

Maturing Satellite Kinematics into a Competitive Probe of the Galaxy-Halo Connection

Johannes U. Lange^{1,2*}, Frank C. van den Bosch¹, Andrew R. Zentner³
and Surhud More⁴

¹*Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06511, USA*

²*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA*

³*Department of Physics and Astronomy & Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA*

⁴*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes of Advanced Study, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa 277-8583, Japan*

Accepted xxx. Received xxx

ABSTRACT

The kinematics of satellite galaxies moving in a dark matter halo are a direct probe of its underlying gravitational potential. Thus, the phase-space distributions of satellites represent a powerful tool to determine the galaxy–halo connection from observations. By stacking the signal of a large number of satellite galaxies this potential can be unlocked even for haloes hosting a few satellites galaxies on average. In this work, we test the impact of various modelling assumptions on constraints derived from analysing satellite phase-space distributions in the non-linear, 1–halo regime. We discuss their potential to explain the discrepancy between average halo masses derived from satellite kinematics and gravitational lensing previously reported. Furthermore, we develop an updated, more robust analysis to extract constraints on the galaxy–halo relation from satellite properties in spectroscopic galaxy surveys such as the SDSS. We test the accuracy of this approach using a large number of realistic mock catalogues. Furthermore, we find that constraints derived from such an analysis are complimentary and competitive with respect to the commonly used galaxy clustering and galaxy–galaxy lensing observables.

Key words: methods: statistical – galaxies: kinematics and dynamics – galaxies: groups: general – cosmology: dark matter

1 INTRODUCTION

In 1933, Fritz Zwicky was one of the first astronomers to find evidence for the existence of dark matter. He based his conclusions on the observed velocity dispersion of galaxies in the Coma cluster. Because galaxy kinematics directly probe the underlying gravitational potential, Zwicky (1933) concluded that most of the matter inside Coma could be dark.

Today, the Λ cold dark matter (Λ CDM) paradigm is our current best cosmological model for the formation and evolution of cosmic structure. In this scenario, dark matter haloes provide the potential wells that enable the formation of galaxies. Hence, constraining the link between galaxies and dark matter haloes, which goes by the catch-all name ‘halo-occupation modelling’, can provide valuable insight regarding galaxy formation. Our current best constraints on this galaxy–dark matter connection come from

galaxy clustering (e.g., van den Bosch et al. 2007; Zehavi et al. 2011; Hearin & Watson 2013; Guo et al. 2015a,b, 2016; Zentner et al. 2016; Wibking et al. 2017; Campbell et al. 2017; Moster et al. 2017), galaxy–galaxy lensing (e.g., Sheldon et al. 2009a,b; Zu & Mandelbaum 2015, 2016; Mandelbaum et al. 2016; Sonnenfeld & Leauthaud 2017), and combinations thereof (e.g., Cacciato et al. 2009; Leauthaud et al. 2012; Cacciato et al. 2013; DES Collaboration et al. 2017). Another popular method used in halo occupation modelling is abundance matching, which postulates a tight, monotonic relation between halo mass and galaxy luminosity or stellar mass. Abundance matching can be applied to individual galaxies, in which case one talks of subhalo abundance matching (e.g., Vale & Ostriker 2004, 2006; Guo et al. 2010; Hearin & Watson 2013), or to galaxy groups (e.g., Yang et al. 2005, 2007, 2008, 2009; Weinmann et al. 2006), in which case one only uses host haloes.

Somewhat surprisingly, despite being the original method used by Zwicky to reveal the presence of dark mat-

* email: johannesulf.lange@yale.edu

ter, satellite kinematics has hitherto been little utilized in halo occupation modelling. Although various studies have used the kinematics of satellite galaxies as tracers of their dark matter potential wells (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; Conroy et al. 2007; Norberg et al. 2008; More et al. 2009b,a, 2011; Li et al. 2012; Wojtak & Mamon 2013), very few have done so within the framework of halo occupation modelling. One reason for this is that most previous studies applied highly restrictive isolation criteria to select central (or ‘host’) and satellite galaxies and, consequently, these samples can no longer be considered representative of the galaxy population as a whole. In addition, the strictness severely limited the total number of central-satellite pairs, and hence the signal-to-noise and dynamic range of the data. The goal of this paper is to rectify this situation and to promote satellite kinematics as a halo-occupation modelling tool on par with clustering and lensing.

Following pioneering efforts by Erickson et al. (1987) and Zaritsky et al. (1993, 1997), the first attempt to use satellite kinematics to infer the galaxy–halo connection from a large redshift survey was by McKay et al. (2002), who used the SDSS to study the velocity distribution of 1225 satellites around 618 central galaxy candidates. The main finding was that the average host halo masses scale roughly linearly with the luminosity of the central galaxies, confirming previous results from McKay et al. (2001) using galaxy–galaxy lensing. A similar study was performed by Prada et al. (2003) utilizing up to 2734 satellites in the SDSS. They found that the velocity dispersion of satellites decreases with the distance to the central candidate, in good agreement with predictions from dark matter simulations. They also confirmed the results of McKay et al. (2002) that dark matter halo mass increases with the luminosity of the central. Brainerd & Specian (2003) analysed the kinematics of 1556 satellites in the Two-Degree Field Galaxy Redshift Survey (2dFGRS). They mostly confirmed previous results by McKay et al. (2002) and additionally found that the dark matter halo mass seems to be independent of luminosity for spiral galaxies. Conroy et al. (2007) expanded the use of satellite kinematics to higher redshifts, $z \sim 1$, using ~ 1000 satellites in the DEEP2 Galaxy Redshift Survey (Davis et al. 2003). By comparing results from both DEEP2 and SDSS at lower redshifts they found the stellar-to-halo mass ratio to be mostly independent of stellar mass and redshift. Norberg et al. (2008) analysed the motions of 1003 satellites around 571 central candidates in 2dFGRS. By comparing their findings to the previous studies, they pointed out that most satellite kinematics studies thus far had produced quantitatively very discrepant results. They argued that these discrepancies arise from the use of different isolation criteria, and different estimators of the satellite velocity dispersion, and concluded that the interpretation of satellite kinematics data remains “questionable” unless one applies the same methods to a set of realistic mock galaxy catalogues. Finally, Wojtak & Mamon (2013) were the first to try and constrain the anisotropy of satellite orbits; by modelling both the line-of-sight velocities and projected distances of more than 10,000 satellite galaxies in the SDSS Data Release 7 (DR7 Abazajian et al. 2009), with respect to their centrals, they were able to simultaneously constrain the halo masses and the orbital anisotropy of the satellites (but see below).

A commonality among the satellite kinematic studies listed above is that they have been extremely conservative in their selection of centrals and satellites. In particular, centrals were selected to be significantly brighter than any other galaxy within their neighborhood (typically defined as a cylindrical region in redshift-space, see below), and satellites typically had to be at least 4 to 8 times fainter than their corresponding central. This not only severely limits the number of central-satellite pairs, it also introduces a selection bias against haloes with bright satellites. This in turn could introduce a significant bias in halo mass. These shortcomings were circumvented by van den Bosch et al. (2004), who introduced an iterative, adaptive method to select centrals and satellites. This drastically increased the number of central-satellite pairs, while simultaneously decreasing the fraction of interlopers (galaxies unassociated with the dark matter halo of the central) and increasing the dynamic range of the galaxy-halo connection probed.

Another important point made by van den Bosch et al. (2004) is the impact of mass-mixing: because there is no one-to-one relation between central luminosity or stellar mass and dark matter halo mass, stacking implies that one combines the kinematics of satellites orbiting in haloes of different masses. If this is not properly accounted for in the analysis, one can make large errors in the inferred (mean) masses, or, as in the case of Wojtak & Mamon (2013), the inferred orbital anisotropy (see also Becker et al. 2007; Norberg et al. 2008). More et al. (2009b) developed a method that accounts for this mass mixing when stacking central galaxies. In particular, they demonstrated that one can actually measure the amount of scatter in halo mass by comparing the satellite kinematics obtained using different weighting schemes.

One of the most comprehensive studies to date constraining the global galaxy-halo connection from satellite kinematics is that by More et al. (2011). This study is a continuation of the work by van den Bosch et al. (2004) and More et al. (2009a,b). Using data on ~ 6000 satellite galaxies selected from a volume-limited subsample of the SDSS DR4 (Adelman-McCarthy et al. 2006), covering the redshift range $0.02 \leq z \leq 0.071$, More et al. (2011) obtained constraints on the galaxy-dark matter connection that were qualitatively in good agreement with other investigations based on clustering and/or lensing. Furthermore, More et al. (2011) confirmed previous findings by Conroy et al. (2007) that red centrals live in more massive haloes than blue centrals of the same luminosity (also see Norberg et al. 2008, for splits by morphological type). They also found the same trend using stellar mass samples, a result that has since been verified by other studies using gravitational lensing (Velandier et al. 2014; Mandelbaum et al. 2016; Zu & Mandelbaum 2016). However, despite this success, on a more quantitative level, the results of More et al. (2011) imply average halo masses that are significantly different (by a factor of two to three) than those inferred from clustering and/or galaxy–galaxy lensing (see e.g., Dutton et al. 2010; Leauthaud et al. 2012; Mandelbaum et al. 2016). Such a level of disagreement is well beyond what is considered acceptable for numerous applications of the galaxy-halo connection [XXX REFS XXX], and is in dire need of an explanation.

This is the first paper in a series in which we seek to significantly improve upon previous studies of satellite kinematics, and to develop a methodology that is both

complementary to, and competitive with, other probes of the galaxy-dark matter connection, such as clustering and galaxy-galaxy lensing. In this first paper we identify potential shortcomings in the analysis by More et al. (2009a,b, 2011), which we take as the starting point of our investigation. We use detailed mock galaxy redshift surveys to compute a proper covariance matrix, to be used in the analysis, and to demonstrate the importance of accounting for redshift incompleteness, in particular fibre-collisions, something that was not accounted for in any previous satellite kinematic study with the exception of Wojtak & Mamon (2013). We also improve upon the treatment of interlopers, and use forward modelling to quantify small biases in the analytical model. Using realistic mock catalogues, we demonstrate that these biases can be properly accounted for, resulting in a method for analyzing satellite kinematics that yields unbiased constraints on the galaxy-dark matter connection that are both competitive with and complimentary to studies based on clustering or lensing. In Paper II (Lange et al., in prep.), we apply this improved method to the SDSS DR7, finding constraints on the galaxy-halo connection that are in good agreement with other independent studies, abating the tension with the analysis by More et al. (2011) eluded to above.

This paper is organized as follows. In §2 we start with a broad outline of the methodology. §3 describes our model for how galaxies occupy dark matter haloes, while §4 describes our algorithm for creating realistic mock SDSS-like catalogues. The observables that we use to constrain the galaxy-halo connection are detailed in §5 and a simplified analytical model for those observables, for a given galaxy-halo connection, is described in §6. §7 describes the full analysis procedure, which we apply to mock catalogues with known input parameters in §8. Finally, in §9 we summarize our findings.

Throughout this work we assume a flat Λ CDM cosmology with $\Omega_m = 0.307$, $\Omega_b = 0.048$, $\sigma_8 = 0.8228$, and $h = 0.678$. Here, $h = H_0/(100 \text{ km/s/Mpc})$ and H_0 is the Hubble constant (Planck Collaboration et al. 2014). All magnitudes are given in the AB magnitude system. When appropriate, quantities like comoving distance, halo mass or luminosity are scaled by h to make results independent of H_0 . Finally, throughout this paper we use r to denote 3D radii, and R for projected 2D radii.

2 METHODOLOGY

This section sketches the rough outline of our methodology. We first describe how to measure satellite kinematics from a galaxy redshift survey, and then detail how to use such data to constrain the galaxy-dark matter connection.

The first step in measuring satellite kinematics from a galaxy redshift survey is the selection of centrals and satellites. The standard method of selecting centrals is based on a cylindrical isolation criterion; i.e., a galaxy is identified as a central if it is the brightest galaxy in some specified, cylindrical volume whose symmetry axis falls along the line-of-sight. Satellite galaxies are defined as those galaxies that fall within a similar cylindrical volume, centred on the central galaxy, that meet a certain number of criteria. Ideally this selection of centrals and satellites is optimized to maximize completeness and minimize impurities (see §5.1 for details).

Using the centrals and satellites thus identified, the next step is to quantify the kinematics of satellite galaxies within the host haloes of their associated centrals. Since the typical number of satellite galaxies per central is small, except in nearby clusters, this requires stacking whereby one co-adds all central-satellite pairs for centrals specified by n galaxy properties $\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n$ (e.g., luminosity, color, size). The satellite kinematics are then specified by the line-of-sight velocity distribution (LOSVD) $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, where

$$\Delta v_z = c \frac{(z_{\text{sat}} - z_{\text{cen}})}{1 + z_{\text{cen}}}, \quad (1)$$

with z_{sat} and z_{cen} the observed redshifts of the satellite and central, and c the speed of light. The summary statistic that is most often used in the study of satellite kinematics is the satellite velocity dispersion, $\sigma_{\text{sat}}(\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, which characterizes the second moment of this LOSVD.

The main goal of the analysis of satellite kinematics envisioned in this paper, is to constrain the galaxy-dark matter connection, which can be quantified in terms of the conditional probability $P(\mathcal{H} | \mathcal{G}_1, \dots, \mathcal{G}_n)$. Here \mathcal{H} is the halo property of interest, which typically will be some measure of halo mass. Hence, the final step will be to translate $\sigma_{\text{sat}}(\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, or $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, into corresponding constraints $P(\mathcal{H} | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$. Ideally this is done using forward modelling based on mock galaxy redshift surveys. This process consists of the following steps:

- (i) Using Monte Carlo techniques, populate the dark matter halos in a dark-matter-only numerical simulation of a cosmological volume according to $P(\mathcal{H} | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$.
- (ii) Place a virtual observer in the cosmological volume, define an angular coordinate system with respect to that observer, and compute for each mock galaxy the corresponding right ascension and declination, as well as its redshift and apparent magnitude.
- (iii) Apply the survey mask and sample selection of the redshift survey, and mimic fibre collisions and other survey incompleteness effects.
- (iv) Apply the same central/satellite selection criteria as for the real data, and compute the corresponding $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$.
- (v) Compare to the data, and embed in a Bayesian inference engine to constrain the posterior distributions of the parameters that quantify $P(\mathcal{H} | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$.

This forward modelling approach has a number of advantages. First of all, it uses the full information from the LOSVD, $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, rather than just its second moment. Secondly, with forward modelling it is straightforward to accurately account for biases that may arise from details related to the selection of centrals and satellites, and for a variety of observational complications, such as fibre collisions.

Unfortunately, a full forward modelling approach as outlined above is currently computationally infeasible. Tests show that we need $\sim 10^8$ model evaluations in order to reliably constrain the posteriors of our model parameters that quantify the galaxy-dark matter connection (see §3). Furthermore, each model evaluation requires of order 100 mocks to suppress the realization noise from a single SDSS-size mock. The construction of a single mock, and the subsequent

computation of $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, is orders of magnitude too slow to allow for the construction of $\mathcal{Q}(10^{10})$ mocks, and we therefore have to rely on a less computationally-intensive model.

As in previous studies (e.g. van den Bosch et al. 2004; More et al. 2009a, 2011; Wojtak & Mamon 2013), we therefore use an analytical model which relies on the Jeans equations to predict the satellite kinematics as function of halo mass. We improve upon previous studies in a number of ways: (i) we use a proper covariance matrix for all the data, (ii) we account for fibre collisions and other survey selection effects, and (iii) we use a new method to correct for interlopers that does not rely on some ad-hoc assumption regarding the functional form of $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$. We also use detailed forward modelling as described above to gauge the reliability of the analytical model. In particular, we characterize its systematic bias, and use that to correct the model. Finally, once we have obtained the posterior distribution for the model, we construct a set of mocks for the best-fit model, and use forward modelling to verify that no systematic error has seeped into our approach. If necessary, we modify our treatment for the bias and iterate this entire procedure until convergence.

Our model based on the Jeans equations only predicts the second moment of the full LOSVD. Comparison with the data therefore requires that we compute $\sigma_{\text{sat}}(\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$ from the observed $P(\Delta v_z | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, which, as we discuss in §6.5, is hampered by a systematic bias. Although our method is codified to correct for this bias, an obvious downside is that our analytical model does not use all information that is available from the detailed *shape* of the LOSVD. However, we can access some of that information by using a clever method, first proposed by More et al. (2009b): by using two different weights for the central-satellite pairs (host-weighting and satellite-weighting, see §5.4), we obtain two different estimates of the second moment, $\sigma_{\text{sat}}(\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$. Their difference can be used to constrain higher-order moments of $P(\mathcal{H} | \mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$, such as the amount of scatter in the relation between halo mass and galaxy luminosity). In addition, we also use the set of mocks constructed from our best-fit model to check to what level the model is able to reproduce the full LOSVD.

