter than themselves in the inner cylinder defined by the selection criteria.

distributions are clearly biased towards the low mass end, especially, in the bright luminosity bins. This owes to the fact that Norberg et al. (2008) adopt $f_{\rm h} = 2$, which preferentially selects centrals that do not have satellite galaxies of comparable brightness. This biases the distributions towards the low mass end. Note, though, that this is not a critique regarding their selection criteria; after all, as Norberg et al. (2008) clearly described in their paper, their principal goal is to study the kinematics around *isolated* galaxies. However, it does mean that it is not meaningful to compare their MLR, which is only applicable to isolated galaxies, to that obtained here, which is representative of the entire central galaxy population.

Chapter 4

The Halo Mass−Luminosity Relationship

The contents of this chapter are partially based upon the article More et al. (2009c) published in the Monthly Notices of the Royal Astronomical Society. The reference is
More, S., van den Bosch, F. C., Cacciato, M., et al. 2009, MNRAS, 392, 801
The original article does not contain the results of the colour dependence of the Halo mass−Luminosity relationship of central galaxies. This analysis is part of a manuscript in preparation. The manuscript will be submitted as:
Satellite Kinematics III: Colour and Stellar Mass Dependence
More, S., van den Bosch, F. C., Cacciato, M., et al. 2009

4.1 Introduction

According to the standard framework of galaxy formation, dark matter haloes form gravitational potential wells in which baryons collapse, dissipate their energy and form stars and galaxies (White & Rees 1978; Blumenthal et al. 1984). The complex process of galaxy formation and evolution is believed to be governed by the mass of the dark matter halo in which it occurs. To understand the halo mass dependence of this process, it is important to statistically relate the observable properties (e.g. luminosity) of galaxies to the masses of their dark matter haloes. The kinematics of satellite galaxies, stacked according to the property of their centrals (e.g. luminosity), can be used to determine the scaling relation between central galaxy properties and halo mass. However, the stacking procedure complicates the interpretation of the kinematic signal in terms of the halo mass.

The scaling relation between halo mass and the luminosity of the central galaxies (halo mass−luminosity relation MLR) can, more generally, be specified in terms of the conditional probability $P(M|L_{\rm c})$, which describes the probability for a central galaxy with luminosity $L_{\rm c}$ to reside in a halo of mass M. For a completely deterministic relation between halo mass and

central luminosity, $P(M|L_{\rm c}) = \delta(M - M_0)$, where δ denotes the Dirac−delta function and M_0 is a characteristic halo mass corresponding to centrals of luminosity $L_{\rm c}$. The velocity dispersion, $\sigma(L_{\rm c})$, measured by stacking centrals with luminosity $L_{\rm c}$, then translates into a mass M_0 according to the scaling relation $\sigma^3 \propto M$. However, galaxy formation is a stochastic process and the distribution $P(M|L_{\rm c})$ is expected to have non-zero scatter. If this scatter is appreciable then the stacking procedure results in combining the kinematics of haloes spanning a wide range in halo mass. This complicates the interpretation of the velocity dispersion. We addressed this issue in Chapter 2 (More et al. 2009c); where we investigated a method to measure both the mean and the scatter of the MLR of central galaxies[1] using satellite kinematics. We outlined two different weighting schemes to measure the velocity dispersion of satellites, *satellite-weighting* and *host-weighting*, and showed that the mean and the scatter of the MLR can be inferred by modelling the velocity dispersion measurements in these two schemes simultaneously.

In Chapter 3, we carried out a series of tests on a realistic mock galaxy catalogue to validate our method to infer the MLR of central galaxies from a redshift survey using satellite kinematics. In particular, we showed that our central−satellite selection criterion and the method to measure the kinematics reliably recovers the true kinematics present in the mock catalogue. We also showed that the mean and the scatter of the MLR inferred from the kinematics match the corresponding true relations in the mock catalogue.

In this chapter, we apply this method to the spectroscopic galaxy catalogue from SDSS (Data Release 4) in order to determine both the mean and the scatter of the MLR of central galaxies. In addition, we also analyse the kinematics of the satellites of central galaxies separated on the basis of their colour. This allows us to investigate the impact of the scatter in the colour of central galaxies at a particular luminosity on the scatter in the MLR of central galaxies.

This chapter is organized as follows. In Section 4.2, we describe the SDSS data and apply the iterative criteria described in Section 3.3 to identify central and satellite galaxies for our analysis of the kinematics. We form three different samples to analyse the kinematics of satellites around (i) all central galaxies, (ii) red central galaxies only and (iii) blue central galaxies only. The measurements of the velocity dispersion as a function of the luminosity for central galaxies in each of these samples are presented in Section 4.3. The number density distribution of satellites is an essential ingredient to infer the MLR from the kinematic measurements of the satellites around their centrals. In Section 4.4, we present the projected number density distribution of satellites around centrals from the SDSS data. The results obtained from the MCMC analysis of the kinematics of satellites are presented and compared with results from independent studies in Section 4.5. A summary of all the results is presented in Section 4.6.

For the analysis presented in this chapter, we assume the cosmological parameters from the 3 year data release of WMAP (Spergel et al. 2007), $\Omega_{\rm m} = 0.238$, $\Omega_\Lambda = 0.762$, $h =$

[1]The term MLR refers to the distribution $P(M|L_{\rm c})$. The mean and the scatter of the MLR refer to the average of this distribution and its scatter respectively.

$H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}} = 0.734$, the spectral index of initial density fluctuations $n_\mathrm{s} = 0.951$ and normalization $\sigma_8 = 0.744$.

4.2 Central and Satellite Samples from the SDSS

The SDSS (York et al. 2000) is a joint five−passband (u, g, r, i and z) imaging and medium resolution ($R \sim 1800$) spectroscopic survey. The observations are carried out using a dedicated 2.5-m telescope at the Apache Point Observatory in New Mexico. The SDSS is designed to cover one quarter of the entire sky and obtain images of around 100 million objects and obtain the spectra of around 1 million objects. Data are made available to the scientific community through periodic data releases.

In this study, we use the New York University Value Added Galaxy Catalogue (Blanton et al. 2005, hereafter NYU−VAGC), which is based upon SDSS Data Release 4 (Adelman-McCarthy et al. 2006) but includes a set of significant improvements over the original pipelines. The magnitudes and the colours of the galaxies are based upon the standard SDSS Petrosian technique and have been k−corrected and evolution corrected to $z = 0.1$ using the method described in Blanton et al. (2003a,b). The notations ${}^{0.1}(g-r)$ and ${}^{0.1}M_r - 5\log h$ are used to denote the resulting $(g-r)$ colour and the absolute magnitude of the galaxies. The magnitude limit of the spectroscopic sample is 17.77 in the ${}^{0.1}r$ band.

From this catalogue, we select all galaxies in the main galaxy sample with redshifts in the range $0.02 \leq z \leq 0.072$ and with a redshift completeness limit $\mathcal{C} > 0.8$. We construct a volume limited sample of galaxies with r-band luminosities above $L_\mathrm{min} = 10^{9.5}\,h^{-2}\mathrm{L}_\odot$. This sample consists of $57,593$ galaxies and is henceforth denoted as SDSSV.

In Fig. 4.1, we show the scatter plot of the ${}^{0.1}(g-r)$ colour and the luminosity of galaxies in this sample. The galaxy distribution follows a bimodal distribution in the colour−luminosity scatter plot (Baldry et al. 2004; Blanton et al. 2005; Li et al. 2006). The solid line shows the separation criteria between red and blue galaxies obtained in Yang et al. (2009) by fitting a binormal distribution to the ${}^{0.1}(g-r)$ colour as a function of luminosity. The separation criteria is given by

$$ {}^{0.1}(g-r) = 1.022 - 0.0651\,x - 0.00311\,x^2 \tag{4.1}$$

where $x = {}^{0.1}\,M_r - 5\log h + 23.0$. The sample SDSSV consists of $30,383$ red galaxies and $27,210$ blue galaxies.

We apply the iterative criteria (ITER) outlined in Section 3.3 of Chapter 3 to select centrals and their satellites from SDSSV for the analysis of the MLR of central galaxies without a split in colour. For every iteration, the selection criteria is adjusted based upon the velocity dispersion estimate in the previous iteration. The sample of centrals and satellites identified in this manner is denoted by LA. For the analysis of the MLR around red (blue) central galaxies, a separate sample of central and satellite galaxies is formed by applying a selection criteria which is iteratively tuned using the estimate of velocity dispersions around the red (blue) central

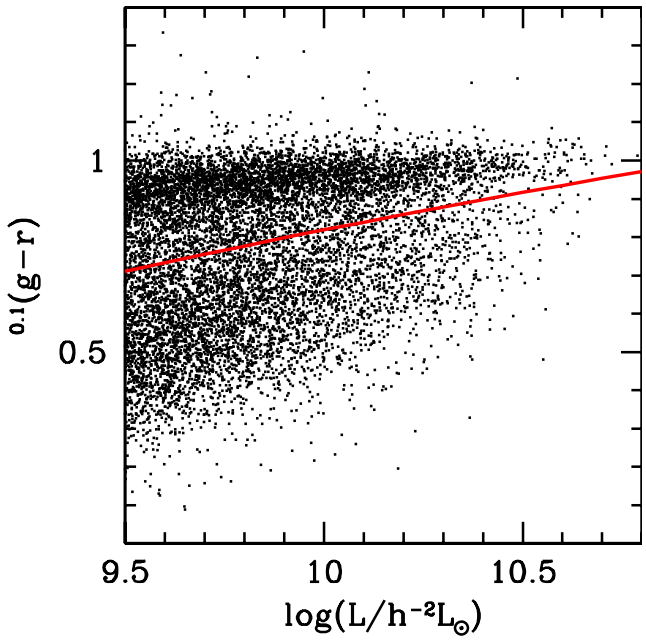

Figure 4.1: Scatter plot of the $^{0.1}(g-r)$ colour and the absolute luminosities of galaxies in the SDSS. To avoid overcrowding, a random subset of 10,000 galaxies was used to make this plot. The red line shows the separation criterion from Yang et al. (2009) that was used to classify the galaxies into red and blue.

galaxies only. The sample that contains the red (blue) central galaxies and their satellites is denoted by LR (LB).

The parameters (a, b, c) in Eq. (3.5) that define the aperture used in the final iteration of the central−satellite selection for samples LA, LR and LB are listed in Table 4.1. Sample LA consists of $3,863$ central galaxies that host at least one satellite galaxy. The total number of satellite galaxies in Sample LA is $6,101$. The number of red central galaxies that host at least one satellite in Sample LR (LB) is $2,503$ $(1,221)$. The number of satellites in Sample LR (LB) is $4,599$ $(1,449)$.

4.3 Velocity Dispersion−Luminosity Relation

The scatter plot of the velocity difference, ΔV, between the satellites and the corresponding centrals as a function of the central luminosity in Sample LA is shown in Fig. 4.2. The scatter plots obtained from Samples LR and LB are shown in the left and the right panels of Fig. 4.3 respectively. The scatter in the velocities of satellites with respect to their centrals clearly increases with central galaxy luminosity for all the three samples. The scatter plot of Sample LB is markedly different from that of Sample LR, not only in terms of the number of satellites but also the amount of scatter in the velocity differences, ΔV at fixed luminosity.

To quantify the velocity dispersion−luminosity relation, we obtain the $P(\Delta V)$ distribu-

Table 4.1: Selection criteria parameters

Parameters	Samples		
	LA	LR	LB
a	2.20	2.25	2.12
b	0.38	0.37	0.44
c	0.33	0.31	0.32

The parameters $a,\ b$ and c that define the criteria used to select central and satellite galaxies for the three samples used in this Chapter.

Table 4.2: Sample LA: Velocity dispersion measurements

$\log(L_{\rm c})$	$\sigma_{\rm sw}$	$\Delta\sigma_{\rm sw}$	$\sigma_{\rm hw}$	$\Delta\sigma_{\rm hw}$
$h^{-2}{\rm L}_{\odot}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$
9.61	108	20	107	20
9.73	148	26	146	25
9.88	159	19	155	18
10.03	162	14	159	12
10.17	214	43	203	20
10.31	254	11	220	11
10.45	272	15	247	13
10.59	412	35	287	39
10.72	470	28	378	73
10.87	650	54	574	254

The velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample LA.

tions in both the satellite-weighting and the host-weighting schemes by combining the velocity differences, ΔV, of satellites within luminosity bins of uniform width $\Delta\log[L_{\rm c}] = 0.15$ for Samples LA and LR, and 0.13 for Sample LB. The satellite-weighted and the host-weighted velocity dispersions are estimated from these distributions by fitting a model that consists of two Gaussians and a constant as described in Section 3.4.2. Fig. 4.4 shows these dispersions as a function of central luminosity for Sample LA. The values of $\sigma_{\rm sw}$, $\sigma_{\rm hw}$ and their associated errors are listed in Table 4.2. Both the satellite-weighted and the host-weighted velocity dispersions increase with the luminosity of the central galaxy. Note that the satellite-weighted velocity dispersions are systematically larger than the host-weighted velocity dispersions. As is evident from Eqs. (2.20) and (2.21), this is a sufficient condition to indicate the presence of scatter in the MLR of central galaxies (see Chapter 2 for a detailed discussion).

The left (right) panel of Fig. 4.5 shows the comparison between the satellite-weighted (host-weighted) velocity dispersions around red and blue central galaxies separately. Both the satellite-weighted and the host-weighted velocity dispersions of satellites around red centrals

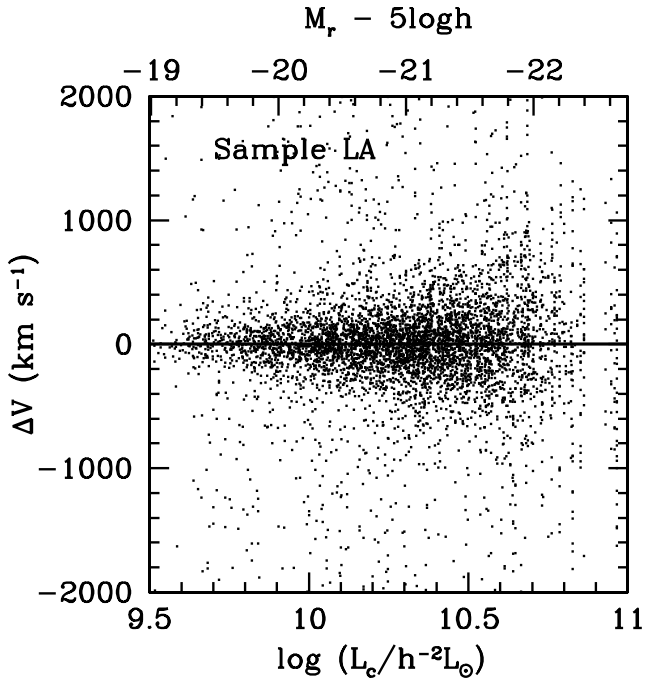

Figure 4.2: Scatter plot of the velocity difference, ΔV, between the satellites and their central galaxies in sample LA as a function of the central galaxy luminosity.

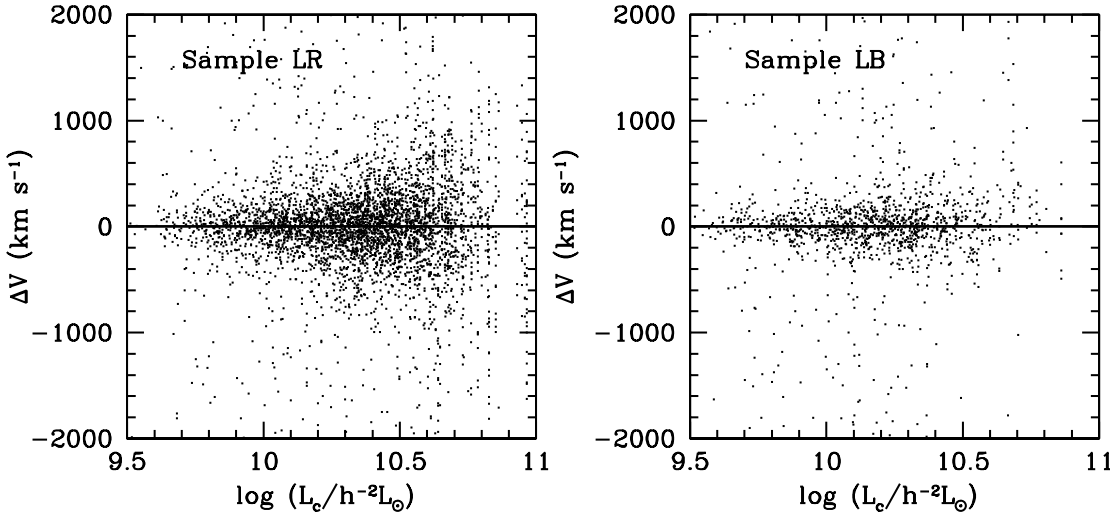

Figure 4.3: Scatter plot of the velocity difference, ΔV, between the satellites and their central galaxies in sample LR sample LB as a function of the central galaxy luminosity are shown in the left and the right panels respectively.

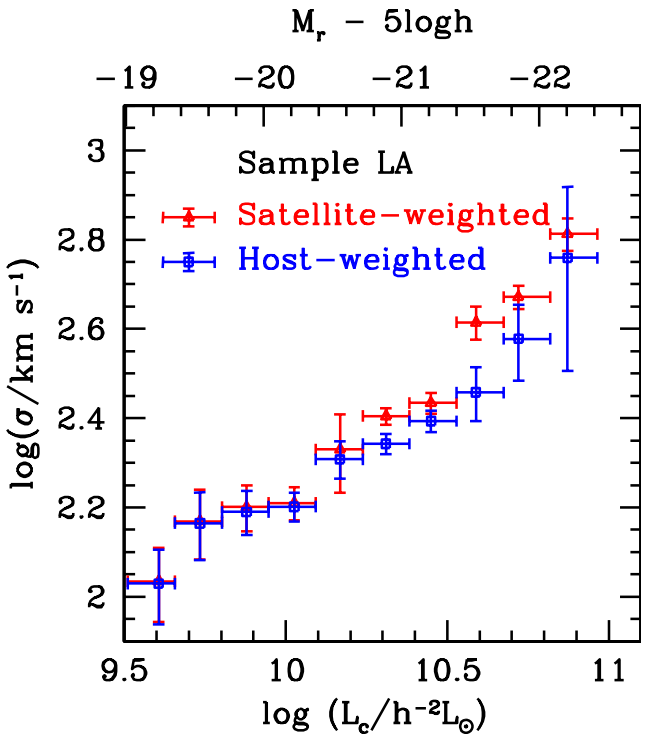

Figure 4.4: The satellite-weighted (red triangles) and the host-weighted (blue squares) velocity dispersions as a function of the central galaxy luminosity obtained from satellites in the sample LA.

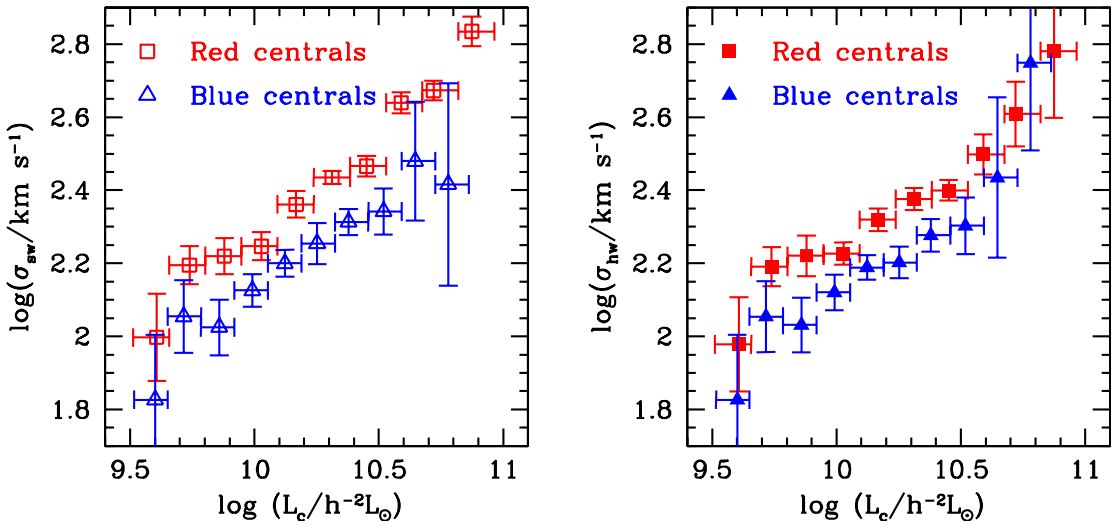

Figure 4.5: Comparison between the satellite-weighted (host-weighted) velocity dispersions as a function of the luminosity of the central galaxy obtained from Sample LR and Sample LB is shown in the left (right) panel.

Table 4.3: Sample LR: Velocity dispersion measurements

$\log(L_{\rm c})$	$\sigma_{\rm sw}$	$\Delta\sigma_{\rm sw}$	$\sigma_{\rm hw}$	$\Delta\sigma_{\rm hw}$
$h^{-2}{\rm L}_\odot$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$
9.61	99	27	95	28
9.74	157	19	155	19
9.88	166	19	166	21
10.03	177	16	169	12
10.17	230	19	209	15
10.31	272	11	238	17
10.45	292	19	251	16
10.59	436	29	315	39
10.72	471	29	406	83
10.87	683	64	603	253

The velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample LR.

Table 4.4: Sample LB: Velocity dispersion measurements

$\log(L_{\rm c})$	$\sigma_{\rm sw}$	$\Delta\sigma_{\rm sw}$	$\sigma_{\rm hw}$	$\Delta\sigma_{\rm hw}$
$h^{-2}{\rm L}_\odot$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$	$\rm km\,s^{-1}$
9.60	67	27	67	27
9.72	113	26	113	25
9.86	106	19	107	19
9.99	134	14	132	15
10.12	159	13	154	12
10.25	180	23	159	16
10.38	206	17	189	19
10.52	220	32	201	36
10.65	302	113	272	137
10.78	260	166	561	310

The velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample LB.

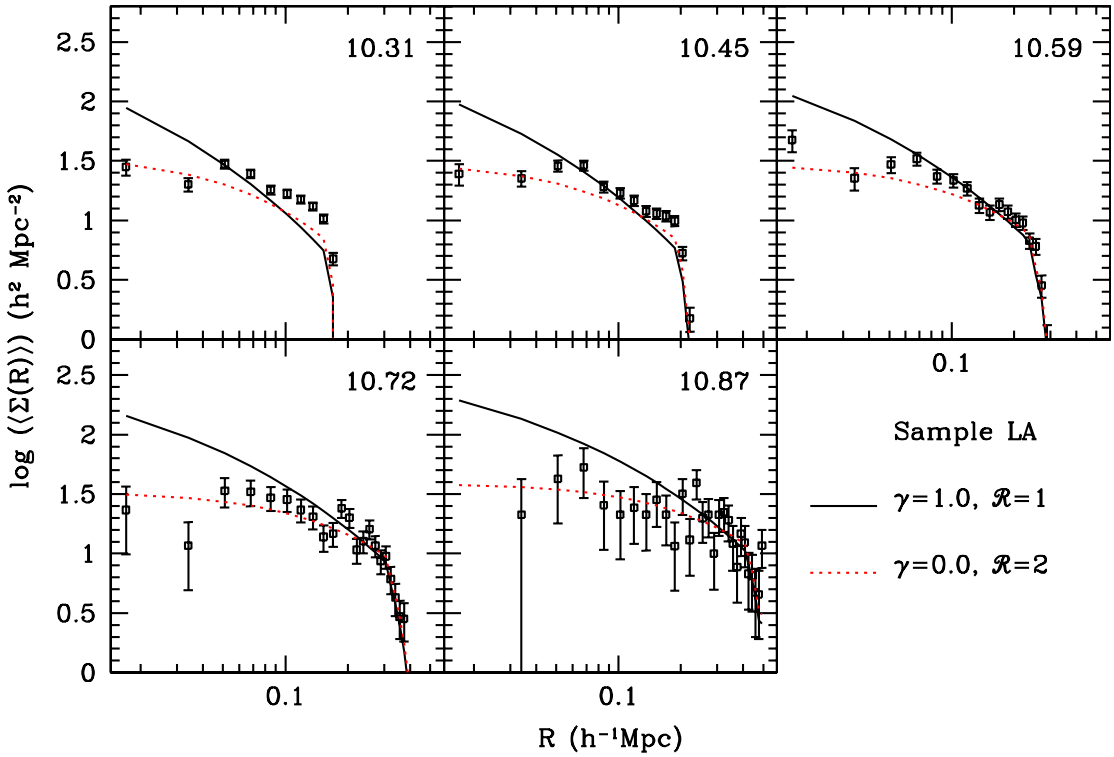

Figure 4.6: The projected number density distributions of satellites around centrals in Sample LA for the five bright luminosity bins (average $\log(L_{\rm c}/\ h^{-2}{\rm L}_{\odot})$ in right corner). The (black) solid curves indicate the expected distributions if the number density distribution of satellites follows the dark matter density, i.e. $\mathcal{R} = \gamma = 1$ in Eq. (2.24). The (red) dotted curves in turn indicate the expected distributions for a model in which the satellite galaxies are a factor 2 less concentrated than dark matter and have a central core in the number density distribution i.e. $\mathcal{R} = 2$, and $\gamma = 0$.

are larger than those around blue centrals. This figure captures the difference between Sample LR and LB observed in the scatter plots of Fig. 4.3. The velocity dispersion measurements for Sample LR and for Sample LB are listed in Tables 4.3 and 4.4, respectively.

4.4 Number Density Distribution of Satellites in SDSS

The model to infer the MLR of central galaxies from the kinematics requires the radial number density distribution of satellites, $n_{\rm sat}(r)$, as an input. In Chapter 3, to infer the MLR of central galaxies from the mock catalog, we used a model of $n_{\rm sat}(r)$ that follows the density distribution of the dark matter in an unbiased manner, i.e., $\gamma = \mathcal{R} = 1$ in Eq. (2.24). However with SDSS, it is not clear what functional form of $n_{\rm sat}(r)$ should be used. In fact, various studies have shown that the satellite galaxies are spatially antibiased with respect to the dark matter (Yang et al. 2005b; Chen 2007, 2008). Rather than including γ and $\mathcal{R}$ as free parameters in our model, we seek to constrain these parameters using the observable $\Sigma(R|L_{\rm c})$. Fig. 4.6 shows the projected

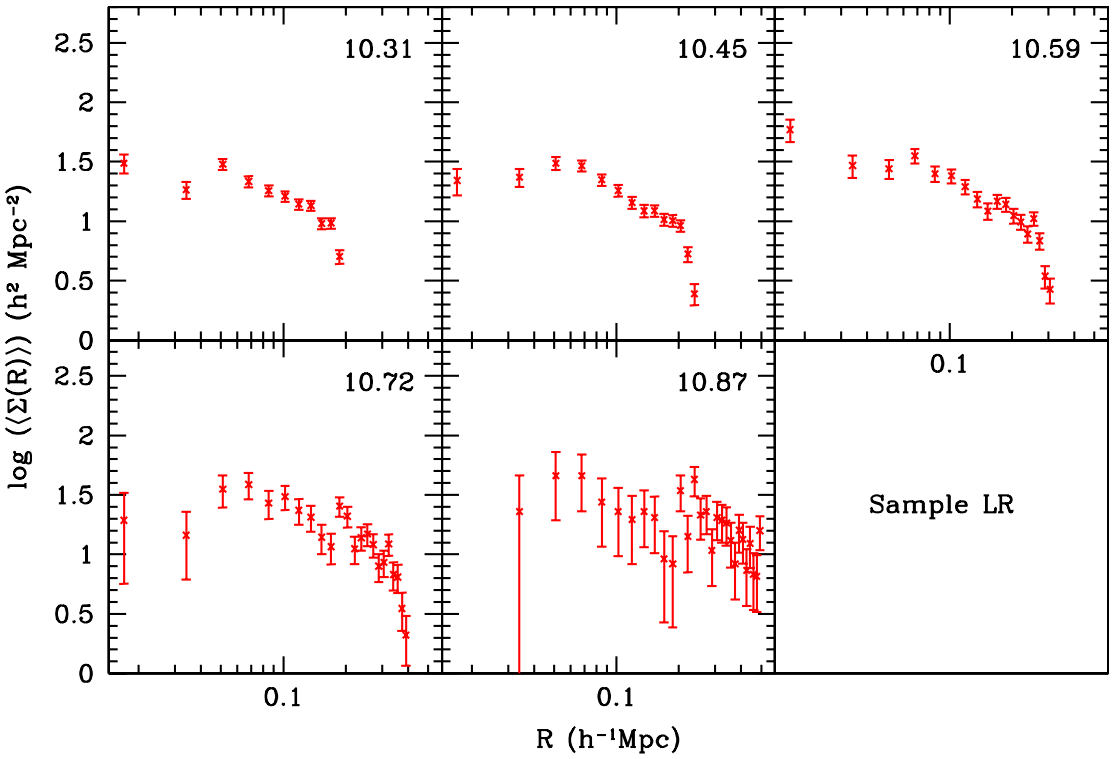

Figure 4.7: The projected number density distributions of satellites around red centrals from Sample LR for the five bright luminosity bins (average $\log(L_{\rm c}/\ h^{-2}{\rm L}_{\odot})$ in right corner).

