

The galaxy–halo connection from satellite kinematics

– II. Application to SDSS

Johannes U. Lange^{1*}, Frank C. van den Bosch¹ and Andrew R. Zentner²

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

²*Department of Physics and Astronomy & Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260, USA*

Accepted xxx. Received xxx

ABSTRACT

XXX

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

1 INTRODUCTION

2 OBSERVATIONAL DATA

Observational constraints in this work come from the New York University Value-Added Galaxy Catalog (VAGC; Blanton et al. 2005). This catalogue is derived from the Seventh Data Release of the Sloan Digital Sky Survey (SDSS DR7; Abazajian et al. 2009). Specifically, we use the **bright0**¹ of the VAGC. This includes roughly $\sim 570,000$ galaxies with a limiting Petrosian magnitude of $m_r < 17.6$. k -corrections and evolution corrections to redshift $z = 0.1$ have been applied to all galaxies in the sample. We then require an r -band luminosity of $L > 10^{9.5} h^{-2} L_\odot$. We apply the same $g-r$ colour cut to divide galaxies into red and blue as used by Zehavi et al. (2011). Note that this cut is slightly different than the cut More et al. (2011) applied. The colour cut can be written as

$$(g-r)_{\text{cut}}^{0.1} = 0.21 - 0.03(M_r^{0.1} - 5 \log h). \quad (1)$$

Due to mechanical limitations, spectroscopic fibres cannot be placed simultaneously for objects separated by less than $55''$ on the sky. Due to these fibre collisions around $\sim 6\%$ of the targets lack a spectroscopic redshift. In this case, we use the nearest-neighbour redshift assignment scheme (Zehavi et al. 2005).

2.1 Sample Selection

As described in Paper I, we use a cylindrical isolation criterion to identify central candidates. A galaxy is considered isolated, a primary, if it does not have another brighter galaxy within a cylindrical volume defined by depth $(\Delta v)_h$ and radius R_h . Any other fainter galaxy inside this cylinder is removed from the list of potential primaries. The

cylinder dimensions, as described below, are a function of the luminosity and colour of the galaxy in question. Additionally, we require primaries to have a spectroscopic redshift, i.e. no fibre collision, and to lie in the redshift range $0.02 < z < 0.067$. Furthermore, we remove galaxies close to the angular survey edge. This is characterized by putting a ring around each galaxy corresponding to the angular size of R_h at that galaxy’s redshift and demanding that not more than 20% of it lies outside the survey area. Finally, primaries are required to lie in a survey region with at least 80% spectroscopic completeness. Afterwards, satellite candidates, so called secondaries, lying inside a cylinder defined by $(\Delta v)_s$ and R_s are associated to each primary. Contrary to van den Bosch et al. (2004) and More et al. (2009a,b, 2011), primaries are not removed if they have no secondary.

We follow van den Bosch et al. (2004) and make the cylinder sizes dependent on the galaxy properties. This typically increases completeness and purity of the resulting sample of central and satellite candidates. Specifically, we choose $(\Delta v)_h = 1000\sigma_{200}(L, C)$, $R_h = 0.5\sigma_{200}(L, C)h^{-1}$ Mpc, $(\Delta v)_s = 4000 \text{ km s}^{-1}$ and $R_s = 0.15\sigma_{200}(L, C)h^{-1}$ Mpc. Here, $\sigma_{200} = \sigma/200 \text{ km s}^{-1}$ is an estimate for the velocity dispersion of satellites as a function of primary luminosity L and colour C . Because this estimate is a priori unknown, the cylinder sizes must be extracted from the data itself in an iterative fashion. We start with an estimate for $\sigma_{200}(L, C)$ and extract primaries and secondaries. An estimate for $\sigma_{200}(L, C)$ is extracted from the data by fitting the distribution of secondaries in the Δv_z – L_{pri} plane. Here, Δv_z is the line-of-sight velocity difference of the secondary and its primary and L_{pri} the luminosity of the primary. The model is

$$P(\Delta v_z, L_{\text{pri}}) = \frac{f_{\text{int}}}{2z_s} + \frac{1 - f_{\text{int}}}{\sqrt{2\pi}\sigma^2 \text{erf}(z_h/\sqrt{2}\sigma)} \exp\left[-\frac{\Delta v_z^2}{2\sigma^2}\right], \quad (2)$$

where both the interloper fraction f_{int} and the velocity dis-

* email: johannesulf.lange@yale.edu

¹ <http://sdss.physics.nyu.edu/lss/dr72/bright/0/>

persion σ are a function of the primary luminosity. Specifically, f_{int} is assumed to be a linear function of $\log L_{\text{pri}}$ and

$$\log \sigma / \text{km s}^{-1} = a + b(\log L_{\text{pri}} - 10) + c(\log L_{\text{pri}} - 10)^2, \quad (3)$$

where a , b and c are constants. The best fit is found by maximizing the likelihood,

$$\mathcal{L} \propto \prod_{\text{secondaries}} P(\Delta v_{z,i}, L_{\text{pri},i})^{w_{\text{sw},i}}, \quad (4)$$

where $w_{\text{sw},i}$ is a weight for each primary–secondary pair, as discussed in the next section. We then update our estimate for $\sigma_{200}(L, C)$ and extract a new set of primaries and secondaries. Using the best-fit values of More et al. (2011) as a starting point, we find our algorithm to arrive at $a = 2.210$, $b = 0.478$, $c = 0.275$ and $a = 2.142$, $b = 0.402$, $c = -0.170$ for red and blue galaxies, respectively.

Figure 1 shows the line-of-sight velocity distribution of secondaries around primaries of a given luminosity. We show the distribution around red and blue primaries separately. From this Figure it is apparent that the velocity dispersion is a strong function of primary luminosity for red primaries and less so for blue ones. We will analysis this more quantitatively in the coming sections.

2.2 Constraints

We seek to constrain the galaxy–halo connection using a set of observables extracted from the above data set. In short, those constraints are the overall number density of galaxies, the red fraction of primaries, the number of same-halo secondaries around primaries and their velocity dispersion. All quantities are measured in 10 bins in $\log L / h^{-2} L_{\odot}$ from 9.5 to 11.0.