3 GALAXY-HALO CONNECTION

When modelling satellite kinematics, and when constructing mock galaxy catalogues from dark-matter-only simulations, one needs to characterize the abundance, demographics, and phase-space distributions of central and satellite galaxies as function of halo mass. In this section we describe the model that we use to parametrize this ‘galaxy-halo connection’.

3.1 Galaxy Occupation Statistics

We model the galaxy occupation statistics using the conditional luminosity function (CLF; Yang et al. 2003), $\Phi(L|M)dL$, which describes the average number of galaxies with luminosities in the range $L \pm dL/2$ that reside in a dark matter halo of virial mass M . The CLF is composed of a central and a satellite part,

$$\Phi_{\text{tot}}(L|M) = \Phi_{\text{c}}(L|M) + \Phi_{\text{s}}(L|M), \quad (2)$$

where subscripts ‘c’ and ‘s’ refer to central and satellite, respectively. In what follows we describe each of these two components in turn.

3.1.1 Centrals

Each halo hosts at most one central galaxy, which can be either red or blue. We assume that the colour distribution is governed by halo mass (see e.g. More et al. 2011; Mandelbaum et al. 2016; Zu & Mandelbaum 2016), such that

$$\Phi_{\text{c}}(L|M) = f_{\text{r}}(M)\Phi_{\text{c,r}}(L|M) + f_{\text{b}}(M)\Phi_{\text{c,b}}(L|M), \quad (3)$$

where subscripts ‘r’ and ‘b’ refer to red and blue, respectively. Here $f_{\text{r}}(M) = 1 - f_{\text{b}}(M)$ is the probability that a central galaxy in a halo of mass M is red. Throughout we assume that f_{r} scales linearly with halo mass according to

$$f_{\text{r}}(M) = \min[1, \max(0, f_0 + \alpha_f \log M_{12})], \quad (4)$$

where $M_{12} = M/(10^{12} h^{-1} M_{\odot})$. The CLFs of red and blue centrals are parametrized separately by independent log-normal distributions,

$$\Phi_{\text{c}}(L|M)dL = \frac{\log e}{\sqrt{2\pi\sigma_c^2}} \exp\left(-\frac{[\log L - \log \tilde{L}_{\text{c}}(M)]^2}{2\sigma_c^2}\right) \frac{dL}{L}. \quad (5)$$

The median luminosity $\tilde{L}_{\text{c}}(M)$, in turn, is parametrized by a broken power-law:

$$\tilde{L}_{\text{c}}(M) = L_0 \frac{(M/M_1)^{\gamma_1}}{(1 + M/M_1)^{\gamma_1 - \gamma_2}}. \quad (6)$$

Altogether, we have 12 free parameters to describe how central galaxies populate dark matter haloes; f_0 , α_f , $L_{0,\text{r}}$, $M_{1,\text{r}}$, $\gamma_{1,\text{r}}$, $\gamma_{2,\text{r}}$, σ_{r} , $L_{0,\text{b}}$, $M_{1,\text{b}}$, $\gamma_{1,\text{b}}$, $\gamma_{2,\text{b}}$, and σ_{b} . This parametrization is identical to the one used in More et al. (2011).

3.1.2 Satellites

Each halo can have an arbitrary number of satellite galaxies. We assume that the occupation is governed by a Poisson distribution with expectation value

$$N_{\text{s}}(M) = \int_{L_{\text{th}}}^{\infty} \Phi_{\text{s}}(L|M)dL, \quad (7)$$

where L_{th} is the luminosity threshold. We model the satellite CLF as a modified Schechter function:

$$\Phi_{\text{s}}(L|M) = \frac{\phi_{\text{s}}^*(M)}{L_{\text{s}}^*(M)} \left(\frac{L}{L_{\text{s}}^*(M)}\right)^{\alpha_{\text{s}}(M)} \exp\left[-\left(\frac{L}{L_{\text{s}}^*(M)}\right)^2\right]. \quad (8)$$

In essence, the luminosities of satellites are drawn from a power-law with slope α_{s} with an exponential cut-off above a critical luminosity, $L_{\text{s}}^*(M)$. The luminosity cut-off, $L_{\text{s}}^*(M)$, is related to the characteristic luminosity of red, central galaxies in haloes of the same mass according to

$$L_{\text{s}}^*(M) = 0.562 \tilde{L}_{\text{c,r}}(M). \quad (9)$$

This scaling is a good fit to the luminosity distribution of central and satellite galaxies as inferred from the SDSS

galaxy group catalogue of Yang et al. (2009). Finally, the normalization $\phi_s^*(M)$ is parametrized by

$$\log [\phi_s^*(M)] = b_0 + b_1 \log M_{12} + b_2 (\log M_{12})^2. \quad (10)$$

As described in detail in §4, we assume that no satellite galaxy is brighter than its corresponding central. For the colour of satellite galaxies we choose the same parametrization as for the centrals. We find that with $f_{0,\text{sat}} = 0.44$ and $\alpha_{f,\text{sat}} = 0.14$ this is a very good approximation of the data reported in Yang et al. (2008) as a function of both halo mass and satellite luminosity. Since we are mainly interested in central galaxy colour, and ignore the colours of satellites candidates, this only matters for the small amount of impurity in our central sample, as discussed in §5.1. We make no attempt to fit $f_{0,\text{sat}}$ and $\alpha_{f,\text{sat}}$ with our data and therefore keep their values fixed. This leaves us with 4 free parameters to describe the satellite CLF: α_s , b_0 , b_1 , and b_2 . This parametrization is identical to the one used by van den Bosch et al. (2013) and Cacciato et al. (2013).

3.1.3 Summary

Table 1 lists all parameters that describe the galaxy occupation. It also lists their prior ranges used when fitting data and their default values used to create and analyse mock catalogues in §8. These default values are taken from More et al. (2011) and Cacciato et al. (2013) for centrals and satellites, respectively. We note that in this work we use the virial mass definition whereas those two studies used slightly different ones. For the purpose of testing our analysis technique on complicated mock catalogues, these slight inconsistencies are not relevant.

3.2 Phase-space distributions

The CLF described above characterizes the abundance of central and satellite galaxies as function of halo mass. We now describe our model for the phase-space coordinates of these galaxies with respect to their host halo.

3.2.1 Centrals

Throughout we assume that central galaxies reside at rest at the centre of their host haloes, which are assumed to be spherically symmetric. While it is known that centrals are not perfectly at rest relative to the centre-of-mass of their host halo, (van den Bosch et al. 2005b; Behroozi et al. 2013; Guo et al. 2015a,b, 2016; Ye et al. 2017), this motion does not have a significant impact. The typical speed of a central is much lower than the average speed of its satellites, of the order of $\sim 10\%$. Thus, neglecting the motion of centrals will underestimate the central-satellite velocity dispersion by $\sqrt{1 + 0.1^2} - 1 = 0.5\%$. This is much smaller than our measurement uncertainties and other systematic effects.

3.2.2 Satellites

Satellite galaxies are assumed to follow a radial profile, $n_s(r)$, that is also spherically symmetric. Since we stack a large number of dark matter haloes with random orientations, this assumption of spherical symmetry is fair, despite

the fact that individual haloes, and their satellite populations, are known to be aspherical (e.g., Wang et al. 2008). We follow van den Bosch et al. (2004) and assume that the radial profile of satellite galaxies, as a function of the radial distance r from the halo centre, is given by a generalized Navarro–Frenk–White (NFW) profile (Navarro et al. 1996),

$$n_s(r) \propto \left(\frac{r}{\mathcal{R}r_s}\right)^{-\gamma} \left(1 + \frac{r}{\mathcal{R}r_s}\right)^{\gamma-3}. \quad (11)$$

Here \mathcal{R} and γ are free parameters and r_s is the scale radius of the dark matter halo, which is related to the halo virial radius via the concentration parameter $c_{\text{vir}} = r_{\text{vir}}/r_s$. By definition, no satellites are located outside the virial radius. Throughout this work we consider three different radial profiles for satellites. The first has $\gamma = \mathcal{R} = 1$, which implies that satellites follow a NFW profile with the same concentration parameter as the dark matter; in other words, satellite galaxies are unbiased tracers of the dark matter within their host haloes. We also consider a profile with $\gamma = 1$ and $\mathcal{R} = 2.0$, for which the satellites still follow an NFW profile, but with a concentration parameter that is twice that of the dark matter. The final profile that we consider has $\gamma = 0$ and $\mathcal{R} = 2.5$, which implies that the distribution of satellites has a core of constant number density, and a concentration parameter that is 2.5 times larger than that of its dark matter. As we show in §8.3, this latter profile is a good description of the distribution of subhaloes in dark matter simulations. Altogether, these profiles roughly bracket the range of observational constraints on the radial distribution of satellite galaxies in groups and clusters (e.g., Carlberg et al. 1997; van der Marel et al. 2000; Lin et al. 2004; Yang et al. 2005; Chen 2008; More et al. 2009a; Guo et al. 2012; Cacciato et al. 2013).

In our analytical model, we assume that the velocities of satellites are isotropic, and obey the spherically symmetric Jeans equation (Binney & Tremaine 1987). For satellite galaxies with a radial profile given by Eq. (11), located in a halo with virial velocity V_{vir} and concentration parameter c_{vir} , this implies a velocity dispersion given by

$$\sigma^2(r|V_{\text{vir}}, c_{\text{vir}}) = \frac{c_{\text{vir}} V_{\text{vir}}^2}{\mathcal{R}^2 g(c_{\text{vir}})} \left(\frac{r}{\mathcal{R}r_s}\right)^{\gamma} \left(1 + \frac{r}{\mathcal{R}r_s}\right)^{3-\gamma} \int_{r/r_s}^{\infty} \frac{g(y)dy}{(y/\mathcal{R})^{\gamma+2}(1+y/\mathcal{R})^{3-\gamma}}, \quad (12)$$

where

$$g(y) = \ln(1+y) - \frac{y}{1+y}. \quad (13)$$

Figure 1 shows how our choice of the radial profile affects the velocity dispersion. For each of our three radial profiles we plot the average velocity dispersion inside a fixed aperture centred on the halo. The concentration parameter is set to $c_{\text{vir}} = 7.0$. By scaling the aperture radius by r_{vir} and the dispersion by V_{vir}^2 , the relation becomes independent of halo mass. Note that less centrally concentrated satellite profiles imply a larger velocity dispersion. However, the differences decrease as the aperture encloses larger parts of the satellite population, declining from $\sim 50\%$ to $\sim 20\%$ when $R_{\text{ap}}/r_{\text{vir}}$ increases from 0.1 to 1.0.

Our model for the velocities of satellite galaxies makes

Name	Description	Prior	Default
$\log L_{0,r}/(h^{-2}L_{\odot})$	normalization of mass–luminosity relation for red centrals	[9.0, 12.0]	9.99
$\log M_{1,r}/(h^{-1}M_{\odot})$	characteristic mass of mass–luminosity relation for red centrals	[10.0, 13.0]	11.50
$\gamma_{1,r}$	low-mass slope of mass–luminosity relation for red centrals	[0.0, 5.0]	4.88
$\gamma_{2,r}$	high-mass slope of mass–luminosity relation for red centrals	[0.0, 5.0]	0.31
σ_r	logarithmic scatter in luminosity at fixed halo mass for red centrals	[0.05, 0.5]	0.20
$\log L_{0,b}/(h^{-2}L_{\odot})$	normalization of mass–luminosity relation for blue centrals	[9.0, 12.0]	9.55
$\log M_{1,b}/(h^{-1}M_{\odot})$	characteristic mass of mass–luminosity relation for blue centrals	[10.0, 13.0]	10.55
$\gamma_{1,b}$	low-mass slope of mass–luminosity relation for blue centrals	[0.0, 5.0]	2.13
$\gamma_{2,b}$	high-mass slope of mass–luminosity relation for blue centrals	[0.0, 5.0]	0.34
σ_b	logarithmic scatter in luminosity at fixed halo mass for blue centrals	[0.05, 0.5]	0.24
f_0	red fraction of centrals for haloes of mass $10^{12} h^{-1}M_{\odot}$	[0.0, 1.0]	0.70
α_f	linear dependence of the red fraction of centrals on halo mass	[0.0, 2.0]	0.15
α_s	low-luminosity slope of satellites of the satellite CLF	[−1.5, −0.9]	−1.18
b_0	determines the normalization of the satellite CLF	[−3.0, 3.0]	−1.17
b_1	determines the normalization of the satellite CLF	[−3.0, 3.0]	1.53
b_2	determines the normalization of the satellite CLF	[−3.0, 3.0]	−0.217

Table 1. Variable governing the occupation of dark matter haloes with galaxies. See the text for details.

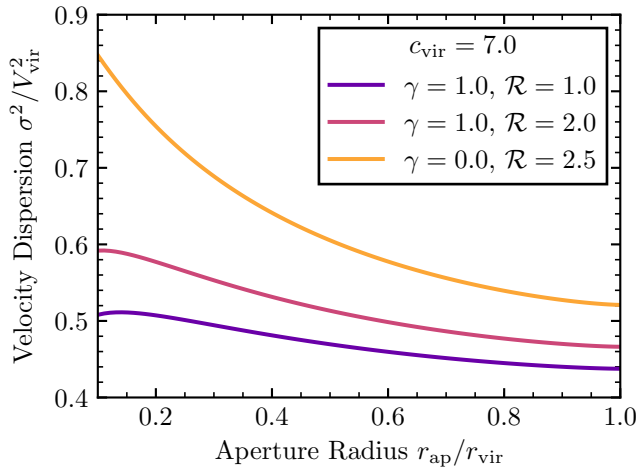


Figure 1. The average line-of-sight velocity dispersion of satellites within a projected aperture of radius R_{ap} implied by the spherically symmetric Jeans equation as a function of R_{ap} in units of the host halo’s virial radius, r_{vir} . Generally, the velocity dispersion decreases for larger aperture radii, as does the difference between the radial profiles. More extended radial profiles lead to a higher overall velocity dispersion.

two crucial assumptions: isotropy and virial equilibrium. It is well known that dark matter subhaloes and satellite galaxies have kinematics that are slightly radially anisotropic (e.g., Diemand et al. 2004; Wojtak & Mamon 2013). However, velocity anisotropy mainly affects the *radial profile* of $\sigma^2(r)$; not the *total* velocity dispersion averaged over the entire satellite population (the latter appears in the virial theorem, which is independent of velocity anisotropy). Since we are measuring, and modelling, the velocity dispersion of satellite galaxies inside a relatively large aperture centred on the central galaxy, neglecting anisotropy only affects the resulting velocity dispersions by a few percent (see van den Bosch et al. 2004; ?; Conroy et al. 2007, for details). The assumption of virial equilibrium is also subject to concern. After all, dark matter haloes are constantly accreting matter, and new satellite galaxies, from their surroundings. Hence, some

satellite galaxies will be on first infall, and their kinematics will not be in perfect virial equilibrium. As shown by Ye et al. (2017) ignoring the resulting time-dependence in the Jeans equations (which are based on the assumption of a steady-state), may result in small biases in the inferred velocity dispersion. We address this issue in §8.3 by analysing mock catalogues where satellite galaxies are placed on subhaloes. Those tests suggest that deviations from the steady-state assumption do not cause a strong bias in our inferred model parameters.

4 MOCK SDSS CATALOGUES

In order to test and validate our analytical model, we construct mock SDSS-like galaxy redshift surveys. By treating these mocks exactly as we treat the real data, we can test the accuracy to which we can recover the galaxy–dark matter connection, $P(\mathcal{H}|\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_n)$.

Our mocks are constructed using the SMDPL simulations (Klypin et al. 2016), which traces the dark matter distribution in a cosmology that is based on the Planck Collaboration et al. (2014) cosmological parameters. It traces the dark matter distribution inside a cubic volume of $(400 h^{-1}\text{Mpc})^3$ using 3840^3 particles, resulting in a particle mass of $1.0 \times 10^8 h^{-1}M_{\odot}$. We construct mock redshift surveys using the publicly available¹ $z = 0.05$ ROCKSTAR halo catalogues (Behroozi et al. 2013), and use `halotools` (Hearin et al. 2016) to populate all dark matter haloes with $M_{\text{vir}} > 3 \times 10^{10} h^{-1}M_{\odot}$ with mock galaxies according to our fiducial CLF model (see §3.1 and Table 1).