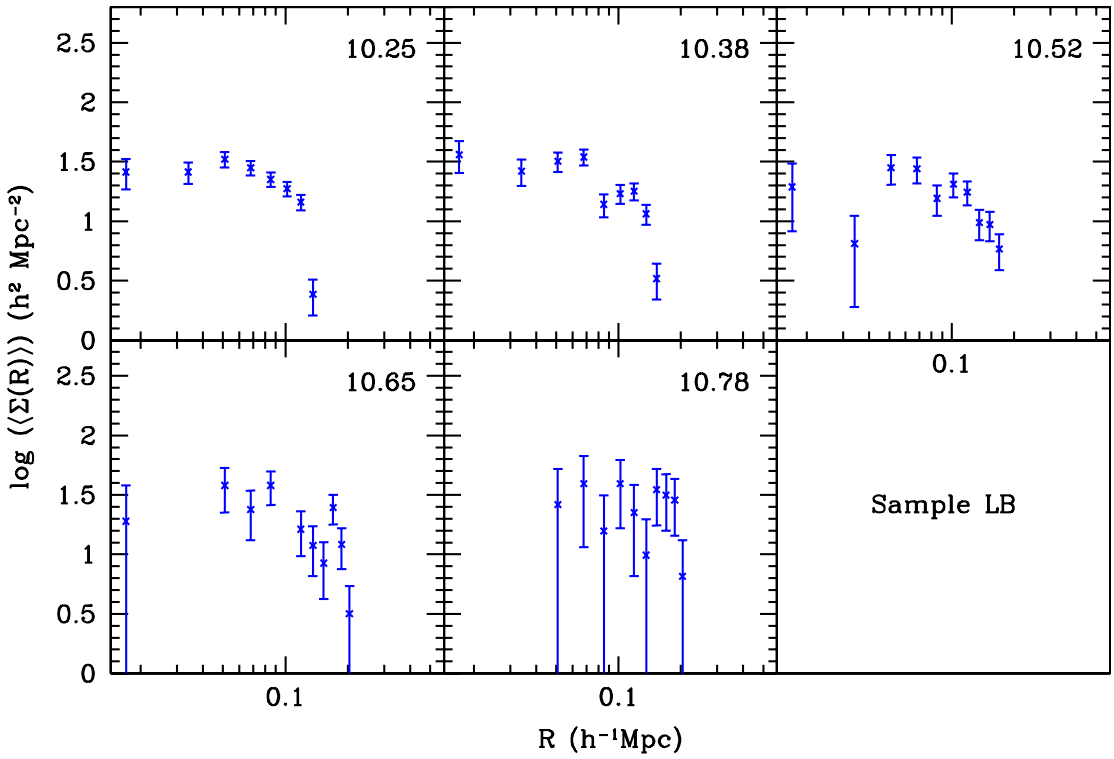

Figure 4.8: The projected number density distributions of satellites around blue centrals from Sample LB for the five bright luminosity bins (average $\log(L_{\rm c}/\ h^{-2}{\rm L}_{\odot})$ in right corner).

number density distributions of the satellites in Sample LA for the five brightest luminosity bins. As can be seen from the analytical expressions that describe $\Sigma(R|L_{\rm c})$ (see Eq. 3.15), predicting $\Sigma(R|L_{\rm c})$ requires the knowledge of $P(M|L_{\rm c})$, which is the principle goal of our study. Furthermore, it also requires the knowledge of $\langle N_{\rm sat}\rangle(M)$. Both these quantities are unknown.

To proceed, we use the $P(M|L_{\rm c})$ and $\langle N_{\rm sat}\rangle(M)$ from the CLF model of Cacciato et al. (2009) which was also used to populate the mock catalogue in Chapter 3. We explore two different models for $n_{\rm sat}(r)$, one with $\mathcal{R} = \gamma = 1$, where the number density distribution of satellites follows the dark matter density distribution, and the other with $\mathcal{R} = 2$ and $\gamma = 0$, where the number density distribution of satellites is spatially antibiased with respect to the dark matter distribution. The former model is shown as a (black) solid line while the latter with a (red) dotted line in Fig. 4.6. Clearly, the latter model is favoured by the data. Therefore, we use $\mathcal{R} = 2$ and $\gamma = 0$ to specify $n_{\rm sat}(r)$ for the analysis of the velocity dispersions to infer the MLR of central galaxies from Sample LA.

The projected number density distributions of satellites in Sample LB and LR are shown in Figs. 4.7 and 4.8, respectively. These distributions also show a similar behaviour with a flattening of the projected number density distributions at the center. We use the same values of $\mathcal{R} = 2$ and $\gamma = 0$ for the analysis of both Sample LR and Sample LB.

4.5 Results from the MCMC Analysis

Next, we use the parametric model described in Section 3.6.1 and constrain it using the measured velocity dispersions, $\sigma_{\rm sw}$ and $\sigma_{\rm hw}$, and the number of satellites per central, $\langle N_{\rm sat}\rangle$ as a function of the luminosity of centrals for the Samples LA, LR and LB. As shown in Chapter 3, this allows us to determine both the mean and the scatter of the MLR of central galaxies in these three samples. We use a Monte−Carlo Markov Chain to recover the relations $\langle \log M\rangle(L_{\rm c})$, $\sigma_{\log M}(L_{\rm c})$ and $\langle \log L_{\rm c}\rangle(M)$.

4.5.1 The Halo Mass−Luminosity Relation

First, we present the results obtained from the MCMC analysis of the velocity dispersions measured from Sample LA. The $16^{\rm th}, 50^{\rm th}$ and the $84^{\rm th}$ percentiles of the distributions of the parameter values that describe the distribution $P(L_{\rm c}|M)$ for the central galaxies in this sample are listed in Table 4.5. Fig. 4.9 shows the results of the MCMC analysis. The upper row of panels shows the data that was used to constrain the parameters. In the bottom row, panel (d) shows the confidence levels on the relation $\tilde{L_{\rm c}}(M)$ while panels (e) and (f) show the confidence levels on the inferred mean and scatter of the distribution $P(M|L_{\rm c})$ (i.e the MLR), respectively. The values of $\langle \log M\rangle(L_{\rm c})$ and $\sigma_{\log M}(L_{\rm c})$ together with their 1-σ errors are listed in Table 4.6.

Clearly, the average masses of dark matter haloes increase with central galaxy luminosity, as expected. Interestingly, the scatter in halo masses also increases systematically with the lu-

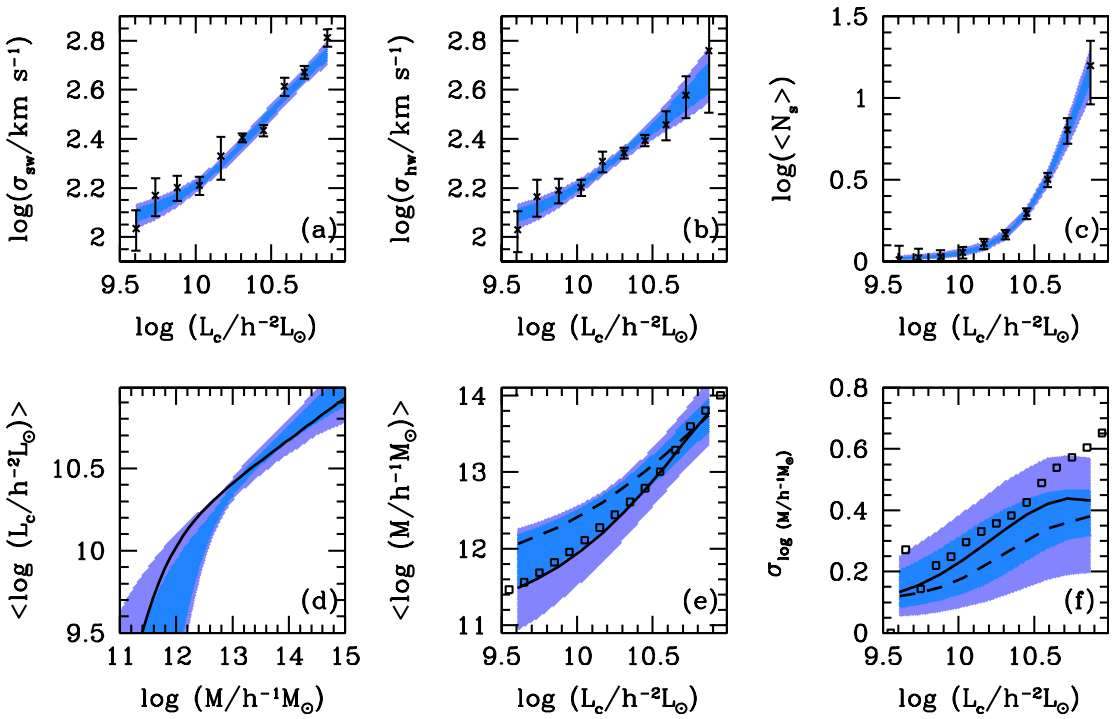

Figure 4.9: Results of the MCMC analysis of the velocity dispersions obtained from Sample LA. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of luminosity in panel (c), all measured by using the satellites in Sample LA. The blue and purple bands represent the 68% and 95% confidence regions respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(L_{\rm c})$ is in panel (d), and the mean and the scatter in the MLR of central galaxies in panels (e) and (f) respectively. The median relations obtained from the MCMC are shown using dashed lines. The relations obtained from the best-fit CLF model of Cacciato et al. (2009), are shown using solid lines. The squares in panels (e) and (f) indicate the values obtained from the semianalytical model of Croton et al. (2006).

Table 4.5: MLR of central galaxies: Parameters recovered from the MCMC

Sample	Parameter	16%	50%	84%
LA	$\log(L_0)$	9.69	10.05	10.33
	$\log(M_1)$	10.85	11.74	12.01
	γ_2	0.19	0.28	0.35
	$\sigma_{\log L}$	0.12	0.16	0.19
	N_s	0.09	0.14	0.21
	α_s	1.17	1.32	1.46
LR	$\log(L_0)$	9.58	9.78	10.05
	$\log(M_1)$	10.75	11.35	11.83
	γ_2	0.26	0.31	0.37
	$\sigma_{\log L}$	0.17	0.20	0.21
	N_s	0.07	0.11	0.17
	α_s	1.27	1.41	1.54
LB	$\log(L_0)$	9.63	9.91	10.27
	$\log(M_1)$	10.87	11.34	11.71
	γ_2	0.10	0.26	0.47
	$\sigma_{\log L}$	0.18	0.27	0.35
	N_s	0.29	0.47	0.68
	α_s	0.56	0.83	1.07

The $16^{\rm th}$, $50^{\rm th}$ and $84^{\rm th}$ percentiles of the distributions of our model parameters for the three samples analysed in this chapter.

Table 4.6: Sample LA: MLR of central galaxies

$\log(L_{\rm c})$	$\langle \log {\rm M} \rangle$	$\Delta \langle \log {\rm M} \rangle$	$\sigma_{\log {\rm M}}$	$\Delta \sigma_{\log {\rm M}}$
$h^{-2}{\rm L}_\odot$	$h^{-1}{\rm M}_\odot$	$h^{-1}{\rm M}_\odot$	$h^{-1}{\rm M}_\odot$	$h^{-1}{\rm M}_\odot$
9.61	12.06	0.35	0.12	0.06
9.73	12.16	0.32	0.13	0.07
9.88	12.28	0.29	0.15	0.08
10.03	12.44	0.26	0.18	0.10
10.17	12.60	0.23	0.22	0.10
10.31	12.80	0.21	0.26	0.11
10.45	13.01	0.19	0.30	0.10
10.59	13.24	0.19	0.34	0.09
10.72	13.47	0.21	0.36	0.08
10.87	13.74	0.23	0.38	0.07

The mean and scatter of the halo masses as a function of the central galaxy luminosity inferred from the MCMC analysis. The errors on each of the inferred quantities correspond to the 68% confidence levels.

minosity of the central galaxy. At the bright end, this scatter is roughly half a dex. Therefore, stacking central galaxies by luminosity amounts to stacking haloes that cover a wide range in masses. This justifies the need to account for this scatter in the analysis of the satellite kinematics. Neglecting this scatter leads to an overestimate of the halo mass at a given central luminosity (see Chapter 2). Most previous studies dealing with satellite kinematics have neglected this scatter which has resulted in a biased estimate of the halo mass-luminosity relationship. As we have shown in the Appendix 3.A of Chapter 3, their use of strict selection criteria to identify the centrals and satellites have further biased their estimate of the MLR of central galaxies.

In a recent study, Cacciato et al. (2009) have constrained the CLF using the abundance and clustering of galaxies in SDSS. They have shown that this CLF is also able to reproduce the galaxy−galaxy lensing signal and is further consistent with the MLR obtained from a SDSS group catalogue (Yang et al. 2008). As a consistency check, we compare the results of their study with the results obtained here from satellite kinematics. In panels (e) and (f) of Fig. 4.9, dashed lines show the mean and scatter of the distribution $P(M|L_{\rm c})$ (MLR), obtained from satellite kinematics, while the solid lines show the relations obtained from the best−fit CLF model of Cacciato et al. (2009). The agreement with the results obtained here using the kinematics of satellite galaxies is not perfect. However, given the errorbars it is certainly consistent with 68% confidence. Amongst others this consistency provides further support that the halo mass assignment in the SDSS group catalogue of Yang et al. (2007) is reliable (Wang et al. 2008).

Since the mean and scatter of the MLR reflect the physics, and in particular the stochasticity, of galaxy formation, it is interesting to compare the results obtained here to predictions from semi-analytical models (SAM) of galaxy formation. To that extent we use the SAM of Croton et al. (2006), which has been shown to match the observed properties of the local galaxy population with reasonable accuracy[2]. Using a volume limited sample of galaxies selected from the SAM with the same luminosity and redshift cuts as SDSSV, we measure the mean and the scatter of the distributions of halo masses for central galaxies in several bins of r-band luminosity. The results are shown in panels (e) and (f) of Fig. 4.9 as open squares. The agreement with our constraints from the satellite kinematics is remarkably good. It is both interesting and encouraging that a semi-analytical model, which uses simple, physically motivated recipes to model the complicated baryonic physics associated with galaxy formation, is able to reproduce not only the mean of the MLR of central galaxies but also the correct amount of stochasticity in this relation.

In our model, the stochasticity of galaxy formation is best described by the parameter $\sigma_{\log L}$, which indicates the amount of scatter in the luminosity of central galaxies given the mass of a halo, i.e. the scatter in the distribution $P(L_{\rm c}|M)$. The histogram in Fig. 4.10 shows the posterior probability of $\sigma_{\log L}$, obtained from our MCMC, which yields $\sigma_{\log L} = 0.16 \pm 0.04$ (68%

[2]Note that Croton et al. (2006) adopted a slightly different cosmology than the one used in our data analysis which can have a small impact on the MLR.

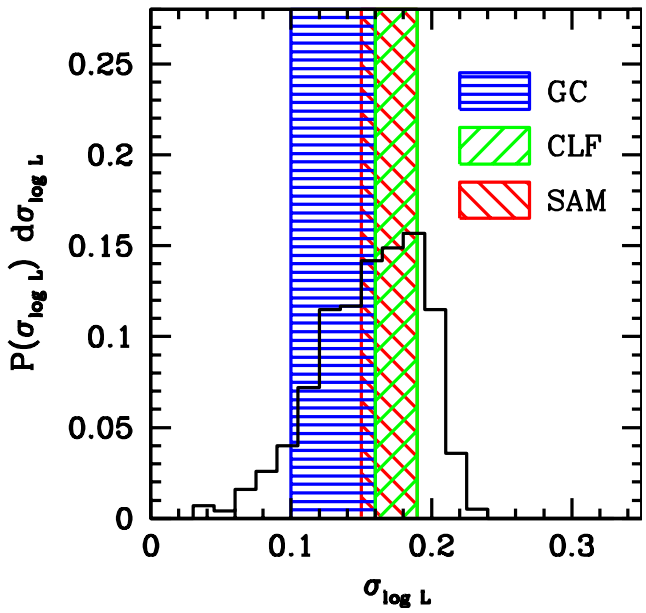

Figure 4.10: Posterior distribution of the parameter $\sigma_{\log L}$ as obtained from the MCMC analysis of the satellite velocity dispersions. The 1-σ constraints on the parameter $\sigma_{\log L}$ obtained from other independent methods are shown as shaded regions. Region GC indicates the SDSS group catalogue result by Yang et al. (2008), region CLF indicates the result obtained by Cooray (2006) with an independent CLF analysis and region SAM shows our measurement of $\sigma_{\log L}$ from the semi-analytical model of Croton et al. (2006).

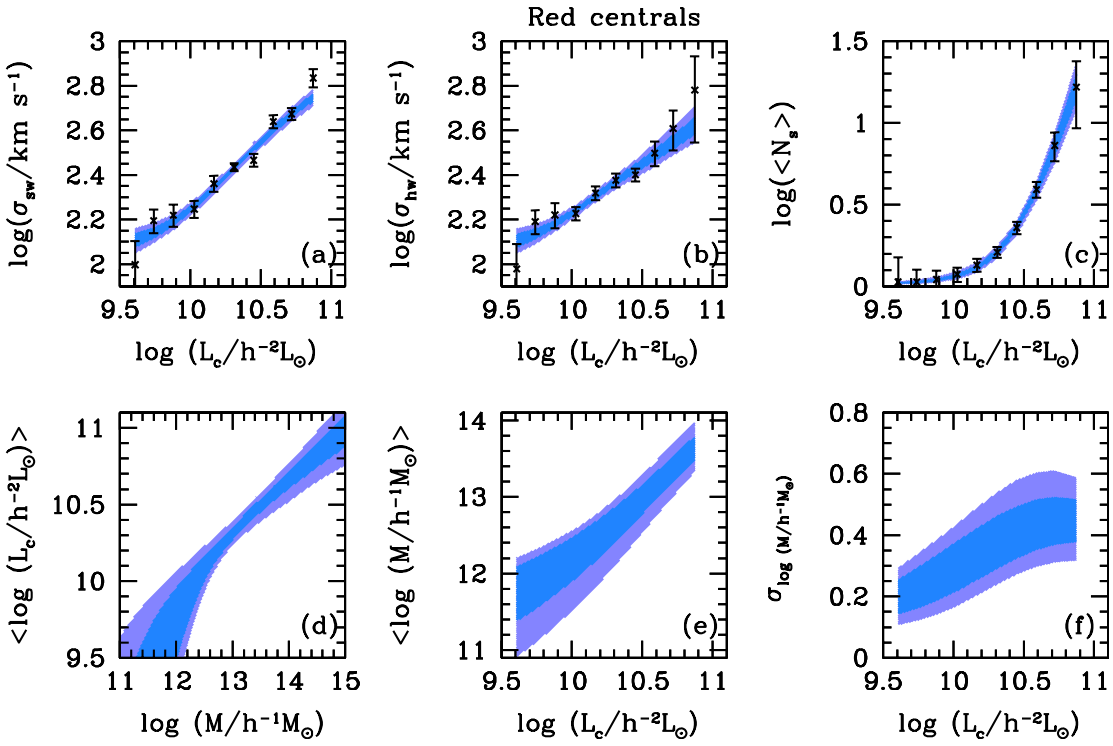

Figure 4.11: Results of the MCMC analysis of the velocity dispersions obtained from Sample LR. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of luminosity in panel (c), all measured by using the satellites from Sample LR. The blue and purple bands represent the 68% and 95% confidence regions respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(L_{\rm c})$ is in panel (d), and the mean and the scatter in the MLR of central galaxies in panels (e) and (f), respectively.

confidence levels). Note that we have made the assumption that $\sigma_{\log L}$ is independent of halo mass. The same assumption was made by Cooray (2006), who obtained that $\sigma_{\log L} = 0.17^{+0.02}_{-0.01}$ using the luminosity function and clustering properties of SDSS galaxies (see also Cacciato et al. 2009). Using a large SDSS galaxy group catalogue, Yang et al. (2008) obtained direct estimates of the scatter in $P(L_{\rm c}|M)$, and found that $\sigma_{\log L} = 0.13 \pm 0.03$ with no obvious dependence on halo mass. Finally, we also determined $\sigma_{\log L}$ in the SAM of Croton et al. (2006): using several bins in halo mass covering the range $10^{10}\ h^{-1}{\rm M}_{\odot} \leq {\rm M} \leq 10^{16}\ {\rm h}^{-1}{\rm M}_{\odot}$, we find that $\sigma_{\log L} = 0.17 \pm 0.02$, once again with virtually no dependence on halo mass. All these results are summarized in Fig. 4.10. Not only do they support our assumption that $\sigma_{\log L}$ is independent of halo mass, they also are in remarkable quantitative agreement with each other.

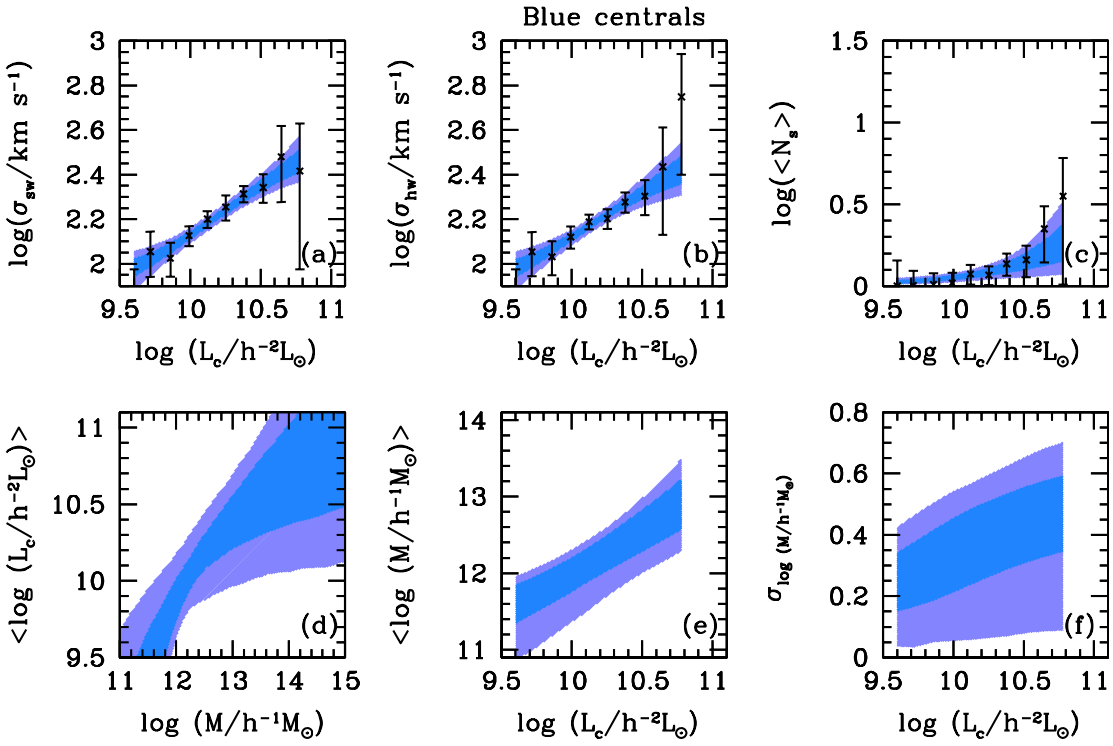

Figure 4.12: Results of the MCMC analysis of the velocity dispersions obtained from Sample LB. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of luminosity in panel (c), all measured by using the satellites from Sample LB. The blue and purple bands represent the 68% and 95% confidence regions respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(L_c)$ is in panel (d), and the mean and the scatter in the MLR of central galaxies in panels (e) and (f), respectively.

Table 4.7: Sample LR: MLR of red central galaxies

$\log(L_{\rm c})$	$\langle \log {\rm M} \rangle$	$\Delta \langle \log {\rm M} \rangle$	$\sigma_{\log {\rm M}}$	$\Delta \sigma_{\log {\rm M}}$
$h^{-2}{\rm L}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$
9.61	11.81	0.36	0.20	0.06
9.74	11.95	0.33	0.22	0.07
9.88	12.11	0.31	0.26	0.07
10.03	12.30	0.28	0.30	0.08
10.17	12.49	0.25	0.33	0.08
10.31	12.70	0.23	0.37	0.09
10.45	12.92	0.20	0.40	0.08
10.59	13.15	0.18	0.42	0.08
10.72	13.38	0.16	0.43	0.08
10.87	13.64	0.15	0.44	0.07

The mean and scatter of the halo masses as a function of the central galaxy luminosity inferred from the MCMC analysis of red central galaxies. The errors on each of the inferred quantities correspond to the 68% confidence levels.

Table 4.8: Sample LB: MLR of blue central galaxies

$\log(L_{\rm c})$	$\langle \log {\rm M} \rangle$	$\Delta \langle \log {\rm M} \rangle$	$\sigma_{\log {\rm M}}$	$\Delta \sigma_{\log {\rm M}}$
$h^{-2}{\rm L}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$
9.60	11.66	0.25	0.25	0.10
9.72	11.76	0.24	0.27	0.10
9.86	11.88	0.23	0.31	0.11
9.99	12.02	0.22	0.34	0.12
10.12	12.15	0.22	0.38	0.12
10.25	12.28	0.23	0.41	0.12
10.38	12.42	0.25	0.43	0.12
10.52	12.57	0.27	0.46	0.12
10.65	12.70	0.30	0.48	0.12
10.78	12.84	0.33	0.49	0.12

The mean and scatter of the halo masses as a function of the central galaxy luminosity for blue central galaxies inferred from the MCMC analysis. The errors on each of the inferred quantities correspond to the 68% confidence levels.

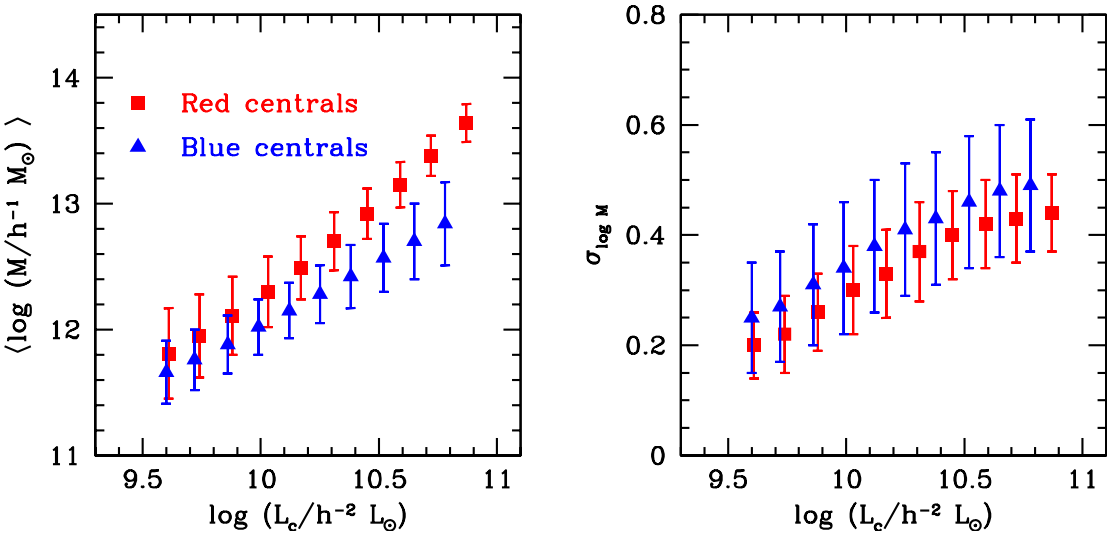

Figure 4.13: Comparison of the average halo masses and the scatter in halo masses of central galaxies split by colour as a function of their luminosity. The results for red central galaxies are shown with (red) squares and those for blue central galaxies are shown with (blue) triangles.