The number density of galaxies is estimate via

$$n_{\text{gal}} = \frac{\sum w_{s,i}}{\Omega_{\text{SDSS}} (d_{\text{com}}^3(z = 0.067) - d_{\text{com}}^3(z = 0.02))}, \quad (5)$$

where the sum goes over all galaxies with a spectroscopic redshift and $\Omega_{\text{SDSS}} = 2.273 \text{sr}$ is the effective survey area of SDSS. Finally, $w_{s,i}$ is a weight designed to correct spectroscopic incompleteness, as discussed in Paper I. For each galaxy, we count the number of neighbouring targets within $55''$. We then assign a weight that is the inverse of the fraction of targets with the same number of neighbours that have been assigned a spectroscopic redshift. Similarly, the red fraction of primaries is defined by

$$f_{\text{pri},r} = \frac{\sum_{\text{red primaries}} w_{s,i}}{\sum_{\text{all primaries}} w_{s,i}}. \quad (6)$$

For the rest of the constraints, we need to take into account that not all secondaries are satellites of the same halo as the primary. This is particularly important when evaluating the velocity dispersion.

As discussed in detail in Paper I, we fit the Δv_z – R_p distribution of secondaries in each bin of primaries with a flat interloper model $P_{\text{int}}(\Delta v_z, R_p)$ and a model for same-halo satellites $P_{\text{sat}}(\Delta v_z, R_p)$. The phase-space distribution of satellites in a halo of fixed mass is discussed in the appendix. Furthermore, we assume that the host halo masses of the satellites are drawn from a log-normal distribution

with mean and spread as free parameters. We determine the fit that maximises the likelihood and assign a membership probability

$$p_{\text{mem}}(\Delta v_z, r_p) = \frac{P_{\text{sat}}(\Delta v_z, r_p)}{P_{\text{sat}}(\Delta v_z, r_p) + P_{\text{int}}(\Delta v_z, r_p)} \quad (7)$$

for each secondary. When determining the fit that maximizes the likelihood, we assign a weight to each primary–secondary pair,

$$w_{\text{sw}} = w_{\text{s,pri}} w_{\text{s,scd}}, \quad (8)$$

the product of the spectroscopic weight of the primary and the secondary. Except for the fibre collision correction, this assign equal weights to all secondaries. Thus, quantities derived from this weighting scheme are satellite-weighted. We then estimate the number of same-halo secondaries via

$$n_{\text{s}} = \frac{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i}}{\sum_{\text{primaries}} w_{s,i}}, \quad (9)$$

where w_{sw} is the product of the spectroscopic weight of each secondary and its according primary.

Similarly, an estimate for the satellite-weighted velocity dispersion is obtained via

$$\sigma_{\text{sw}}^2 = \frac{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i} \Delta v_{z,i}^2}{\sum_{\text{secondaries}} p_{\text{mem},i} w_{\text{sw},i}}. \quad (10)$$

Finally, we wish to estimate the average velocity dispersion when weighting each primary or host instead of each secondary. We do so by first assigning each secondary a weight of

$$w_{\text{hw}} = \frac{w_{\text{s,pri}} w_{\text{s,scd}}}{N_{\text{scd}}}, \quad (11)$$

where N_{scd} is the number of secondaries hosted by each primary. σ_{hw} is then calculated in complete analogy to σ_{sw} .

Ultimately, we have up to 80 constraints, n_{gal} , $f_{\text{pri},r}$, $n_{\text{s},r}$, $n_{\text{s},b}$, $\log \sigma_{\text{hw},r}$, $\log \sigma_{\text{hw},b}$, $\sigma_{\text{hw},r}^2 / \sigma_{\text{sw},r}^2$ and $\sigma_{\text{hw},b}^2 / \sigma_{\text{sw},b}^2$, each measured in 10 bins of primary luminosity and n_{s} , $\log \sigma_{\text{hw}}$ and $\sigma_{\text{hw}}^2 / \sigma_{\text{sw}}^2$ for red and blue primaries separately. However, we only consider data points for which we estimate to have at least 10 primaries with at least 1 same-halo secondary and for which we can create an uncertainty estimate, as described in the next section. In Table 1 we show an overview of the data used in this study. Note that this requires an assumed radial profile to get membership probabilities. For this data we assumed that satellite follow an NFW profile that is half as concentrated as that of dark matter.

3 MODEL FOR CONSTRAINTS

Ideally, one would compute all constraints in mock catalogues and directly compare those results to SDSS. Unfortunately, this is prohibitively expensive for our current analysis. Instead, we will employ a novel approach where we use a large number of mock catalogues to find the best-fit model and an analytical approach to evaluate uncertainties. In the following, we will briefly summarize the analytical model.