For each host halo above this mass limit, we first decide on the colour of the central by drawing a random number between 0 and 1. The central is assumed to be red (blue) if this random number is smaller (larger) than $f_r(M)$. Next we draw the corresponding luminosity from $\Phi_{c,r}(L|M)$, if red, or $\Phi_{c,b}(L|M)$, if blue. Any central galaxy with a luminosity below a threshold of $10^9 h^{-2}L_{\odot}$ is removed from the catalogue. The remaining central galaxies are given the

¹ <http://yun.ucsc.edu/sims/SMDPL/trees/index.html>

position and velocity of their halo cores, which is defined as the region that encloses the innermost 10 percent of the halo virial mass. These positions and velocities are calculated by ROCKSTAR as detailed in Behroozi et al. (2013).

Next we draw the number of satellite galaxies, under the assumption that $P(N_s|M)$ follows a Poisson distribution, whose mean is given by Eq. (7). Throughout, we adopt a luminosity threshold of $10^9 h^{-2} L_\odot$ for our mock galaxies. For each of the N_s satellites in the halo in question we then draw a luminosity from the satellite CLF, $\Phi_s(L|M)$, given by equation (8) and a colour based on $f_{r,\text{sat}}(M)$. After drawing satellites from the above CLF we remove all satellites that are brighter than their respective centrals.² For the positions and velocities we assume the analytical model for the phase-space coordinates of satellite galaxies described in §3.2.2 above; we draw the positions from a spherical distribution with radial profile $n_s(r)$, given by Eq. (11), and one-dimensional velocities from a Gaussian distribution with dispersion $\sigma(r)$, given by Eq. (12). Both the positions and the velocities are with respect to the core of the host halo.³

After populating the dark matter haloes with mock galaxies, we place a virtual observer inside the simulation box. Both the position and the orientation are chosen randomly. We then proceed to calculate angular coordinates, redshifts and apparent magnitudes for all galaxies within $z < 0.15$. If necessary, the simulation box is periodically repeated out to that redshift. The apparent magnitudes include K and evolution correction effects. Next, an apparent magnitude cut of $m_r \leq 17.6$ is applied, after which we add redshift space distortions to the cosmological redshifts. In addition, we also add a random scatter in redshift of the order of $35 \text{ km s}^{-1}/c$ to simulate redshift uncertainties in SDSS (More et al. 2009a). Subsequently, we apply the SDSS DR7 survey mask⁴, rejecting galaxies outside of the survey window or inside of masked regions.

Finally, we simulate the effect of spectroscopic incompleteness, particularly fibre collisions. In the SDSS, on a single plate spectroscopic fibres cannot be placed simultaneously for objects separated by less than $55''$ (Blanton et al. 2003). However, some galaxies are observed with multiple plates, allowing spectroscopic redshifts even for close pairs. This fibre collision effect leads to some objects missing spectroscopic redshifts. We follow the common practice of assigning such ‘fibre-collided’ galaxies the redshift of its nearest neighbour (Zehavi et al. 2005). To simulate this type of incompleteness, we first construct a maximal ‘decollided’ set of galaxy targets (Blanton et al. 2003). Within this set of decollided galaxies, no two targets are within $55''$ of each other.

² By removing bright satellites the satellite occupation is slightly lowered and has a small dependence on the actual central luminosity. However, we find this decrease and its impact on satellite kinematics to be negligible. This is because it is generally unlikely that any given satellite is brighter than the central in CLF models (Skibba et al. 2011; Lange et al. 2018). In our fiducial mock, this only removes 1% of all satellites.

³ A more realistic approach would be to express all phase-space positions of satellites with respect to the bulk of the halo. However, this information is not available in the public halo catalogues. As discussed previously, the impact of this simplification on the velocity dispersion should be negligible.

⁴ <http://sdss.physics.nyu.edu/lss/dr72>

The remaining galaxies are potentially collided targets that would not be assigned a redshift in SDSS if they would not lie in regions of overlapping spectroscopic plates. We then randomly choose 65% of those galaxies to be missing a spectroscopic redshift. This value is very close to what we observe in the New York University Value-Added Galaxy Catalog (VAGC; Blanton et al. 2005). See the appendix for more details. Additionally, we randomly remove an additional 1% of all spectroscopic redshifts to simulate other failure modes (Blanton et al. 2003). While this algorithm reproduces the most salient features of fibre collisions in SDSS, it misses less important details like the tiling algorithm and the according correlation of fibre collision probability with large-scale structure (Blanton et al. 2003). We test the impact of this correlation on our observables in the appendix but find it negligible. Thus, the fibre collision algorithm simulated here suffices for our mock simulations.

5 OBSERVATIONS

In this section we first describe the detailed criteria used to select primary (assumed to be central) and secondary (assumed to be satellite) galaxies from a galaxy redshift survey. We then describe how we use these galaxies to measure the host- and satellite-weighted velocity dispersions and other observational constraints that we use to model the galaxy-dark matter connection.

Throughout this study, we only use galaxies (both in our mocks and in the SDSS data) in a volume-limited sample specified by $0.02 \leq z \leq 0.067$, which, for our adopted apparent magnitude limit of $m_r = 17.6$, is complete down to a r -band luminosity of $10^{9.5} h^{-2} L_\odot$. The main reason for using a volume-limited sample is to facilitate constraining the scatter in the galaxy-halo connection via a comparison of host- and satellite-weighted velocity dispersions (see More et al. 2009b, for details). In addition, if one were to use a flux-limited sample, one also needs to model potential luminosity segregation, since haloes at a lower redshift, are now sampled down to lower satellite luminosity. Mainly for these two reasons, we accept the price of a reduced dynamic range and a reduced signal-to-noise that comes with using only a volume-limited subsample of the full SDSS data. We hope to extend our methodology to flux-limited samples in future work.

5.1 Sample selection criteria

We use cylindrical isolation criteria to select samples of centrals and satellites. Due to interlopers and other impurities, not every central (satellite) thus selected is indeed a central (satellite). We therefore refer to galaxies that are selected as centrals and satellites as primaries and secondaries, respectively.

A galaxy is considered a primary if it is at least f_h times brighter than any other galaxy within a cylindrical volume specified by $R_p < R_h$ and $|\Delta v_z| < (\Delta v)_h$. Here R_p is the separation projected on the sky at the distance of the primary candidate and Δv_z is the line-of-sight (los) velocity difference (Eq. [1]). We apply this isolation criterion in a rank-ordered fashion, starting with the brightest galaxy. If

a galaxy is selected as a primary, we remove all other galaxies within its $[R_h, (\Delta v)_h]$ cylinder from the list of potential primary candidates. Around each primary galaxy thus selected, secondaries are defined as those galaxies that are at least f_s times fainter than their primary and located within a cylindrical volume with $R_p < R_s$ and $|\Delta v_z| < (\Delta v)_s$. Furthermore, to avoid complications due to survey edge effects, we remove all primaries for which more than 20% of a ring with radius R_h centred on the primary falls outside of the survey window or inside of the mask.

In total, the selection of primaries and secondaries thus depends on six free parameters: R_h , $(\Delta v)_h$ and f_h to specify the population of primaries, and R_s , $(\Delta v)_s$ and f_s to specify the secondaries. These parameters determine the completeness and purity of the sample, both in terms of primaries and secondaries. They also determine the interloper fraction, which is defined as the fraction of secondaries that are not satellite galaxies within the same halo as the corresponding primary. Minimizing the number of interlopers requires sufficiently small R_s and $(\Delta v)_s$, while maximizing the purity of the primaries (i.e., minimizing the number of satellites that are erroneously identified as centrals) requires large R_h , $(\Delta v)_h$ and f_h . Of course, each of these restrictions dramatically reduces the completeness of the samples, thereby making the measurements more noisy.

Most previous studies of satellite kinematics have adopted extremely conservative selection criteria, with both f_h and f_s significantly larger than unity (e.g., Zaritsky et al. 1993, 1997; McKay et al. 2002; Brainerd & Specian 2003; Prada et al. 2003; Conroy et al. 2007; Norberg et al. 2008; Wojtak & Mamon 2013). In addition, all these studies adopted fixed values for the selection criterion parameters, independent of the luminosity of the galaxy under consideration. 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 primary's halo. They used an iterative criterion whereby the cylindrical aperture scales with the velocity dispersion around the primary obtained in the previous iteration. By using such an adaptive, iterative selection criterion, they were able to increase the number of secondaries by almost an order of magnitude (using $f_h = f_s = 1$), while drastically reducing the actual interloper fraction.

The same iterative method was later used in More et al. (2009a) and More et al. (2011), who found that the satellite velocity dispersion around red primaries scales with luminosity L as

$$\log \sigma_{200} \equiv \log \left(\frac{\sigma}{200 \text{ km s}^{-1}} \right) = a + b \log L_{10} + c \log^2 L_{10}, \quad (14)$$

where $L_{10} = L/10^{10} h^{-2} L_\odot$ and $a = -0.07$, $b = 0.38$ and $c = 0.29^5$. We base our selection of primaries on this scaling relation. We follow van den Bosch et al. (2004) and More et al. (2011) by setting $f_h = f_s = 1$, and scale the dimensions of the cylinder used to select primaries as $R_h = 0.5\sigma_{200} h^{-1} \text{Mpc}$ and $(\Delta v)_h = 1000\sigma_{200} \text{ km s}^{-1}$. This

value for R_h roughly corresponds to 1.25 times the virial radius. We have experimented with other cylinder sizes for the isolation criterion, but typically find that increasing completeness results in a larger contamination (i.e., reduced purity), which in turn implies a more significant bias in our results. In the end, we decided on values that result in at least $\sim 99\%$ purity in centrals for the primary sample. Note that, contrary to the approach adopted here, More et al. (2009a, 2011) did not remove galaxies located inside the larger cylinder of a brighter galaxy from the list of potential primary candidates. On the other hand, they used $R_h = 0.8\sigma_{200} h^{-1} \text{Mpc}$ instead of $0.5\sigma_{200} h^{-1} \text{Mpc}$. Coincidentally, we find that both strategies result in very similar completeness and purity. Contrary to More et al. (2011), we choose the same cylinder sizes for red and blue primaries. This is done primarily to not bias our selection towards either blue or red primaries. In order to avoid excessive values for σ_{200} , which can occur for fibre-collided galaxies that are accidentally assigned a too high redshift, we limit σ_{200} to a maximum value of 5, i.e. $\sigma \leq 1000 \text{ km s}^{-1}$.

The criterion for secondaries is that have a spectroscopic redshift (i.e., they can not be fibre-collided galaxies) and that they lie in a cylinder centred on a primary with $R_s = 0.15\sigma_{200} h^{-1} \text{Mpc}$ and $(\Delta v)_s = 4000 \text{ km s}^{-1}$. This value for R_s roughly corresponds to 0.375 times the virial radius, while the value for $(\Delta v)_s$ is large enough to include the vast majority of all satellites, even in massive clusters. Note that we do not scale this parameter with σ_{200} ; although this implies an increasing fraction of interlopers with decreasing central luminosity, these interlopers are easily identified as such (see §5.2). Also, contrary to the selection of primaries, this time we do allow the cylinder sizes to be different for red and blue primaries: for red primaries we use equation (14) with the parameters listed above, while for blue primaries we use $a = -0.19$, $b = 0.46$ and $c = -0.16$ (see More et al. 2011). Finally, in order to alleviate the impact of imperfect correction for fibre-collisions (see §5.3), we exclude secondaries that have a projected distance to their primary $R_p \leq R_c \equiv 60 h^{-1} \text{Mpc}$.

We explore the efficiency of our selection criterion in Figure 2. All quantities are inferred from SDSS-like mock catalogues, constructed using the method described in §4 in which satellites follow an NFW profile. The left-hand panel displays the central completeness, defined as the fraction of centrals that are ultimately selected as primaries. Ideally, we want this to be close to unity. Our selection criteria achieve a central completeness that increases from $\sim 60\%$ for the faintest centrals to 80% at the bright end, with virtually no dependence on galaxy colour. The solid curve in the middle panel shows that the central contamination, defined as the fraction of primaries that are not centrals, is extremely low, less than 1% for all primary luminosities. The dotted line shows the same quantity in a mock with 100% spectroscopic completeness (i.e., no fibre collisions), which results in a very similar contamination fraction. Hence, central contamination is not dominated by catastrophic failures in the nearest neighbour redshift assignment of fibre-collided galaxies, which can happen if the fibre-collided central of a halo is assigned the redshift of a fore- or background galaxy. Finally, the dashed line shows the central contamination if one removes fibre collided galaxies *before* identifying primaries. This results in a contamination fraction that is roughly 5

⁵ Note that Eqs. (11) in More et al. (2009a) and (3) in More et al. (2011) are incorrect, and inconsistent with the values of a quoted in those papers. The equation presented here is the one actually used by More et al. (2011) for his LR sample. (Surhud More, *priv. comm.*)

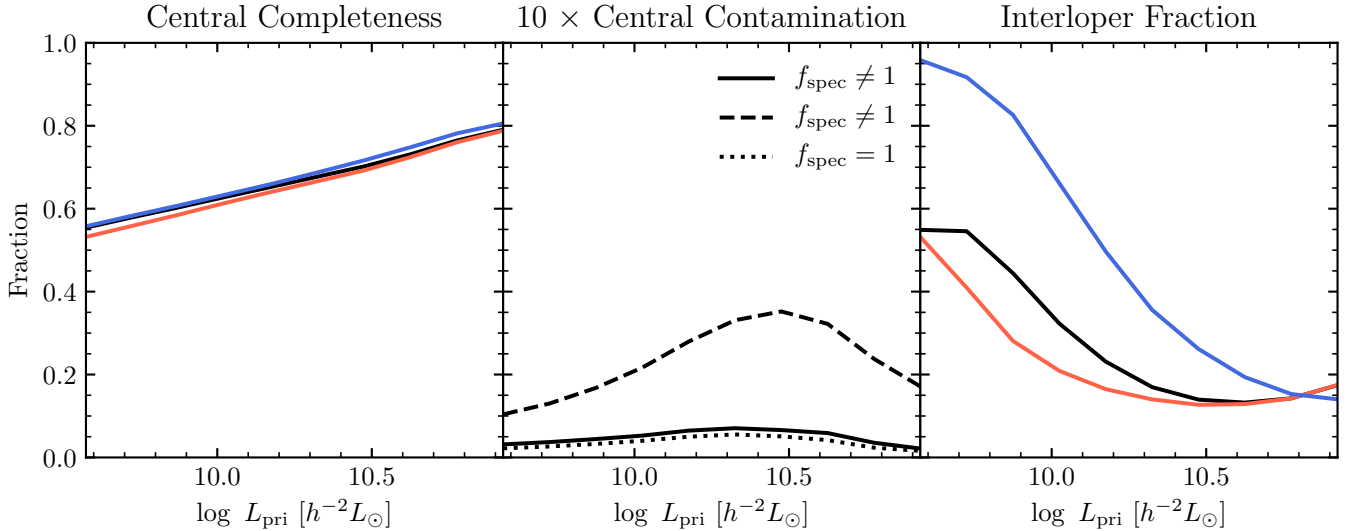


Figure 2. This plot shows the efficiency of our selection criterion as inferred from mock catalogues built with the default model. We display the central completeness, the fraction of centrals selected as primaries (left panel), the central contamination, the probability of a primary to not be a central (middle panel), and the interloper fraction, which is defined as the fraction of secondaries not residing in the same halo as the primary (right panel). Note that we also include primaries without any secondaries when calculating central completeness and purity. All probability are displayed as a function of primary luminosity. Additionally, we show them for red primaries (red), blue primaries (blue) and all (black). The dashed line in the middle panel shows the contamination if there was 100% spectroscopic completeness, e.g. no fibre collisions, and the dashed line the fraction if fibre collided galaxies are removed before applying the selection criterion. Note that the central contamination is multiplied by 10 to fit into the same range as the other ratios.

times larger. This demonstrates that assigning fibre-collided galaxies the redshift of its nearest neighbour is reasonably accurate (see also Zehavi et al. 2011), and that one should only remove fibre collided galaxies *after* the primaries have been identified.