4.5.2 The Colour Dependence of the Halo Mass–Luminosity Relation

Next, we present the results of the MCMC analysis of the velocity dispersions obtained from Samples LR and LB respectively. The $16^{\rm th}, 50^{\rm th}$ and $84^{\rm th}$ percentiles of the distributions of the parameter values that describe the distribution $P(L_{\rm c}|M)$ for the red central galaxies and for blue central galaxies are listed in Table 4.5. In Fig. 4.11, we show the results of the MCMC analysis of velocity dispersions around red central galaxies while in Fig. 4.12 we show the corresponding results for blue central galaxies. The values of $\langle \log M \rangle (L_{\rm c})$ and $\sigma_{\log M}(L_{\rm c})$ together with their 1-σ errors for samples LR and LB are listed in Table 4.7 and Table 4.8, respectively.

The mean luminosity of red central galaxies scales with halo mass as $L_{\rm c} \propto M^{0.31\pm0.06}$ while that of blue central galaxies scales as $L_{\rm c} \propto M^{0.3\pm0.2}$ at the bright end (see Panel (d) in Figs. 4.11 and 4.12). The scaling between the luminosity and halo mass is poorly constrained for the blue galaxies at the bright end as the data does not contain many bright blue centrals. The value of the scatter in the distribution $P(L_{\rm c}|M)$ for red central galaxies is $\sigma_{\log L} = 0.20^{+0.02}_{-0.02}$ and that for blue central galaxies is $\sigma_{\log L} = 0.27^{+0.09}_{-0.07}$. The value of the scatter for blue central galaxies is rather poorly constrained due to the poor quality of the velocity dispersion data (smaller satellite sample).

The average halo mass of central galaxies increases with the luminosity of the central galaxy irrespective of their colour (Panel e in Figs. 4.11 and 4.12). However, the MLR of blue central galaxies differs from the MLR of red central galaxies. The blue central galaxies exhibit a shallower scaling relation than their red counterparts (see Fig 4.13). This shows that the luminosity of a central galaxy alone is not a good proxy for its halo mass. Furthermore, the scatter in halo masses at fixed luminosity is non-negligible, and increases as a function of the luminosity

for both the red and the blue central galaxies. Thus, even after central galaxies are stacked on the basis of their luminosities and colours, the MLR shows a considerable amount of scatter. This further emphasizes the need to model this scatter when interpreting results of studies that involve stacking.

Next, we compare the results obtained using satellite kinematics with results from the group catalogue of Yang et al. (2007). These authors have assigned group halo masses based upon either the total stellar mass or the total luminosity content of each group. We use this group catalog and investigate the MLR of central galaxies, with and without the split into red and blue by colour. The solid lines in Fig. 4.14 correspond to the MLR of central galaxies in their group catalogue where the halo masses have been assigned using the total stellar mass content of the group. The dashed lines, in turn, show the MLR of central galaxies in the group catalogue where the halo masses have been assigned according to the luminosity content of the group. The MLR for central galaxies obtained in this chapter are shown with shaded areas (68 percent confidence region).

The MLR of all central galaxies (top left panel in Fig. 4.14) inferred from satellite kinematics is in good agreement with the results from the group catalog. There exists a slight tension at the bright end between the MLR obtained from the group catalogue and that from satellite kinematics for red central galaxies and this problem is worse for the blue central galaxies. Since the central galaxy colour information was not used in the construction of the galaxy group catalogue, we believe this to be an artefact in the group catalogue. The group catalogue can reliably reproduce the average properties of the entire galaxy sample, however it may not be reliable to deduce the average properties of a subset of galaxies.

Galaxy−galaxy lensing is yet another technique to probe the halo masses and hence the MLR of central galaxies. Mandelbaum et al. (2006) presented the weak lensing signal around galaxies stacked by luminosity. The galaxies were split into early (red) and late (blue) types based upon their morphology. The results obtained from their analysis of the weak lensing signal are shown in Fig. 4.14 as squares with errorbars. These results are in excellent agreement with our results from satellite kinematics. The potential disagreement for the brightest luminosity bin of red centrals from their sample is most likely a result of the different criteria used to separate red and blue central galaxies.

4.6 Summary

The kinematics of satellite galaxies have been widely used to statistically relate the mean halo masses of central galaxies to their luminosities (Zaritsky et al. 1993; Zaritsky & White 1994; Zaritsky et al. 1997; McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; Norberg et al. 2008). These studies use strict criteria to identify central and satellite galaxies that reside preferentially in isolated environments. van den Bosch et al. (2004) have advocated the use of a relaxed but adaptive selection criterion to identify centrals and their satellites, not only in

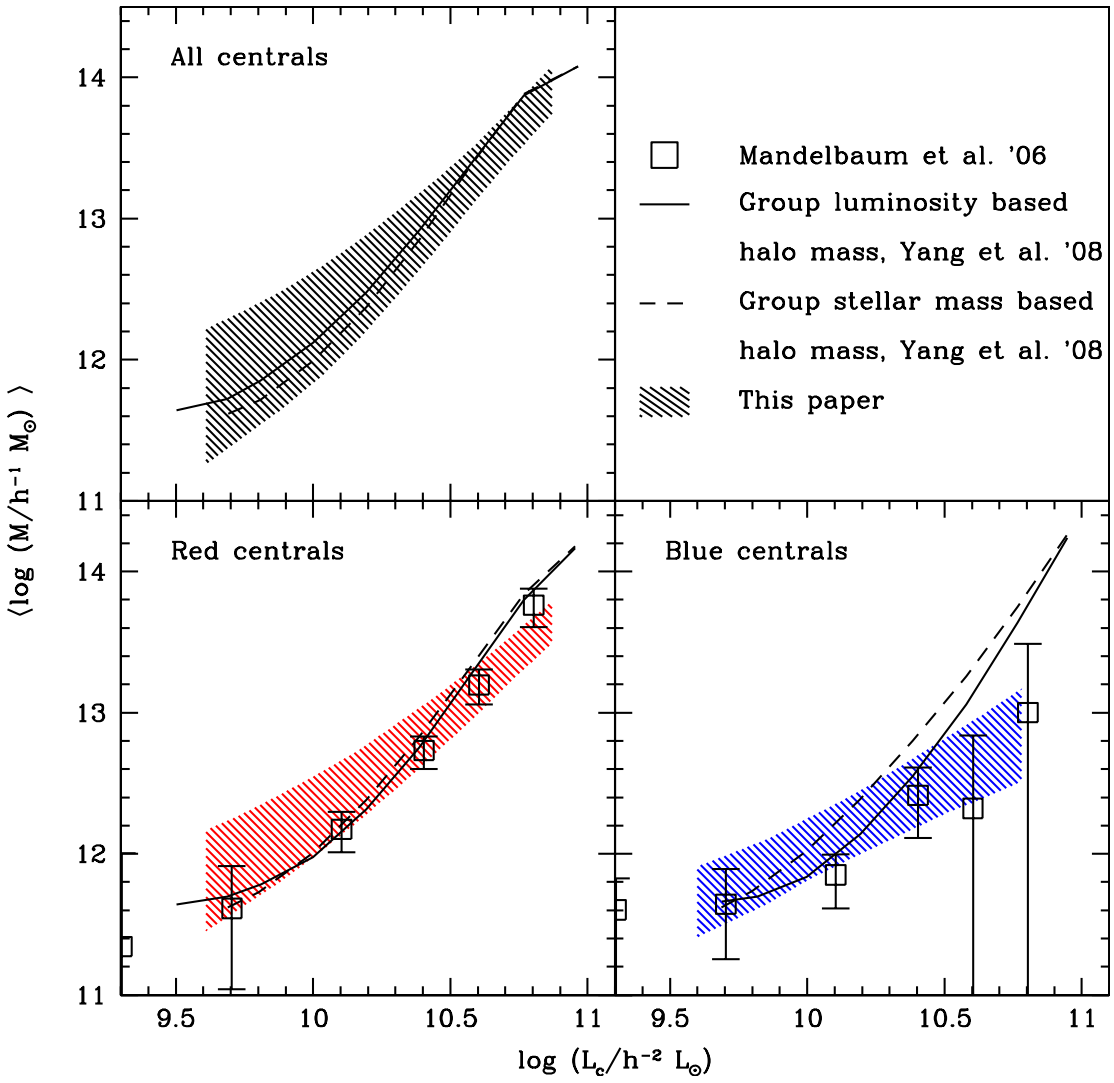

Figure 4.14: Comparison of the MLR of central galaxies by different methods. The results obtained from the weak lensing analysis are shown as squares with errorbars. The results obtained from the group catalogue are shown with solid and dashed lines. The shaded areas represent the results obtained in this chapter using the kinematics of satellites.

isolated environments but also in massive groups and clusters. This has the potential to allow the study of the kinematics of satellites over a wide range of central galaxy luminosity. In this chapter, we applied a relaxed but adaptive selection criterion to a volume limited sample from SDSS to identify centrals and their satellites, not only in isolated environments but also in massive groups and clusters which allowed us to study the kinematics of satellites over a wide range of central galaxy luminosity. We inferred both the mean and the scatter of the mass−luminosity relationship of central galaxies from the kinematics of satellite galaxies both with and without a split in the colour of central galaxies.

The analysis of the kinematics of satellites around centrals stacked without a split in their colour shows that the mean of the mass−luminosity relation increases as a function of the central host luminosity indicating that, as expected, brighter centrals reside in more massive haloes. This result is in quantitative agreement with a recent study by Cacciato et al. (2009), who use the abundance and the clustering properties of galaxies in SDSS to constrain the CLF, and with the SAM of Croton et al. (2006). The satellite kinematics obtained in our study are consistent with a model in which $P(L_{\rm c}|M)$ has a constant scatter, $\sigma_{\log L}$, independent of the halo mass M. We obtain $\sigma_{\log L} = 0.16 \pm 0.04$ in excellent agreement with other independent measurements suggesting that the amount of stochasticity in galaxy formation is similar in haloes of all masses. This is also suggested by the SDSS group catalogue of Yang et al. (2008) and by the SAM of Croton et al. (2006). It is important to note that a constant scatter in the distribution $P(L_{\rm c}|M)$ leads to a scatter in the distribution $P(M|L_{\rm c})$ that systematically increases with luminosity (see Chapter 2).

We also analysed the kinematics of satellite galaxies around red and blue central galaxies separately to investigate the colour dependence of the MLR. We found that the MLR of central galaxies is different for central galaxies of different colours. Red central galaxies on average occupy more massive haloes than blue central galaxies. This shows that the scatter in the colour of central galaxies is, to a certain extent, responsible for the scatter in halo masses at fixed central galaxy luminosity. However, we also found that both red and blue central galaxies of a given luminosity reside in haloes with a large scatter in their masses (~ 0.4 dex) especially at the bright end. Hence it is imperative to account for the scatter in any analysis that involves stacking.

We compared the average MLR inferred from satellite kinematics with those inferred from the SDSS group catalogue of Yang et al. (2008) for central galaxies. The average MLR of central galaxies without a split in colour is in excellent agreement with the group catalogue results. The average MLRs of central galaxies split by colour are in slight tension with those present in the group catalogue which we believe to be an artefact of the group catalogue. The group catalogue may not be reliable for deducing the average properties of subsets of galaxies. We also compared the average MLRs of red and blue central galaxies obtained by us with the results obtained using galaxy−galaxy lensing by Mandelbaum et al. (2006b) and showed that these results are in excellent quantitative agreement with each other.

Chapter 5

The Halo Mass−Stellar Mass Relationship

The contents of this chapter are based upon an article which is in preparation. The article will be submitted to the Monthly Notices of the Royal Astronomical Society as:
Satellite Kinematics III: Colour and Stellar Mass Dependence
More, S., van den Bosch, F. C., Cacciato, M., et al. 2009.

5.1 Introduction

Establishing scaling relations between the properties of central galaxies and their dark matter halo properties is central to understanding the process of galaxy formation. In this thesis, we have presented a systematic method that can be used to probe the halo masses of galaxies that reside at the centre of dark matter haloes using the kinematics of their central galaxies (Chapters 2 and 3). In Chapter 5, we have established the scaling relation between halo mass and luminosity of central galaxies and the scatter in this scaling relation. In this chapter, we infer the halo mass−stellar mass relationship of central galaxies (hereafter MSR) with and without a split in their colour.

This chapter is organized as follows. In Section 5.2, we describe the samples of central and satellite galaxies used in this chapter. In Section 5.3, we present the measurement of the kinematics of the satellites as a function of the stellar mass of the central galaxy. In Section 5.4, we present the model we use to infer the MSR of central galaxies from the kinematics of satellites. We present our results in Section 5.5 and also compare them with results from other independent studies. Finally, we summarize our findings in Section 5.6.

For the analysis presented in this chapter, we assume the cosmological parameters from the 3 year data release of WMAP (Spergel et al. 2007), $\Omega_{\rm m} = 0.238$, $\Omega_\Lambda = 0.762$, $h = H_0/100\ {\rm km\,s^{-1}\ Mpc^{-1}} = 0.734$, the spectral index of initial density fluctuations $n_{\rm s} = 0.951$ and normalization $\sigma_8 = 0.744$. We consider the dark matter halo mass to be M_{180}, i.e. the mass

enclosed within a spherical overdensity $\delta\rho/\bar{\rho} = 180$, where $\bar{\rho}$ denotes the mean matter density of the universe.

5.2 Central and Satellite Samples from the SDSS

As in Chapter 4, we again use data from the NYU−VAGC (Blanton et al. 2005) which is based upon Data Release 4 (Adelman-McCarthy et al. 2006) of the SDSS. We start from sample SDSSV, the volume limited sample of galaxies described in Section 4.2. The galaxies in this sample were assigned a colour (red or blue), according to the separation criteria given by Eq. (4.1) which was based upon the bimodal distribution of galaxies in the ${}^{0.1}(g-r)$ colour at a given luminosity. The stellar masses of these galaxies (denoted as M_*) are computed using the relation between the stellar mass-to-light ratio and the ${}^{0.0}(g-r)$ color provided by Bell et al. (2003),

$$\log\left[\frac{M_*}{h^{-2}\mathrm{M}_\odot}\right] = -0.306 + 1.097\,[{}^{0.0}(g-r)] - 0.10 - 0.4\,({}^{0.0}M_r - 5\log h - 4.64). \quad (5.1)$$

Here, ${}^{0.0}(g-r)$ and ${}^{0.0}M_r - 5\log h$ denote the $(g-r)$ colour and the r-band absolute magnitude of galaxies k−corrected and evolution corrected to $z = 0.0$; 4.64 is the r-band magnitude of the Sun in the AB system; and the -0.10 term is a result of adopting the Kroupa (2001) initial mass function (see Borch et al. 2006).

For the analysis of the MSR, we use a slightly modified version of the iterative criteria (ITER) outlined in Section 3.3 of Chapter 3 to identify centrals and their satellites from SDSSV. Firstly, the modified criteria requires the central galaxy to be the largest in terms of the stellar mass (instead of the brightest in terms of luminosity) in its neighbourhood (specified by $(\Delta V)_\mathrm{h}$ and R_h). Secondly, the unbinned estimate of the velocity dispersion as a function of the *stellar mass* of the central galaxy is used to refine the search cylinders of the iterative criteria. We use the method described in Section 3.4.1 to measure the unbinned estimate of the velocity dispersion. The velocity dispersion (in units of 200 $\mathrm{km\,s^{-1}}$), σ_{200}, is parametrized as,

$$\sigma_{200}(\log M_{*10}) = a + b\,(\log M_{*10}) + c\,(\log M_{*10})^2\,. \quad (5.2)$$

Here M_{*10} denotes the stellar mass of the galaxy in units of $10^{10}h^{-2}\mathrm{M}_\odot$. During every iteration, the parameters (a, b, c) are fit by using the velocity difference information between the satellites and their centrals (see Section 3.4.1 for details). The parameters that define the search cylinder are scaled for the subsequent iteration based upon the estimate of $\sigma_{200}(M_{*10})$ (see criteria ITER in Table 3.1).

Similar to Chapter 4, we form three different samples of centrals and their satellites for the analysis of the MSR. The sample used to analyse the MSR around all central galaxies (without

Table 5.1: Selection criteria parameters

Parameters	Samples		
	SA	SR	SB
a	2.06	2.12	1.97
b	0.22	0.18	0.45
c	0.20	0.21	-0.07

The parameters a, b and c that are used to define the criteria used to select central and satellite galaxies for the samples used in this chapter (see Table 3.1 for details of the selection criteria).

a split by colour) is denoted by SA. The samples used to analyse the MSR around red and blue central galaxies are denoted by SR and SB, respectively. Note that the difference between the three samples is the estimate of the velocity dispersion used to scale the search cylinders. The search criteria used for sample SR (SB) are tuned based upon the estimate of the velocity dispersion around the red (blue) central galaxies only. The parameters (a, b, c) in Eq. (5.2) that define the aperture used in the final iteration of the central−satellite selection for samples SA, SR and SB are listed in Table 5.1.

The number of central galaxies with at least one satellite is $3,778$ for Sample SA, $2,877$ for Sample SR and 805 for Sample SB. The number of satellite galaxies in Sample SA is $6,104$ while that in Sample SR and SB are $5,061$ and 912, respectively.

5.3 Velocity Dispersion−Stellar Mass Relation

The scatter plot of the velocity difference, ΔV, between the satellites and the corresponding centrals as a function of the central galaxy stellar mass in Sample SA is shown in Fig. 5.1, while those in Samples SR and SB are shown in the left and the right panels of Fig. 5.2, respectively. For all the three samples, the scatter in the velocities of satellites with respect to their centrals increases with central galaxy stellar mass. Its relatively clear from these figures that the scatter of the velocities of satellites in Sample SB is smaller than that of satellites in Samples SR and SA.

The $P(\Delta V)$ distributions for all the three samples, in both the satellite-weighting and the host-weighting schemes are obtained in 10 bins of uniform width. For Sample SA the bin-width is $\Delta \log[M_{*c}] = 0.18$ while for Samples SR and SB we choose the bin widths to be 0.16 and 0.15, respectively. All the $P(\Delta V)$ distributions are fitted using a model that consists of two Gaussians and a constant, for reasons discussed in Section 3.4.2. The resultant velocity dispersions and their errors for Sample SA are listed in Table 5.2 and those for Samples SR and SB in Tables 5.3 and 5.4, respectively. The satellite-weighted and the host-weighted velocity dispersions as a function of stellar mass for Sample SA are shown in

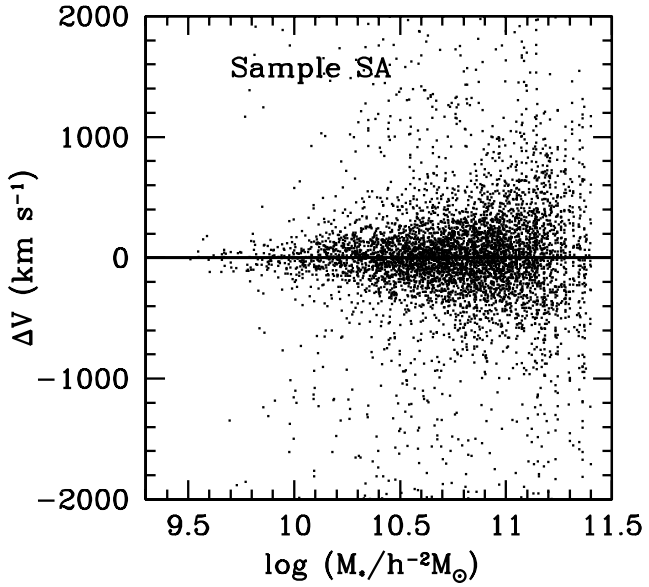

Figure 5.1: Scatter plot of the velocity difference, ΔV, between the satellites and their central galaxies in sample SA as a function of the central galaxy stellar mass.

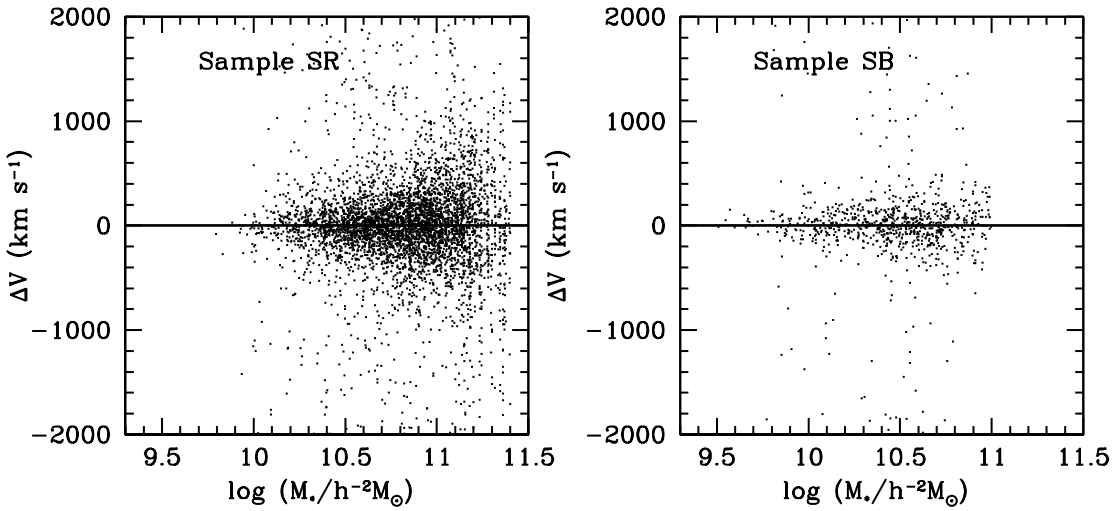

Figure 5.2: Scatter plot of the velocity difference, ΔV, between the satellites and their central galaxies in sample SR sample SB as a function of the central galaxy stellar mass are shown in the left and the right panels respectively.

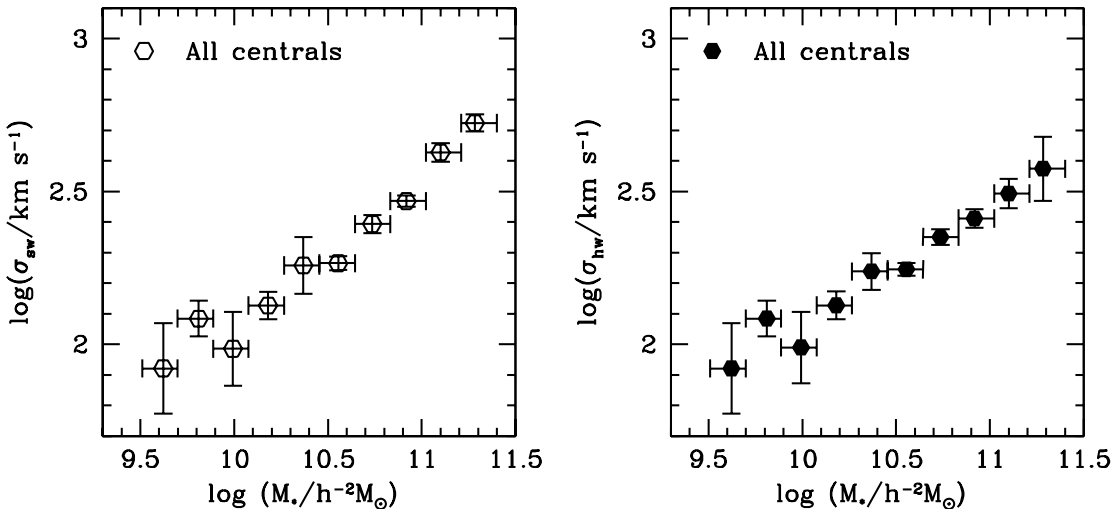

Figure 5.3: The satellite-weighted (open hexagons) and the host-weighted (filled hexagons) velocity dispersions as a function of the central galaxy stellar mass obtained from satellites in the sample SA.

Fig. 5.3. The velocity dispersion−stellar mass relation shows a similar behaviour to the velocity dispersion−luminosity relation. The satellite-weighted velocity dispersioins are systematically larger than the host-weighted velocity dispersions at any given stellar mass, which signals the presence of a non-negligible scatter in the MSR as well. The comparison between the velocity dispersions around red and blue centrals is shown in Fig. 5.4. The velocity dispersions around red centrals in both the schemes are systematically larger than the velocity dispersions around blue centrals. As in Chapter 3, we also measure the average number of satellites as a function of stellar mass and use it to constrain the MSR.

5.4 The Model

We use a very similar model to the one used in Chapters 3 and 4 to determine the MSR from the measured velocity dispersions. The analytical expressions that describe the velocity dispersions and the the average number of satellites as a function of stellar mass are the same as Eqs. (3.12), (3.13) and (3.17) with $P(M|L_{\rm c})$ replaced by $P(M|M_{*\rm c})$, i.e.,

$$\sigma^2_{\rm sw}(M_{*\rm c}) = \frac{\int_0^\infty P(M|M_{*\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|M_{*\rm c})\,\langle N_{\rm sat}\rangle_{{\rm ap},M}{\rm d}M}\,, \tag{5.3}$$

$$\sigma^2_{\rm hw}(M_{*\rm c}) = \frac{\int_0^\infty P(M|M_{*\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,\langle\sigma^2_{\rm sat}\rangle_{{\rm ap},M}\,{\rm d}M}{\int_0^\infty P(M|M_{*\rm c})\,\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M})\,{\rm d}M}\,, \tag{5.4}$$

$$\langle N_{\rm sat}\rangle(M_{*\rm c}) = \frac{\int_0^\infty P(M|M_{*\rm c})\langle N_{\rm sat}\rangle_{{\rm ap},M}{\rm d}M}{\int_0^\infty P(M|M_{*\rm c})\mathcal{P}(\langle N_{\rm sat}\rangle_{{\rm ap},M}){\rm d}M}\,. \tag{5.5}$$

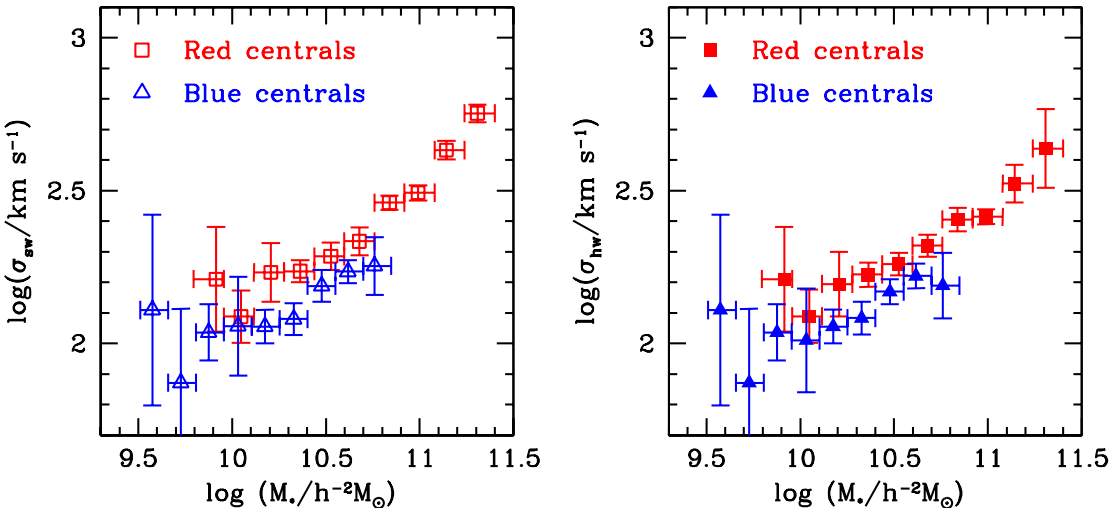

Figure 5.4: Comparison between the satellite-weighted (host-weighted) velocity dispersions as a function of the stellar mass of the central galaxy obtained from Sample SR and Sample SB is shown in the left (right) panel.

Table 5.2: Sample SA: Velocity dispersion measurements

$\log(M_{*c})$	σ_{sw}	$\Delta\sigma_{sw}$	σ_{hw}	$\Delta\sigma_{hw}$
$h^{-2}L_{\odot}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$
9.62	83	28	83	28
9.81	121	16	121	16
9.99	97	27	98	26
10.18	134	14	134	14
10.37	181	39	173	24
10.55	184	10	176	9
10.74	248	17	225	13
10.92	294	12	258	18
11.10	425	30	312	34
11.28	530	34	376	90

Velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample SA.

Table 5.3: Sample SR: Velocity dispersion measurements

$\log(M_{*c})$	σ_{sw}	$\Delta\sigma_{sw}$	σ_{hw}	$\Delta\sigma_{hw}$
$h^{-2}L_{\odot}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$
9.91	162	64	162	64
10.05	122	24	123	25
10.21	171	38	157	38
10.36	172	14	168	15
10.53	193	20	182	15
10.68	216	23	209	17
10.84	289	15	255	22
10.99	311	18	260	14
11.14	429	30	334	47
11.31	567	38	434	129

Velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample SR.