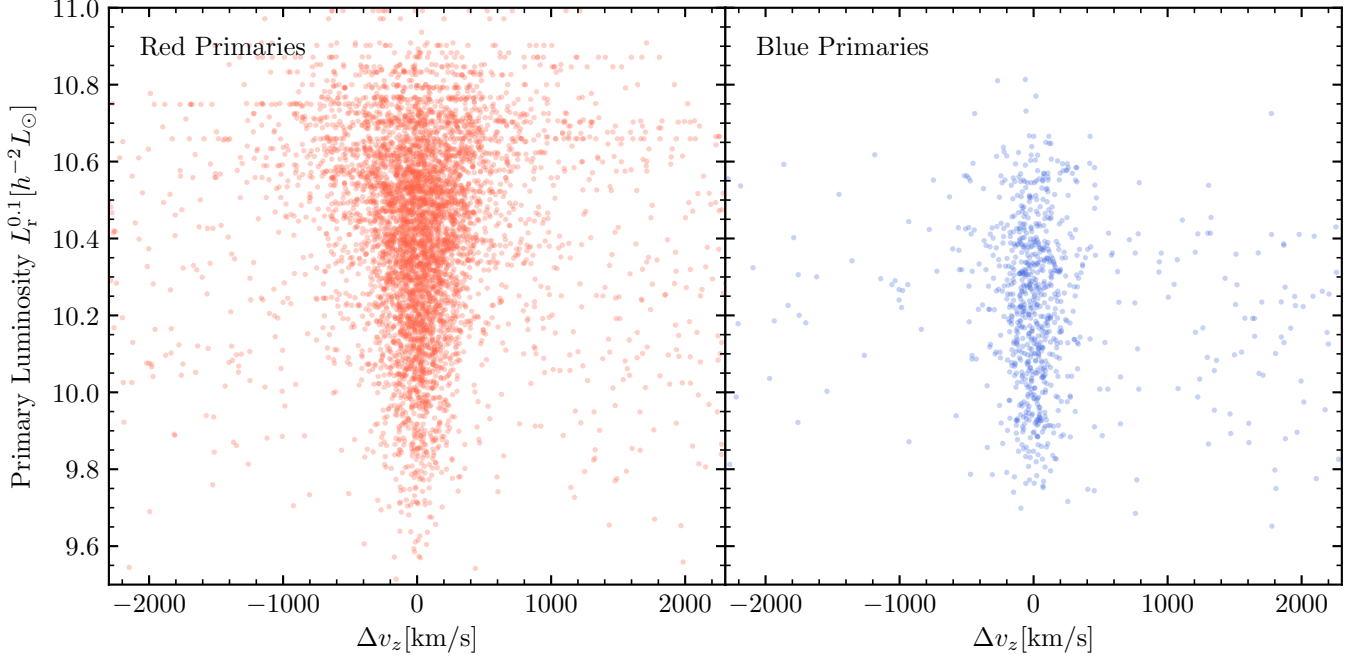


Figure 1. The velocity distribution of secondaries in SDSS as a function of the luminosity of the associated primary. We split the secondaries by the colour of the primary, red (left panel) or blue (right panel).

Description	Bins										Total
$\log L_{\text{pri,min}}$	9.50	9.65	9.80	9.95	10.10	10.25	10.40	10.55	10.70	10.85	9.50
$\log L_{\text{pri,max}}$	9.65	9.80	9.95	10.10	10.25	10.40	10.55	10.70	10.85	11.00	11.00
Red Primaries	2693	3138	3430	3186	2715	1871	1042	408	95	11	18589
Associated Secondaries	10	62	178	410	706	1060	1313	1230	734	152	5855
Blue Primaries	7160	6029	4888	3548	2317	1351	526	126	17	0	25962
Associated Secondaries	8	59	151	179	240	267	132	63	7	0	1106
$\log n_{\text{gal}}[h^3 \text{Mpc}^{-3}]$	-2.405	-2.458	-2.526	-2.641	-2.806	-3.038	-3.381	-3.882	-4.549	-5.565	-1.803
$f_{\text{pri,r}}$	0.273	0.342	0.413	0.474	0.543	0.586	0.675	0.776	0.866	1.000	
$n_{\text{s,r}}$	–	0.013	0.041	0.095	0.223	0.514	1.259	3.423	8.324	–	
$n_{\text{s,b}}$	–	–	0.019	0.038	0.081	0.167	0.204	0.439	–	–	
$\sigma_{\text{hw,r}}[\text{km s}^{-1}]$	–	150.6	157.4	152.3	224.7	227.7	257.4	389.1	423.6	–	
$\sigma_{\text{hw,b}}[\text{km s}^{-1}]$	–	–	90.0	158.7	135.3	172.3	218.0	188.7	–	–	
$\sigma_{\text{hw,r}}^2/\sigma_{\text{sw,r}}^2$	–	–	–	–	0.861	0.772	0.725	0.450	0.559	–	
$\sigma_{\text{hw,b}}^2/\sigma_{\text{sw,b}}^2$	–	–	–	–	–	0.687	0.837	0.728	–	–	

Table 1. Overview of the data used in this analysis. We have defined 10 bins in the luminosity of the primary $\log L_{\text{pri}}$. The first two rows show the bin edges and the next four rows the number of primaries and secondaries. Finally, the last eight rows show the constraints derived from the data that we use to constrain the model. Constraints that could not be measured or for which no reliable uncertainties could be derived from mocks are omitted.

We refer the reader to Paper I to a more detailed discussion of this model and how it fares in comparison to mock catalogues.

The number density of galaxies in a luminosity range $[L_1, L_2]$ can be inferred from the CLF combined with the halo mass function via

$$n_{\text{gal}}(L_1, L_2) = \int_{L_1}^{L_2} \int_0^\infty \Phi_{\text{tot}}(L|M) n_{\text{h}}(M) dM dL. \quad (12)$$

Here, $\Phi_{\text{tot}}(L|M)$ is the combined CLF of central and satel-

lites for a halo of virial mass M . The halo mass function $n_{\text{h}}(M)$ is computed directly from the SMDPL simulation output. See the appendix for a discussion of the CLF parametrization.

The number red fraction of primaries can be approximated as the red fraction of centrals,

$$f_{\text{pri,r}}(L_1, L_2) \approx \frac{n_{\text{c,r}}(L_1, L_2)}{n_{\text{c,r}}(L_1, L_2) + n_{\text{c,b}}(L_1, L_2)}. \quad (13)$$

The number density of centrals, similarly n_{gal} above, is com-

puted by integrating the central CLF over the halo mass function in SMDPL.

The expected number of secondaries around primaries can be very well described by the corresponding number of satellites around centrals,

$$n_s(L_1, L_2) \approx \frac{\int_{L_1}^{L_2} \int_0^\infty N_s(M) n_h(M) \Phi_c(L|M) f_{\text{ap}}(L, M) dM dL}{n_c(L_1, L_2)}. \quad (14)$$

Here, $N_s(M)$ is the expected number of satellites for a halo of mass M and Φ_c is the central CLF. $f_{\text{ap}}(L, M)$ describes the fraction of satellites in a halo of mass M with a central of luminosity L that are expected to lie inside the cylinder defined by R_s ,

$$f_{\text{ap}}(L, M) = \int_0^\infty 4\pi r^2 \bar{n}_{\text{sat}}(r|M) (\zeta(r, R_s(L)) - \zeta(r, R_c)) dr, \quad (15)$$

where \bar{n}_{sat} is the normalized radial profile of satellites in a halo of mass M ,

$$\int_0^\infty 4\pi r^2 \bar{n}_{\text{sat}}(r|M) dr = 1, \quad (16)$$

and

$$\zeta(r, r_s) = \begin{cases} 1 & \text{if } r \leq r_s \\ 1 - \sqrt{1 - r_s^2/r^2} & \text{otherwise.} \end{cases} \quad (17)$$

Note that we have excluded those satellites lying inside $R_c = 60 h^{-1} \text{kpc}$ to correct for fibre collisions.