Finally, we note that in our mocks centrals are always the brightest galaxies in their halo. This is contrary to various studies showing that in a small fraction of haloes a satellite will be the brightest galaxy (Skibba et al. 2011; Hikage et al. 2013; Wang et al. 2014; Hoshino et al. 2015; Lange et al. 2018). We have tested that, if that is the case, our sample of primaries should be very pure in brightest halo galaxies (hereafter BHGs), rather than central galaxies. One might worry that some of our supposed centrals being satellites would systematically bias our inferences. For example, one might naively expect that the measured velocity dispersion is $\sqrt{2}$ times higher when measured around satellites. In Lange et al. (2018), we have shown that ignoring this complication is unlikely to have a significant impact on our inferences (see also Skibba et al. 2011).

5.2 Interloper rejection

One of the main challenges in measuring satellite kinematics is that not all secondaries reside in the same halo as their primary. Such interlopers have a very different velocity distribution than true satellites, making it crucial that one corrects for this interloper contamination. The right-hand panel of Figure 2 shows the fraction of interlopers among secondaries. For this figure we only counted secondaries with $|\Delta v_z| < 600\sigma_{200}$. Secondaries with larger velocity differences are trivially identified as interlopers. We first note the ex-

tremely high interloper fraction for low-luminosity blue primaries. That is mainly due to the fact that those galaxies have very few same-halo satellites in our model. However, for the rest of the primaries, the selection of secondaries results in interloper fractions as low as $\sim 10\%$.

The Δv_z distribution of secondaries consists of a satellite component centred on $\Delta v_z = 0$ and an interloper component that is close to flat. Most previous studies modelled the Δv_z distribution of satellites as a single Gaussian and the interloper distribution as flat (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; van den Bosch et al. 2004; Conroy et al. 2007). In this case, the velocity dispersion of the best-fit Gaussian is the estimate for the dispersion of same-halo satellites. However, Becker et al. (2007) showed that a single Gaussian fit significantly underestimates the dispersion of same-halo satellites. More et al. (2009a) advocated using a fit to a double-Gaussian plus a flat background model to estimate the velocity dispersion of satellites. Still, contrary to the claims by More et al. (2009a), we also find this dispersion estimate to be biased. Similar to what was estimated by Becker et al. (2007), we find that it underpredicts the dispersion by roughly 5%. We have also tested the bi-weight estimator (Beers et al. 1990), but encountered similar biases. Here, we use a different method to reject interlopers. While this new technique does not significantly improve the accuracy in the velocity dispersion estimate (i.e., it does not significantly reduce the bias), it does improve the precision of the estimator slightly. We later correct for the bias numerically using a large number of mock observations for which the true velocity dispersions are known.

The main idea of our interloper rejection is to use the distribution of galaxies in the $\Delta v_z - R_p$ plane instead of just using velocity information. Thus, we model the total phase-

space distribution of secondaries around primaries as the sum of an interloper component $P_{\text{int}}(\Delta v_z, R_p)$ and a model for the phase-space distribution of satellites around centrals $P_{\text{sat}}(\Delta v_z, R_p)$. Thus,

$$P_{\text{tot}}(\Delta v_z, R_p) = f_{\text{int}} P_{\text{int}}(\Delta v_z, R_p) + (1 - f_{\text{int}}) P_{\text{sat}}(\Delta v_z, R_p). \quad (15)$$

For the interloper model we assume a flat distribution in Δv_z and a constant projected number density, i.e.,

$$P_{\text{int}}(\Delta v_z, R_p) \propto R_p. \quad (16)$$

Based on tests with our mock catalogues, we find the assumption regarding the R_p distribution to be fairly accurate (see also Wojtak & Mamon 2013; Zheng & Guo 2016). On the other hand, similar to van den Bosch et al. (2004) and More et al. (2009a), we find an increased number of interlopers at $\Delta v_z \sim 0$. However, this only has a minor impact on the estimate of the satellite velocity dispersion, which anyways is accounted for in our bias-treatment discussed in §7.

Regarding the model for the satellites, we assume that they have a spherically symmetric density profile and obey the Jeans equation. See §3.2 for more details. Furthermore, we assume that the host halo masses of satellites are drawn from a log-normal distribution and that the concentration parameter c_{vir} at each host halo mass is the median concentration for that mass extracted from SMDPL.

In each of the 20 bins of colour and luminosity of primaries, we fit the Δv_z - R_p distribution of secondaries with a 3-parameter model: the interloper fraction f_{int} , the median halo mass \tilde{M} and the spread in halo masses $\sigma_{\log M}$. We determine the fit that maximizes the likelihood,

$$\mathcal{L} \propto \prod_i^n P_{\text{tot}}^{w_i}(\Delta v_{z,i}, r_{p,i} | f_{\text{int}}, \tilde{M}, \sigma_M), \quad (17)$$

where the product runs over all secondaries in the sample and w denotes an optional weight assigned to each secondary. The choice of weights is discussed in the next two sections. Afterwards, a membership probability p_{mem} is assigned to each secondary,

$$p_{\text{mem}}(\Delta v_z, R_p) = \frac{(1 - f_{\text{int}}) P_{\text{sat}}(\Delta v_z, R_p)}{P_{\text{tot}}(\Delta v_z, R_p)}, \quad (18)$$

based upon the best-fit model.

Those membership probabilities can then be used in a straightforward manner to estimate the velocity dispersion of same-halo satellites, excluding interlopers. An example is shown in Figure 3. For this plot we have populated the SMDPL with galaxies and used the distance observer approximation. No fibre collision effects were included. We show the line-of-sight velocity difference of secondaries around red centrals with $10.70 < \log(L/h^{-2} L_\odot) < 10.85$. The black solid line is the distribution of all secondaries, whereas the blue solid line only shows same-halo satellites. The blue dashed line is our estimate for the latter by weighting each secondary by its membership probability.

Finally, while the 3 best-fit parameters of the fit contain information about average halo masses and their spread, the true distribution in halo masses is more complicated and will be discussed below. Thus, those best-fit parameters are disregarded and not used for further analysis. In other

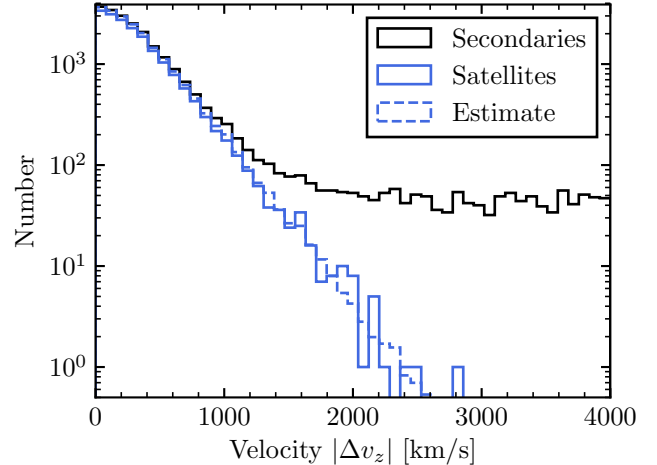


Figure 3. The distribution of line-of-sight velocities of secondaries around red primaries with $10.70 < \log L < 10.85$. We show all secondaries (black, solid), same-halo secondaries (blue, solid) and our estimate for the latter (blue, dashed) using the distribution of all secondaries. See the text for details. **[Surhud: Is this bias due to the Delta vz cut?] [Johannes: What makes you think so? I don't immediately see a reason for that.]**

words, the model for $p_{\text{mem}}(\Delta v_z, R_p)$ discussed here is solely used to estimate membership probabilities.

5.3 Fibre collision correction

When running our selection criterion to identify primaries and secondaries, we assign each galaxy missing a spectroscopic redshift that of its closest neighbour (see Blanton et al. 2005; Zehavi et al. 2005). However, we stress that in our subsequent analysis only primary-secondary pairs with spectroscopic redshifts for both are used. Additionally, we assign each galaxy a spectroscopic weight w_{spec} to counteract the effect of fibre collisions. For each galaxy we count the number of galaxies inside a radius of $55''$ around it. We then determine for the entire survey what fraction f_{spec} of galaxies with exactly that number of neighbours within $55''$ have been assigned a spectroscopic redshift. The weight designated to the galaxy is then simply $w_s = 1/f_{\text{spec}}$. As discussed below, this correction works very well but not perfectly below the fibre collision scale. Thus, in order to be conservative, we only consider secondaries separated from the primary by at least $R_c = 60 h^{-1} \text{kpc}$, roughly the fibre collision scale at the maximum redshift of our analysis. This removes less than 10% of secondaries around primaries with $L_{\text{pri}} > 10^{10} h^{-2} L_\odot$.

We test the fidelity of our correction using mock catalogues. Particularly, for this test we assume that satellites follow the dark matter density distribution. In Figure 4 we show the surface density of secondaries around primaries with $10.6 < \log L/(h^{-2} L_\odot) < 10.8$ in 100 simulated SDSS-like surveys. Since we know true halo associations, we only show same-halo satellites. The black line is the idealized case of a spectroscopic survey without fibre collisions. The blue dashed line is for a survey with fibre collisions where we only consider galaxies with spectroscopic redshifts. We see an overall decrease in the surface density. Furthermore,

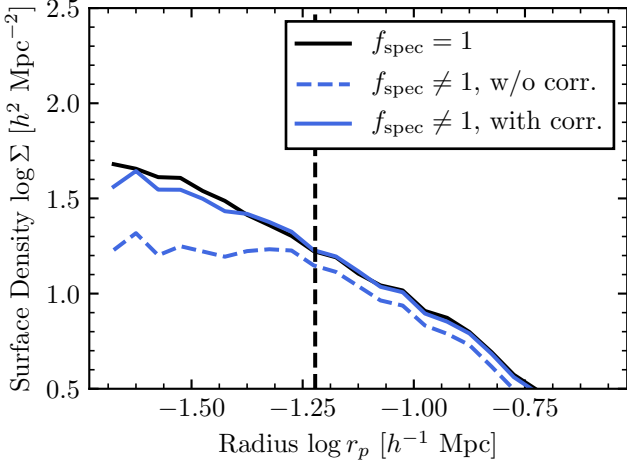


Figure 4. The surface density of same-halo secondaries around primaries with $10.6 < \log L / (h^{-2} L_{\odot}) < 10.8$. The surface densities have been averaged over 100 mock SDSS-like catalogues. We show the true surface density for surveys without fibre collisions (black), surveys having fibre collisions but no correction (blue, dashed) and surveys with fibre collisions and a correction applied to it (blue, solid). Secondaries at separations less than $60h^{-1}\text{kpc}$, as shown by the black dashed line, are excluded from our analysis. See the text for details.

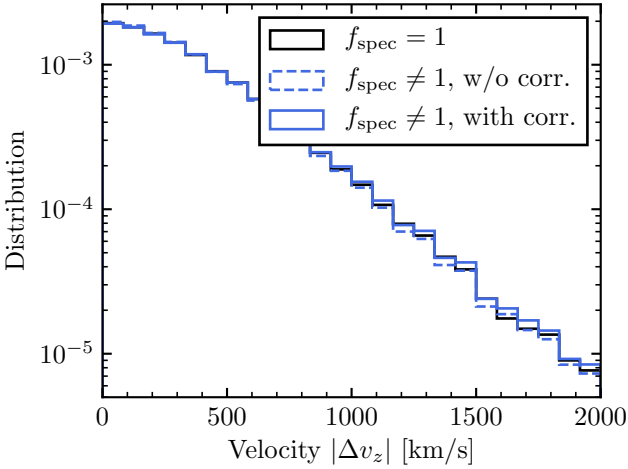


Figure 5. The line-of-sight velocity distribution of secondaries around primaries with $10.6 < \log L / (h^{-2} L_{\odot}) < 10.8$. The primaries and secondaries are the same as in Figure 4, except for excluding those secondaries with $R_p < 60h^{-1}\text{kpc}$. As in Figure 4, we show the distributions for surveys without fibre collisions (black), surveys having fibre collisions and no correction (blue, dashed) and surveys with collisions and a correction (blue, solid).

there is a flattening of the profile around the fibre collision scale which is at $\sim 60h^{-1}\text{kpc}$ at the maximum redshift of our survey and shown by the black dashed line. Finally, the blue solid line is the surface density estimate if the weight of each secondary is set to the product of w_s of itself and its primary and the weight of each host is the primary weight. We see that this correction rectifies the overall decrease and also most of the effects below the fibre collision scale.

In addition to biasing the observed radial profile, fibre collisions also slightly alter the average host halo masses of observed primaries. The reason is that more massive haloes have more satellites and thus a higher probability to have a fibre collision between a primary and a secondary. Since we exclude primaries with fibre collisions, this biases our halo mass distribution at a fixed primary luminosity to lower masses. This effect is shown in Figure 5 where we show the line-of-sight velocities of the same secondaries as in Figure 4, excluding those below $60h^{-1}\text{kpc}$. The measured velocity dispersion of secondaries if not correcting for fibre collisions is $\sim 375\text{km/s}$, whereas it is $\sim 388\text{km/s}$ without fibre collisions or with our correction applied. Thus not correcting for fibre collisions results in a velocity dispersion that is artificially lowered by around 5%. For other luminosity bins, we find results that are qualitatively and quantitatively similar.

Finally, we note that the corrections discussed above have not been applied in most other satellite kinematic studies (McKay et al. 2002; Prada et al. 2003; Brainerd & Specian 2003; Conroy et al. 2007; More et al. 2009a, 2011). From Figure 4 it is apparent that neglecting those corrections can lead to wrong inferences about the surface density profile of satellites. Particularly, this casts doubt on the result by More et al. (2009a) that satellites are strongly anti-biased with respect to dark matter. This is significant because the surface density profile is also strongly correlated with the velocity dispersion at a fixed halo mass via the Jeans equation. See the discussion in §8 for more details on how the assumed radial profile affects the observables. Thus, the more extended radial profile assumed by More et al. (2009a, 2011) could systematically underestimate halo masses. Additionally, as we have shown there is an additional bias in the velocity dispersion of $\sim 5\%$ if not correcting for fibre collisions. This will further lead to an underestimation of average halo masses. Altogether, not correcting for fibre collisions likely causes systematic errors in the analysis of More et al. (2009a, 2011).

5.4 Definition of constraints

We now describe the various observational constraints that we use in our analysis of satellite kinematics. In addition to the host- and satellite-weighted velocity dispersions around red and blue primaries, this includes the abundances of secondaries around red and blue primaries, the fraction of primaries that are red, and the overall number density of galaxies. The latter is important to constrain the overall normalization of the CLF. For example, the best-fit model of More et al. (2011), who did not include the overall abundance of galaxies as a constraint, implies⁶ a number density of galaxies brighter than $10^{9.5}h^{-2}L_{\odot}$ that is $\sim 2.8 \times 10^{-2}h^3\text{Mpc}^{-3}$. However, the observed number density is only $\sim 1.6 \times 10^{-2}h^3\text{Mpc}^{-3}$ (Zehavi et al. 2011; Guo et al. 2015b).

Given our volume-limited sample, the number density

⁶ We have computed this using `HMFcalc` (Murray et al. 2013), using the mass function of Tinker et al. (2008) and the cosmology adopted by More et al. (2011).

in a luminosity bin defined by L_1 and L_2 is

$$n_{\text{gal}} = \frac{\sum w_{s,i}}{\frac{\Omega_{\text{SDSS}}}{3} [d_{\text{com}}^3(z_{\text{max}}) - d_{\text{com}}^3(z_{\text{min}})]}. \quad (19)$$

Here $d_{\text{com}}(z)$ is the comoving distance out to redshift z , the sum goes over all galaxies with $L_1 \leq L \leq L_2$, the angular sky coverage of our sample is $\Omega_{\text{SDSS}} = 2.27$ steradians, and $w_{s,i}$ is the weight of galaxy i used to compensate for spectroscopic incompleteness, as discussed above. Throughout this paper we use 10 luminosity bins of 0.15 dex width covering the range $\log[L/(h^{-2} L_{\odot})] = [9.5, 11.0]$.