Table 5.4: Sample SB: Velocity dispersion measurements

$\log(M_{*c})$	σ_{sw}	$\Delta\sigma_{sw}$	σ_{hw}	$\Delta\sigma_{hw}$
$h^{-2}L_{\odot}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$	$km\,s^{-1}$
9.57	129	92	129	92
9.73	74	41	74	41
9.88	109	23	109	23
10.03	114	42	102	40
10.18	114	14	114	14
10.33	120	14	121	15
10.48	154	18	148	14
10.62	172	15	166	16
10.76	179	39	155	38
10.91	192	24	174	35

Velocity dispersion measurements in the satellite-weighted and host-weighted schemes together with the associated errors for sample SB.

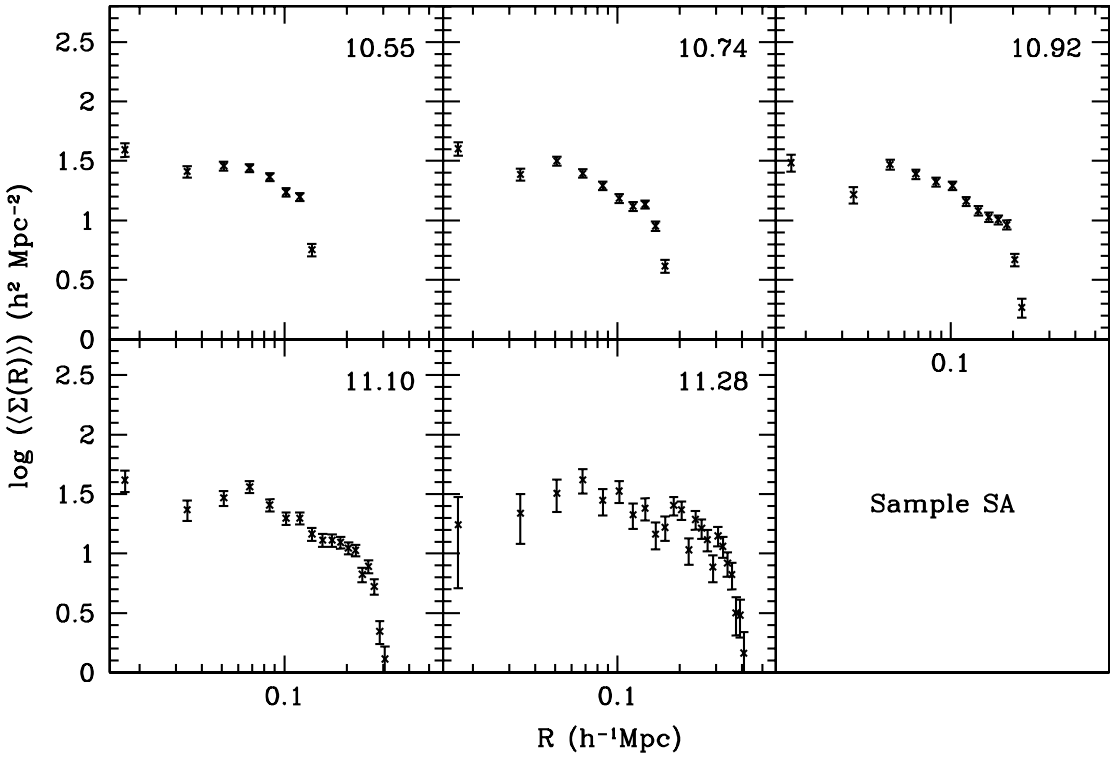

Figure 5.5: Projected number density distributions of satellites around centrals in Sample SA for the five massive stellar mass bins (average $\log(M_{*\rm c}/\ h^{-1}{\rm M}_\odot)$ in right corner).

The notations $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ and $\langle \sigma^2_{\rm sat}\rangle_{{\rm ap},M}$ describe the average number of satellites in the aperture used to select the satellites and their aperture−averaged velocity dispersion in a halo of mass M used to select the satellites. The factor $\mathcal{P}$ adjusts the distribution $P(M|M_{*\rm c})$ to account for those centrals that do not host any satellite. We use the expressions derived in Section (2.4) to calculate these quantities. Note that the calculation of $\langle N_{\rm sat}\rangle_{{\rm ap},M}$ and $\langle \sigma^2_{\rm sat}\rangle_{{\rm ap},M}$ requires us to specify the number density distribution of satellites in the halo. We use the generalised NFW profile given by Eq. (2.24) for this purpose, keeping in mind that the parameters, $\mathcal{R}$ and γ have to be fixed using the projected number density distribution, $\Sigma(R)$, of satellites around their centrals. Figs. (5.5)−(5.7) show the distributions $\Sigma(R)$ of satellites in the samples SA, SR and SB, respectively. These distributions resemble the projected number density distribution shown in Fig. 4.6 in that they show a flattening of the distribution in the central parts. Based upon this behaviour, we use the values of $\mathcal{R} = 2$ and $\gamma = 0$ in the further analysis.

We use the Bayes' theorem to relate the distribution $P(M|M_{*\rm c})$ to $P(M_{*\rm c}|M)$,

$$P(M|M_{*\rm c}) = \frac{n(M)P(M_{*\rm c}|M)}{\int n(M)P(M_{*\rm c}|M){\rm d}M}, \tag{5.6}$$

where $n(M)$ is the halo mass function. The distribution $P(M_{*\rm c}|M)$ is modelled as a lognormal

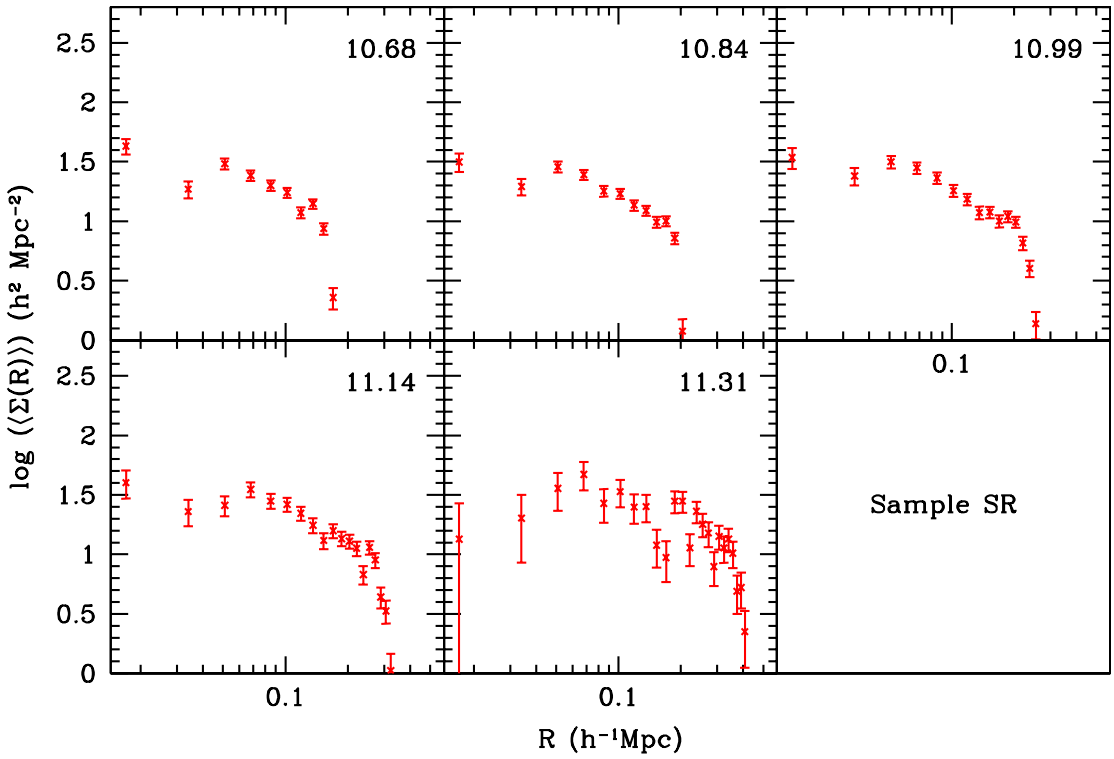

Figure 5.6: Projected number density distributions of satellites around red centrals from Sample SR (average $\log(M_{*c}/\ h^{-1}\mathrm{M}_{\odot})$ in right corner).

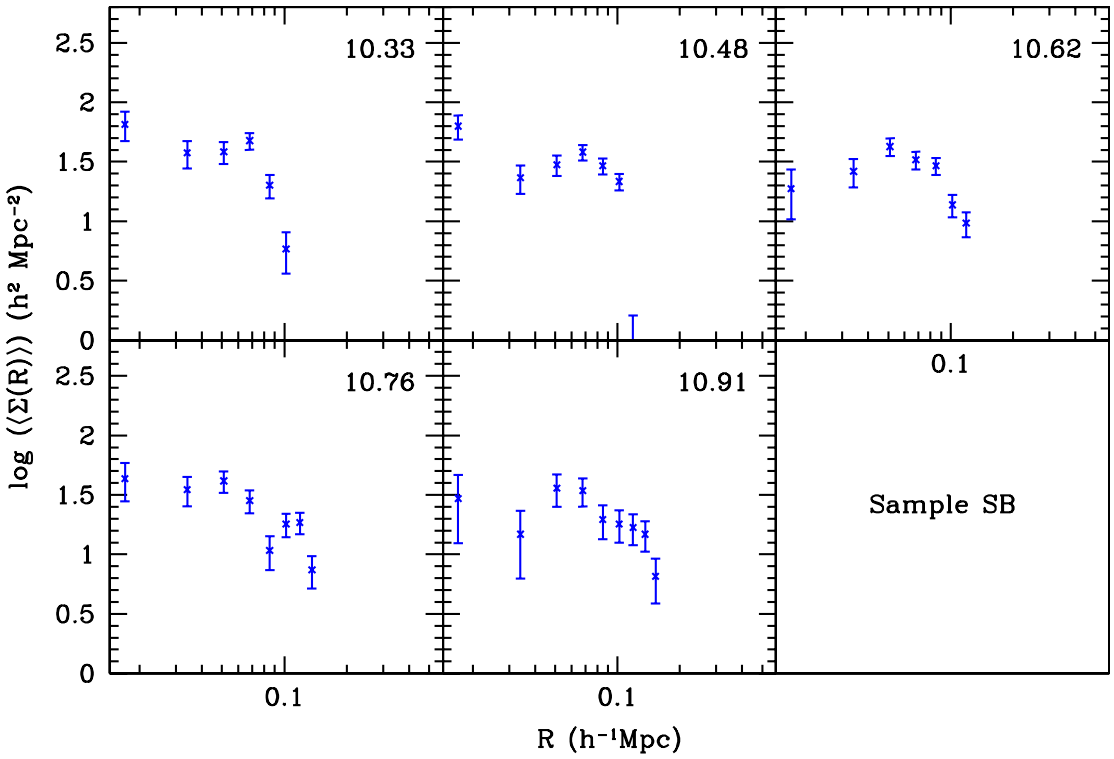

Figure 5.7: Projected number density distributions of satellites around blue centrals from Sample SB (average $\log(M_{*\mathrm{c}}/\ h^{-1}\mathrm{M}_{\odot})$ in right corner).

distribution in $M_{*\rm c}$,

$$P(M_{*\rm c}|M)\mathrm{d}M_{*\rm c} = \frac{\log(e)}{\sqrt{2\pi}\sigma_{\log M_*}} \exp\left[-\frac{(\log[M_{*\rm c}/\tilde{M}_{*\rm c}])^2}{2\sigma^2_{\log M_*}}\right] \frac{\mathrm{d}M_{*\rm c}}{M_{*\rm c}}\,. \tag{5.7}$$

The quantities $\log \tilde{M}_{*\rm c}(M)$ and $\sigma_{\log M_*}$ are the mean and the scatter of the lognormal distribution at a particular halo mass M. The scatter $\sigma_{\log M_*}$ is assumed to be independent of the halo mass M and the relation $\log \tilde{M}_{*\rm c}(M)$ is specified using four parameters: a low mass end slope γ_1, a high mass end slope γ_2, a characteristic mass scale M_1, and a normalisation M_{*0};

$$\tilde{M}_{*\rm c} = M_{*0}\frac{(M/M_1)^{\gamma_1}}{[1+(M/M_1)]^{\gamma_1-\gamma_2}}. \tag{5.8}$$

Thus, the distribution $P(M|M_{*\rm c})$ is completely specified by the five parameters ($\sigma_{\log M_*}$, M_{*0}, M_1, γ_1, γ_2) and the halo mass function. The last ingredient of our model is the satellite occupation number, $\langle N_{\rm sat}\rangle(M)$. We use a simple power law distribution given by,

$$\langle N_{\rm sat}\rangle(M) = N_s\left(\frac{M}{10^{12}\,h^{-1}\mathrm{M}_\odot}\right)^{\alpha}\,. \tag{5.9}$$

to specify the occupation number of satellites. Thus in total our model has seven free parameters. We use a MCMC to constrain the parameters and infer the MSR. Since the velocity dispersions at the low stellar mass are not well measured, we do not expect to properly constrain the low mass end slope, γ_1. Therefore, we impose a flat prior on γ_1 and allow it to vary in the interval $[2.0, 4.0]$.

5.5 Results

5.5.1 The Halo Mass−Stellar Mass Relationship

We analyse the kinematics of satellite galaxies around their centrals in Sample SA to infer the MSR for all central galaxies. Samples SR and SB are analysed to infer the MSR of central galaxies split by colour into red and blue, respectively. The results from the MCMC analysis are shown in Figs. 5.8, 5.9 and 5.10. The 16, 50 and 84 percentiles of the posterior distributions of the parameters obtained in the MCMC analysis are listed in Table 5.5.

The stellar mass of all central galaxies scales as $M^{0.38^{+0.08}_{-0.07}}$ at the high mass end. This is in excellent agreement with recent results from Moster et al. (2009) who fit the stellar mass function of SDSS galaxies using a halo occupation model and obtain that the stellar mass of central galaxies scales as $M^{0.370\pm0.014}$ (see table 3 in Moster et al. 2009)[1]. In the case of red central galaxies, the stellar mass scales as $M_* \propto M^{0.39\pm0.08}$ and in the case of blue central galaxies the stellar mass scales as $M_* \propto M^{0.7^{+0.3}_{-0.2}}$. The halo mass−stellar mass scaling for

[1]Note that the parameter γ_c in the analysis of Moster et al. 2009 is related to the parameter γ_2 in our analysis such that $\gamma_c = \gamma_2 - 1$.

Table 5.5: MSR of central galaxies: Parameters recovered from the MCMC

Sample	Parameter	16 percent	50 percent	84 percent
SA	$\log(M_{*0})$	9.57	9.97	10.39
	$\log(M_1)$	10.25	10.83	11.68
	γ_2	0.31	0.38	0.46
	$\sigma_{\log P}$	0.18	0.21	0.23
	N_s	0.11	0.15	0.20
	α_s	1.19	1.28	1.37
SR	$\log(M_{*0})$	9.63	9.99	10.36
	$\log(M_1)$	10.28	10.82	11.55
	γ_2	0.31	0.39	0.47
	$\sigma_{\log P}$	0.21	0.23	0.25
	N_s	0.06	0.09	0.13
	α_s	1.34	1.47	1.62
SB	$\log(M_{*0})$	9.35	9.98	10.60
	$\log(M_1)$	10.73	11.29	11.78
	γ_2	0.33	0.68	0.89
	$\sigma_{\log P}$	0.13	0.27	0.41
	N_s	0.16	0.22	0.28
	α_s	0.69	1.10	1.48

The 16, 50 and 84 percentile values of the posterior distribution for the parameters of our model obtained from the MCMC analysis of the velocity dispersion data from Sample SA, Sample SR and Sample SB.

Table 5.6: Sample SA: MSR of central galaxies

$\log(L_{\rm c})$	$\langle \log {\rm M} \rangle$	$\Delta \langle \log {\rm M} \rangle$	$\sigma_{\log {\rm M}}$	$\Delta \sigma_{\log {\rm M}}$
$h^{-2}{\rm L}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$
9.62	11.18	0.44	0.17	0.05
9.81	11.35	0.41	0.20	0.07
9.99	11.54	0.37	0.23	0.08
10.18	11.77	0.32	0.27	0.09
10.37	12.03	0.27	0.32	0.09
10.55	12.30	0.23	0.36	0.09
10.74	12.59	0.20	0.39	0.09
10.92	12.90	0.17	0.41	0.09
11.10	13.22	0.15	0.42	0.08
11.28	13.53	0.14	0.43	0.07

The mean and scatter of the halo masses as a function of the central galaxy stellar mass for all central galaxies inferred from the MCMC analysis. The errors on each of the inferred quantities correspond to the 68% confidence levels.

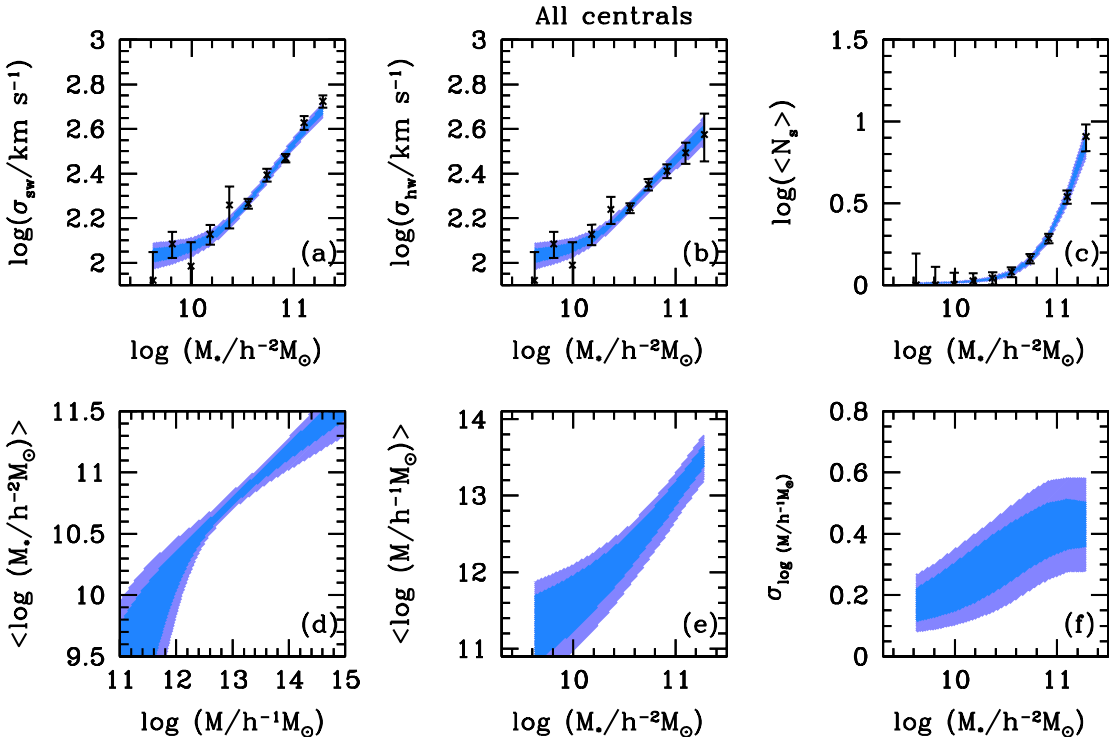

Figure 5.8: The results of the MCMC analysis of the velocity dispersions obtained from Sample SA. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of stellar mass in panel (c), all measured by using the satellites in Sample SA. The blue and purple bands represent the 68% and 95% confidence regions, respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(M_{*\mathrm{c}})$ is in panel (d), and the mean and the scatter in the MSR of central galaxies in panels (e) and (f), respectively.

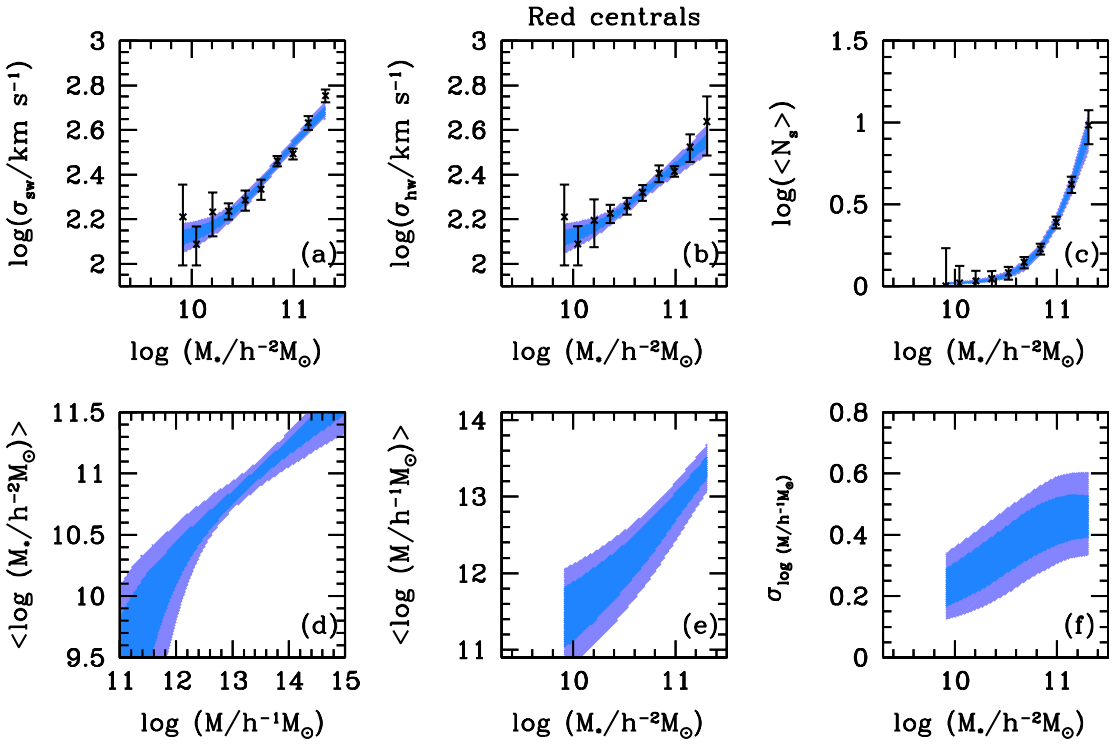

Figure 5.9: The results of the MCMC analysis of the velocity dispersions obtained from Sample SR. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of stellar mass in panel (c), all measured by using the satellites in Sample SR. The blue and purple bands represent the 68% and 95% confidence regions, respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(M_{*\mathrm{c}})$ is in panel (d), and the mean and the scatter in the MSR of central galaxies in panels (e) and (f), respectively.

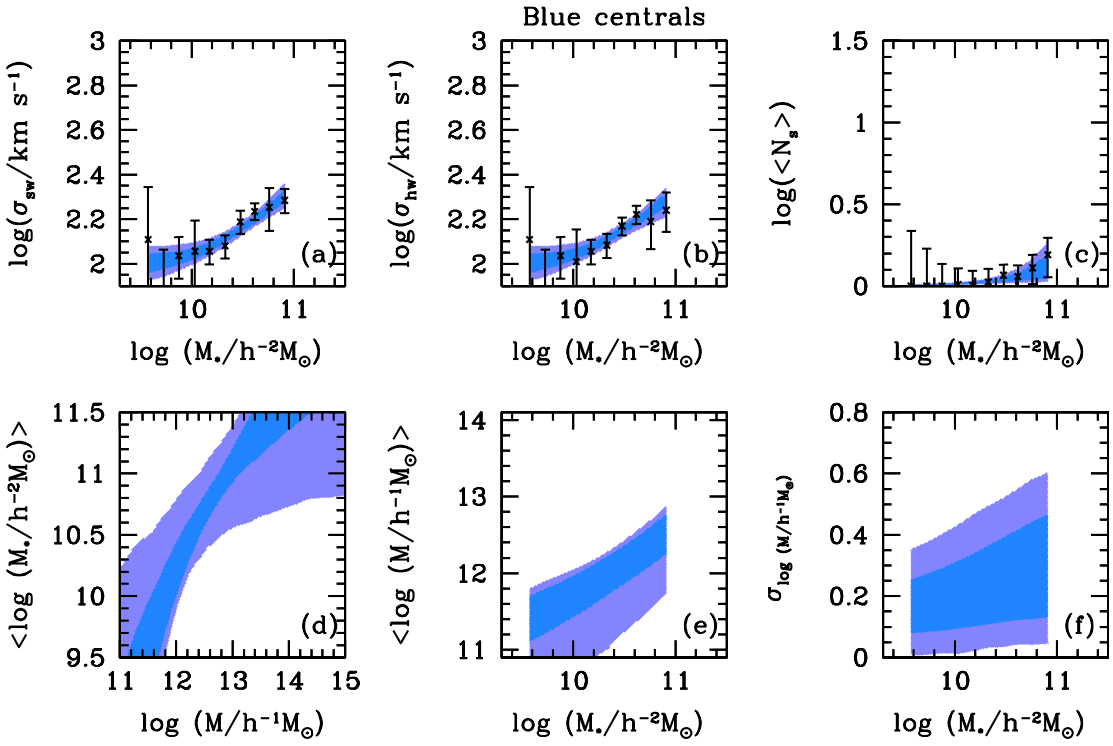

Figure 5.10: The results of the MCMC analysis of the velocity dispersions obtained from Sample SB. Crosses with errorbars in the upper panels show the data points used to constrain the MCMC; the satellite-weighted velocity dispersions in panel (a), the host-weighted velocity dispersions in panel (b) and the mean number of satellites per central as a function of stellar mass in panel (c), all measured by using the satellites in Sample SB. The blue and purple bands represent the 68% and 95% confidence regions, respectively. The bottom panels show the relations inferred from the MCMC; the average $\log(M_{*\mathrm{c}})$ is in panel (d), and the mean and the scatter in the MSR of central galaxies in panels (e) and (f), respectively.

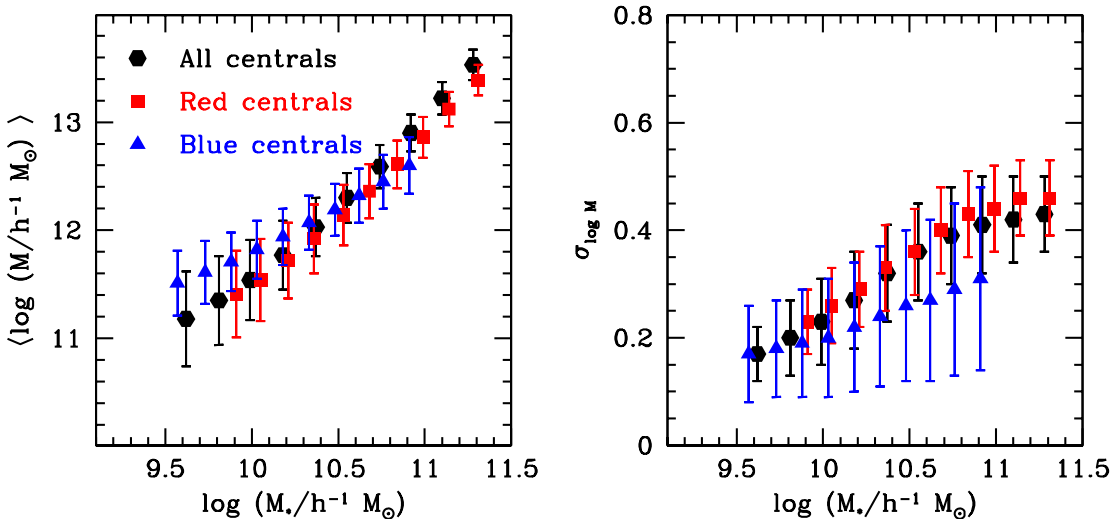

Figure 5.11: Comparison of the average halo masses and the scatter in halo masses of central galaxies split by colour as a function of their stellar mass. The results for red central galaxies are shown with (red) squares and those for blue central galaxies are shown with (blue) triangles.

Table 5.7: Sample SR: MSR of red central galaxies

$\log(L_c)$	$\langle \log M \rangle$	$\Delta\langle \log M \rangle$	$\sigma_{\log M}$	$\Delta\sigma_{\log M}$
$h^{-2}L_\odot$	$h^{-1}M_\odot$	$h^{-1}M_\odot$	$h^{-1}M_\odot$	$h^{-1}M_\odot$
9.91	11.41	0.40	0.23	0.06
10.05	11.54	0.38	0.26	0.07
10.21	11.72	0.35	0.29	0.07
10.36	11.92	0.32	0.33	0.08
10.53	12.14	0.28	0.36	0.08
10.68	12.36	0.25	0.40	0.08
10.84	12.61	0.22	0.43	0.08
10.99	12.86	0.19	0.44	0.08
11.14	13.12	0.16	0.46	0.07
11.31	13.39	0.14	0.46	0.07

The mean and scatter of the halo masses as a function of the central galaxy stellar mass inferred from the MCMC analysis of red central galaxies. The errors on each of the inferred quantities correspond to the 68% confidence levels.