In a similar fashion, the velocity dispersion is approximated as the expected velocity dispersion of satellites around all centrals in a given luminosity range,

$$\sigma^2(L_1, L_2) \approx \frac{\int_{L_1}^{L_2} \int_0^\infty w(L, M) \sigma_{\text{ap}}^2(L, M) n_h(M) \Phi_c(L|M) dM dL}{\int_{L_1}^{L_2} \int_0^\infty w(L, M) n_h(M) \Phi_c(L|M) dM dL}. \quad (18)$$

$w(L, M)$ is a weight set to

$$w_{\text{sw}}(L, M) = f_{\text{ap}}(L, M) N_s(M) \quad (19)$$

for the satellite-weighted velocity dispersion and

$$w_{\text{hw}}(L, M) = 1 - \exp(-f_{\text{ap}}(L, M) N_s(M)) \quad (20)$$

for the host-weighted velocity dispersion. Finally, $\sigma_{\text{ap}}^2(L, M)$ is the expected average velocity dispersion of all satellites living inside a halo of mass M and being inside the cylinder defined by $R_s(L)$

$$\sigma_{\text{ap}}^2(L, M) = \frac{\int_0^\infty r^2 \bar{n}_{\text{sat}}(r|M) \sigma^2(r|M) (\zeta(r, R_s(L)) - \zeta(r, R_c)) dr}{\int_0^\infty r^2 \bar{n}_{\text{sat}}(r|M) (\zeta(r, R_s(L)) - \zeta(r, R_c)) dr}. \quad (21)$$

The expected velocity dispersion $\sigma^2(r|M)$ at any given radial distance r and halo mass M is computed from the spherical Jeans equation without anisotropy. The expected velocity dispersion also needs the radial profile of satellites \bar{n}_{sat} as an input. Throughout this work, we test three different choices for the radial profile. One assumes that the satellite distribution can be described by an NFW profile with the same scale radius as the dark matter in each halo. Another choice is that satellites follow an NFW profile with a scale radius that is twice as large as that of the dark matter. Finally, we consider satellites having a cored profile that very well fits the distribution of resolved subhaloes in SMDPL (see Paper I). We will call these three choices NFW, bNFW and Cored, respectively.

4 ANALYSIS PROCEDURE

The goal of this work is to constrain the galaxy-halo connection parametrized by the model vector $\theta = \{L_{0,r}, M_{1,r}, \gamma_{1,r}, \dots\}$ by our observational data $D = \{n_{\text{gal}}, f_{\text{pri}}, \dots\}$. To do this we assume a likelihood of the form

$$\mathcal{L}(\theta|D) \propto \exp \left[-\frac{(\mathbf{M}^*(\theta) - D)^t \Psi (\mathbf{M}^*(\theta) - D)}{2} \right], \quad (22)$$

where Ψ is the precision matrix, i.e. the inverse of the covariance matrix, and $\mathbf{M}^*(\theta) = \langle D(\theta) \rangle$ the model prediction for the data vector. Both the model prediction and the covariance matrix are derived from detailed SDSS-like mock catalogues. However, the following simplifications are made in order to make the likelihood evaluation computationally feasible.

In the previous section, we have introduced a simple analytic model $\mathbf{M}(\theta)$ for the full forward-modelling prediction $\mathbf{M}^*(\theta)$. In Paper I, we have shown that this simple model is able to reproduce all qualitative features of the forward-modelling prediction. At the same time, small biases between that model and results from mock catalogues of the order of $\sim 1\sigma$ exists. Thus, we calibrate the analytic model such that it reproduces the forward-modelling prediction for a given set of parameters θ_0 . Specifically, we introduce a bias vector $\mathbf{B}(\theta_0) = \mathbf{M}^*(\theta_0) - \mathbf{M}(\theta_0)$. We then approximate the forward-modelling prediction for an arbitrary model θ as

$$\mathbf{M}^*(\theta) \approx \mathbf{M}(\theta) + \mathbf{B}(\theta_0). \quad (23)$$

An estimate $\hat{\mathbf{B}}(\theta_n)$ for the bias vector is obtained from 1000 mock catalogues created from model θ_0 . \mathbf{M}^* in this case is simply the average of the 1000 data vectors.

As with the bias vector, we assume that the covariance matrix does not change significantly around any given set of parameters θ_0 . An unbiased estimate for the covariance matrix is

$$\hat{\mathbf{C}} = \frac{1}{N_S - 1} \sum_i^{N_S} (\mathbf{D}_i - \langle D \rangle) (\mathbf{D}_i - \langle D \rangle)^t, \quad (24)$$

where \mathbf{D}_i is the data vector of the i -th out of N_S mock catalogues and $\langle D \rangle$ the average of all. Finally, an unbiased estimate for the precision matrix is

$$\hat{\Psi} = \frac{N_S - N_D - 2}{N_S - 1} \hat{\mathbf{C}}^{-1}, \quad (25)$$

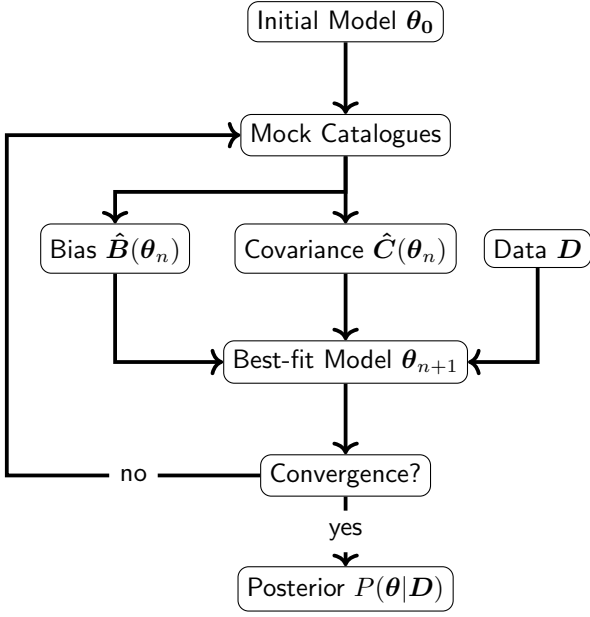


Figure 2. Diagram outlining our analysis procedure of the SDSS data. We start out with an initial model θ_0 and create mock catalogues. These mock catalogues are used to estimate the covariance of the data and the offset between the analytical model for the constraints and the forward-modelling results. Using these estimates, the analytical model and the actual SDSS data, a new best-fitting model θ_1 is obtained. This process is repeated until the sequence of best-fitting models $\theta_{0,1,\dots,n}$ is reasonably converged. Afterwards, the full posterior is estimated using the analytical model and the bias and covariance from the latest mock catalogues.