The fraction of red primaries in each luminosity bin, $f_{\text{pri,r}}$, is calculated according to

$$f_{\text{pri,r}} = \frac{\sum_{\text{red primaries}} w_{s,i}}{\sum_{\text{all primaries}} w_{s,i}}. \quad (20)$$

An important constraint on the CLF is the number of satellite galaxies per central. This is related to the average number of same-halo secondaries per primary, which we estimate as

$$n_s = \frac{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i}}{\sum_{\text{primaries}} w_{s,i}}. \quad (21)$$

Here the sum in the numerator goes over all secondaries having primaries in that particular bin of luminosity and colour, while the sum in the denominator goes over all primaries of that bin. $p_{\text{mem},i}$ is the membership probability of secondary i , obtained from fitting the $\Delta v_z - R_p$ distribution in each primary bin using the method described in §5.2, and w_{sw} is the weight assigned to each secondary, which is equal to the product of the spectroscopic incompleteness weights of the secondary and the corresponding primary, i.e.,

$$w_{\text{sw}} = w_{s,p} w_{s,s}, \quad (22)$$

Following More et al. (2009b), we quantify the kinematics of satellite galaxies using both the satellite- and host-weighted velocity dispersions, σ_{sw} and σ_{hw} , respectively. The former is computed by giving equal weight to each central-satellite pair, while the latter assigns equal weight to each central. Since more massive haloes, on average, contain more satellites, the ratio $\sigma_{\text{sw}}/\sigma_{\text{hw}}$ contains information regarding the scatter in the galaxy-halo relation. In particular, $\sigma_{\text{sw}}/\sigma_{\text{hw}}$ increases with increasing amount of scatter. We estimate the satellite-weighted velocity dispersion as

$$\sigma_{\text{sw}}^2 = \frac{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i} \Delta v_{z,i}^2}{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i}} - \sigma_{\text{err}}^2. \quad (23)$$

Here $\sigma_{\text{err}} = \sqrt{2} \times 35 \text{ km s}^{-1}$ is the error on the velocity difference between primary and secondary arising from the redshift uncertainties in SDSS (e.g., More et al. 2009a). In the case of host-weighting, we use the same equation, but with w_{sw} replaced by a new weight

$$w_{\text{hw}} = \frac{w_{s,p} w_{s,s}}{N_{\text{scd}}}. \quad (24)$$

Here N_{scd} is the number of secondaries hosted by each primary. Note that, when using these weights, we also use new membership probabilities, $p_{\text{mem},i}$, that are appropriate for calculating host-weighted quantities. These are computed

using equation (18), but with the weights in equation (17) replaced by w_{hw} .

Finally, we note that we only use the resulting n_s , σ_{sw}^2 and σ_{hw}^2 for primary bins with at least an estimated 10 satellites, i.e., with $\sum p_{\text{mem}} > 10$.

6 ANALYTICAL MODEL FOR CONSTRAINTS

Ideally one would want to forward-model all relevant observational effects to compute the expectation values of the observables for a given set of model parameters. However, as discussed in §2, this is computationally infeasible given the large number of likelihood evaluations needed to estimate parameter uncertainties. Instead, we estimate the expectation values for a given set of model parameters using an analytical model that is similar to that used in van den Bosch et al. (2004) and More et al. (2009b,a, 2011), and described below.

6.1 Galaxy number density

The total number density of galaxies in a given luminosity bin defined by L_1 and L_2 can be straightforwardly calculated from the CLF according to

$$n_{\text{gal}}(L_1, L_2) = \int_{L_1}^{L_2} \int_0^{\infty} \Phi_{\text{tot}}(L|M) n_h(M) dM dL, \quad (25)$$

where n_h is the halo mass function. In this work, we extract the mass function directly from the SMDPL $z = 0.05$ simulation in bins of 0.1 dex.

6.2 Red fraction of primaries

We estimate the expectation value for the red fraction of primaries as the red fraction of centrals:

$$f_{\text{pri,r}}(L_1, L_2) \approx \frac{n_{c,r}(L_1, L_2)}{n_{c,r}(L_1, L_2) + n_{c,b}(L_1, L_2)}. \quad (26)$$

Here $n_{c,r}(L_1, L_2)$ and $n_{c,b}(L_1, L_2)$ are the total number density of red and blue centrals, respectively, which we compute using equation (25), but with $\Phi(L|M)$ replaced with $f_r(M)\Phi_{c,r}(L|M)$ and $f_b(M)\Phi_{c,b}(L|M)$, respectively.

This estimate underlies the assumptions that all primaries are centrals, and that the completeness of red primaries is identical to that of blue primaries. As we have shown in the middle panel of Fig. 2, the impurity of our sample, defined as the fraction of primaries that are not centrals, is very low, typically below 1%. In addition, the left-hand panel of that figure shows that indeed the completeness of centrals has only an extremely mild colour dependence. Hence, equation (26) should give a reliable estimate for the red fraction of primaries.

6.3 Number of same-halo secondaries

Using the same arguments as for the red fraction of primaries, we can approximate the average number of same-

halo secondaries around primaries by the number of satellites around centrals. Thus,

$$n_s(L_1, L_2) \approx \frac{\int_{L_1}^{L_2} \int_0^\infty N_s(M) n_h(M) \Phi_c(L|M) f_{\text{ap}}(L, M) dM dL}{n_c(L_1, L_2)} \quad (27)$$

where $f_{\text{ap}}(L, M)$ is the probability of halo members to fall within the aperture, i.e. inside the hollow cylinder defined by R_s , R_c and $(\Delta v)_s$, as discussed in §5.1. Under the assumption that the primary is indeed the central, this probability is simply the expected fraction of satellites that lie in the cylinder centred on the halo. Given that $(\Delta v)_s$ is much bigger than the extent of the halo in redshift space, we have that

$$f_{\text{ap}}(L, M) = 4\pi \int_0^\infty \bar{n}_{\text{sat}}(r|M) [\zeta(r, R_s(L)) - \zeta(r, R_c)] r^2 dr, \quad (28)$$

with $\bar{n}_{\text{sat}}(r|M)$ the average radial profile of satellites around haloes of mass M , normalized such that

$$4\pi \int_0^\infty \bar{n}_{\text{sat}}(r|M) r^2 dr = 1, \quad (29)$$

and

$$\zeta(r, R) = \begin{cases} 1 & \text{if } r \leq R \\ 1 - \sqrt{1 - R^2/r^2} & \text{otherwise.} \end{cases} \quad (30)$$

Note that this neglects the (small) possibility that some primaries are satellites. Lange et al. (2018) have shown that this leads to negligible differences. Finally, $\bar{n}_{\text{sat}}(r|M)$ is the average radial profile around halos of mass M , which we approximate as a generalized NFW profile (Equation [11]) with the median concentration, c_{vir} , as measured in the appropriate halo catalogue.

6.4 Velocity dispersion

The observed velocity dispersion of same-halo secondaries can be calculated analogously to their number. We approximate

$$\sigma^2(L_1, L_2) \approx \frac{\int_{L_1}^{L_2} \int_0^\infty w(L, M) \sigma_{\text{ap}}^2(L, M) n_h(M) \Phi_c(L|M) dM dL}{\int_{L_1}^{L_2} \int_0^\infty w(L, M) n_h(M) \Phi_c(L|M) dM dL}. \quad (31)$$

w is a weight function. In the case of the satellite-weighted velocity dispersion the weight function is

$$w_{\text{sw}}(L, M) = f_{\text{ap}}(L, M) N_s(M), \quad (32)$$

the expected number of satellites inside the aperture. For the host-weighted velocity dispersion the weight function, assuming a Poisson distribution for the satellite occupation, is

$$w_{\text{hw}}(L, M) = 1 - \exp[-f_{\text{ap}}(L, M) N_s(M)], \quad (33)$$

the probability of having at least one same-halo secondary inside the aperture. Finally, σ_{ap}^2 is the average velocity dispersion of satellites inside the aperture and depends on L , and M . This can be calculated similarly to $f_{\text{ap}}(L, M)$ in equation (28) using equation (12),

$$\sigma_{\text{ap}}^2(L, M) = \frac{\int_0^\infty \bar{n}_{\text{sat}}(r|M) \sigma^2(r|M) [\zeta(r, R_s(L)) - \zeta(r, R_c)] r^2 dr}{\int_0^\infty \bar{n}_{\text{sat}}(r|M) [\zeta(r, R_s(L)) - \zeta(r, R_c)] r^2 dr}. \quad (34)$$

Again, we have neglected the spread in halo concentrations and have used the median halo concentration for each halo mass.

6.5 Accuracy of the analytic model

We can test the accuracy of our analytic model by comparing it to Monte-Carlo simulations. We populate the SMDPL simulation with galaxies according to our default galaxy occupation model and a satellite phase-space distribution implied by $\gamma = \mathcal{R} = 1$. The following conclusions are qualitatively very similar for the other two radial profiles. We create mock galaxy surveys according to the recipe described in §4, which we analyze using the procedure outlined in §5. We do this 1000 times and calculate the average and spread for all constraints among these different realizations. For each realization we completely re-draw the galaxy occupation and luminosities for each halo, calculate new phase-space positions and place the observer at a new random position inside the simulation box. We only use those constraints which could be measured for all 1000 realizations. Furthermore, we exclude measurements for $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$ if $n_s < 0.01$ on average. The reason is that the resulting ratios are very non-Gaussian and could bias our inferences if we were to approximate them as being Gaussian. In any way, $\sigma_{\text{hw}}/\sigma_{\text{sw}} \approx 1$ if $n_s \ll 1$, so we do not expect these measurements to contain much information.

Figure 6 compares the average values of the constraints obtained in this forward-modelling fashion with the analytic predictions. The main panels directly show differences between the two, whereas the smaller panels show the absolute values. Error bars denote the 68% scatter in those 1000 mock catalogues. We first note that the overall qualitative agreement is good as all salient trends are recovered by the analytic prediction. In the following, we will briefly discuss some of the shortcomings.

The number density of galaxies is basically recovered perfectly in the analytic model. This is not surprising as both the analytic model and the mock catalogues use the halo mass function of SMDPL. The analytic model slightly, but not significantly, overpredicts the red fraction of primaries. The reason is that red central galaxies reside, on average, in more massive haloes, which are more strongly clustered. This explains why red centrals have a slightly lower probability than blue centrals to be considered isolated (cf. left-hand panel of Fig. 2). Also, the analytic model underpredicts the average number of same-halo secondaries for both red and blue primaries. The main reason for that is that our analysis procedure in §5 does not perfectly reject interlopers.

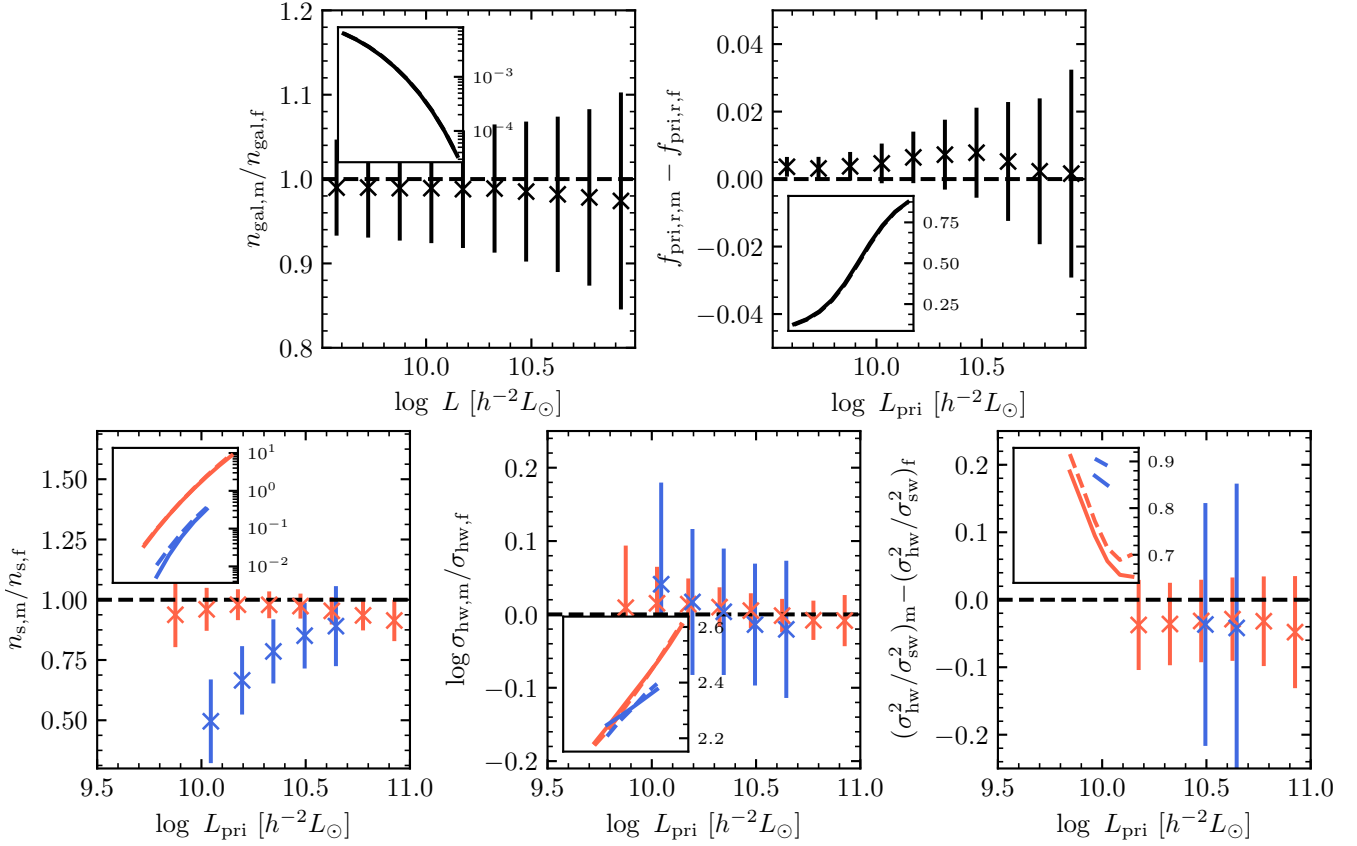


Figure 6. Comparison of the different constraints derived from direct forward modelling and our simple analytical model. The main panels directly highlight differences between the two. Error bars show the 1σ spread observed in 1000 mock catalogues. The satellite observables in the lower row have been measured for red and blue primaries separately, as indicated by the colour of the error bars. In all cases, the dashed line indicates perfect agreement between the analytic model and the full forward-modelling approach. Finally, each smaller panel shows the absolute values of those observables for the analytic model (solid) and the forward-modelling result (dashed).

Particularly, we assume $P(\Delta v_z)$ of interlopers to be flat, whereas they are slightly more abundant at $\Delta v_z \sim 0$ (see van den Bosch et al. 2004; More et al. 2009a). Therefore, the membership probabilities for $\Delta v_z \sim 0$ secondaries and, accordingly, the overall number of satellites are overestimated. Given their much higher interloper fraction (cf. right-hand panel of Fig. 2), it is also evident why this effect is strongest for faint blue primaries. The velocity dispersions are generally recovered well qualitatively except for low-luminosity blue primaries. The underlying problem here is that the majority of secondaries around these primaries are actually interlopers instead of true halo members. Since their $P(v_z)$ is not flat, as discussed above, we cannot expect the analysis procedure in §5 to recover the true velocity dispersion. Finally, the model generally predicts slightly higher velocity dispersions than the one observed. As discussed in §5, this is primarily a shortcoming of our interloper rejection and not due to the analytic model. Finally, the ratio of host- to satellite-weighted velocity dispersion is underestimated by at most around $\sim 0.5\sigma$.

When using the analytic model to fit observations, we apply a correction to account for observational biases and biases in the model. This is discussed in more detail in the next section.

7 ANALYSIS PROCEDURE

When applying the analysis procedure described in §5 to a spectroscopic catalogue we extract several observables. Particularly, we use n_{gal} , $f_{\text{pri},r}$, n_s , $\log \sigma_{\text{hw}}$ and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$ as our data points from which we construct the data vector \mathbf{D} . Note that the latter 3 quantities are measured for red and blue primaries separately. We have chosen $\log \sigma_{\text{hw}}$ and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$, as opposed to for example σ_{hw} and σ_{sw} , because their distributions in a large number of mocks with the same input models can be better, but not perfectly, described by Gaussian distributions.