Table 5.8: Sample SB: MSR of blue central galaxies

$\log(L_{\rm c})$	$\langle \log {\rm M} \rangle$	$\Delta \langle \log {\rm M} \rangle$	$\sigma_{\log {\rm M}}$	$\Delta \sigma_{\log {\rm M}}$
$h^{-2}{\rm L}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$	$h^{-1}{\rm M}_{\odot}$
9.57	11.51	0.30	0.17	0.09
9.73	11.61	0.29	0.18	0.09
9.88	11.71	0.27	0.19	0.10
10.03	11.82	0.27	0.20	0.11
10.18	11.94	0.26	0.22	0.12
10.33	12.07	0.25	0.24	0.13
10.48	12.19	0.24	0.26	0.14
10.62	12.32	0.25	0.27	0.15
10.76	12.45	0.25	0.29	0.16
10.91	12.60	0.26	0.31	0.17

The mean and scatter of the halo masses as a function of the central galaxy stellar mass for blue central galaxies inferred from the MCMC analysis. The errors on each of the inferred quantities correspond to the 68% confidence levels.

blue galaxies at the massive end is poorly constrained as the sample does not consist of massive blue central galaxies. The value of the scatter in stellar masses of all central galaxies at fixed halo mass is $0.21^{+0.03}_{-0.02}$. The value of the scatter is $0.23^{+0.02}_{-0.02}$ for red central galaxies while it is $0.27^{+0.14}_{-0.13}$ for blue central galaxies.

The average halo mass of all central galaxies and that of red and blue central galaxies increases as a function of the stellar mass and so does the scatter in halo masses. In Fig. 5.11, we overplot the mean and the scatter of the MSR from Samples SA, SR and SB together. The difference in the scaling relation between the average halo mass and the stellar mass for red and blue galaxies is much less pronounced than that seen in the MLR. This demonstrates that when stacked by stellar mass central galaxies on average occupy haloes of similar mass independent of their colour and in this respect stellar mass is a better proxy for halo mass. We find that the scatter in halo masses also increases as a function of the stellar mass, similar to the behaviour seen as a function of the luminosity, and reaches about ~ 0.4 dex at the massive end.

Finally, we would like to point out here that the velocity dispersion of satellite galaxies around red central galaxies is always systematically higher than that around blue central galaxies of the same stellar mass. However, the average MSR inferred from the analysis is roughly similar in both cases. This shows that it is not straightforward to use the velocity dispersion as a proxy for halo mass and that inferring the average halo masses demands a careful modelling similar to the one presented in this thesis.

5.5.2 Comparison of MSR with Other Studies

In this section, we compare the MSR obtained using satellite kinematics with those obtained with other independent methods. We first investigate the MSR of central galaxies, with and

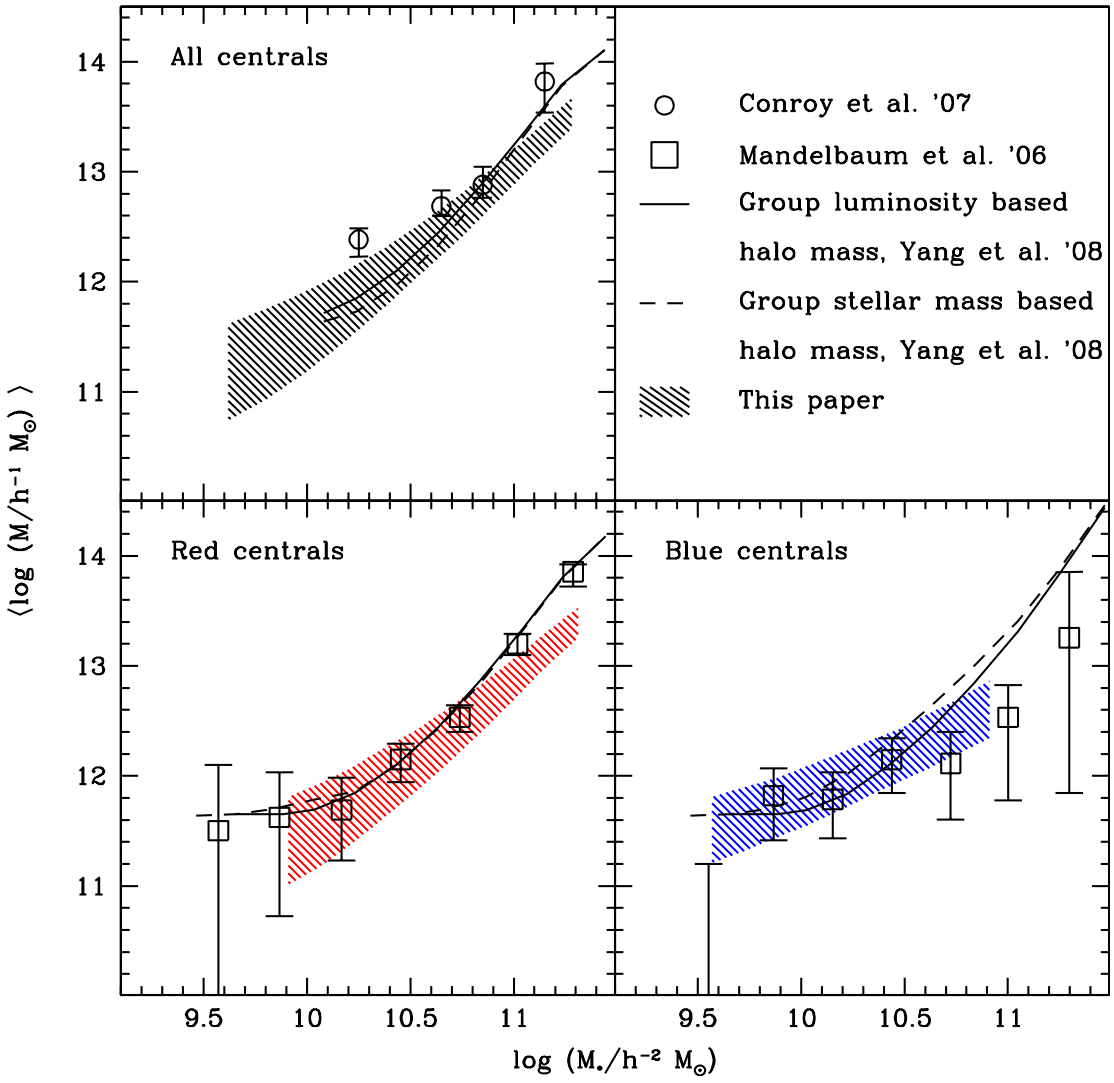

Figure 5.12: Comparison of the MSR of central galaxies by different methods. The shaded areas represent the results obtained in this chapter using the kinematics of satellites. The results obtained by analysing the group catalogue are shown with solid and dashed lines. The results obtained from the weak lensing analysis are shown as squares with errorbars. The results obtained by the analysis of satellite kinematics by Conroy et al. (2007) are shown using circles with errorbars.

without the split into red and blue colours in the SDSS group catalogue of Yang et al. (2007). The solid lines in Fig. 5.12 correspond to the MSR of central galaxies in their group catalogue where the halo masses have been assigned using the total stellar mass content of the group. The dashed lines, in turn, show the MSR of central galaxies in the group catalogue where the halo masses have been assigned according to the luminosity content of the group. The MSR for central galaxies obtained in this chapter are shown with shaded areas (68 percent confidence region). The MSR of all central galaxies (top left panel in Fig. 5.12) inferred from satellite kinematics is overall in good agreement with the results from the group catalog. The MSR for red centrals obtained in our study are in slight tension at the massive end while those for blue centrals are in a fair agreement. As mentioned in the previous chapter, one has to be careful while interpreting results from the group catalogue for a subset of galaxies because it may not be reliable to deduce the average properties of a subset of galaxies (such as central galaxies split by their colour).

The MSR of central galaxies from the galaxy−galaxy lensing by Mandelbaum et al. (2006b) are shown in Fig. 5.12 as squares with errorbars. These results are also in fair agreement with the results from satellite kinematics. The potential disagreement for the most massive stellar mass bin from their sample is most likely a result of the different criteria used to separate red and blue central galaxies. Unlike our sample, Mandelbaum et al. split the galaxies into early (red) and late (blue) types based upon their morphology.

Using data from the SDSS and the DEEP2 survey, Conroy et al. (2007) used the kinematics of satellite galaxies to determine the evolution of the stellar mass-to-light ratio of central galaxies from $z \sim 1$ to $z \sim 0$. They measured and modelled the radial dependence of the velocity dispersion to infer the average mass of the halo as a function of the stellar mass of the central galaxy. We compare the MSR at $z \sim 0$ obtained by Conroy et al. (2007) with our results. Note that the halo mass definition used by Conroy et al. (2007) corresponds to M_{200}, i.e., the mass of the halo within a radius which encloses an average density which is 200 times the critical density of the Universe. We have converted their definition of the halo mass to our definition of the halo mass following the procedure outlined in Hu & Kravtsov (2003). The stellar masses of the galaxies were calculated using a stellar mass-to-light ratio based on a Chabrier initial mass function. We have also converted these stellar masses to the ones based on the initial mass function obtained by Kroupa (2001) for consistency using results from Bell et al. (2003). The halo mass−stellar mass relationship for all central galaxies obtained from the analysis of Conroy et al. (2007) is shown in Fig. 5.12 using circles with errorbars. The average halo masses obtained by their analysis are in agreement at the 2-σ level with the ones obtained in this chapter. However, their halo mass measurements are always systematically larger than those obtained by us. This could be a result of the fact that the scatter in the halo mass−stellar mass relationship is assumed to be negligible in their analysis. This tends to overestimate the halo masses of central galaxies (see Chapter 2).

5.6 Summary

In this chapter, we used the kinematics of satellites galaxies to investigate the scaling relation between the halo mass and the stellar mass of a central galaxy (the halo mass−stellar mass relationship or MSR). We also investigated the dependence of the MSR on the colour of central galaxies. The MSR shows similar trends as seen in the halo mass−luminosity relationship (MLR, see Chapter 4). Both the mean and the scatter of the MSR of central galaxies increase with the stellar mass of the central galaxy. However, when split by colour, the difference between the mean MSR of the red and the blue central galaxies is less pronounced than that seen in the MLR. This implies that when stacked by stellar mass, the red and the blue central galaxies on average occupy similar mass haloes. We also found that the MSR of both red and blue central galaxies, individually, have an appreciable scatter at the massive end.

We compared the average MSR of central galaxies obtained by our analysis of the kinematics of satellite galaxies with other independent studies. The average MSR we obtain is in good quantitative agreement with results from galaxy−galaxy lensing (Mandelbaum et al. 2006b) and with results from the SDSS group catalogue of Yang et al. (2007). This shows that the average scaling relations that relate the stellar mass of galaxies to their dark matter halo masses are well established and are supported by several different astrophysical probes.

The scatter in the MLR and the MSR of central galaxies that we have inferred from satellite kinematics reflect the stochasticity of galaxy formation. In the next chapter, we investigate the physical processes that are responsible for this scatter. With the help of a semi-analytical model of galaxy formation, we investigate the scatter in the merger histories of dark matter haloes and its effect on the properties of the central galaxies that form at their centres.

Chapter 6

On the Stochasticity of Galaxy Formation

The contents of this chapter are based upon an article which is in preparation. The article will be submitted to the Monthly Notices of the Royal Astronomical Society as:
Stochasticity of Galaxy Formation: Insights from Galaxy Formation Models
More, S., More, A., van den Bosch, F. C., et al. 2009

6.1 Introduction

Dark matter is the most abundant gravitationally unstable component in the Universe and is therefore responsible for the formation of structure. The tiny initial fluctuations in the dark matter density field grow over time by the action of gravity and form bound structures (haloes). The baryons within these haloes undergo cooling and gradually transform into stars and form galaxies (White & Rees 1978). The ratio of the mass in baryons to the mass in dark matter is universal and each halo collapses with its "fair share" of baryons. If the efficiency with which the baryons transform into stars is independent of the halo mass, it is a natural expectation that extensive properties of galaxies such as the luminosity and the stellar mass correlate positively with halo mass. However, each halo is unique, each has its own merger history and its own star formation history and thus, haloes with similar mass need not harbour galaxies with the same properties. It is certainly interesting to quantify and study this difference as it directly reflects the stochasticity in the physics of galaxy formation.

In this thesis, we have used the kinematics of satellite galaxies to constrain the scaling relations between halo mass and central galaxy properties. In Chapter 4, we inferred the mean and scatter of the relationship between halo mass and the luminosity of central galaxies (MLR) which occupy the centres of dark matter haloes. In Chapter 5, we inferred the mean and scatter of the relationship between halo mass and the stellar mass of central galaxies (MSR). The results show that the average halo mass increases with both the luminosity and the stellar mass of

a central galaxy. In addition, the *scatter* in halo masses for galaxies of a particular luminosity (stellar mass) also increases as a function of the galaxy luminosity (stellar mass). Understanding the origin of this scatter can help us learn more about the physics of galaxy formation.

There exists a large scatter in the assembly histories of different dark matter haloes (see e.g., Lacey & Cole 1993; Kauffmann & White 1993; Lemson & Kauffmann 1999; van den Bosch 2002; Gao et al. 2005; Wechsler et al. 2006; Li et al. 2007, 2008). This scatter in the assembly history is also reflected in the clustering properties of dark matter haloes such that haloes that assemble earlier are more strongly clustered than those which assemble later (Gao et al. 2005; Wechsler et al. 2006; Harker et al. 2006; Jing et al. 2007; Gao & White 2007). The scatter in the assembly history of haloes possibly also reflects in a scatter in the star formation histories of the galaxies, which may finally result in a scatter in their properties. Therefore, the scatter in the assembly history of haloes is a plausible origin of the stochasticity of galaxy formation. A semi-analytical model of galaxy formation (SAM) is an excellent tool to investigate the origin of the scatter in the MLR/MSR of galaxies. A SAM considers the hierarchical merger histories of dark matter haloes and implements simple physical recipes to model the astrophysical processes that affect the baryons in these haloes (see e.g., White & Frenk 1991; Kauffmann & White 1993). The most recent versions of these models include processes such as radiative cooling, star formation, energetic feedback from supernovae and active galactic nuclei, and galactic outflows (see Section 6.2).

In this chapter, we only focus on the MSR of central galaxies and investigate the origin of the scatter in the MSR using the semi-analytical model of De Lucia & Blaizot (2007, hereafter DL07). This chapter is organized as follows. In Section 6.2, we explain the ideology behind the use of semi-analytic methods to model galaxy formation and briefly explain the processes that are commonly modelled in SAMs. In Section 6.3, we describe the numerical simulation used by DL07 and describe the procedure used to construct the merger histories of haloes from this simulation. The merger histories of haloes can be characterized by the formation times of the haloes. In Section 6.4, we present various definitions for the formation times of haloes based upon their merger histories. In Section 6.5, we investigate the effect of the scatter in the merger histories of similar mass haloes on the stellar mass of central galaxies that form in these haloes. We conclude in Section 6.6 with a discussion and a short summary of the results.

6.2 Semi-analytical Models of Galaxy Formation

The last decade has seen a remarkable progress on the observational front which has helped us further our understanding of galaxy formation. Detailed observations of the local galaxy population (redshift $z \sim 0.0$) via large scale galaxy redshift surveys coupled with observations of galaxies at high redshifts (redshift $1 \leq z \leq 5$) in various bands have resulted in a wealth of observational results: e.g. the luminosity functions of local galaxies (Blanton et al. 2003b; Norberg et al. 2002b) and galaxies at high redshifts (Drory et al. 2003; Gabasch et al. 2004; Drory

et al. 2005; Faber et al. 2007), the spatial clustering of galaxies (Norberg et al. 2002a; Tegmark et al. 2004; Zehavi et al. 2005), the stellar mass−metallicity relation of galaxies (Tremonti et al. 2004; Gallazzi et al. 2005), the bimodal distribution of galaxies in the colour−magnitude plane (Baldry et al. 2004), the correlation between the mass of the central black hole and the bulge in a galaxy (Häring & Rix 2004), the Tully−Fisher relation (Giovanelli et al. 1997) and the the star formation rate history of the Universe (Madau et al. 1996). To understand and connect these observations in the cosmological framework and to sketch a picture of the formation and evolution of galaxies through cosmic time is a holy grail for the modern astrophysicist.

The initial conditions that describe a ΛCDM Universe are fairly simple and can be specified by a handful of parameters (Dunkley et al. 2009): the energy density parameters of the different components of the Universe (dark matter, dark energy and baryons), the parameters that specify the power spectrum of small fluctuations in the initial density field (the normalization and the spectral index), and the present rate of the expansion of the Universe (specified by the Hubble parameter). The fluctuations in the dark matter component evolve purely due to gravity and can be followed in a straightforward manner by using a numerical simulation (Davis et al. 1985). However, following the evolution of baryons and resolving the formation of individual galaxies in a cosmological volume is computationally very expensive. SAMs use a hybrid approach, in which the evolution of the dark matter skeleton is followed by a numerical simulation and the evolution of baryons is followed by using simple analytical recipes, to model the formation of galaxies in a cosmological context (see e.g., White & Frenk 1991; Kauffmann & White 1993; Cole et al. 1994; Kauffmann 1996; Kauffmann et al. 1997; Baugh et al. 1998; Somerville & Primack 1999; Cole et al. 2000; Benson et al. 2002; Springel et al. 2005; Croton et al. 2006; De Lucia & Blaizot 2007).

The internal structure of dark matter haloes, specified by the density profile and the angular momentum, is an important ingredient to model the galaxies that form within these haloes. High resolution simulations have shown that the density profile of dark matter haloes is universal and can be described by Eq. 2.22 (Navarro et al. 1997) which has only one free parameter for a given mass, the concentration c. The concentration depends very weakly on the mass (Bullock et al. 2001; Macciò et al. 2007), however this relation has a considerable scatter which is correlated with the formation histories of haloes (Navarro et al. 1997; Wechsler et al. 2002; Zhao et al. 2003). Large scale tidal torques impart angular momentum to the dark matter haloes. The angular momentum of a halo is quantified by the dimensionless spin parameter, $\lambda = J|E|^{0.5}/GM^{2.5}$, where J, M and E are the angular momentum, mass and the energy of the halo. The spin parameter follows a lognormal distribution with a mean and scatter which is relatively independent of the mass of the halo (Barnes & Efstathiou 1987; Cole & Lacey 1996; Lemson & Kauffmann 1999; Bullock et al. 2001; Macciò et al. 2007).

The amount of baryons in each dark matter halo is initially a fixed fraction of the halo mass and is present in a diffuse form. The elemental composition of this gas in the halo is set by the big bang nucleosynthesis. As haloes merge, the infalling gas is shock−heated. Numerical

simulations show that this hot gas settles down to a isothermal density profile which has a core at roughly one third the scale radius of the halo and with a temperature which is close to the virial temperature (Navarro et al. 1995; Eke et al. 1998; Frenk et al. 1999). The rate of cooling of gas depends upon the cooling function which is a function of the temperature, density and the metallicity of the gas (Sutherland & Dopita 1993). By the virtue of its angular momentum, the cold gas settles down into a disc which then fragments due to gravitational instabilities and leads to the formation of stars (Kennicutt 1998). The initial mass function specifies the distribution of masses of the newly formed stars (Salpeter 1955; Chabrier 2003; Kroupa 2001). Stars are reservoirs of energy that constantly heat up their surroundings. They also lose a fraction of their mass to the surroundings by stellar winds. In addition, young massive stars undergo supernovae explosions and heat up the surrounding cold gas. The supernova ejecta also enrich the metallicity of the surrounding gas. These baryonic processes which take place in haloes are modelled using simple analytical prescriptions in a SAM.

Simple recipes are also required to determine the fate of the baryons in two haloes that merge. Mergers are commonly classified as major or minor based upon the ratio of masses of haloes that merge together. In case of a minor merger, the central galaxies and satellites in the smaller halo become satellites of the bigger halo. Their individual (sub)haloes can be resolved in a numerical simulation if they are above the detection limit of substructure specified in the substructure finding routines. Once the subhalo masses fall below this limit, these galaxies are merged with the central galaxy on a dynamical friction timescale (Chandrasekhar 1943). Any cold gas in the satellite galaxies is added to the disc of the central galaxy. Minor mergers of galaxies which have some cold gas are accompanied by minor bursts of star formation. Major mergers between galaxies are more violent and can result in an extreme episode of star formation if any of the galaxies have a reservoir of cold gas. Such mergers destroy the discs present in any of the haloes and the stars so formed are distributed in a spheroid.

The final ingredient we discuss are the supermassive blackholes which reside at the centres of galaxies. The process of gas accretion on to supermassive blackholes results in the release of a significant amount of energy. A high accretion rate, which can be a result of perturbations such as bar instabilities induced during mergers, can cause the blackhole to be in a quasar phase. In this phase, the blackhole emits a tremendous amount of energy which is transferred to the surrounding gas via massive jets. A more quiescent accretion of the hot gas by the black hole can cause low energy radio activity which pumps energy into the surroundings (Tabor & Binney 1993; McNamara et al. 2000, 2005). Both feedbacks have been shown to be effective in shaping the bright end of the luminosity function of galaxies and preventing star formation in clusters by suppressing cooling flows (Croton et al. 2006).

This concludes our very brief description of the various processes that are commonly modelled in a SAM. We refer the reader to Cole et al. (2000) for an excellent overview on this subject and to the papers by Springel et al. (2005), Croton et al. (2006) and De Lucia & Blaizot (2007) for the details of the specific model that we use in our analysis. With this description,

we hope to have conveyed the diversity of processes that shape the properties of galaxies and that results in the stochasticity of galaxy formation in haloes of similar masses.

6.3 Numerical Simulation and Halo Merger Trees

The semi-analytical recipes that model the processes described in the previous section are implemented in halo merger trees. DL07 use the halo merger trees derived from the Millennium simulation which was carried out by the Virgo consortium (Springel et al. 2005). The Millennium simulation tracks the positions and velocities of dark matter particles in a comoving cube with length equal to $500\ h^{-1}\mathrm{Mpc}$ on each side. The simulation is carried out assuming a flat ΛCDM cosmology with the following cosmological parameters: density parameter for the cosmological constant $\Omega_\Lambda = 0.75$, matter density parameter $\Omega_\mathrm{m} = 0.25$, baryon density parameter $\Omega_\mathrm{b} = 0.045$, $h = H_0/100\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, the linearly extrapolated root mean squared variance of density fluctuations on a scale of $8\ h^{-1}\mathrm{Mpc}$ $\sigma_8 = 0.9$ and the spectral index of the initial density fluctuations $n_s = 1$. The number of particles used to carry out the simulation is 2160^3 and the particle mass is $8.6 \times 10^8\ h^{-1}\mathrm{M}_\odot$.

The particle positions and velocities from the Millennium simulation are stored at 64 different epochs. For each snapshot, a friends-of-friends (FOF) algorithm with a linking length of 0.2 times the mean inter-particle separation is used to construct a catalogue of haloes. The *SUBFIND* algorithm (Springel et al. 2001) is run on each of these FOF haloes to identify the substructure in each halo. The smooth background halo is also identified as a substructure by this algorithm. This catalogue which groups the simulation particles as subhaloes is the input for the construction of the merger trees. Hereafter, we denote the smooth background haloes identified by SUBFIND as central haloes and the rest of the substructure as satellite haloes. The number of particles in each of these haloes is used to calculate the mass of the halo. Whenever required, we explicitly use the notation "FOF haloes" to denote the haloes identified originally by the FOF algorithm.

For every halo at a given snapshot, one must identify its progenitors at an earlier snapshot to construct a merger tree. A halo H_1 at redshift z_1 is considered to be a progenitor of a halo H_0 at a latter epoch with redshift z_0, if a certain fraction of the most bound particles of the halo H_1 are part of the halo H_0. In this manner, any given halo at redshift 0 branches into its progenitors at the previous redshift and these progenitors subsequently branch into their own progenitors at higher redshifts. By following the growth of mass in dark matter haloes, merger trees capture the assembly history of these haloes.

The merger trees of haloes from the Millennium simulation can be accessed online at the URL `http://www.g-vo.org/Millennium` from a database table using the Sequential Query Language (SQL). Another database table contains the output of the SAM of DL07. The query execution times on the database are limited to 30 (500) seconds for unregistered (registered) users. The website also offers similar tables for a *milli-version* of the entire simulation

which is $8 \times 8 \times 8$ smaller than the original simulation but has the same mass resolution as the larger simulation. The SQL queries that are run on the milli-version require a shorter execution time compared to that on the original simulation because of the smaller volume. Therefore, we carry out our analysis only using the milli-version. Eventually, we would like to use the entire simulation to capture the full diversity of the merging histories of haloes.

6.4 Halo Formation Timescales

The merger trees of haloes with the same masses can be quite different in appearance. Our aim is to quantify if the scatter in the properties of the central galaxies such as their luminosities or stellar masses at fixed halo mass correlate with the scatter in the merger histories of their haloes. One way to characterise the merger history of a halo is to consider its formation time. The formation time of a halo is generally defined as the time when the halo has assembled a fixed fraction of its final mass.

The growth of assembled mass in any merger tree can be studied in two different ways. Starting from a halo at present time one can consider its most massive progenitor in the previous snapshot (see the red arrows in Fig. 6.1. This most massive progenitor has its own most massive progenitor and so on. The linked list of the most massive progenitors of the most massive progenitors which lead to the formation of the final halo is called the main branch of the merger tree. Alternatively, one can consider the linked list which connects the most massive progenitor at every snapshot with the most massive progenitor at an earlier snapshot (see the blue arrows in Fig. 6.1). Such a linked list may not always be a continous merging branch. This list represents the growth of the "maximum progenitor". Li et al. (2008) have presented eight different definitions of the formation times for any given halo based either upon the main branch or the linked list of the maximum progenitors and shown how these formation times vary with the mass of the halo.

We use the following 4 different definitions of the formation redshift of the halo for the purpose of our analysis. We have chosen to use these definitions for the relative ease of extracting these formation redshifts from the database with the help of SQL queries.

1. $z_{1/2,\mathrm{mp}}$: This is the highest redshift at which the "maximum progenitor" has accumulated a mass which is greater than or equal to one half of the final halo mass.

2. $z_{1/2,\mathrm{t1}}$: This is the highest redshift at which the progenitors with at least 2% of the final mass have collectively assembled a mass which is greater than or equal to one half of the final halo mass. As noted by Li et al. (2008), such a definition of the formation time of the halo has also been used by Navarro et al. (1997) to study the dependence of halo concentrations on their formation times.