where N_D is the number of data points, i.e. the dimensionality of \mathbf{D} (Taylor et al. 2013). We note that we only use data points which can be measured in all 1000 mock catalogues and $\sigma_{\text{hw}}/\sigma_{\text{sw}}$ only if $n_{\text{sat}} > 0.1$ on average. The latter constraint is mainly due to the distributions of $\sigma_{\text{hw}}/\sigma_{\text{sw}}$ in the mocks being very non-Gaussian otherwise.

Figure 2 outlines how we find that best-fitting model and corresponding uncertainties. We start out with an initial guess for the galaxy–halo connection given by θ_0 . In this work, the initial guess is the default parameter set of Paper I. We then create mock catalogues for this particular choice of parameters and construct the bias vector $\hat{\mathbf{B}}$, precision and covariance matrix from them. These are then used to find a new best-fit model θ_1 using the simplified analytic model that maximizes the likelihood in equation (22). From this model we create another series mock catalogues and recompute the bias and covariance. Afterwards, a new best-fitting model θ_2 is found. This process is repeated until a convergence criterion is reached, signalling that the best-fit models do not change significantly. At this point the last bias and covariance estimates are used to compute the entire posterior using the analytic model.

How do we assess the convergence of this algorithm? We have a series of best-fit models $\theta_{0,1,\dots,n}$ based upon different estimates for the bias and covariance. We now take the latest estimate for the latter and compute the χ^2 of all $\theta_{0,1,\dots,n}$.

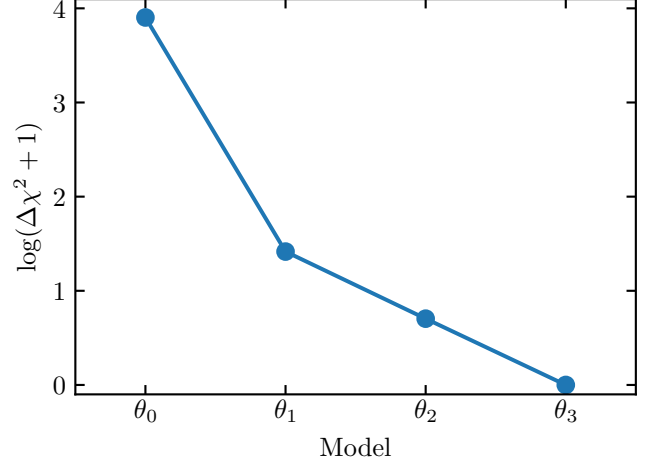


Figure 3. Series of best-fitting models when fitting SDSS with the NFW satellite profile. We show the χ^2 difference of each iteration with respect to the latest, θ_3 . The model covariance and bias for this plot have been obtained from mocks constructed with θ_2 . See the text for details.

By construction, θ_n will have the lowest χ^2 . On the other hand, previous iterations will have higher values because they used different estimates for the bias and covariance for the minimization. If the posterior of θ were described by a multivariate Gaussian distribution with 16 degrees of freedom, 68% of all models would lie within $\Delta\chi^2 = 18$ of the best-fit model. Thus, if $\chi^2(\theta_{n-1}) - \chi^2(\theta_n) < 18$, the best-fitting model did not change significantly from the previous iteration and we regard the result as converged. An example is shown in Figure 3 where we show the χ^2 series when fitting a model with the NFW profile to the SDSS data. Convergence is achieved at the third iteration. The initial model θ_0 gives an insufficient fit to the data with $\Delta\chi^2 \sim \chi^2 \sim 10000$. Thus, we cannot expect estimates for the bias and covariance matrix estimated from such mock catalogues to be very realistic with respect to SDSS. Nevertheless, the new best-fitting model θ_1 estimated from these estimates is only $\Delta\chi^2 \sim 25$ [Johannes: (Might be updated later.)] from the final best-fit model θ_3 away. Finally, θ_2 found using the bias and covariance estimate from θ_1 has $\Delta\chi^2 \sim 4$ [Johannes: (Might be updated later.)] when compared to θ_3 which is constructed from the bias and covariance estimates from θ_2 .

We note that the bias estimate $\hat{\mathbf{B}}(\theta)$ of the analytic model used to find θ_3 has been calibrated at θ_2 . Specifically, that means that the expected constraints have been fully forward-modelled for θ_2 . For our full posterior, only $\sim 1\%$ [Johannes: (Might be updated later.)] of all models have a χ^2 as good or better than θ_2 . Thus, a model very close to the best-fitting parameter has been verified with a large number of mock catalogues.

For finding the best-fitting model, as well as estimating the posterior, we make use of MultiNest (Feroz & Hobson 2008; Feroz et al. 2009) and its Python implementation PyMultiNest (Buchner et al. 2014). In all cases, we use a set of 10,000 live points and a stopping criterion of $\Delta \ln \mathcal{Z} = 10^{-4}$, where \mathcal{Z} is the estimate for the global Bayesian evidence. In case of finding a best-fitting model,

we run **MultiNest** in constant efficiency mode and a target efficiency of 50%. On the other hand, when estimating the posterior, we de-activate the constant efficiency mode and set a target efficiency of 1%. Although not its main purpose, **MultiNest** is an extremely robust global minimization scheme that does not get stuck in local minima. To verify that our estimate for the posterior is fully converged, we have re-run **MultiNest** with 50,000 live points and a target efficiency of 0.5% and getting consistent results. **[Johannes: Again, I'll do it in the end. I firmly believe that this will be the result.]**

5 RESULTS

We have performed the analysis procedure described in the above section for the three different satellite radial profiles separately. In each case, we used the specific radial profile to determine likely interlopers, produce mock catalogues and fit the data. We find that the observational constraints we used so far cannot reliably distinguish between the different models, all giving very similar χ^2 . Thus, we first analyse the observed radial profile of secondaries around primaries to distinguish between the different possible radial profiles. Afterwards, we will explore the fit to the data for the best-fitting radial profile in more detail.