We construct covariance matrices using 1000 mock catalogues derived from the same input model. In this paper, the covariance matrix is always constructed using the default parameters described in Table 1. An estimate $\hat{\mathbf{C}}$ for the covariance matrix is calculated via

$$\hat{\mathbf{C}} = \frac{1}{N_S - 1} \sum_i^{N_S} (\mathbf{D}_i - \langle \mathbf{D} \rangle)(\mathbf{D}_i - \langle \mathbf{D} \rangle)^t, \quad (35)$$

where $N_S = 1000$ is the number of samples, \mathbf{D}_i is the data vector of the i -th mock and $\langle \mathbf{D} \rangle$ the average for all mock catalogues. In Figure 7 we show the correlation matrix for all observables. For this particular matrix, we used mocks where satellites follow an NFW profile, $\gamma = \mathcal{R} = 1$. Covariance

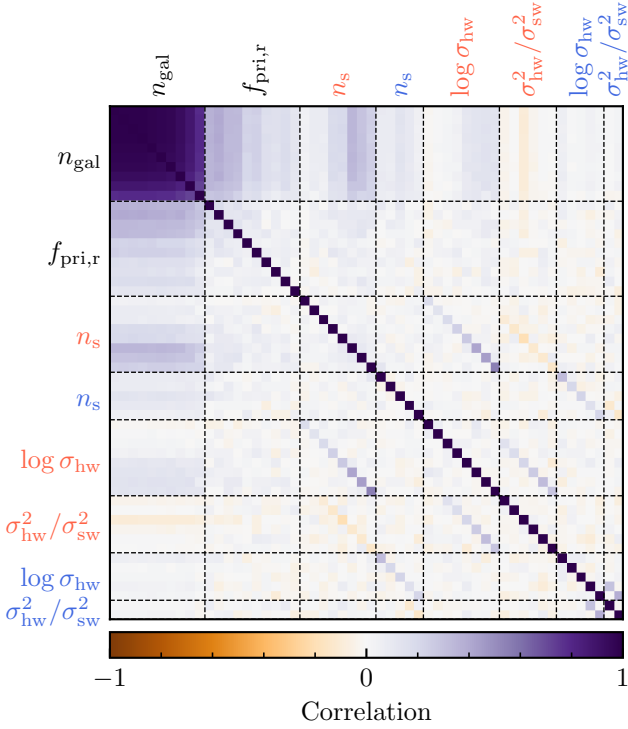


Figure 7. The correlation matrix of different observables when populating the mocks according to the default model and having satellites follow an NFW profile. The correlation matrix has been obtained from 1000 mock catalogues. The dashed lines separate the different kinds of observables, whereas each single square represents this kind of observable for a certain bin in primary luminosity. The satellite statistics are measured for red and blue primaries separately, as indicated by the colour of the labels.

matrices for other profiles are very similar. Note that we only show observables that can be measured for all mocks.

The first thing to note is the very strong correlation for the n_{gal} measurements in different bins. This well known cosmic variance (see e.g. Moster et al. 2011; Smith 2012) implies that the overall number density of galaxies of all luminosities has a large uncertainty, whereas the “shape” of the luminosity function is very well constrained. In fact, when fitting a model to observational data or mock catalogues, we will add an additional 2% error, i.e. $(0.02n_{\text{gal}})^2$, to the diagonal elements of the luminosity function covariance matrix. This reduces the strong covariance and allows the abundances of galaxies of different luminosities to vary independently. We do this primarily to not bias our inference on the galaxy-halo connection in case our parametrization, as described in §3, cannot perfectly reproduce the observed luminosity function. This is primarily an issue for observational data where we also have systematic uncertainties like the photometric pipeline that can affect these observables. Next, the first column in Figure 7, which depicts the correlation of n_{gal} with other observables, shows that n_{gal} is weakly, positively correlated with n_s and $\log \sigma_{\text{hw}}$, and weakly anti-correlated with $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$. These trends are expected because higher values of n_{gal} should be associated with over-dense regions. Consequently, the average dark matter halo masses

will be slightly higher (e.g., Mo et al. 2004), even at fixed primary luminosity. Thus, n_s and $\log \sigma_{\text{hw}}$ are expected to increase and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$ to decrease. Finally, there is also a correlation of n_s with $\log \sigma_{\text{hw}}$ and an anti-correlation of n_{gal} and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$ at fixed colour and luminosity of the primary. If all primaries of a given colour and luminosity live in slightly more massive haloes, we expect n_s and $\log \sigma_{\text{hw}}$ to be higher and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$ to be lower.

We note that σ_{hw} and σ_{sw} in the same primary bin are always highly positively correlated (not shown). For example, for low-luminosity samples with primaries having only 0 or 1 secondaries, $\sigma_{\text{hw}} \equiv \sigma_{\text{sw}}$. Thus, there should be no extra information in measuring σ_{sw} over just σ_{hw} alone. In their analysis, however, More et al. (2009a, 2011) implicitly assumed those measurements to be independent. Thus, especially for low-luminosity primaries where σ_{hw} and σ_{sw} are highly correlated, this should lead to an underestimation of the uncertainty in the average halo mass. On the other hand, the uncertainties in the spread in halo mass at fixed luminosity which is characterized by the ratio of σ_{hw} and σ_{sw} is likely overestimated. Thus, we expect better constraints on the scatter in luminosity at fixed halo mass, σ_r and σ_b by using the full covariance matrix.

The precision matrix Ψ is the inverse of the covariance matrix C . In case of a noisy estimate of the latter, the unbiased estimator $\hat{\Psi}$ for the precision matrix becomes

$$\hat{\Psi} = \frac{N_S - N_D - 2}{N_S - 1} \hat{C}^{-1}, \quad (36)$$

where N_D is the number of data points (Taylor et al. 2013).

We have shown in §6 that our analytic model has some small biases with respect to the full-forward modelling approach for a given set of parameters. We assume that those biases do not change substantially for models allowed by the data. Then we can correct for those systematics by calculating the expectation value for a given set of parameters θ for all observables via

$$M^*(\theta) = M(\theta) + B(\theta_0), \quad (37)$$

where

$$B(\theta_0) = \langle D \rangle(\theta_0) - M(\theta_0) \quad (38)$$

is the bias between the average data vector of the 1000 mock catalogues and the simple analytic model for the default parameter set θ_0 .

Now, given a set of observables D extracted from a mock catalogue we can calculate the likelihood via

$$\mathcal{L} \propto \exp \left[-\frac{(M^*(\theta) - D)^t \Psi (M^*(\theta) - D)}{2} \right] = \exp \left[-\chi^2/2 \right]. \quad (39)$$

We use **MultiNest** to calculate the resulting posterior. We assume flat priors in all parameters with the ranges shown in Table 1. α_s which regulates the shape of the satellite CLF are kept fixed as it is basically unconstrained by our data. We run **MultiNest** with 10,000 live points and a target sampling efficiency of 1%. Constant efficiency mode is turned off. We use $\Delta \ln \mathcal{Z} = 10^{-4}$ as a stopping criterion, where $\Delta \ln \mathcal{Z}$ is the uncertainty in the estimate of the global Bayesian evidence. We have tested that our results are converged by running **MultiNest** once with 50,000 live points

and a target efficiency of 0.5% and getting comparable results. **[Johannes: This still needs to be done. It's running right now. 04/24/18]**

Finally, we note that we have created an estimate for the covariance matrix and the bias for the default parameter set in Table 1. This makes sense since the mock catalogues that we test our method on, as described in the next section, have the same parameters. However, when applying our method to observations, the input parameters are unknown. This problem can be solved in an iterative fashion. One can start out with the default model used here, calculate an estimate for the bias and the covariance and use this to get a new best-fit model. Afterwards, one then calculates a new bias and covariance for this new model and so on. We find that this algorithm converges sufficiently after only 3 iterations when applied to SDSS (Lange et al., in preparation).

8 APPLICATION TO MOCKS

We now validate our analysis procedure by applying it to mock catalogues with known input parameters. In all cases, those input values are the ones listed in Table 1.

8.1 Mocks with known analytical phase-space distributions

First, we apply our analysis procedure to mock catalogues where satellites obey analytical radial density profiles and the spherically averaged Jeans equation without anisotropy, as described in §3.2. Furthermore, we assume that we know the true underlying radial profile of satellites perfectly. Thus, we always use the same radial profile to model the data as we used to create the mock catalogue in question.

In Figure 8 we show the best-fit model when fitting to a mock SDSS catalogue where satellites follow an NFW profile, i.e. $\mathcal{R} = \gamma = 1$. As shown, we can fit all observations in the mock catalogue very well with an overall χ^2 of 28.7 for $55 - 15 = 40$ degrees of freedom. The posterior predictions for the galaxy-halo connection inferred from the same mock catalogue are shown in Figure 9. We show the input model as dashed lines and error bands denote the 68% and 95% posterior prediction. We see that this analysis succeeded in recovering the input model with all posteriors including the input to within $\sim 2\sigma$. We have also looked at all 18 free model parameters and their $18 \times 17/2 = 153$ two-dimensional distributions and also found a good agreement. Finally, we have applied our analysis procedure to mocks produced with the other two radial profiles, $\gamma = 1$, $\mathcal{R} = 2.0$ and $\gamma = 0$, $\mathcal{R} = 2.5$, and were also able to recover the input parameters to within reasonable uncertainties.

Let us further investigate how competitive our constraints are. Cacciato et al. (2013) constrained both the galaxy-halo connection and cosmological parameters using observations of the luminosity function, galaxy clustering and galaxy-galaxy lensing. For example, they were able to constrain the scatter in luminosity at fixed halo mass to within ± 0.007 dex. This is comparable to our constraint on the scatter for red centrals of ± 0.010 dex. Finally, Cacciato et al. (2013) constrain the median luminosity at $\log M/(h^{-1}M_\odot) = 13$ to within ~ 0.02 dex, similar to our constraints. Note however that this comparison

is only approximate given that Cacciato et al. (2013) also constrained cosmological parameters and that our mock catalogues might be substantially different from the actual Universe. Nevertheless, overall our constraints using only satellite kinematics and galaxies with $0.02 \leq z \leq 0.067$ and $\log L > 9.5$ seem to be competitive to results from clustering and lensing in the range $0.011 \leq z \leq 0.245$ and $\log L > 8.7$ in SDSS.

8.2 Mocks with unknown analytical phase-space distributions

So far, we have assumed that the underlying radial profile of satellites is known a priori. This will not be the case for actual observations as there is still debate whether satellite galaxies follow the dark matter distribution (e.g., van der Marel et al. 2000; van den Bosch et al. 2005a; Tal et al. 2012; Cacciato et al. 2013; Guo et al. 2015a) or are more radially biased (e.g., Yang et al. 2005; Chen 2008; More et al. 2009a; Hoshino et al. 2015).

We show in Figure 10 how the different constraints are impacted by the choice of the radial profile. For this plot all parameters describing the galaxy occupation are fixed to their default values. Only \mathcal{R} and γ describing the radial profile of satellites have been changed. For this plot, all observables have been predicted from the simplified analytical model. The number density of galaxies and the red fraction of primaries are not shown in this comparison as they have no dependence on the radial profile of satellites. The number of satellites inside the aperture generally decreases for more extended radial profiles, e.g. $\mathcal{R} = 2.5$ and $\gamma = 0$. The only exception is for low-luminosity primaries where a substantial amount of satellites are within $60 h^{-1}\text{kpc}$ and not counted due to potential issues with fibre collisions, as discussed in §5.3. The velocity dispersion always increases for more extended radial profiles, as expected from Figure 1. The difference is ~ 0.05 dex between an unbiased profile, $\mathcal{R} = \gamma = 1$, and the most radially anti-biased profile, $\mathcal{R} = 2.5$ and $\gamma = 0$, irrespective of the luminosity or colour of the primary. Finally, the ratio of host- and satellite-weighted velocity dispersion is generally lower for radially more concentrated profiles. The reason is that for those profiles the high-mass halos *at a fixed luminosity* have a larger fraction of their satellites inside the aperture, thereby increasing the satellite-weighted velocity dispersion. Altogether, the different radial profiles predict significantly different observables for a fixed galaxy-occupation model. Thus, choosing a realistic model for the spatial distribution of satellites is important to get an unbiased prediction for the galaxy-halo connection.

So far, we have not used observations that would directly constrain the radial profile of satellites. Additionally, we keep γ and \mathcal{R} fixed when fitting the data. The reason for this is mainly that if one changes the radial profile, one would also need to re-evaluate the Jeans equation to solve for the velocity dispersion. This would be impractical for the large number of likelihood evaluations that are needed to constrain the model. However, once good fitting models for all three different radial profiles in this analysis have been found, we can create mock catalogues and directly compare the resulting observed radial profile to the observations.

We test this approach with a mock catalogue where

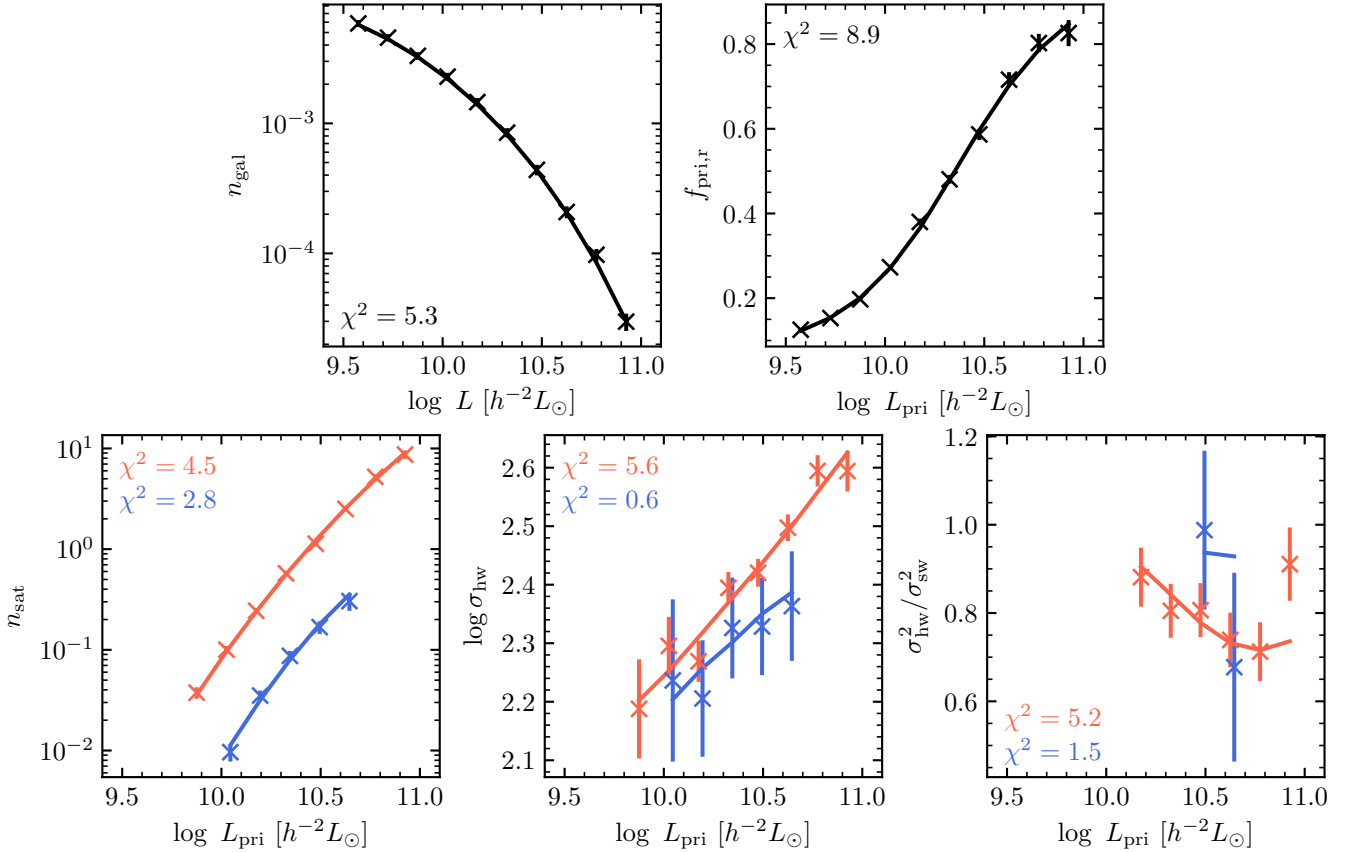


Figure 8. The prediction of the best-fitting model against the measured observables in a mock catalogue where satellites follow an NFW profile and predictions of the isotropic Jeans equation. Error bars denote the 1σ uncertainty derived from mock catalogues. In each panel, we also show the χ^2 of each set of observables. Note however that there is a weak correlation between different observables and the total χ^2 is not the some of all individual ones.

we assume that the underlying radial profile of satellites is unknown. For simplicity, it is the same one as analysed in the previous subsection (i.e., with $\gamma = \mathcal{R} = 1$). In Figure 11, we show the projected surface number density of secondaries around primaries of different luminosities, where blue data points are values from that specific mock catalogue. We compare this to mock catalogues with known radial distributions of satellites as bands. Because we find that the best-fitting parameters in all cases are relatively close to the input parameters, we have not created new mock catalogues and show the 68% range from the default parameter set, instead. [Frank: This is becoming a bit confusing. I am left pondering what exactly has been done. We should rewrite this somewhat] We make no attempt to correct for interlopers, instead we only consider secondaries with $|\Delta z| < 3\sigma$, where σ is given by equation (14). A correction for fibre collisions has been applied in all mock catalogues and secondaries with $R_p < 60 h^{-1}\text{kpc}$ are also included in this analysis. Based on this analysis, we would infer that galaxies in this mock catalogue follow the NFW profile of the dark matter, i.e. $\mathcal{R} = \gamma = 1$. Thus, we would only consider the best-fitting parameters and uncertainties for the galaxy-halo connection under this assumption. Thus, as in the previous subsection, we would get a good agree-

ment between the derived and the input parameters for the galaxy-halo relation.