3. $z_{1/2,\mathrm{t2}}$: This is the highest redshift at which half of the final halo mass has assembled into progenitors above a fixed minimum mass, M_c. We use $M_c = 10^{11.5}\,h^{-1}\mathrm{M}_\odot$ because

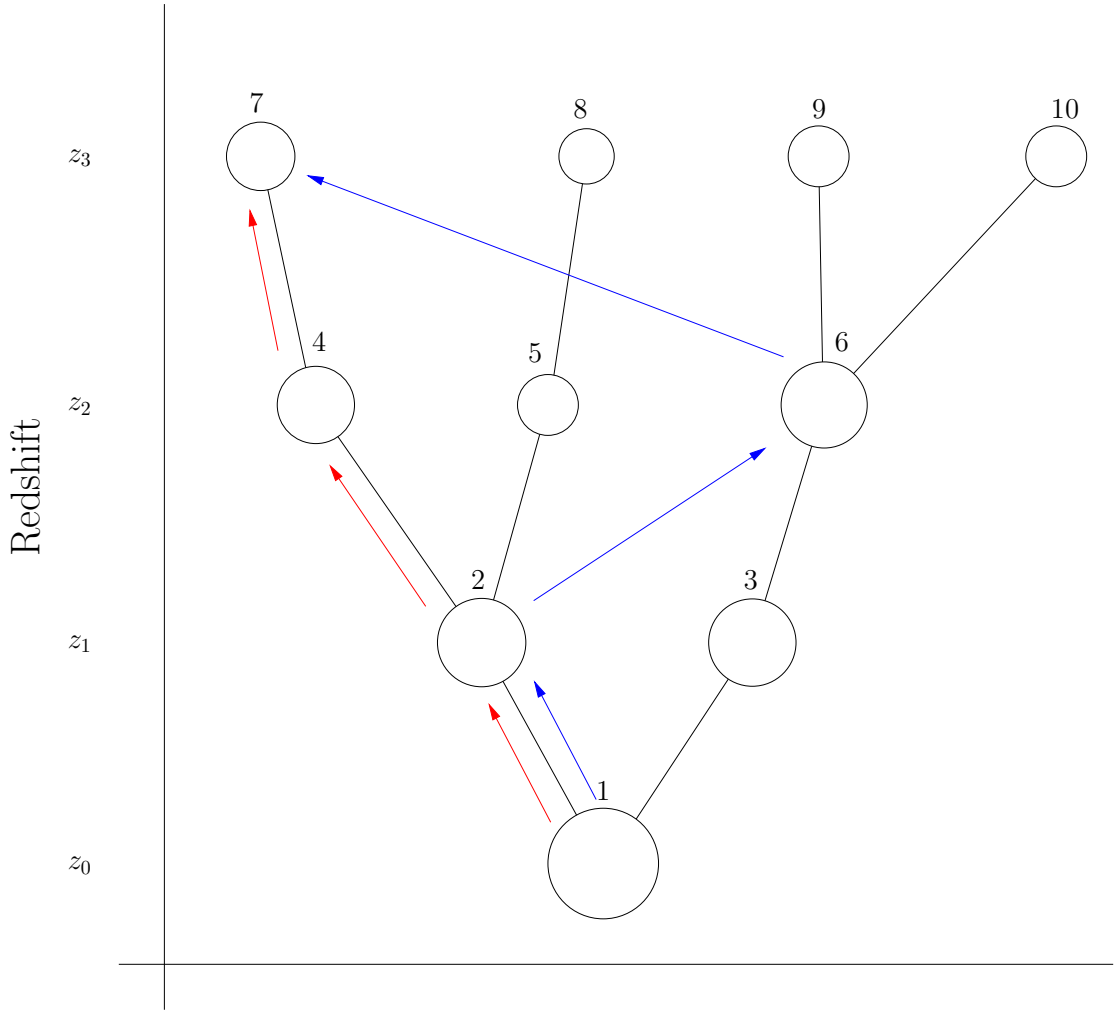

Figure 6.1: Schematic diagram of a merger tree. The haloes at different epochs are shown as circles with radii roughly proportional to the cuberoot of the mass. The x-axis in this figure is arbitrary. Each merger event is represented by a black solid line. The main branch of the merger tree is shown with red arrows while the most massive progenitors at each redshift are linked with blue arrows.

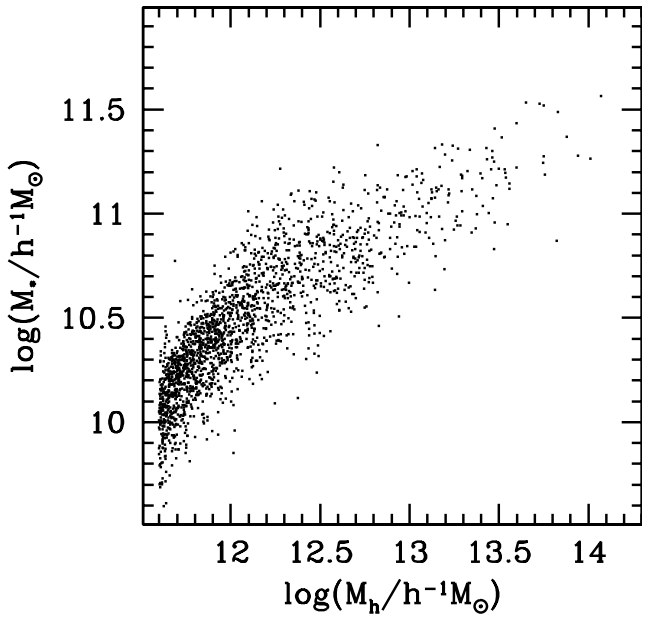

Figure 6.2: Scatter plot of the stellar mass of central galaxies against their halo masses at redshift $z = 0$ predicted by the SAM of DL07.

star formation is the most efficient at this mass scale and this time marks the beginning of the epoch when star formation prevails in the assembly history of a halo.

4. $z_{\rm core}$: This is the highest redshift at which the maximum progenitor has assembled a mass equal to M_c. This formation redshift thus marks the time when the maximum progenitor becomes capable of hosting a bright central galaxy.

We use these definitions for the formation redshift of each halo to quantify its merger history and investigate how it shapes the stellar mass of the central galaxy that forms in it.

6.5 Results

In Fig. 6.2, we first show a scatter plot of the stellar masses of central galaxies with respect to the masses of their corresponding dark matter haloes predicted by the SAM of DL07. Note that we are using the galaxies from the milli-Millennium database only. The use of a small volume results in the plot being sparsely populated at the massive end. The stellar mass of the central galaxy increases with the halo mass with roughly a constant scatter of $\sim 0.16 \pm 0.06$ dex at fixed halo mass. This is in fair agreement with results for the MSR inferred from the kinematics of satellite galaxies presented in Chapter 5.

The different formation redshifts of the halo defined in the previous section show a varied dependence on the halo mass. In Fig. 6.3, we show the scatter plot of the logarithm of the scale factor at the formation redshift (multiplied by -1) against the final mass of the halo for all haloes

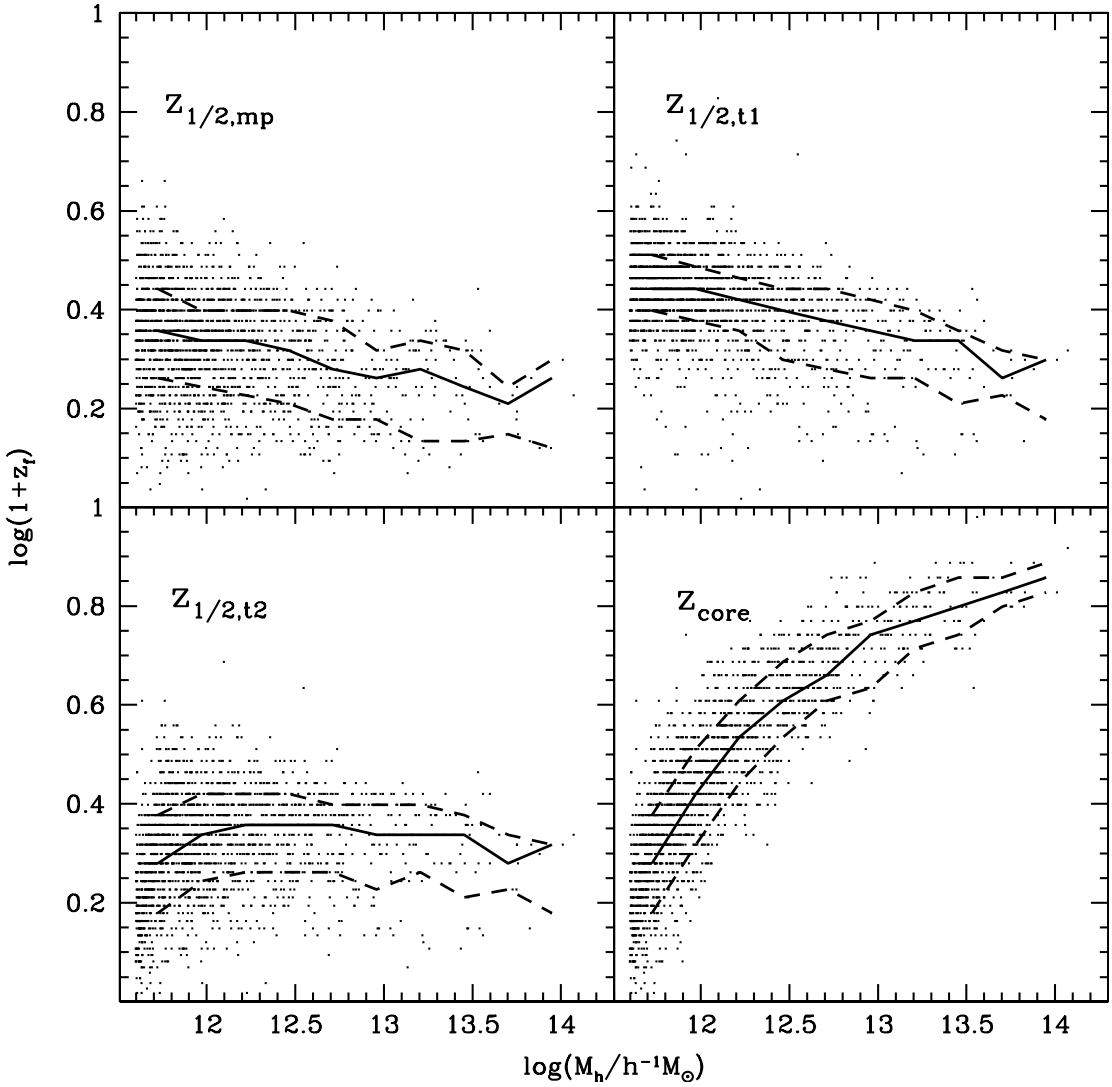

Figure 6.3: Dependence of various formation redshifts of dark matter haloes on their final mass. The solid curves show the median relation between the formation redshifts and the final mass of the halo while the dashed curves show the 20 and 80 percentiles.

in the milli-Millennium database. The median of the distribution of formation redshifts as a function of the halo mass is shown with a solid line in each panel. The dashed lines show the 20 and 80 percentiles of the distribution of the formation redshifts obtained by using bins of equal width in the logarithm of the halo mass $\log M_h$. The different formation times of the haloes are quite different from each other and each of these capture different aspects of the formation of haloes.

The formation redshifts $z_{1/2,\mathrm{mp}}$ and $z_{1/2,\mathrm{t1}}$ show a decreasing trend with the mass of a halo, i.e. according to these definitions the most massive haloes on average form later than the low mass haloes. The formation redshift, $z_{1/2,\mathrm{t2}}$ on the other hand shows a relatively flat behaviour while the formation time z_{core} is highly correlated with the mass of the halo with a scatter which is slightly less than that seen in the other definitions of the formation redshift.

In Fig. 6.4, we show the scatter plot of the logarithm of the scale factor at the formation redshifts against the stellar mass of the central galaxy at the final epoch. The formation redshifts $z_{1/2,\mathrm{mp}}$, $z_{1/2,\mathrm{t1}}$ and $z_{1/2,\mathrm{t2}}$ show a very little correlation with the stellar mass of the central galaxy. However the formation redshift z_{core} is strongly correlated with the final stellar mass of the central galaxy. This shows that the earlier the maximum progenitor reaches the mass M_c, the more stars the central galaxy has at redshift zero.

We now investigate a possible correlation between the scatter in the formation redshifts of the haloes and the scatter in the stellar masses of the central galaxies that form in these haloes. Fig. 6.5 shows the scatter plot of the residuals around the mean relation between the stellar mass and halo mass against the residuals around the mean relation between formation redshifts and halo mass. The correlation between these residuals is quantified by the correlation coefficient which is denoted at the bottom right corner of each panel in the figure. This figure shows that at fixed halo mass the stellar mass of central galaxies is correlated with the formation time of the halo. The formation redshift z_{core} shows the highest correlation with the formation time of the halo. This correlation of the residuals can also be seen from Fig. 6.6, where we show the stellar mass halo mass relation. In this figure, the points in each panel are colour coded according to the residuals in the corresponding formation redshift halo mass plot. The red points which correspond to haloes that form earlier lie preferentially above the mean relation between halo mass and stellar mass while blue points which correspond to haloes that form later lie below the mean relation. This shows that at fixed halo mass, haloes that form early typically host central galaxies that have more stellar mass. However, as expected, the correlation is not perfect. This implies that the scatter in the formation times of haloes is not the only reason for the stochasticity of galaxy formation. In future, we plan to investigate this issue with the help of semi-analytical modelling to find out the main reasons for the presence of scatter in the galaxy property-halo mass scaling relations.

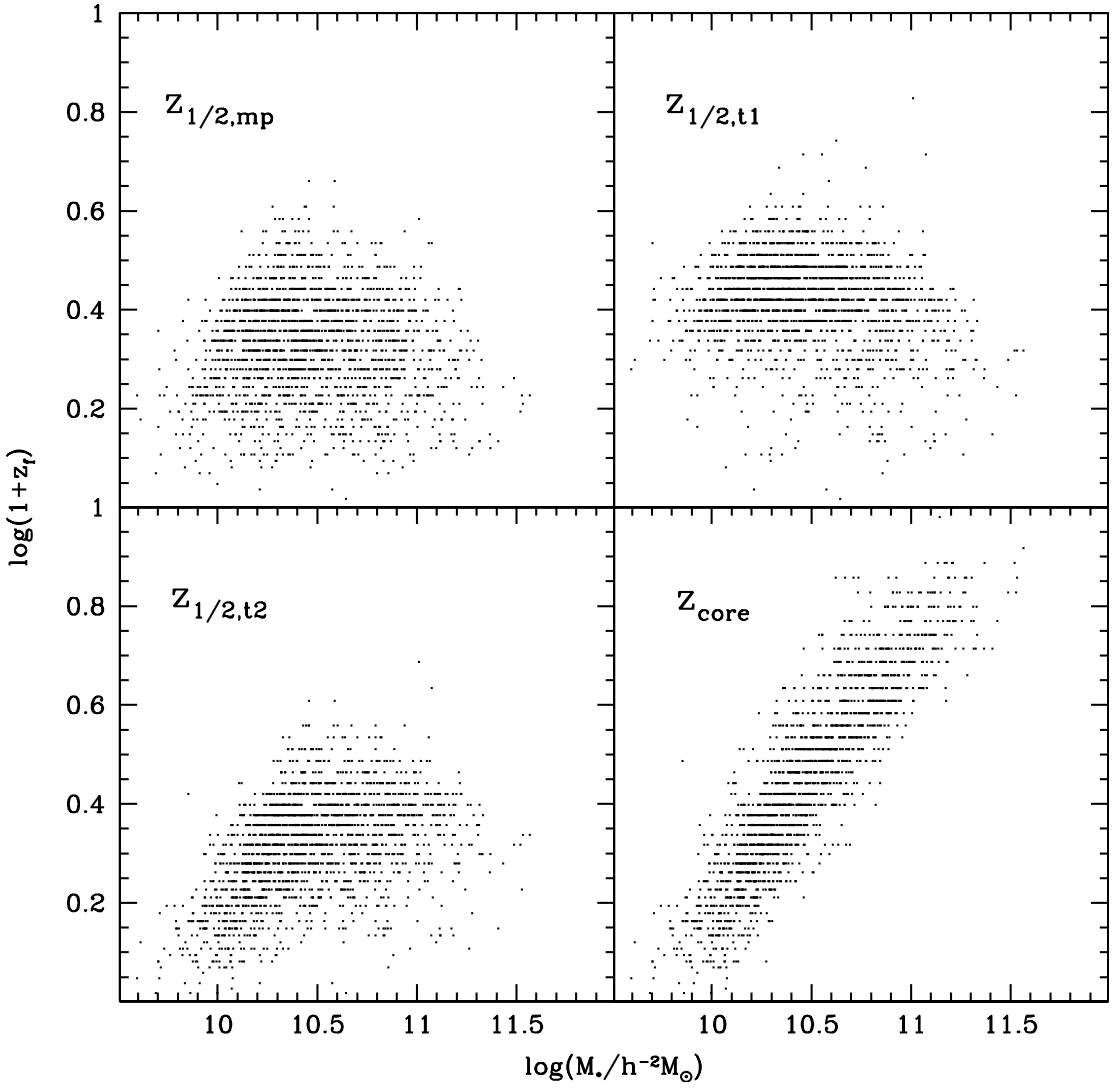

Figure 6.4: Scatter plot of the stellar mass of central galaxies versus the various formation redshifts of their haloes.

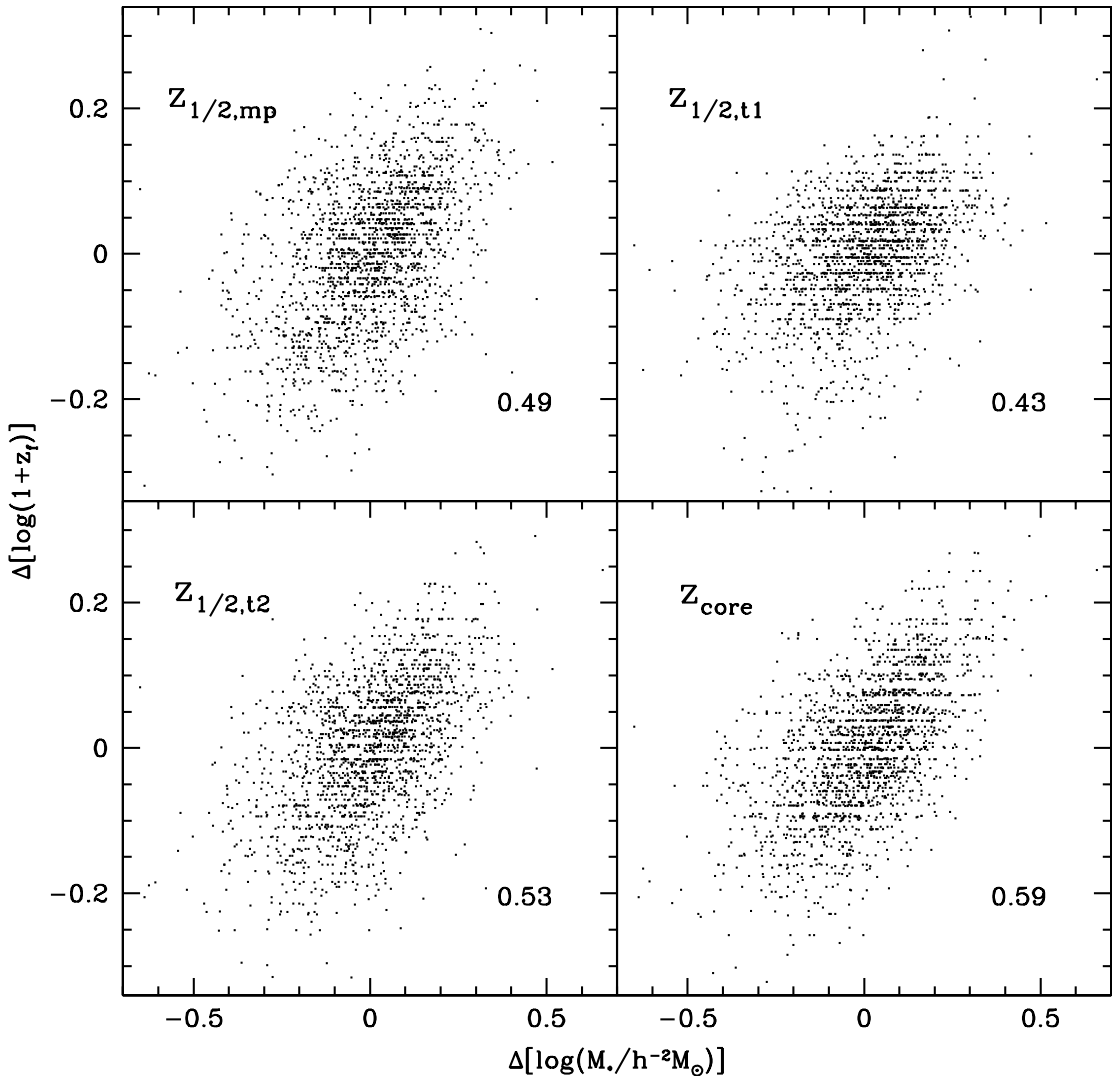

Figure 6.5: Scatter plot of residuals in the formation redshift−halo mass relation versus the residuals in the stellar mass−halo mass relation. The correlation coefficients between these residuals are shown at the bottom right corner in each panel.

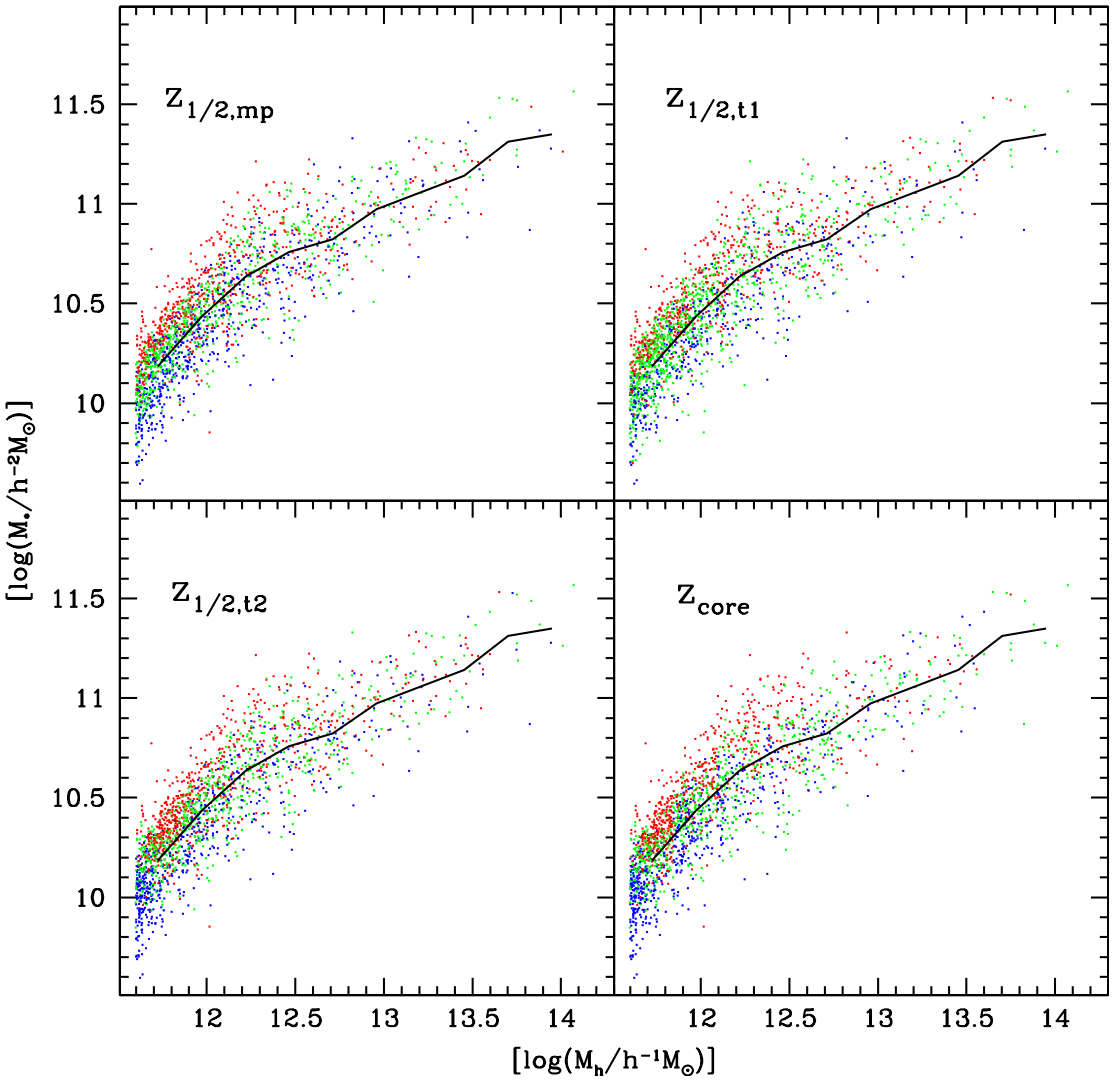

Figure 6.6: Scatter plot of the stellar mass−halo mass relation colour coded according to the residuals in the formation redshift−halo mass relation. The points in red correspond to haloes with residuals $\Delta[\log(1+z_f)]$ less than -0.05, green for haloes with residuals in the range $[-0.05, 0.05]$ and blue for haloes with residuals greater than 0.05. The red (blue) points lie preferentially above (below) the mean relation.

6.6 Summary

Extensive properties of galaxies such as the luminosity and the stellar mass correlate with the mass of the halo in which they reside. However, there is a certain amount of scatter in these scaling relations which reflects the stochasticity in the galaxy formation and evolution processes. We investigated the origin of the scatter in the halo mass−stellar mass relationship of galaxies that reside at the centre of dark matter haloes using the SAM of galaxy formation from DL07.

The stellar mass of galaxies can grow via two processes: (i) by the accretion of the stellar content of satellite galaxies that merge with it and, (ii) by the formation of stars from its own reservoir of cold gas and that brought in by its satellites. Mergers of galaxies are often accompanied with bursts of star formation. Therefore, merger histories of haloes play a central role in building up the stellar mass of central galaxies. We explored the effect of the scatter in the merger histories of haloes of similar masses on the properties of the central galaxies that form in these haloes.

The merger history of a halo can be characterized by its formation time. We used several different definitions of the halo formation time and investigated their behaviour with the halo mass. The halo formation times show different behaviours as a function of mass depending upon the definition used. However, at fixed halo mass, there is a large scatter in the formation times in all the definitions. We correlated the residuals around the mean relation of halo formation time and halo mass with the residuals around the halo mass−stellar mass relationship. We found a positive correlation between these residuals which suggests that, on average, haloes that form earlier harbour central galaxies that have more stellar mass. The correlation we find between these residuals is certainly dependent on the particular SAM that we have used. In future, we would like to investigate the aspect by considering results from several different SAMs and by comparing them to each other.

We conclude by pointing out that a correlation between two variables does not necessarily imply a causal relationship between the two. To give a definite answer to what causes the scatter in the properties of central galaxies at fixed halo mass, it would be interesting to turn off by hand the various stochastic components in a SAM (e.g., use the same spin parameter for each halo) and explicitly quantify the effect each component has on the halo mass−stellar mass relationship of central galaxies. We aim to perform such an analysis and present its results in a forthcoming paper.

Chapter 7

Cosmic Transparency

This research work originated from a Galaxy Coffee talk given by David Hogg when he was visiting MPIA last summer. The contents of this chapter are entirely based upon the article More et al. (2009a) which has been accepted for publication in the Astrophysical Journal. The reference is

More, S., Bovy, J., & Hogg, D. 2009, ApJ, in press, arXiv:0810.5553

7.1 Introduction

The transparency of the Universe is extremely good. A typical astronomical camera has a shutter whose thickness is measured in microns; that shutter is far more opaque than the entire line of sight to the majority of extragalactic sources, even at extremely high redshifts, despite—in many cases—considerable column depths of dark matter, plasma, gas, and dust. There are, however, very few quantitative measures of the transparency with contemporary astronomical data.

There are several sources for photon attenuation that are clustered with matter. For example, as stars eject heavy elements, they also eject photon-absorbing ash (called "dust"). The gas and plasma in and around galaxies absorbs, scatters, and re-emits at longer wavelengths some fraction of incident radiation. More speculatively, if the dark matter is an axion or axion-like particle, it will in general have photon interactions, which can in principle produce effective absorption of photons in regions of high dark matter density and high magnetic fields. The sources of attenuation—such as these—that are clustered with matter will be correlated with galaxies and large-scale structure, and can be found with "angular difference" measurements that compare the apparent properties of sources whose lines of sight have different impact parameters with the correlated structure.

The Sloan Digital Sky Survey has permitted very sensitive angular difference measurements, which find that the attenuation correlated with large-scale structure is very small and consistent with being caused by dust, presumably the dust emitted by the stars in the galaxies that populate the structure. Measurements in the literature constrain this in visible bandpasses

at the part in 10^3 level (Ménard et al. 2008; Bovy et al. 2008). To be specific, these studies constrain *differences* in opacity along different lines of sight caused by absorbers correlated with galaxies.

It is possible, however, that there might be unclustered or "monopole" sources of attenuation, that affect all lines of sight equally, for example if the non-matter contributors to the cosmological energy-momentum tensor (the "dark energy" in modern parlance) have interactions with photons, or if there are small violations of Lorentz invariance on cosmological scales. These sources of attenuation are much harder to detect with differential experiments, but they *can* be detected by "radial difference" experiments that compare cosmological sources of radiation of known physical properties at different redshifts or radial distances.

A number of different mechanisms have been proposed during the last decade to explain the observed dimming of type-Ia supernovae (SNeIa; Riess et al. 1998; Perlmutter et al. 1999) without cosmic acceleration by employing exactly such unclustered sources of attenuation. The mixing of photons with axions in extragalactic magnetic fields could lead to photons oscillating into axions with a non-negligable probability over cosmological distances, thus reducing the flux of SNeIa at large distances (Csáki et al. 2002). Alternatively, "gray" intergalactic dust could be so gray as to evade detection through its *reddening*, while still being cosmologically important because of its overall *opacity* (Aguirre 1999). In order to account for the observed SNeIa dimming, these models predict violations of transparency at the order-unity level out to redshifts of unity (Mörtsell et al. 2002).