5.1 Radial Profile

Figure 4 shows the predictions for the projected number density of primaries around secondaries. Each panel shows the distribution for different primary luminosities without distinction by the colour of primaries. Blue crosses denote the SDSS data and bands the predictions from each of the three posteriors. The bands present the 68% scatter from 1000 mock catalogues. Since for each mock we picked a random model from the posterior, bands show the combination of posterior uncertainty and statistical fluctuations due to cosmic variance. For both the mocks and the SDSS data, we have reduced the impact of interlopers by only considering secondaries within $600\sigma_{200}(L, C)\text{km s}^{-1}$. On the other hand, we have increased R_s from $0.15\sigma_{200} h^{-1}\text{Mpc}$ to $0.5\sigma_{200} h^{-1}\text{Mpc}$ to increase the range that is probed. A correction for fibre collisions has been applied by weighing each primary and secondary by $w_{s,\text{pri}}$ and $w_{\text{sw}} = w_{s,\text{pri}}w_{s,\text{scd}}$, respectively.

As expected, more radially extended profiles also predict the most shallow surface number densities. For example, the Cored profile predicts the lowest surface density at low radii and the largest at large radii. Overall, the model bNFW where satellites follow an NFW profile with a scale radius twice that of dark matter seems to fit the SDSS data the best. Thus, in the following we will concentrate on empirical results on the galaxy–halo connection derived from this radial profile.

5.2 Model for Satellite Kinematics

In Figure 5 we show the measurement values previously reported in Table 1, together with 1σ error bars derived from mock catalogues and the best-fitting model. Qualitatively,

Parameter	Posterior	Best-fit
$\log L_{0,r}$	$9.877^{+0.062}_{-0.045}$	9.904
$\log M_{1,r}$	$11.06^{+0.23}_{-0.13}$	11.15
$\gamma_{1,r}$	$3.52^{+0.95}_{-1.20}$	3.01
$\gamma_{2,r}$	$0.232^{+0.011}_{-0.011}$	0.230
σ_r	$0.1692^{+0.0055}_{-0.0056}$	0.1690
$\log L_{0,b}$	$10.238^{+0.070}_{-0.088}$	10.312
$\log M_{1,b}$	$11.42^{+0.14}_{-0.17}$	11.56
$\gamma_{1,b}$	$3.27^{+0.91}_{-0.55}$	2.73
$\gamma_{2,b}$	$0.082^{+0.055}_{-0.048}$	0.037
σ_b	$0.1780^{+0.0075}_{-0.0081}$	0.1787
f_0	$0.3908^{+0.0095}_{-0.0097}$	0.3929
α_f	$0.446^{+0.014}_{-0.013}$	0.444
b_0	$-0.460^{+0.027}_{-0.028}$	-0.460
b_1	$0.538^{+0.060}_{-0.058}$	0.537
b_2	$0.059^{+0.023}_{-0.024}$	0.061

Table 2. Posterior predictions and best-fit values of all parameters described the galaxy–halo connection. We assumed the bNFW satellite profile, i.e. that satellites follow an NFW profile with a scale radius twice as large dark matter scale radius.

the model excellently reproduces most trends in the observational data. The only exception is the ratio of host to satellite-weighted velocity dispersion for low-luminosity blue primaries. We believe that this is a purely statistical outlier based on two reasons. First, we see in the mock catalogues that the data points for a fixed model are only approximately distributed as Gaussian. Thus, such a 3σ outlier is not as unlikely as for a Gaussian distribution. Secondly, More et al. (2011) did not observe such a low ratio blue primaries in the same luminosity range. Thus, we regard the best-fit model as a qualitatively very good description of the SDSS data. However, expressed via the χ^2 , the best-fit model is not a good fit to the data with $\chi^2 = 84.9$ for 61 – 16 degrees of freedom. Thus, given that the model reproduces all qualitative trends of the data, our observational constraints seem to suggest a more complicated model for the galaxy–halo connection then assumed thus far. We will return to this point later in this section.

In Figure 6 we show the full velocity distribution of secondaries with respect to interlopers. We concentrate on six different bins in primary luminosity, differentiating between secondaries around red and blue primaries. Data points come from SDSS and bands denote the 68% range from 1000 mock catalogues derived from the best-fitting model. Again, error bars are purely indicative of cosmic variance fluctuations. We have mitigated the effect of fibre collisions by weighting each primary-secondary pair by $w_{\text{sw}} = w_{s,\text{pri}}w_{s,\text{scd}}$. No correction for interlopers has been applied. We note that the model has only been fitted to reproduce the second moment of the velocity distribution of satellites. On the other hand, higher order moments and the contribution of interlopers are all implicit predictions. Reassuringly, the best-fit model seems to accurately predict all of those.