Finally, we note that when analysing the radial profile we assumed that BHGs always reside in the centre. This assumption is not true in general and one might worry that violations of this assumption alter the observed radial profile of secondaries around primaries, i.e. BHGs. However, Lange et al. (2018) have shown that the effect on the projected number density of secondaries is negligible.

8.3 Mocks with unknown complex phase-space distributions

We have so far assumed that satellites have a spherically isotropic distribution and obey the Jeans equation without anisotropies. In many ways, this is a gross simplification. For example, subhaloes and satellite populations can have various degrees of non-sphericity and there is a substantial amount of substructure in phase-space. Furthermore, dark matter haloes are not fully relaxed. [Frank: We need to add citations here to studies that have explicitly demonstrated this] Thus, there is no a priori reason to assume that they would obey the Jeans equation. Additionally, we have neglected higher order moments of the velocity distribution and velocity anisotropy in our calculation. Here

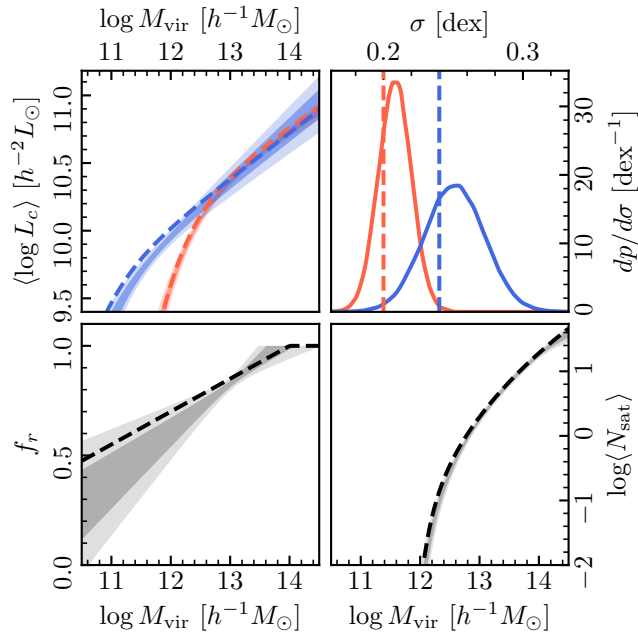


Figure 9. The posterior predictions for the galaxy-halo connection when analysing a mock catalogue where satellite galaxies follow an NFW profile and obey the isotropic Jeans equation. Bands show the 1 and 2σ regions and the dashed lines input values. The solid lines in the upper right panel is a Gaussian kernel density estimate of the posterior. We show the average luminosity as a function of halo mass for red and blue centrals (top, left), the logarithmic spread in luminosity at fixed halo mass (top, right), the red fraction of centrals as a function of mass (bottom, left) and the satellite occupation (bottom, right). Note that we eliminated satellites that are brighter than centrals. Thus, the actual average satellite occupation will have a small dependence on the colour of the central. We have omitted that complication for clarity and calculated the satellite occupation as if satellites brighter than centrals were not removed.

we test to whether these complications influence our inferences regarding the galaxy-halo connection at a level that is relevant compared to current statistical uncertainties.

To evaluate this we create a mock catalogue where satellite galaxies are placed on resolved dark matter subhaloes in SMDPL. This addresses all the potential issues discussed above. We use the `SubhaloPhaseSpace` module of `halotools` to place satellites on subhaloes. We first determine for each dark matter halo the number of satellites that it hosts, according to the recipe described in §3 and regardless of how many dark matter subhaloes we actually find in SMDPL. Then we place the satellites on those subhaloes with the highest M_{peak} , the maximum dark matter halo mass achieved over the lifetime of each subhalo. In rare cases where we have more satellites than subhaloes in the same halo, we use the relative phase-space positions of random subhaloes hosted by haloes of a similar mass. We then proceed to generate a mock SDSS-like catalogue and analyse it in the same way as the one in the previous section.

We start by analysing the radial profile of satellites, as shown by the red crosses in Figure 11. When comparing the mock catalogue to other mock catalogues with analytical phase-space profiles, we see that it most closely resembles

the most radially anti-biased profile, $\mathcal{R} = 2.5, \gamma = 0$. This is another manifestation of the well-known result that subhaloes are spatially anti-biased with respect to dark matter (see e.g. Diemand et al. 2004; van den Bosch et al. 2016). Thus, for our posterior prediction on the galaxy-halo connection, we would choose results when assuming $\mathcal{R} = 2.5, \gamma = 0$. In this case, we find a good fit with $\chi^2 = 40.9$ for $59 - 15 = 44$ degrees of freedom.

In Figure 12 we show the posterior predictions when analysing the above mentioned mock catalogue assuming $\mathcal{R} = 2.5, \gamma = 0$. This Figure is analogous to Figure 9. Also, similar to Figure 9, we find a good agreement of our posterior prediction with the input model. We have repeated this above experiment two more times (not shown), finding similar results. However, our results suggest that average luminosity of centrals living in high-mass haloes might be slightly underestimated. We find those results in our posterior at the ~ 2 , ~ 2 and $\sim 3\sigma$ level for the three runs.

8.4 Caveats

Altogether, we have thus shown that certain simplifications in our model, particularly assuming the radially symmetric Jeans equation without anisotropy, might only have a very small effect on our inferred galaxy-halo relation. We note however, that the very shallow radial profile implied by placing satellites on subhaloes seems at odds with observations. Indeed, Lange et al. (in preparation) find that satellites are much less radially biased compared to dark matter. Thus, it is not guaranteed that satellite galaxies obey the Jeans equation to a similar degree (see also Ye et al. 2017). It is also worth keeping in mind that other simplifications in our model, for example that all BHGs are centrals, might cause small biases in our inferences (Lange et al. 2018). Additionally, it is worth noting that we only analyse a subset of all centrals (compare left panel of Fig. 2). Particularly, our isolation criterion will likely miss centrals in high density regions. Faltenbacher (2010) have shown that the dynamics of subhaloes vary systematically with environment at fixed halo mass. Haloes at fixed mass in low density regions have a lower subhalo velocity dispersion. Confirming the results by Faltenbacher (2010), we find that haloes whose central is identified as primaries have a slightly lower velocity dispersion for their subhaloes. However, this effect is only significant for lower halo masses and practically disappears for $M_{\text{vir}} > 10^{13} h^{-2} M_{\odot}$. We also stress that our final posteriors need to take into account uncertainties in the observed radial profile. This could be done by considering the most shallow and steep profiles still allowed by the data and combining their posteriors. Finally, given how competitive our constraints are compared to a combination of clustering and lensing (Cacciato et al. 2013), it seems likely that our results are also impacted by cosmological parameters (see also Li et al. 2012). However, a framework that also allows for variations in the cosmology is beyond the scope of this work.

9 CONCLUSION

Satellite kinematics are a powerful probe of the galaxy-halo connection, which have hitherto been rarely utilized, except

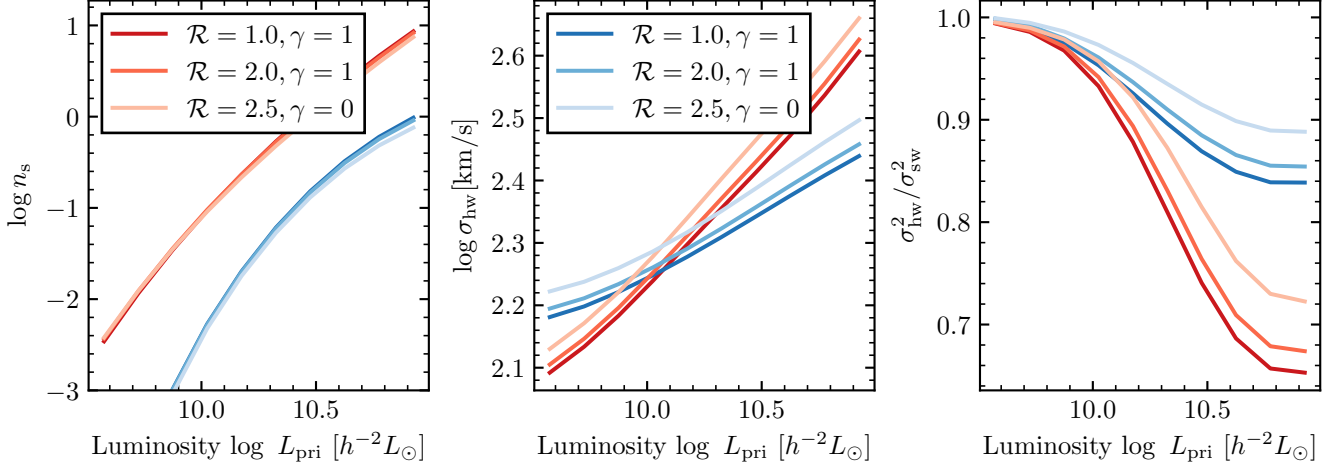


Figure 10. The dependence of the constraints used in this work on the radial profile if satellites obey the spherically averaged Jeans equation without anisotropy. All observables have been predicted from the analytic model. In each panel, we show the predictions with lines of different saturation. Darker lines indicate more radially anti-biased profiles. The number density of galaxies and the red fraction of primaries are not shown since they have no direct dependence on the radial profile of satellites.

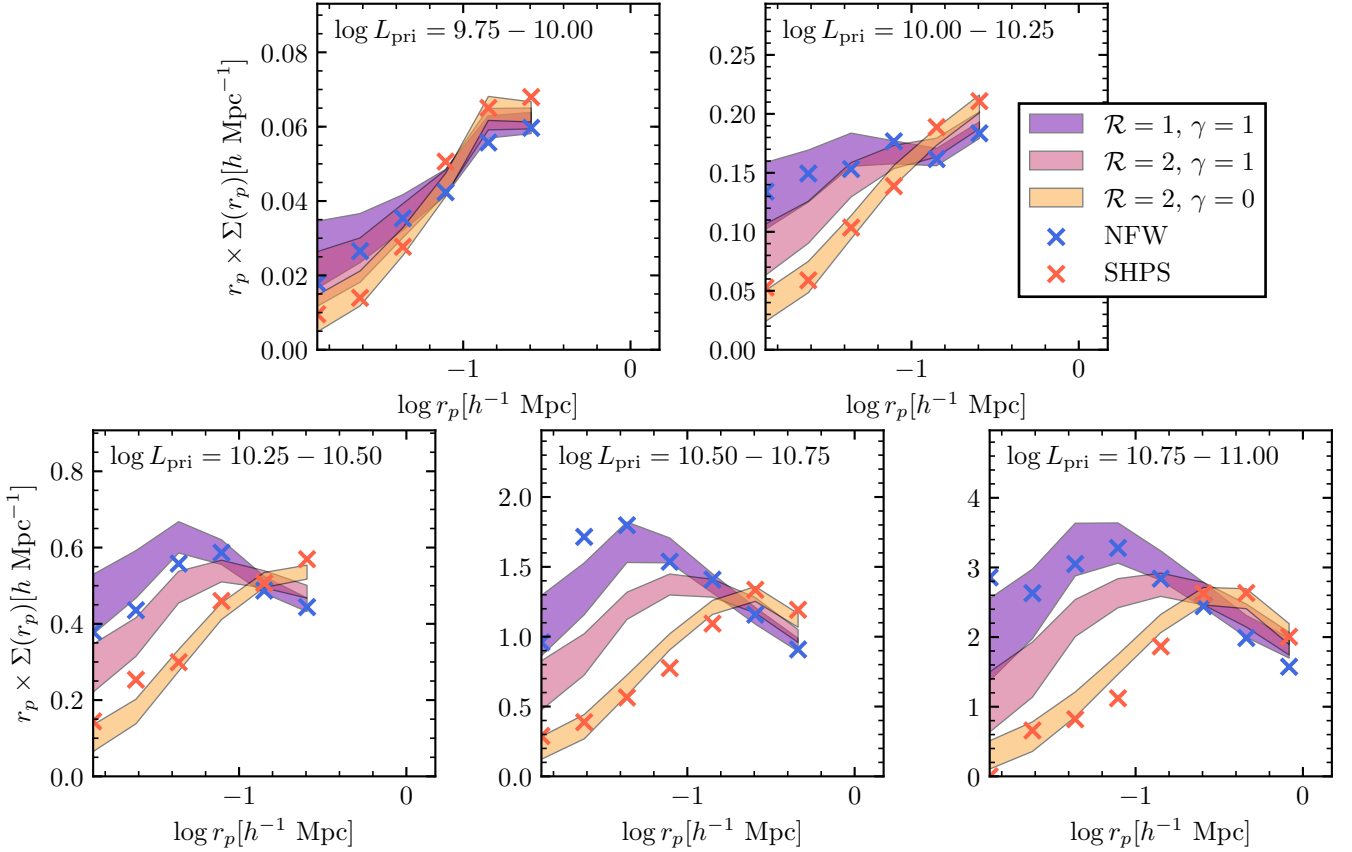


Figure 11. The observed radial profile in different mock catalogues. In each panel we show the projected surface density of secondaries around primaries in a given luminosity bin, as indicated in the top right corner. We only include secondaries with $|\Delta z| < 3\sigma$, where σ is given by equation (14). The shaded bands show the 68% spread of 100 mock catalogues with different input radial profiles for satellites. The crosses indicate the values obtained for the specific mock catalogue that is being analysed. We show a mock catalogue where satellite galaxies follow the best-fit NFW profile of the dark matter (blue, NFW) and a catalogue where those galaxies follow the phase-space distribution of subhaloes (red, SHPS).

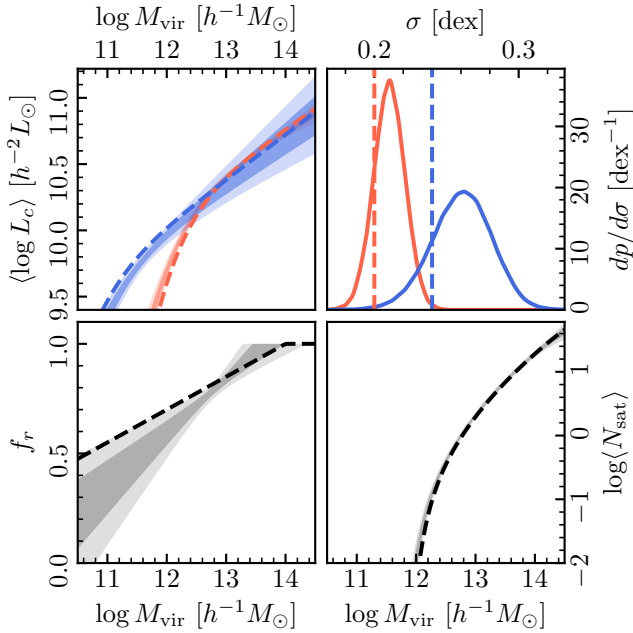


Figure 12. Similar to Figure 9, but now showing the posterior predictions for the galaxy-halo connection when analysing a mock catalogue where satellite galaxies are placed on subhaloes. Bands show the 1 and 2 σ regions and the dashed lines input values. The solid lines in the upper right panel is a Gaussian kernel density estimate of the posterior.

for massive clusters with large numbers of satellites. We developed a new method to analyse satellite kinematics that aims to rectify this shortcoming. This method represents a continuation of the work by van den Bosch et al. (2004) and More et al. (2009a,b, 2011). We test and validate our analysis framework using mock catalogues of increasing complexity. We demonstrate the need to accurately account for interlopers and fibre-collision-induced incompleteness in the spectroscopic survey. Additionally, we use forward-modelling to construct covariance matrices, and to calibrate our analytical method. We highlight a few improvements with respect to More et al. (2011), which, as we demonstrate in Paper II, alleviate the tension between their results and alternative constraints on the galaxy-halo connection from galaxy abundances, galaxy-galaxy lensing, clustering and group catalogues.