Furthermore, even if there are no exotic absorbers in the Universe, it is difficult (and usually model-dependent) to infer the total mean opacity from any absorbers that have been found by angular difference experiments. Radial difference and angular difference experiments are complementary, even when the absorbers are mundane; although radial difference experiments are usually less precise, they provide irreplaceable information for measuring total opacity.

Radial difference experiments are sometimes known as "Tolman tests" because they are variants of the test proposed by Tolman (1930) of the expansion of the Universe: a test of the $(1+z)^{-4}$ (where z is redshift) dimming of bolometric intensity (energy per unit time per unit area per unit solid angle; also called "bolometric surface brightness") with redshift. The intensity is closely related to the phase-space density of photons, which is conserved (in a transparent medium) along the light path; that conservation plus Lorentz invariance implies the $(1+z)^{-4}$ relation: one factor of $(1+z)^{-1}$ comes from the decrease in energy of each photon due to the redshift, another factor comes from the decrease in photons per unit time, and two more factors arise from the solid-angle effects of relativistic aberration. The Tolman test does not really test for the expansion of the Universe—the result does not depend on cosmological model, or even assumptions of isotropy or homogeneity—but rather for the combination of conservation of photon phase-space density and Lorentz invariance.

In addition to models that violate transparency, there are models that violate Lorentz invariance. Generically these models produce an energy-dependent speed of light and birefringence,

breaking the perfect non-dispersiveness of the vacuum (Amelino-Camelia et al. 1998; Gambini & Pullin 1999). These effects generally become larger with increasing energy, and observations of high-energy sources such as active galactic nuclei (Biller et al. 1999; Aharonian et al. 2008; Albert et al. 2008) and gamma-ray bursts (Schaefer 1999; Ellis et al. 2006) have shown that the linear dispersion relation for photons is preserved to good accuracy at these energies. Therefore, while these models do fail the Tolman test because of their non-trivial dispersion relations, the effect will be unmeasurably small for low-energy (visible-band) photons.

By far the most precise radial difference test to date has been performed in the radio with the cosmic background radiation. In contemporary cosmological models, the CBR comes from redshift ~ 1100 and is a near-perfect blackbody. The COBE DMR experiment established that the spectrum and amplitude of this radiation is consistent with the blackbody expectation at the $< 10^{-2}$ level at 95-percent confidence (Mather et al. 1994). A source of attenuating material, unless in perfect thermal equilibrium with the CBR, would tend to change either the spectrum or the amplitude, so this result provides a very strong constraint on the transparency at cm wavelengths (Mirizzi et al. 2005). Another test of transparency at cm wavelengths is the increase in the CMB temperature $T_{\rm CMB}$ according to the relation $T_{\rm CMB} \propto (1+z)$. Srianand et al. (2000) find consistency with a transparent Universe by measuring the CMB temperature at $z = 2.3$. Of course, many sources of attenuation are expected to be wavelength-dependent, so these beautiful results may not strongly constrain the opacity in the visible.

At visible wavelengths, there have been much less precise radial difference tests performed with galaxies, whose properties would deviate from naive predictions under extreme attenuation. After correcting for the evolution of stellar populations in galaxies, these studies find consistency with transparency at the 0.5 mag level at 95-percent confidence (Pahre et al. 1996; Lubin & Sandage 2001), which correspond to optical depth limits < 0.5 out to redshift $z \sim 1$. Unfortunately, the precision of these tests is not limited by the precision of the measurements, but rather by the precision with which the evolution of galaxy stellar populations is known; the results will not be improved substantially with additional or more precise observations.

Another test of transparency at visible wavelengths involves the measurement of the Cosmic Infrared Background (hereafter CIB). The absorption of visible photons by a diffuse component of intergalactic dust and its re-emission in the infrared contributes to the CIB. The amount of dust required to explain the systematic dimming of SNeIa would produce most of the observed CIB (Aguirre & Haiman 2000). However, discrete sources (*e.g.*, dusty star-forming galaxies) also emit in the infrared and account for almost all of the CIB, strongly constraining the role of dust in the dimming of SNeIa. Any constraint on the transparency from the CIB requires a careful subtraction of the discrete sources (Hauser & Dwek 2001).

The Tolman test can be re-written as a relationship among cosmological distance measures. There are several empirical definitions of distance in cosmology (*e.g.*, Hogg 1999); the most important for contemporary observables are the luminosity distance $D_{\rm L}$ and the angular diameter distance $D_{\rm A}$. The luminosity distance $D_{\rm L}$ to an object is defined to be the distance that relates

bolometric energy per unit time per unit area S (flux) received at a telescope to the energy per unit time L (luminosity or power) of the source, or

$$S = \frac{L}{4\pi\, D_{\rm L}^2} \quad . \tag{7.1}$$

The angular diameter distance $D_{\rm A}$ is the distance that relates the observed (small) angular size Θ measured by a telescope to the proper size R of an object, or

$$\Theta = \frac{R}{D_{\rm A}} \quad . \tag{7.2}$$

Because the ratio of flux to the solid angle is essentially the intensity, the $(1+z)^{-4}$ redshift-dependence of the intensity is reflected in these distance measures by

$$D_{\rm L} = (1+z)^2\, D_{\rm A} \quad . \tag{7.3}$$

Both distance measures are strong functions of world model, but this relationship between them—known sometimes as the "Etherington relation" (after Etherington 1933, who showed that the result is valid in arbitrary spacetimes)—depends only on conservation of phase-space density of photons (transparency) and Lorentz invariance. Fortunately, for some fortuitous types of objects, these distances can be measured nearly independently.

A test of this type for transparency has been proposed and carried out previously (Bassett & Kunz 2004a,b), with luminosity distances from SNeIa and angular diameter distances estimated from FRIIb radio galaxies, compact radio sources, and x-ray clusters (Uzan et al. 2004; Jackson 2008). The results were imprecise because there are many astrophysical uncertainties in the proper diameter estimates of these exceedingly complex astrophysical sources.

In the contemporary adiabatic cosmological standard model, there is a feature in the dark-matter auto-correlation function (or the power spectrum) corresponding to the communication of density perturbations by acoustic modes during the period in which radiation dominates (Peebles & Yu 1970; Eisenstein et al. 2005). This feature has a low amplitude in present-day structure (that is, the distribution of galaxies), but because it evolves little in comoving coordinates, it provides a "standard ruler" for direct measurement of the expansion history. A measurement of the baryon acoustic feature (BAF) in a population of galaxies at a particular redshift provides a combined measure of the angular diameter distance to that redshift (from the transverse size of the feature) and the Hubble Constant or expansion rate at that redshift (from the line-of-sight size of the feature). As we discuss below, as signal-to-noise improves, the BAF can be used to measure the angular diameter independently of the local Hubble rate. Most importantly, because the BAF arises from extremely simple physics in the early Universe when the growth of structure is linear and electromagnetic interactions dominate, the BAF measures the angular diameter distance with far fewer assumptions than any method based on complex astrophysical sources in the highly non-linear regime.

At the same time, SNeIa have been found to be standard — or really "standardizable" — candles, which can be used to make an independent direct measurement of the expansion history (Baade 1938; Tammann 1979; Colgate 1979; Riess et al. 1998; Perlmutter et al. 1999). Up to an over-all scale and some uncertainties about the intrinsic spectra and variability among SNeIa, a collection of SNeIa measure the luminosity distance.

Given overall scale uncertainties, the most robust test of global cosmic transparency that can be constructed from these two distance indicators is a measurement of the ratio of the distances to two redshifts z_1 and z_2. That is, transparency requires

$$\frac{D_{\rm L}(z_2)}{D_{\rm L}(z_1)} = \frac{[1+z_2]^2}{[1+z_1]^2}\,\frac{D_{\rm A}(z_2)}{D_{\rm A}(z_1)} \quad . \tag{7.4}$$

This expression cancels out overall scale issues and is independent of world model. We perform a very conservative variant of this test below, where we measure the left-hand side with SNeIa and the right-hand side with the BAF, marginalizing over a broad range of world models.

The tests presented here are not precise, simply because at the present day BAF measurements are in their infancy, and we make use of no cosmological data other than the BAF and SNeIa. As we discuss below, when these measurements are made at higher redshifts and with higher precisions, our limits on transparency and Lorentz invariance will improve by orders of magnitude. Eventually they may be limited not by the data quality but by the cosmic variance limit on the BAF measurement itself (Seo & Eisenstein 2007).

7.2 Data, Procedure, and Results

In surveys to date, where the BAF is measured at low signal-to-noise, the optimal extraction of the signal best constrains not the angular diameter distance directly, but rather a hybrid distance $D_{\rm V}$

$$D_{\rm V} = \left[\frac{c\,z\,[1+z]^2\,D_{\rm A}^2}{H(z)}\right]^{1/3} \quad , \tag{7.5}$$

where $D_{\rm A}$ is the angular diameter distance and $H(z)$ is the Hubble expansion rate (velocity per unit distance) at redshift z (Eisenstein et al. 2005).

Using data from the Sloan Digital Sky Survey and the Two Degree Field Galaxy Redshift Survey, the power spectrum and BAF have now been measured in samples of massive, red galaxies at two different redshifts: $z = 0.20$ and $z = 0.35$. The measured BAF at each redshift z translates to a distance measure $D_{\rm V}(z)$. Accounting for covariances in the measurements at the two redshifts (which are not based on entirely independent data sets), the ratio of distances is $D_{\rm V}(0.35)/D_{\rm V}(0.20) = 1.812 \pm 0.060$ (68-percent confidence; Percival et al. 2007). .

We formed two samples of SNeIa data from a recent compilation (Davis et al. 2007). "Sample A" consists of all 7 SNeIa in the redshift range $0.15 < z < 0.25$ and "Sample B" consists of all 22 SNeIa in the redshift range $0.3 < z < 0.4$. We estimate the distance-modulus DM at

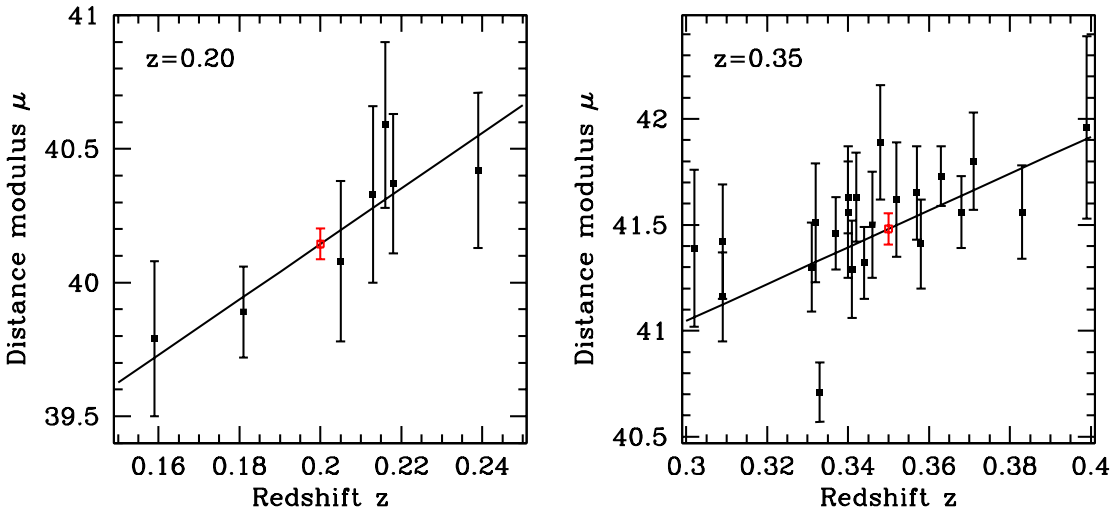

Figure 7.1: The distance-modulus–redshift relation. Filled black squares with uncertainty bars show the SNeIa data (from Davis et al. 2007) used in Samples A (left panel) and B (right panel). Open red squares show the distance moduli $DM(0.20) = 40.14 \pm 0.06$ and $DM(0.35) = 41.48 \pm 0.07$ (68-percent confidence) inferred from the fits to the data.

$z = 0.20$ and $z = 0.35$ by fitting a straight line to Samples A and B separately (Fig. 7.1), and obtain a distance-modulus difference

$$\Delta DM_{\rm obs} = DM_{\rm obs}(0.35) - DM_{\rm obs}(0.20) = [1.34 \pm 0.09]\ {\rm mag} \quad , \tag{7.6}$$

where we are indicating that this is an observed value, and might differ from the true value if there is opacity.

The distance modulus derived from the SNeIa is systematically affected by the presence of any intervening absorber. Let $\tau(z)$ denote the opacity between an observer at $z = 0$ and a source at redshift z due to such extinction effects. The flux received from this source is reduced by the factor $\mathrm{e}^{-\tau(z)}$. The inferred ("observed") luminosity distance differs from the "true" luminosity distance:

$$D^2_{\rm L_{obs}}(z) = D^2_{\rm L_{true}}(z)\, \mathrm{e}^{\tau(z)} \quad . \tag{7.7}$$

The ratio of the luminosity distances at two different redshifts z_1 and z_2 depends upon the factor $\mathrm{e}^{[\tau(0.35)-\tau(0.20)]/2}$. The inferred ("observed") distance modulus differs from the "true" distance modulus:

$$DM_{\rm obs}(z) = DM_{\rm true}(z) + [2.5\ \log e]\, \tau(z) \quad . \tag{7.8}$$

Taking differences of distance moduli at the two redshifts:

$$\Delta DM_{\rm obs} = \Delta DM_{\rm true} + [2.5\ \log e]\, \Delta\tau \quad , \tag{7.9}$$

where $\Delta\tau \equiv [\tau(z_2) - \tau(z_1)]$. If the distance indicator from the BAF is unaffected by the

absorption as we expect, then

$$\Delta\tau = \frac{\ln(10)}{2.5}\left[\Delta DM_{\rm obs} - 7.5\,\log\left(\frac{D_{\rm V}(z_2)}{D_{\rm V}(z_1)}\right) + 2.5\,\log\left(\frac{z_2\,[1+z_1]^2\,H(z_1)}{z_1\,[1+z_2]^2\,H(z_2)}\right)\right] \quad . \tag{7.10}$$

The above equation can be used to determine $\Delta\tau$ from $z = 0.35$ to $z = 0.20$ in light of the ratio of the distances $D_{\rm V}$ obtained from the BAF observations (hereafter B) and the difference in distance moduli obtained from the SNeIa observations (hereafter S) at these redshifts. However, the last term in the above equation makes the result cosmology-dependent. Therefore, we follow a Bayesian approach and assign posterior probabilities to 100 uniformly spaced values of $\Delta\tau \in$ [0,0.5] by marginalising over 100×100 ΛCDM cosmologies uniformly spaced in the $(\Omega_\Lambda, \Omega_M)$ plane with $\Omega_\Lambda \in$ [0,1] and $\Omega_M \in$ [0,1]. Thus,

$$P(\Delta\tau|S,B) = \int_{\Omega_\Lambda}\int_{\Omega_M} P(\Omega_\Lambda,\Omega_M|B)\,P(\Delta\tau,\Omega_\Lambda,\Omega_M|S)\,{\rm d}\Omega_M\,{\rm d}\Omega_\Lambda \quad , \tag{7.11}$$

where $P(\Omega_\Lambda,\Omega_M|B)$ and $P(\Delta\tau,\Omega_\Lambda,\Omega_M|S)$ are the posterior probabilities of the set of model parameters given B and S respectively. We assume that the uncertainties on B and S are Gaussian and calculate the likelihood of B and S for different sets of parameters in the $(\Delta\tau,\Omega_\Lambda,\Omega_M)$ space. Assuming flat priors on Ω_Λ and Ω_M in the ranges $0 < \Omega < 1$, and flat prior on $\Delta\tau$ in the range $0 < \Delta\tau < 0.5$, the posterior probabilities $P(\Omega_\Lambda,\Omega_M|B)$ and $P(\Delta\tau,\Omega_\Lambda,\Omega_M|S)$ are calculated from the likelihoods of the two datasets. Equation (7.11) yields the posterior for $\Delta\tau$, marginalized over all world models. Fig. 7.2 shows the posterior $P(\Delta\tau|S,B)$ for the difference in optical depths between redshifts 0.35 and 0.20 obtained from the procedure outlined above. The posterior peaks at 0 and yields $\Delta\tau < 0.13$ at 95-percent confidence. The result demonstrates the transparency of the Universe between these two redshifts, although not at high precision.

The abundance and absorption properties of absorbers can be constrained using the difference in optical depths measured above. Let $n(z)$ denote the comoving number density of absorbers, each with a proper cross-section $\sigma(z)$ at redshift z. The difference in optical depths between redshifts z_1 and z_2 is then given by

$$\Delta\tau = \int_{z_1}^{z_2} n(z)\,\sigma(z)\,D_{\rm H}\,\frac{(1+z)^2}{E(z)}\,{\rm d}z \quad , \tag{7.12}$$

where $D_{\rm H}$ is c/H_0 and

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{\Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda} \quad . \tag{7.13}$$

In detail, the output of this integral depends on world model. For the concordance model, Hubble Constant $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, and assuming $n(z)$ and $\sigma(z)$ to be independent of redshift, $\Delta\tau$ measured between redshifts 0.35 and 0.20 constrains $n\,\sigma < 2\times10^{-4}\,h$ Mpc^{-1}

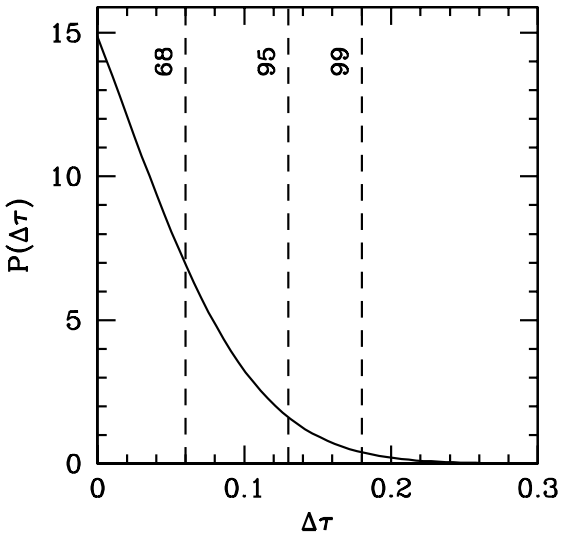

Figure 7.2: Posterior distribution of $\Delta\tau$ between $z = 0.35$ and $z = 0.20$ obtained from the Bayesian analysis described in Section 7.2. The 68, 95 and 99-percent confidence upper limits are indicated by the corresponding dashed lines.

at 95-percent confidence.

A naive calculation of $\Delta\tau$ using Equation (7.10) for the concordance ΛCDM model ($\Omega_M = 0.258, \Omega_\Lambda = 0.742$) obtained from the analysis of the 5-year WMAP data (Dunkley et al. 2009), yields $\Delta\tau = -0.30 \pm 0.26$ at 95-percent confidence. This shows that there is a slight tension between the results of current measurements of the BAF and of the SNeIa under the currently accepted world model. More generally, a similar tension, i.e. a brightening of the SNe, between measurements of the cosmological parameters by using standard rulers and standard candles has been reported before (Bassett & Kunz 2004a,b; Percival et al. 2007; Lazkoz et al. 2008). SNe brightening is not impossible in models that involve axion-photon mixing (Bassett & Kunz 2004b) or chameleon-photon mixing (Burrage 2008) if the corresponding particles are abundantly produced during SNeIa explosions. However, a negative value of $\Delta\tau$ could also indicate the presence of a systematic bias in the distance measurements based upon the SNeIa brightness or the BAF, *e.g.*, overcorrection for extinction in the host galaxy of the SNeIa brightnesses or magnification bias in the SNeIa selection (Williams & Song 2004). Note that the prior, $\Delta\tau > 0$, improves the magnitude of the uncertainty on $\Delta\tau$ (from 0.26 to 0.13). The 95-percent confidence interval shrinks with the prior because we sample only from the rapidly falling tail of the posterior.

7.3 Future Constraints

In the future, the constraints from both the SNeIa and the BAF observations will improve in accuracy and will cover a wider redshift range. The Baryon Oscillation Spectroscopic Survey

(BOSS) is currently underway and plans to measure the BAF in luminous red galaxies at redshifts $z = 0.35$ and $z = 0.6$. The key improvements would be the larger redshift range and the power to resolve the BAF both in the line-of-sight direction (constrains H) and the transverse direction (constrains $D_{\rm A}$). This would remove the weak world-model dependence in our present analysis. The angular diameter distances to these redshifts would be measured to an accuracy of $\sim$ 1 percent (http://www.sdss3.org/). In parallel, the Supernovae Legacy Survey (SNLS), when complete, expects $\sim$ 700 SNeIa in the redshift range $0 < z < 1.7$ (Astier et al. 2006). The uncertainty on the estimate of the distance moduli to redshifts $z = 0.35$ and $z = 0.6$ will be roughly four times better with the increased numbers. Using the test of the duality relation described above, $\Delta\tau$ between $z = 0.35$ and $z = 0.6$ would be constrained to better than 0.07 (95-percent confidence), independent of the adopted cosmological model. The constraint on $n\,\sigma$ would become $n\,\sigma < 5.4 \times 10^{-5}\,h\,{\rm Mpc}^{-1}$.

BOSS will also use the Ly-α forest in the spectra of bright QSOs to measure the BAF at redshift $z \sim 2.5$ with an accuracy of ~ 1.5 percent. No current or planned SNeIa surveys expect to detect SNeIa at such a high redshift. However the highest redshift (~ 1.7) measurements of $D_{\rm L}$ from the SNLS could potentially be used in conjunction with the $D_{\rm A}$ measurement to get a constraint on the transparency of the Universe by marginalizing over different world models. Interestingly, there have been recent efforts to calibrate gamma-ray bursts (hereafter GRBs) as standard candles and to extend the Hubble diagram to higher redshifts (Lazkoz et al. 2008). The SNeIa at low redshift and the GRBs at high redshifts can provide a measurement of the difference between the DM between redshifts 0.35 and 2.5. We optimistically assume that the difference in the DM to these redshifts can be measured with an accuracy of ~ 0.1 similar to the one obtained from the analysis of SNeIa at $z = 0.2$ and $z = 0.35$ in Section 7.2. These measurements shall then constrain $\Delta\tau$ between redshifts 2.5 and 0.35 to an accuracy of 0.2 with 95 percent confidence. This translates into an accuracy on $n\,\sigma$ of $\sim 1.1 \times 10^{-5}\,h\,{\rm Mpc}^{-1}$.

In the optimistic future, the uncertainty on $D_{\rm L}(z)$ could, in the absence of damaging systematics, diminish arbitrarily as the number of observed SNeIa grows. However, the precision of any BAF measurement is limited by sample variance (the number of independent wavelengths of a given fluctuation that can fit in the finite survey volume is limited), even when the uncertainty caused by incomplete sampling of the density field (shot noise) is negligible (Seo & Eisenstein 2007). The sample variance error goes down with the square root of the volume of the survey. To calculate a representative limit, we consider an optimistic all-sky survey covering the redshift range $2.45 < z < 2.55$. Such a survey can be used to determine $D_{\rm A}(z = 2.5)$ to a fractional accuracy of ~ 0.004 (95-percent confidence). This will ultimately constrain the optical depth to redshift $z = 2.5$ to $\tau < 0.008$ and hypothetical absorbers to $n\,\sigma < 4 \times 10^{-7}\,h\,{\rm Mpc}^{-1}$.

7.4 Discussion

We have advocated and analyzed the expected future performance of a simple Tolman test or test of the Etherington relation, that is, that the luminosity distance is larger than the angular diameter distance by two powers $(1 + z)$, using type Ia supernovae to measure the luminosity distance and the baryon acoustic feature to measure the angular diameter distance. We have shown that this test will eventually provide very precise measurements of the conservation of photon phase-space density.

We performed the test with the limited data available at the present day. We used only the ratio of distances at redshifts of $z = 0.20$ and 0.35 to remove uncertainties about the overall scale. We find consistency with a Lorentz-invariant, transparent Universe. Our results are consistent with all other measures of transparency to date. This is in part because they are not extremely precise. Our Tolman test also assumes that the measurements of the SNeIa and of the BAF are not affected by systematic biases with magnitudes that are a significant fraction of the magnitudes of the uncertainties. Our test is limited by the precision of the BAF measurement and the redshift range over which it has been measured. As we have shown, experiments planned and underway will increase the redshift range and improve the overall precision by an order of magnitude.

The most precise transparency measurements at visible wavelengths today are statistical angular difference measurements, which can only constrain attenuation correlated with specific types of absorbing structures in the Universe (*e.g.*, MgII absorbers, Ménard et al. 2008; clusters of galaxies, Bovy et al. 2008). The simple Tolman test performed here limits the full, unclustered, line-of-sight attenuation between two redshifts.

The technique used in this paper provides a test of transparency that is not very sensitive to astrophysical assumptions, both because the BAF has a straightforward origin during an epoch in which growth of structure is linear and the dominant physics is well understood, and because there is no significant "evolution" with cosmic time for which we must account. This is in contrast to other methods for measuring angular diameter distances and brightnesses, where there are no precisely "standard" rulers, and evolution is dramatic with redshift. On the other hand, the ultimate precision of any test of this type may come from the finite comoving volume in the observable Universe. Cosmic variance will dominate the BAF error budget eventually.

The SNeIa samples have been corrected as best as possible for line-of-sight extinction by fitting an empirical correlation of extinction with a change in color. However, there are a few problems with this approach. First, this approach cannot correct for "gray" dust (Aguirre 1999). Second, this approach can also not correct for a monopole component; it only corrects for components that show variations around the mean level. Third, these corrections will be wrong or fooled if there are intrinsic relationships between color and luminosity for SNeIa. Fourth, the empirical corrections found by these projects tend to be odd in the context of what is expected from the reddening and attenuation by dust in the Milky Way (Jöeveer 1982; Conley et al. 2007; Ellis et al. 2008; Nobili & Goobar 2008). The Tolman test is sensitive to any kind of

absorber and makes no assumptions about the wavelength-dependence or fluctuations of the opacity. Given that the SNeIa results have been corrected for a color–brightness relation, the test presented here looks at the mean opacity towards SNeIa of the fiducial color to which the compiled SNeIa have been corrected.

The best-fit value of $\Delta\tau$ obtained from our analysis is negative, i.e. SNeIa are brighter than expected from the angular diameter distance measurements using the BAF. A conversion of dark sector particles into photons could provide a physical explanation for this result. However, a systematic bias in either the SNeIa or BAF experiments cannot be ruled out and the Tolman test is a useful tool to identify such biases.

At present, because the differences among competitive world models are not large over the redshift range $0.20 < z < 0.35$, our test is not yet sensitive enough to rule out extreme axion or "gray" dust models that reconcile SNeIa results with an Einstein-de Sitter Universe by using effective opacity to adjust the inferred redshift–luminosity-distance relation. However, these models will all be severely constrained within the next few years (see also Corasaniti 2006).

Chapter 8

Summary

It is now a well established fact that most of the matter in the Universe is dynamically cold, collisionless and dark. It forms an ever-changing cosmic web of gravitationally bound structures called haloes. Galaxies form and evolve in these dark matter haloes and are shaped by various astrophysical processes which depend upon the properties of the dark matter halo they reside in. A precise knowledge of the connection between galaxies and their dark matter haloes is therefore crucial to improve our understanding of the physics of galaxy formation. In this thesis, we have established scaling relations between the mass of dark matter haloes and the properties of galaxies that reside at their centre.

Satellite galaxies are excellent tracers of the dark matter haloes around their central galaxies. The kinematics of these satellite galaxies reflect the depth of the potential well in which they orbit. Therefore, they can be used to measure the mass of the halo of central galaxies. A large number of satellite galaxies is required to precisely measure the kinematics in individual haloes, a condition which is easily met in cluster-sized haloes, but is rarely satisfied in haloes of low mass. Under the assumption that central galaxies with similar properties reside in similar mass haloes, one can stack central galaxies with similar properties, such as luminosity or stellar mass, and combine the velocity information of their satellites that will allow a precise measurement of the kinematics of satellite galaxies. However, galaxy formation is a stochastic process and one expects a scatter in the relation between central galaxy properties and halo masses. This implies that the stacking procedure results in combining the kinematics of satellite galaxies in a wide range of halo masses. This complicates the interpretation of the kinematics signal.