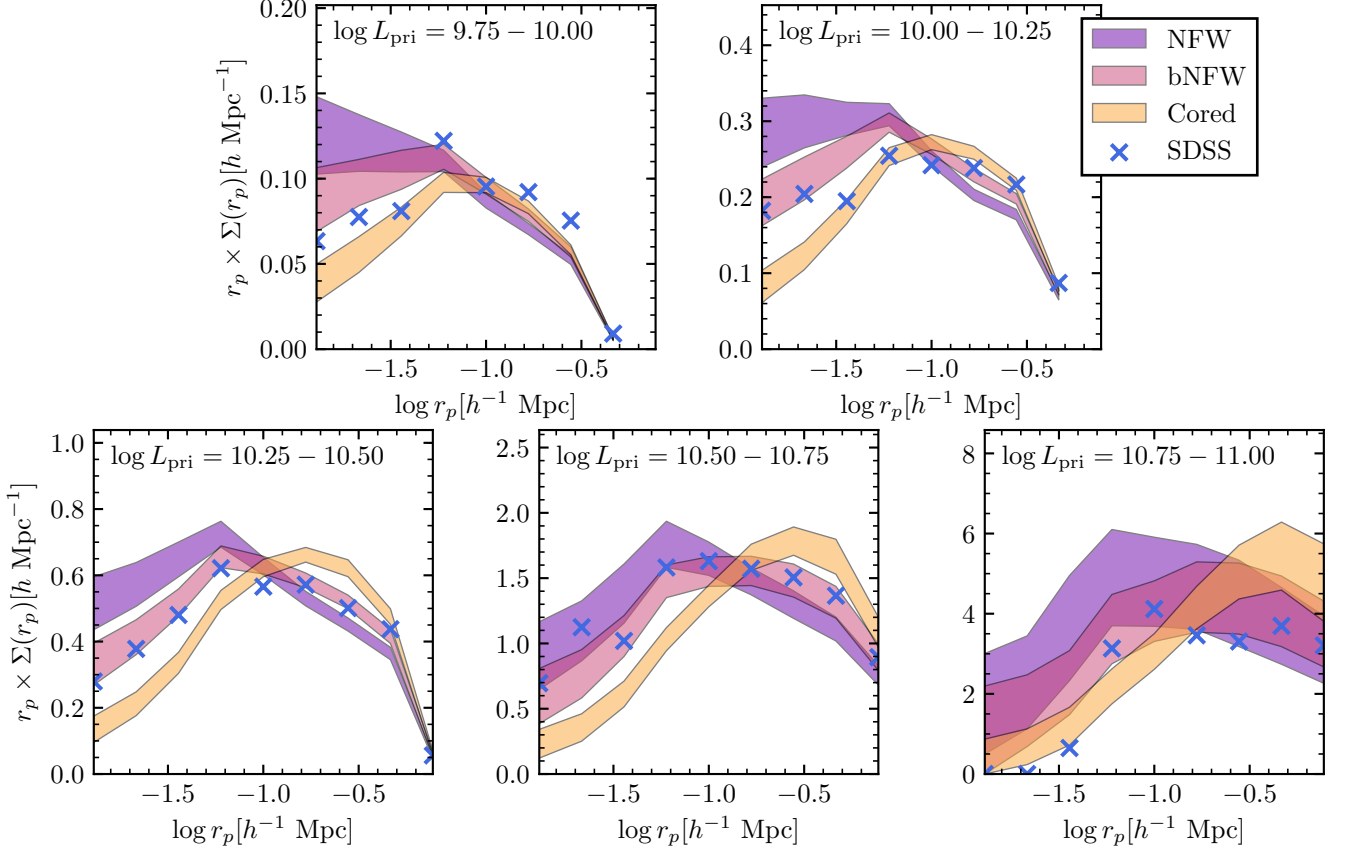


Figure 4. The projected number density of secondaries around primaries of different luminosities. Data points are from SDSS. The bands show the predictions of the best-fitting models for each assumed radial profile. For each best-fitting model, we create 1000 mock catalogues and the bands denote the 68% containment in each bin. Overall, the bNFW profile matches the SDSS data the best.

5.3 Galaxy–Halo Connection

Figure 7 shows marginalized posteriors for a subset of the parameters describing the galaxy–halo connection. Most parameters are tightly constrained with only the low-mass slopes $\gamma_{1,r}$ and $\gamma_{1,b}$ and the low-luminosity slope of the satellite CLF α_s being strongly limited by the prior. We also see tight correlations between certain parameters, for example the characteristic luminosity L_0 and the characteristic mass M_1 of the mass–luminosity relation for red and blue centrals, in good agreement with Cacciato et al. (2013) and Zu & Mandelbaum (2015). We list all posterior predictions on the parameters in describing the galaxy–halo relation in Table 2. We include the median and 1σ ranges as well as the best-fit value for each parameter. Figure 8 summarizes the multidimensional posterior by showing predictions for key statistics of the galaxy–halo relation. Specifically, we show the prediction for the mass–luminosity relation, the scatter in luminosity at fixed halo mass, the red fraction of centrals as a function of mass and the satellite occupation above $10^{9.5} h^{-2} L_\odot$. Interestingly, we find mass–luminosity relations that are very similar for red and blue centrals. The same is true for the scatter in luminosity. We also find that the red fraction increases with halo mass and reaches unity by roughly $\log M_{\text{vir}}/h^{-1} M_\odot \sim 13$. We note that the prediction for the red fraction of lower mass haloes,

$\log M_{\text{vir}}/h^{-1} M_\odot \sim 11$, is purely an extrapolation of our model and is not constrained by data as we do not expect those haloes to host any galaxy above $10^{9.5} h^{-2} L_\odot$. Finally, we find that the average number of satellites above $10^{9.5} h^{-2} L_\odot$ should roughly scale linearly with the halo virial mass. We will compare all those results with previous studies on a more quantitative level in the next section.

5.4 Testing Model Extensions

Here, we test possible extensions to the default parametrization of the galaxy–halo connection. Specifically, we look at models with additional parameters compared to the standard parametrization. Instead of repeating the entire procedure described in §4, we take the estimate of the covariance and bias from the best-fit default model and use the analytical model to evaluate the more complex one. In all cases, we assume the bNFW radial profile for satellites.

To judge the performance of extended models we look at evidence ratio. Note that for nested models one can show that the evidence ratio obeys

$$\frac{\mathcal{Z}(\mathbf{D}|M_n)}{\mathcal{Z}(\mathbf{D}|M_{n+1})} = \frac{P(\theta^{n+1} = 0|\mathbf{D})}{P(\theta^{n+1})}, \quad (26)$$

where \mathcal{Z} denotes the evidence and M_n is the nested model of M_{n+1} . Specifically, M_{n+1} reduces to M_n for θ^{n+1} , the