- We have shown in §5.3 that fibre collisions bias the measured velocity dispersion low by $\sim 10\%$ if not corrected for. Furthermore, it also leads to wrong inferences about the radial profile of satellites which further alters the expected relation between velocity dispersion and halo mass. Both effects will likely lead to an underestimation of the average halo mass at fixed central luminosity. We have introduced a framework to correct for fibre collisions in §5.3. We have demonstrated that this method is highly effective in correcting the spectroscopic incompleteness for all observables considered in this work.

- §7 describes a framework to specifically correct for the biases introduced by the analysis pipeline and the analytic model. Specifically, we use detailed mock SDSS-like mock catalogues including the effects of fibre collisions and the

survey mask to calibrate our model. For example, contrary to the results by More et al. (2009a), we do not find that the velocity dispersion of satellites can be extracted in an unbiased manner in the presence of interlopers. Instead, we find the “measured” velocity dispersion to be biased low by $\sim 5\%$, similar to the results by Becker et al. (2007).

- By using detailed mock catalogues we are also able to create realistic covariance matrices that include the non-negligible correlations between different observables. For example, the host-weighted and satellite weighted velocity dispersion estimates σ_{hw} and σ_{sw} are highly correlated. Thus, their ratio is much better constrained than if they were statistically independent. Since this ratio is a measure of scatter in halo mass at fixed luminosity (More et al. 2009b), this should lead to much stronger constraints on the latter compared to the results by More et al. (2009a, 2011).

- We now use the galaxy luminosity function as a constraint. As shown by Li et al. (2012), the combination of number density and average velocity dispersion can constrain, for example, the scatter in galaxy luminosity or stellar mass at fixed halo mass.

We have tested our method by applying it to detailed mock catalogues with increasing complexity. One might worry that our assumption that satellites obey the spherically symmetric Jeans equation might bias our inferences. For example, subhaloes that host satellites are known to have radial anisotropy, phase-space substructure and are not spherically symmetric. However, we have demonstrated that this does not bias our inferences dramatically, at least in the case of M_{peak} selected subhaloes.

We have also demonstrated that the constraints derived from our analysis are competitive with respect to studies utilizing galaxy clustering, galaxy-galaxy lensing or a combination thereof. A more detailed analysis will be conducted in Paper II by applying the framework developed here to SDSS. Beyond constraining a traditional mass-dependent CLF model, there are many exciting applications of satellite kinematics that would also benefit from an improved understanding of satellite kinematics. For example, galaxy clustering might be affected by assembly bias (Zentner et al. 2014, 2016), the dependence of clustering on halo properties other than mass. While it is not clear whether satellite kinematics will be affected by this, the effect will likely be different. Thus, the combined analysis of satellite kinematics and clustering might further strengthen our observational constraints on the galaxy-halo relationship, particularly assembly bias. Furthermore, Li et al. (2012) showed that satellite kinematics can also be used to constrain cosmological parameters.

ACKNOWLEDGEMENTS

FvdB and JUL are supported by the US National Science Foundation through grant AST 1516962. ARZ is funded by the Pittsburgh Particle physics Astrophysics and Cosmology Center (Pitt PACC) at the University of Pittsburgh and by the NSF through grant NSF AST 1517563. This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1125915. This work was supported by the HPC facilities operated by, and the staff of, the Yale Center for Research Computing.

This work made use of the following software packages: `matplotlib` (Hunter 2007), `SciPy`, `NumPy` (Van Der Walt et al. 2011), `Astropy` (Astropy Collaboration 2013), `Cython` (Behnel et al. 2011), `halotools` (Hearin et al. 2016), `Corner` (Foreman-Mackey 2016), `MultiNest` (Feroz & Hobson 2008; Feroz et al. 2009, 2013), `PyMultiNest` (Buchner et al. 2014), `mangle` (Hamilton & Tegmark 2004; Swanson et al. 2008) and `pymangle`⁷. We also thank the open-source developers behind `Ubuntu`, `GNOME`, `Xfce`, `Spyder`, `JabRef`, `TexStudio` and `Terminator`. All the above mentioned software packages helped to greatly expedite this work.

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE, www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de). The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance AstroComputing Center (UC-HiPACC) and were run at the NASA Ames Research Center.

REFERENCES

- Abazajian K. N., et al., 2009, *ApJS*, 182, 543
- Adelman-McCarthy J. K., et al., 2006, *ApJS*, 162, 38
- Astropy Collaboration 2013, *A&A*, 558, A33
- Becker M. R., et al., 2007, *ApJ*, 669, 905
- Beers T. C., Flynn K., Gebhardt K., 1990, *AJ*, 100, 32
- Behnel S., Bradshaw R., Citro C., Dalcin L., Seljebotn D. S., Smith K., 2011, *Computing in Science & Engineering*, 13, 31
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Binney J., Tremaine S., 1987, *Galactic dynamics*
- Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, *AJ*, 125, 2276
- Blanton M. R., et al., 2005, *AJ*, 129, 2562
- Brainerd T. G., Specian M. A., 2003, *ApJ*, 593, L7
- Buchner J., et al., 2014, *A&A*, 564, A125
- Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, *MNRAS*, 394, 929
- Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, *MNRAS*, 430, 767
- Campbell D., van den Bosch F. C., Padmanabhan N., Mao Y.-Y., Zentner A. R., Lange J. U., Jiang F., Villarreal A., 2017, preprint, ([arXiv:1705.06347](https://arxiv.org/abs/1705.06347))
- Carlberg R. G., Yee H. K. C., Ellingson E., 1997, *ApJ*, 478, 462
- Chen J., 2008, *A&A*, 484, 347
- Conroy C., et al., 2007, *ApJ*, 654, 153
- DES Collaboration et al., 2017, preprint, ([arXiv:1708.01530](https://arxiv.org/abs/1708.01530))
- Davis M., et al., 2003, in Guhathakurta P., ed., *Proc. SPIE Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II*. pp 161–172 ([arXiv:astro-ph/0209419](https://arxiv.org/abs/astro-ph/0209419)), doi:10.1117/12.457897
- Diemand J., Moore B., Stadel J., 2004, *MNRAS*, 352, 535
- Dutton A. A., Conroy C., van den Bosch F. C., Prada F., More S., 2010, *MNRAS*, 407, 2
- Erickson L. K., Gottesman S. T., Hunter Jr. J. H., 1987, *Nature*, 325, 779
- Faltenbacher A., 2010, *MNRAS*, 408, 1113
- Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
- Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint, ([arXiv:1306.2144](https://arxiv.org/abs/1306.2144))
- Foreman-Mackey D., 2016, *The Journal of Open Source Software*, 24
- Guo Q., White S., Li C., Boylan-Kolchin M., 2010, *MNRAS*, 404, 1111
- Guo Q., Cole S., Eke V., Frenk C., 2012, *MNRAS*, 427, 428
- Guo H., et al., 2015a, *MNRAS*, 446, 578
- Guo H., et al., 2015b, *MNRAS*, 453, 4368
- Guo H., et al., 2016, *MNRAS*, 459, 3040
- Hamilton A. J. S., Tegmark M., 2004, *MNRAS*, 349, 115
- Hearin A. P., Watson D. F., 2013, *MNRAS*, 435, 1313
- Hearin A., et al., 2016, preprint, ([arXiv:1606.04106](https://arxiv.org/abs/1606.04106))
- Hikage C., Mandelbaum R., Takada M., Spergel D. N., 2013, *MNRAS*, 435, 2345
- Hoshino H., et al., 2015, *MNRAS*, 452, 998
- Hunter J. D., 2007, *Computing in Science and Engineering*, 9, 90
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, 457, 4340
- Lange J. U., van den Bosch F. C., Hearin A., Campbell D., Zentner A. R., Villarreal A., Mao Y.-Y., 2018, *MNRAS*, 473, 2830
- Leauthaud A., et al., 2012, *ApJ*, 744, 159
- Li C., Jing Y. P., Mao S., Han J., Peng Q., Yang X., Mo H. J., van den Bosch F., 2012, *ApJ*, 758, 50
- Lin Y.-T., Mohr J. J., Stanford S. A., 2004, *ApJ*, 610, 745
- Mandelbaum R., Wang W., Zu Y., White S., Henriques B., More S., 2016, *MNRAS*, 457, 3200
- McKay T. A., et al., 2001, *ArXiv Astrophysics e-prints*, 2001, 10
- McKay T. A., et al., 2002, *ApJ*, 571, L85
- Mo H. J., Yang X., van den Bosch F. C., Jing Y. P., 2004, *MNRAS*, 349, 205
- More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009a, *MNRAS*, 392, 801
- More S., van den Bosch F. C., Cacciato M., 2009b, *MNRAS*, 392, 917
- More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, *MNRAS*, 410, 210
- Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, *ApJ*, 731, 113
- Moster B. P., Naab T., White S. D. M., 2017, preprint, ([arXiv:1705.05373](https://arxiv.org/abs/1705.05373))
- Murray S. G., Power C., Robotham A. S. G., 2013, *Astronomy and Computing*, 3, 23
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Norberg P., Frenk C. S., Cole S., 2008, *MNRAS*, 383, 646
- Planck Collaboration et al., 2014, *A&A*, 571, A16
- Prada F., et al., 2003, *ApJ*, 598, 260
- Sheldon E. S., et al., 2009a, *ApJ*, 703, 2217
- Sheldon E. S., et al., 2009b, *ApJ*, 703, 2232
- Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, *MNRAS*, 410, 417
- Smith R. E., 2012, *MNRAS*, 426, 531
- Sonnenfeld A., Leauthaud A., 2017, preprint, ([arXiv:1710.00007](https://arxiv.org/abs/1710.00007))
- Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, *MNRAS*, 387, 1391
- Tal T., Wake D. A., van Dokkum P. G., 2012, *ApJ*, 751, L5
- Taylor A., Joachimi B., Kitching T., 2013, *MNRAS*, 432, 1928
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Vale A., Ostriker J. P., 2004, *MNRAS*, 353, 189
- Vale A., Ostriker J. P., 2006, *MNRAS*, 371, 1173
- Van Der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, 13, 22
- Velander M., et al., 2014, *MNRAS*, 437, 2111
- Wang Y., Yang X., Mo H. J., Li C., van den Bosch F. C., Fan Z., Chen X., 2008, *MNRAS*, 385, 1511
- Wang L., et al., 2014, *MNRAS*, 439, 611

⁷ <https://github.com/esheldon/pymangle>

- Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2
- Wibking B. D., et al., 2017, preprint, ([arXiv:1709.07099](https://arxiv.org/abs/1709.07099))
- Wojtak R., Mamon G. A., 2013, MNRAS, 428, 2407
- Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057
- Yang X., Mo H. J., van den Bosch F. C., Weinmann S. M., Li C., Jing Y. P., 2005, MNRAS, 362, 711
- Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153
- 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
- Ye J.-N., Guo H., Zheng Z., Zehavi I., 2017, ApJ, 841, 45
- 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
- Zehavi I., et al., 2005, ApJ, 630, 1
- Zehavi I., et al., 2011, ApJ, 736, 59
- Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
- Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A., 2016, preprint, ([arXiv:1606.07817](https://arxiv.org/abs/1606.07817))
- Zheng Z., Guo H., 2016, MNRAS, 458, 4015
- Zu Y., Mandelbaum R., 2015, MNRAS, 454, 1161
- Zu Y., Mandelbaum R., 2016, MNRAS, 457, 4360
- Zwicky F., 1933, Helvetica Physica Acta, 6, 110
- 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., Norberg P., 2005a, MNRAS, 356, 1233
- van den Bosch F. C., Weinmann S. M., Yang X., Mo H. J., Li C., Jing Y. P., 2005b, MNRAS, 361, 1203
- van den Bosch F. C., et al., 2007, MNRAS, 376, 841
- van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725
- van den Bosch F. C., Jiang F., Campbell D., Behroozi P., 2016, MNRAS, 455, 158
- van der Marel R. P., Magorrian J., Carlberg R. G., Yee H. K. C., Ellingson E., 2000, AJ, 119, 2038

APPENDIX: EFFECTS OF THE SDSS TILING ALGORITHM

As detailed in Blanton et al. (2003), the SDSS fibre spectrograph can at most observe 592 science targets on a single spectroscopic plate. Due to the overall large-scale structure of galaxies and the resulting large variance of target density, a naive uniform placement of these tiles on the sky would lead to a significant spectroscopic incompleteness in high density regions. To circumvent this problem, the SDSS tiling algorithm perturbs the tile positions slightly such that the fraction of decollided targets that receive spectroscopic redshifts is maximised. Particularly, this results in high density regions on the sky having a higher density of tiles and a larger fraction of overlapping tile areas. Although this optimal set of tiles is constructed using only the decollided targets (Blanton et al. 2003), it results in a higher spectroscopic completeness for potentially collided galaxies, i.e. galaxies not in the decollided set, in overdense regions.

Implementing the entire SDSS tiling algorithm in our mock catalogues is beyond the scope of this work. We can however estimate its impact by comparing with the actual SDSS data. In Figure 13 we show the fraction of targets with

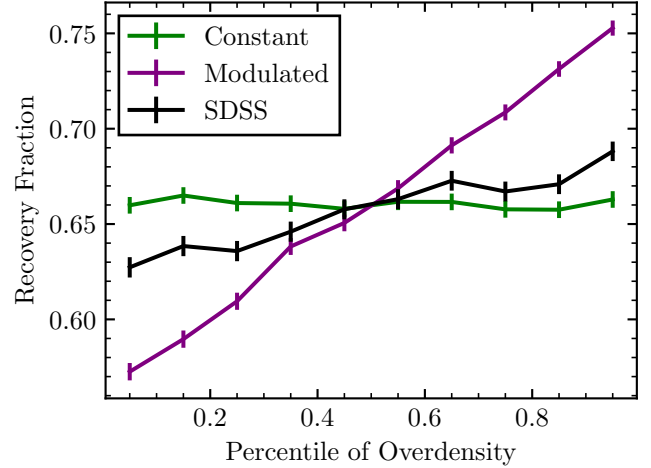


Figure 13. The recovery fraction, defined as the fraction of targets with other targets within $55''$ that receive spectroscopic redshifts, as a function of density. The density is defined as the number of targets within 1.5 deg. We show the results of SDSS (black) against those of a mock catalogue with our standard implementation of fibre collisions (green) and an alternative scheme (purple).

neighbours within $55''$ that receive spectroscopic redshifts. We plot this recovery fraction as a function of the overall large scale density. Particularly, we calculate for each object the number of other targets within 1.5 deg, i.e. the size of the SDSS spectroscopic plates. The recovery fraction is plotted as a function of the overall percentile of this overdensity with respect to all other potentially collided targets. For the SDSS, we see a roughly linear dependence with the recovery fraction increasing from $\sim 63\%$ to $\sim 69\%$ when going from underdense to overdense regions. We tested other radii for measuring the density but found similar or smaller dependences.

This overall scaling of the recovery fraction with large-scale density is not included in our default mock catalogues. The concern is that average halo masses will be positively correlated with this overall large-scale density. Particularly, our fibre collision correction method implicitly assumes the recovered galaxies to be representative of those lost due to fibre collisions. However, the results in Figure 13 suggest that the SDSS tiling algorithm will be slightly biased towards resolving collisions (i.e., assigning fibres to collided galaxies) for high-mass haloes, which, on average, reside in denser environments. We can address this concern using mock catalogues. For our regular mocks we assumed that 35% of all potentially collided targets receive spectroscopic redshifts, irrespective of the overall large-scale density. The resulting recovery fraction is shown by the green line in Figure 13. In a different set of mocks, we assume this fraction to linearly increase with the rank of the overdensity on 1.5 deg from 15% to 55%. The resulting recovery fraction is shown by the purple line and clearly more pronounced than the SDSS data. We create 100 mock catalogues where for each mock we run the two different fibre collision algorithms. Afterwards, we analyse these mocks with our analysis pipeline and compare the differences in the observables, e.g. the velocity dispersion, between these two sets of mocks. We do not find any significant difference in the observables for these two kind

of mock making algorithms. This suggests that the overall modulation of the fibre collision recovery fraction with large-scale density is negligible and does not impact our results significantly.