Most studies that use the kinematics of satellite galaxies to probe the halo masses have made the simplified assumption that the scatter in the halo masses of stacked centrals is negligible. In this thesis, we have demonstrated that the inference of the halo masses from the kinematics of satellite galaxies faces a degeneracy problem. The average relation between the halo mass of central galaxies and their properties cannot be inferred without the knowledge of the scatter in this relation. We have presented a novel method that can break this degeneracy. The method involves the measurement of the kinematics of satellite galaxies using two different weighting schemes: the satellite-weighting scheme and the host-weighting scheme. The ratio of the mea-

surements in these two different schemes is sensitive to the scatter in the masses of the stacked haloes. Therefore, a simultaneous modelling of the kinematics obtained with these two weighting schemes can be used to measure both the average and the scatter in the scaling relation between halo mass and the property of the central galaxy used for stacking.

The interpretation of the kinematics of stacked centrals is also complicated as a result of various selection effects which can bias the final determination of the halo masses. To understand these biases carefully, a realistic mock catalogue of galaxies was constructed and analysed. We compared the strict isolation criteria that have been abundantly used in the literature to select central and satellite galaxies with our adaptive iterative criteria. The comparison shows that the strict isolation criteria result in a preferential selection of low mass haloes. Another selection bias that has been previously neglected is due to the fact that the kinematics of satellite galaxies is always averaged over those haloes that host at least one satellite. This bias preferentially misses low mass haloes. This implies that the kinematic studies carried out previously have selected a sample which is not representative for the entire population of galaxies. The scaling relations derived so far are, at best only valid for the sample of the isolated galaxies. We have improved this situation and presented an analytical model which accounts for all these selection effects. Tests using the mock catalogue have shown that the method proposed and used by us in this thesis can be used to identify a representative population of central galaxies from a redshift survey, quantify the kinematics of their satellites and use this kinematic information to reliably infer the scaling relation between halo mass and central galaxy properties.

This method was then applied to data from the SDSS galaxy catalogue. The velocity dispersion-luminosity relation for central galaxies was measured from the data and used to infer the mean and the scatter of the halo mass−luminosity relation of central galaxies (MLR). The results show that brighter central galaxies on average reside in more massive haloes and that the scatter in halo masses is an increasing function of central galaxy luminosity. The investigation of the colour dependence of the MLR showed that at fixed luminosity, red galaxies on average occupy haloes that are more massive than their blue counterparts. A similar method was also applied to infer the halo mass−stellar mass relationship of central galaxies (MSR) and its colour dependence. Central galaxies that have more stellar mass on average reside in more massive haloes and the scatter in halo masses increases as a function of the stellar mass too. The analysis of satellite kinematics around centrals separated by colour shows that at fixed stellar mass, the red and the blue centrals on average occupy similar mass haloes.

The existence of scatter in halo masses at fixed central galaxy property implies that haloes of equal masses harbour central galaxies with a scatter in their properties. To investigate the origin of this stochasticity in galaxy formation, a semi-analytical model was analysed. Haloes of similar masses show a large scatter in their formation times. Our analysis shows that the residuals around the average formation time−halo mass relation positively correlate with the residuals around the halo mass−stellar mass relationship predicted by the semi-analytical model. This implies that haloes that form early, on average, host central galaxies which contain more stel-

lar mass. Thus, the scatter in the merger histories of haloes is a plausible explanation for the stochasticity of galaxy formation. The dependence of this statement on the particulars of the semi-analytical model are under investigation and a subject of future work.

The transparency of the Universe is also crucial to understand the galaxy-dark matter connection as the observed properties of galaxies can be biased if the Universe is significantly opaque. Transparency can be affected by intergalactic dust or interactions between photons and the dark sector. Such effects cause a deviation from the Etherington relation which relates the distances measured using standard candles to the distances measured using standard rulers at a particular redshift. A test of this relation was carried out by using the currently available observations of these distance measures to obtain a quantitative measure of the transparency of the Universe. With the limited amount of data available, we find consistency with a transparent Universe between redshifts 0.2 and 0.35. We analyzed the expected future performance of this test and showed that as better distance measurements covering a wider range in redshift become available, the test can provide a very precise measurement of the transparency of the Universe. Such precise measurements ultimately can also be used to limit the cross-section of interactions of the photons with the dark sector.

8.1 Future possibilities

8.1.1 Properties of Satellite Galaxies

In this thesis, we have used satellite galaxies to probe the connection between the properties of their central galaxies and dark matter haloes. However, we did not investigate various properties of satellite galaxies themselves. In the hierarchical structure formation scenario, satellite galaxies are interesting in their own right. The satellite galaxies may be affected due to various physical process, such as dynamical friction, effects due to tidal fields, ram pressure stripping, harrassment and strangulation (see e.g., Chandrasekhar 1943; Gunn & Gott 1972; Farouki & Shapiro 1980, 1981; Larson et al. 1980; Byrd & Valtonen 1990). It would be interesting to investigate the effects that these processes have on the properties of satellite galaxies.

The satellite galaxies selected in our samples can be used to study the abundance and radial distribution of satellites in haloes. In Chapter 3, we have shown that our selection criteria is able to recover the projected number density distribution of satellite galaxies around bright centrals quite accurately. An analysis of how these number density distributions depend on the colour of the satellite galaxies and the luminosities/stellar masses of the centrals can be carried out with the help of our sample. Very preliminary analyses indicate that at fixed luminosity of centrals, the number density distribution of blue satellites may be less concentrated than that of the red satellite galaxies. However, the effects of fiber collisions and stacking biases need to be carefully checked before interpreting these distributions. Our sample of satellites can also be used to analyse the dependence of the blue satellite fraction on the colour and luminosity/stellar mass of the centrals.

In this thesis, we have always focussed on the aperture averaged velocity dispersions. However, it would also be interesting to study the radial decline of the velocity dispersions and its dependence on the satellite-colour. Other interesting studies such as the measurements of the higher order moments of the velocity distribution and their colour dependence can also be carried out to investigate the orbital properties of satellite galaxies.

8.1.2 Shapes of Dark Matter Haloes

The existence of triaxial dark matter haloes is a firm prediction of the ΛCDM theory. These triaxial haloes should appear flattened when seen in projection. If observed, the flattening of haloes at large radii has a potential to discriminate between alternative theories of gravity such as MOND because the potentials at large radii predicted by such theories are isotropic far away from the stellar content of galaxies. Statistical measurements of dark matter halo shapes often rely on the statistical alignment of the dark matter halo axis and the ellipticity of the central galaxies that reside in them. There have been recent claims of the detection of ellipticities of dark matter haloes inferred from the modelling of the azimuthal dependence of the weak lensing signal around galaxies that are stacked by aligning their major axes, (see e.g., Hirata et al. 2004; Parker et al. 2007). However, there exist several sources of systematics which can contaminate this signal and lead to a spurious detection (Mandelbaum et al. 2006a). This makes it important to have alternative confirmations of these results. It would be interesting to study the azimuthal dependence of the velocity dispersions around central galaxies stacked by their properties and aligned along their major axis and its interpretation.

8.1.3 Redshift Evolution of the Halo Occupation Distributions

The halo occupation distribution (HOD) is the end result of the complex baryonic physics involved in the formation of galaxies. The halo mass-to-light ratio and the halo mass-to-stellar mass ratio obtained from the HOD quantify the efficiency of dark matter haloes to turn the baryons into stars (see e.g., Yang et al. 2003). The halo occupation distribution of galaxies is expected to evolve with redshift. The time evolution of the HOD can reflect interesting changes in the physics that takes place in dark matter haloes of different masses. Therefore, it is important to obtain observational constraints on such evolution. Such constraints have the ability to uncover new aspects in the physics of galaxy formation.

The satellite kinematics analysis presented in this thesis can be applied to high redshift datasets such as that provided by the DEEP2 redshift survey. Conroy et al. (2007) have already used the kinematics of satellite galaxies to derive the mass-to-light ratio (and the mass-to-stellar mass ratio) of central galaxies in SDSS and DEEP2. They use these results to constrain the evolution of the mass-to-light ratios. In light of the fact that their criteria were tuned to only select isolated central galaxies and their method to analyse the kinematics of satellites may have been systematically biased due to simplistic assumptions (see Section 5.5.2), it would be

certainly interesting to revisit some of their conclusions on the redshift evolution of the mass-to-light ratio. We are currently carrying out such an analysis. Preliminary investigation seems to indicate that at high redshift the small number statistics will be a big problem. It may be possibly to overcome this problem by an appropriate tuning of the parameters of the selection criteria and/or by using flux limited samples.

I would like to conclude by pointing out the quotation from the Brihadaranyaka Upanishad that was used to open this thesis. It is perfectly possible that dark components in our current theory are figments of our imagination and a result of our ignorance about nature. The spirit of curiosity and a constant questioning of our beliefs should always be kept alive to take us from ignorance to the Truth, and from darkness towards Light.

Acknowledgements

I am greatly indebted to my supervisor Frank van den Bosch for being extremely supportive and helpful right through the first day I arrived in Heidelberg. He has been a constant source of energy, ideas and an inspiration for hard work. I am thankful to Hans-Walter Rix for his support, encouragement and guidance during the phase of my PhD. I am extremely grateful to him, to Thomas Henning and the entire staff at MPIA for a wonderful research atmosphere through which I have learned a lot just by interacting with people around me. I am also thankful to all my collaborators for their help in the research work presented in this thesis. Many thanks to David Hogg for excellent discussions during the work on transparency.

It is a pleasure being part of the theory group at the MPIA. It was great to have Ramin Skibba, Andrea Macciò, Kris Blindert and Xi Kang available to answer any questions I had. I enjoyed a lot having particle physics discussions with Suchita Kulkarni. My heartfelt thanks to Anna Pasquali for her generous affection and support. I am also thankful to Svitlana Zhukovska for a careful reading of the introductory and the concluding chapters of this manuscript and to Sebastian Jester for help in translating the abstract into German. I hardly had any astronomy courses before I came to Heidelberg. I thoroughly enjoyed the lecture courses that I took during the time of my PhD and I am thankful to all the lecturers for their efforts. In particular, it was a great experience to learn from Matthias Bartelmann, Hans-Walter and Eric Bell. I must also thank Christian Fendt for his wonderful skills in managing the IMPRS and being always ready for help. I am also thankful to the administration department in the institute, especially Frau Apfel, Frau Seifert and Frau Schleich for always being helpful and making administrative tasks easy.

I had a great partner in Marcello from whom I have learnt a lot both about work and life. I shall always remember and cherish our "back to the basics" discussions on the whiteboard. I appreciate his friendliness, care and support during the not-so-good phases of my time here. I could not have had better officemates than Cassie and Marcello and virtual officemates in Ros and Christian, who maintained constant giggles and laughter in the office and cheered me up. I am also thankful to all the IMPRS guys for a wonderful time during IMPRS seminars, outings, barbeques and Christmas parties.

I am thankful to Osman Oezalp for providing me an accomodation in Kohlhof during the first two years of my PhD. I thoroughly enjoyed my stay in the woods. Many thanks also to my neighbour Frau Solweig for her support and grandmotherly love. A great thanks goes to Kris

Blindert for allowing me to use her guest room. I learnt a great deal of culinary skills from her. It was wonderful being friends with Pepper, who is a thorough entertainment package. He has always surprised me with his smartness.

Yaroslav has been a great companion and teacher for me. I have greatly benefitted from his zeal and enthusiasm for science. I hope we finish writing at least one chapter of our book on fundamental problems. I am sorry that even after his best efforts I could not master the skills of table tennis that he tried so hard to teach me. His constant support and motivation has been precious. I will always cherish our friendship.

I really cannot imagine myself here without all the love, hard work and efforts of my family. Aai and Baba have been my strongest pillars of support. Maushi, Aaji, Maai, Bhaiya, Chamu and my relatives in Dehugaon and Mushet, you all are an inseparable part of me. Mrinalini and Madhu, your love and support has been invaluable. A hearty thanks to my in-laws for their love and belief in me. I am also grateful to Gau, Anu, Aru, Anji, Mani, Jui, Deepu, Sweety, Jyoti and Bharu for the short but sweet time we spent together.

Life can really spring surprises in unexpectedly short amounts of time. Had it not been for the IMPRS Christmas party of 2007, I would never have met my sweetheart, Anupreeta. I feel really lucky to have a loving, caring and supportive $ardhangini$ in her. You make my life meaningful, Nupur!

Bibliography

Abazajian, K., Zheng, Z., Zehavi, I., et al. 2005, ApJ, 625, 613

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2006, ApJS, 162, 38

Aguirre, A. 1999, ApJ, 525, 583

Aguirre, A. & Haiman, Z. 2000, ApJ, 532, 28

Aharonian, F., Akhperjanian, A. G., Barres de Almeida, U., et al. 2008, Physical Review Letters, 101, 170402

Akerib, D. S., Alvaro-Dean, J., Armel, M. S., et al. 2003, PRD, 68, 082002

Akerib, D. S., Alvaro-Dean, J., Armel-Funkhouser, M. S., et al. 2004, Physical Review Letters, 93, 211301

Albert, J., Aliu, E., Anderhub, H., et al. 2008, Physics Letters B, 668, 253

Amelino-Camelia, G., Ellis, J., Mavromatos, N. E., Nanopoulos, D. V., & Sarkar, S. 1998, Nature, 393, 763

Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31

Baade, W. 1938, ApJ, 88, 285

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

Barnes, J. & Efstathiou, G. 1987, ApJ, 319, 575

Bassett, B. A. & Kunz, M. 2004a, ApJ, 607, 661

Bassett, B. A. & Kunz, M. 2004b, PRD, 69, 101305

Baugh, C. M., Cole, S., Frenk, C. S., & Lacey, C. G. 1998, ApJ, 498, 504

Becker, M. R., McKay, T. A., Koester, B., et al. 2007, ApJ, 669, 905

Bekenstein, J. D. 2004, PRD, 70, 083509

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Benoit, A., Bergé, L., Broniatowski, A., et al. 2002, Physics Letters B, 545, 43

Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., & Frenk, C. S. 2002, MNRAS, 333, 156

Berlind, A. A. & Weinberg, D. H. 2002, ApJ, 575, 587

Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, ApJ, 593, 1

Biller, S. D., Breslin, A. C., Buckley, J., et al. 1999, Physical Review Letters, 83, 2108

Blanton, M. R., Brinkmann, J., Csabai, I., et al. 2003a, AJ, 125, 2348

Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003b, ApJ, 592, 819

Blanton, M. R., Schlegel, D. J., Strauss, M. A., et al. 2005, AJ, 129, 2562

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517

Borch, A., Meisenheimer, K., Bell, E. F., et al. 2006, A&A, 453, 869

Bosma, A. 1978, PhD thesis, PhD Thesis, Groningen Univ., (1978)

Bovy, J., Hogg, D. W., & Moustakas, J. 2008, ApJ, 688, 198

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Brainerd, T. G. & Specian, M. A. 2003, ApJL, 593, L7

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

Bullock, J. S., Wechsler, R. H., & Somerville, R. S. 2002, MNRAS, 329, 246

Burrage, C. 2008, PRD, 77, 043009

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

Cacciato, M., van den Bosch, F. C., More, S., et al. 2009, MNRAS, 394, 929

Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 1997, ApJ, 478, 462

Carlberg, R. G., Yee, H. K. C., Ellingson, E., et al. 1996, ApJ, 462, 32

Chabrier, G. 2003, PASP, 115, 763

Chandrasekhar, S. 1943, ApJ, 97, 255

Chen, J. 2007, ArXiv e-prints (arXiv:0712.0003)

Chen, J. 2008, A&A, 484, 347

Clowe, D., Gonzalez, A., & Markevitch, M. 2004, ApJ, 604, 596

Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. 1994, MNRAS, 271, 781

Cole, S. & Lacey, C. 1996, MNRAS, 281, 716

Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168

Colgate, S. A. 1979, ApJ, 232, 404

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039

Conley, A., Carlberg, R. G., Guy, J., et al. 2007, ApJL, 664, L13

Conroy, C., Newman, J. A., Davis, M., et al. 2005, ApJ, 635, 982

Conroy, C., Prada, F., Newman, J. A., et al. 2007, ApJ, 654, 153

Cooray, A. 2006, MNRAS, 365, 842

Cooray, A. & Sheth, R. 2002, Phys. Rept., 372, 1

Corasaniti, P. S. 2006, MNRAS, 372, 191

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

Csáki, C., Kaloper, N., & Terning, J. 2002, Physical Review Letters, 88, 161302

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

Davis, T. M., Mörtsell, E., Sollerman, J., et al. 2007, ApJ, 666, 716

De Lucia, G. & Blaizot, J. 2007, MNRAS, 375, 2

dell'Antonio, I. P. & Tyson, J. A. 1996, ApJL, 473, L17+

Diaferio, A. & Geller, M. J. 1996, ApJ, 467, 19

Diemand, J., Moore, B., & Stadel, J. 2004, MNRAS, 352, 535

Dolgov, A. D. 2004, ArXiv High Energy Physics - Phenomenology e-prints

Drory, N., Bender, R., Feulner, G., et al. 2003, ApJ, 595, 698

Drory, N., Salvato, M., Gabasch, A., et al. 2005, ApJL, 619, L131

Dunkley, J., Komatsu, E., Nolta, M. R., et al. 2009, ApJS, 180, 306

Eisenstein, D. J. & Hu, W. 1999, ApJ, 511, 5

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Eke, V. R., Navarro, J. F., & Frenk, C. S. 1998, ApJ, 503, 569

Ellis, J., Hagelin, J. S., Nanopoulos, D. V., Olive, K., & Srednicki, M. 1984, Nuclear Physics B, 238, 453

Ellis, J., Mavromatos, N. E., Nanopoulos, D. V., Sakharov, A. S., & Sarkisyan, E. K. G. 2006, Astroparticle Physics, 25, 402

Ellis, R. S., Sullivan, M., Nugent, P. E., et al. 2008, ApJ, 674, 51

Erickson, L. K., Gottesman, S. T., & Hunter, Jr., J. H. 1987, Nature, 325, 779

Etherington, I. M. H. 1933, Philosophical Magazine, 15, 761

Evrard, A. E., Summers, F. J., & Davis, M. 1994, ApJ, 422, 11

Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, ApJ, 665, 265

Farouki, R. & Shapiro, S. L. 1980, ApJ, 241, 928

Farouki, R. & Shapiro, S. L. 1981, ApJ, 243, 32

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

Frenk, C. S., White, S. D. M., Bode, P., et al. 1999, ApJ, 525, 554

Gabasch, A., Bender, R., Seitz, S., et al. 2004, A&A, 421, 41

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, MNRAS, 362, 41

Gambini, R. & Pullin, J. 1999, PRD, 59, 124021

Gao, L., Springel, V., & White, S. D. M. 2005, MNRAS, 363, L66

Gao, L. & White, S. D. M. 2007, MNRAS, 377, L5

Giovanelli, R., Haynes, M. P., Herter, T., et al. 1997, AJ, 113, 53

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

Guzik, J. & Seljak, U. 2002, MNRAS, 335, 311

Häring, N. & Rix, H.-W. 2004, ApJL, 604, L89

Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, MNRAS, 367, 1039

Hauser, M. G. & Dwek, E. 2001, ARA&A, 39, 249

Hirata, C. M., Mandelbaum, R., Seljak, U., et al. 2004, MNRAS, 353, 529

Hogg, D. W. 1999, ArXiv Astrophysics e-prints (arXiv:astro-ph/9905116)

Hu, W. & Kravtsov, A. V. 2003, ApJ, 584, 702

Hudson, M. J., Gwyn, S. D. J., Dahle, H., & Kaiser, N. 1998, ApJ, 503, 531

Jackson, J. C. 2008, MNRAS, 390, L1

Jing, Y. P. & Suto, Y. 1998, ApJL, 494, L5+

Jing, Y. P., Suto, Y., & Mo, H. J. 2007, ApJ, 657, 664

Jöeveer, M. 1982, Astrophysics, 18, 328

Katz, N., Hernquist, L., & Weinberg, D. H. 1992, ApJL, 399, L109

Katz, N., Weinberg, D. H., & Hernquist, L. 1996, ApJS, 105, 19

Kauffmann, G. 1996, MNRAS, 281, 487

Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795

Kauffmann, G. & White, S. D. M. 1993, MNRAS, 261, 921

Kennicutt, Jr., R. C. 1998, ApJ, 498, 541

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

Klypin, A. & Prada, F. 2009, ApJ, 690, 1488

Koester, B. P., McKay, T. A., Annis, J., et al. 2007, ApJ, 660, 239

Kowalski, M., Rubin, D., Aldering, G., et al. 2008, ApJ, 686, 749

Kravtsov, A. V. 1999, PhD thesis, AA(NEW MEXICO STATE UNIVERSITY)

Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35

Kroupa, P. 2001, MNRAS, 322, 231

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

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

Lazkoz, R., Nesseris, S., & Perivolaropoulos, L. 2008, Journal of Cosmology and Astro-Particle Physics, 7, 12

Lemson, G. & Kauffmann, G. 1999, MNRAS, 302, 111

Li, C., Kauffmann, G., Jing, Y. P., et al. 2006, MNRAS, 368, 21

Li, Y., Mo, H. J., & Gao, L. 2008, MNRAS, 389, 1419

Li, Y., Mo, H. J., van den Bosch, F. C., & Lin, W. P. 2007, MNRAS, 379, 689

Lubin, L. M. & Sandage, A. 2001, AJ, 122, 1084

Macciò, A. V., Dutton, A. A., van den Bosch, F. C., et al. 2007, MNRAS, 378, 55

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388

Madgwick, D. S., Hawkins, E., Lahav, O., et al. 2003, MNRAS, 344, 847

Mandelbaum, R., Hirata, C. M., Broderick, T., Seljak, U., & Brinkmann, J. 2006a, MNRAS, 370, 1008

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006b, MNRAS, 368, 715

Mather, J. C., Cheng, E. S., Cottingham, D. A., et al. 1994, ApJ, 420, 439

McGaugh, S. S. & de Blok, W. J. G. 1998, ApJ, 499, 66

McKay, T. A., Sheldon, E. S., Johnston, D., et al. 2002, ApJL, 571, L85

McNamara, B. R., Nulsen, P. E. J., Wise, M. W., et al. 2005, Nature, 433, 45

McNamara, B. R., Wise, M., Nulsen, P. E. J., et al. 2000, ApJL, 534, L135

Ménard, B., Nestor, D., Turnshek, D., et al. 2008, MNRAS, 385, 1053

Milgrom, M. 1983a, ApJ, 270, 371

Milgrom, M. 1983b, ApJ, 270, 384

Milgrom, M. 1983c, ApJ, 270, 365

Mirizzi, A., Raffelt, G. G., & Serpico, P. D. 2005, PRD, 72, 023501

More, S., Bovy, J., & Hogg, D. W. 2009a, ApJ, in press, arXiv:0810.5553

More, S., van den Bosch, F. C., & Cacciato, M. 2009b, MNRAS, 392, 917

More, S., van den Bosch, F. C., Cacciato, M., et al. 2009c, MNRAS, 392, 801

Mörtsell, E., Bergström, L., & Goobar, A. 2002, PRD, 66, 047702

Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2009, ArXiv e-prints (arXiv:0903.4682)

Mushotzky, R. F., Serlemitsos, P. J., Boldt, E. A., Holt, S. S., & Smith, B. W. 1978, ApJ, 225, 21

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1995, MNRAS, 275, 720

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Nobili, S. & Goobar, A. 2008, A&A, 487, 19

Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002a, MNRAS, 332, 827

Norberg, P., Cole, S., Baugh, C. M., et al. 2002b, MNRAS, 336, 907

Norberg, P., Frenk, C. S., & Cole, S. 2008, MNRAS, 383, 646

Pahre, M. A., Djorgovski, S. G., & de Carvalho, R. R. 1996, ApJL, 456, L79+

Panter, B., Heavens, A. F., & Jimenez, R. 2004, MNRAS, 355, 764

Parker, L. C., Hoekstra, H., Hudson, M. J., van Waerbeke, L., & Mellier, Y. 2007, ApJ, 669, 21

Peacock, J. A. & Smith, R. E. 2000, MNRAS, 318, 1144

Pearce, F. R., Jenkins, A., Frenk, C. S., et al. 1999, ApJL, 521, L99

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

Percival, W. J., Cole, S., Eisenstein, D. J., et al. 2007, MNRAS, 381, 1053

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Porciani, C., Dekel, A., & Hoffman, Y. 2002, MNRAS, 332, 325

Prada, F., Vitvitska, M., Klypin, A., et al. 2003, ApJ, 598, 260

Preskill, J., Wise, M. B., & Wilczek, F. 1983, Physics Letters B, 120, 127

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Roberts, M. S. & Whitehurst, R. N. 1975, ApJ, 201, 327

Rubin, V. C., Ford, Jr., W. K., Thonnard, N., & Burstein, D. 1982, ApJ, 261, 439

Rubin, V. C., Thonnard, N., & Ford, Jr., W. K. 1978, ApJL, 225, L107

Salpeter, E. E. 1955, ApJ, 121, 161

Sanders, R. H. 2003, MNRAS, 342, 901

Sanders, R. H. & McGaugh, S. S. 2002, ARA&A, 40, 263

Sanglard, V., Benoit, A., Bergé, L., et al. 2005, PRD, 71, 122002

Schaefer, B. E. 1999, Physical Review Letters, 82, 4964

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Springer-Verlag Berlin Heidelberg New York.)

Schneider, P. & Rix, H.-W. 1997, ApJ, 474, 25

Seo, H.-J. & Eisenstein, D. J. 2007, ApJ, 665, 14

Shostak, G. S. 1973, A&A, 24, 411

Sofue, Y. & Rubin, V. 2001, ARA&A, 39, 137

Somerville, R. S. & Primack, J. R. 1999, MNRAS, 310, 1087

Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

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

Srianand, R., Petitjean, P., & Ledoux, C. 2000, Nature, 408, 931

Sutherland, R. S. & Dopita, M. A. 1993, ApJS, 88, 253

Tabor, G. & Binney, J. 1993, MNRAS, 263, 323

Tammann, G. A. 1979, in ESO/ESA Workshop on Astronomical Uses of Space Telescope ed F. Macchetto, F. Pacini, & M. Tarenghi (Geneva: ESO), 329

Tasitsiomi, A., Kravtsov, A. V., Wechsler, R. H., & Primack, J. R. 2004, ApJ, 614, 533

Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, ApJ, 606, 702

Tinker, J. L., Norberg, P., Weinberg, D. H., & Warren, M. S. 2007, ApJ, 659, 877

Tolman, R. C. 1930, Proceedings of the National Academy of Science, 16, 511

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898

Tyson, J. A. 1987, in IAU Symposium, Vol. 117, Dark matter in the universe, ed. J. Kormendy & G. R. Knapp, 241–+

Uzan, J.-P., Aghanim, N., & Mellier, Y. 2004, PRD, 70, 083533

van den Bosch, F. C. 2002, MNRAS, 331, 98

van den Bosch, F. C., Norberg, P., Mo, H. J., & Yang, X. 2004, MNRAS, 352, 1302

van den Bosch, F. C., Yang, X., Mo, H. J., et al. 2007, MNRAS, 376, 841

Wang, Y., Yang, X., Mo, H. J., & van den Bosch, F. C. 2007, ApJ, 664, 608

Wang, Y., Yang, X., Mo, H. J., van den Bosch, F. C., & Chu, Y. 2004, MNRAS, 353, 287

Wang, Y., Yang, X., Mo, H. J., et al. 2008, ApJ, 687, 919

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

Wechsler, R. H., Zentner, A. R., Bullock, J. S., Kravtsov, A. V., & Allgood, B. 2006, ApJ, 652, 71

White, S. D. M. & Frenk, C. S. 1991, ApJ, 379, 52

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

Williams, L. L. R. & Song, J. 2004, MNRAS, 351, 1387

Wilson, G., Kaiser, N., Luppino, G. A., & Cowie, L. L. 2001, ApJ, 555, 572

Yang, X., Mo, H. J., Jing, Y. P., & van den Bosch, F. C. 2005a, MNRAS, 358, 217

Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057

Yang, X., Mo, H. J., & van den Bosch, F. C. 2008, ApJ, 676, 248

Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, ApJ, 695, 900

Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2007, ApJ, 671, 153

Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2005b, MNRAS, 362, 711

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zaritsky, D., Smith, R., Frenk, C., & White, S. D. M. 1993, ApJ, 405, 464

Zaritsky, D., Smith, R., Frenk, C., & White, S. D. M. 1997, ApJ, 478, 39

Zaritsky, D. & White, S. D. M. 1994, ApJ, 435, 599

Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172

Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004, ApJ, 608, 16

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, ApJ, 630, 1

Zhao, D. H., Mo, H. J., Jing, Y. P., & Börner, G. 2003, MNRAS, 339, 12

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791

Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760

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