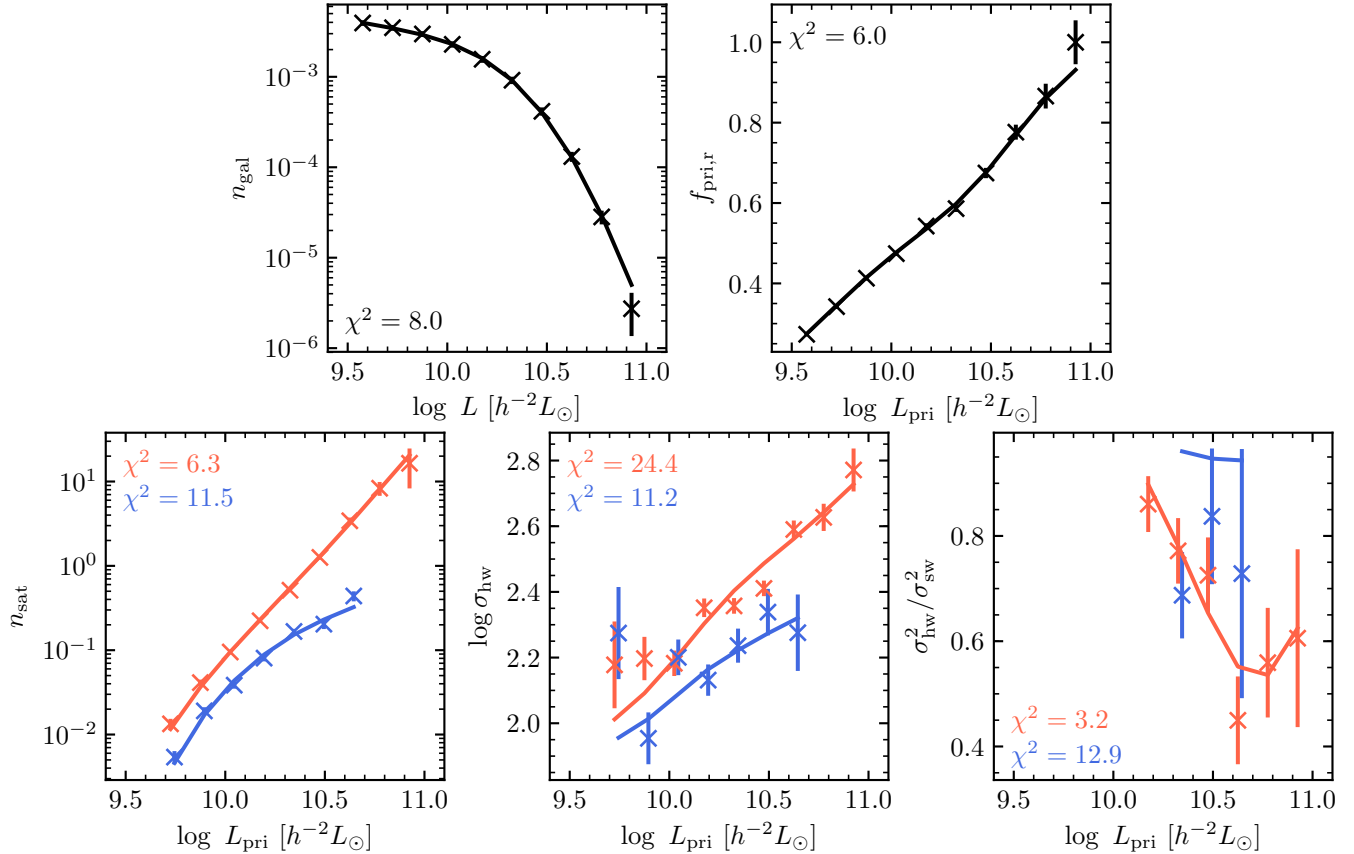


Figure 5. Observational constraints from SDSS (data points) and the best-fitting model (line). We show the number density of galaxies (top, left), the red fraction of primaries (top, right), the estimated number of satellites around primaries (bottom, left), the host-weighted velocity dispersion (bottom, middle) and the ratio of host to satellite-weighted velocity dispersion (bottom, right). All observables in the bottom row are measured separately for red and blue primaries. We also show the χ^2 for each set of observations separately in each panel. Error bars represent the diagonals of the covariance matrix estimated from mock catalogues. The best-fitting model assumed the bNFW satellite profile.

additional parameter of the complex model, being 0. This makes it clear that the evidence ratio can be judged directly from looking at the marginalized posterior probability $P(\theta^{n+1} = 0 | \mathbf{D})$.

5.4.1 Mass-Luminosity Relation

So far, we have parametrized the mass-luminosity relation of red and blue centrals via a broken power-law relation. In $\log M$ – $\log L$ space, this roughly describes a piecewise linear relation with a low and a high-mass slope. We now add to this a quadratic term in $\log M$ – $\log L$ space. This can be written as

$$L_c(M) = L_0 \frac{(M/M_1)^{\gamma_1}}{(1 + M/M_1)^{\gamma_1 - \gamma_2}} \left(\frac{M}{M_1} \right)^{\gamma_3 \log(M/M_1)}. \quad (27)$$

For $\gamma_3 = 0$, this relation reduces to the previous parametrization with low-mass slope γ_1 and high-mass slope γ_2 . Thus, we introduce 2 new free parameters, $\gamma_{3,r}$ and $\gamma_{3,b}$, for red and blue centrals respectively. We use priors in the range $[-0.3, +0.3]$ for both. Applying this more general model to SDSS yields marginalized posteriors of $0.003^{+0.022}_{-0.022}$ and $-0.158^{+0.111}_{-0.086}$ for $\gamma_{3,r}$ and $\gamma_{3,b}$, respectively. Thus, the

data is compatible with $\gamma_3 = 0$ in both cases, indicating that this additional model freedom is not favoured by the SDSS analysis.

5.4.2 Red Fraction of Centrals

So far, we have assumed that the red fraction of centrals $f_r(M)$ increases linearly with $\log M$ of the dark matter halo. We now increase the model complexity by including a quadratic term, such that

$$f_r(M) = f_0 + \alpha_f \log M_{12} + \beta_f (\log M_{12})^2, \quad (28)$$

where $M_{12} = M_{\text{vir}}/10^{12} \ h^{-1} \text{M}_\odot$. Furthermore, we demand that the red fraction increases with halo mass. Thus, for $\beta_f < 0$ and $\log M_{12} > -\alpha_f/(2\beta_f)$, we set $f_r(M) = f_0 - \alpha_f^2/(4\beta_f)$. Conversely, for $\beta_f > 0$ and $\log M_{12} < -\alpha_f/(2\beta_f)$, we fix $f_r(M) = f_0 - \alpha_f^2/(4\beta_f)$. Thus, this more general model also allows for a red fraction that is constant above or below a certain halo mass. Evidently, for $\beta_f = 0$, this model reduces to the default parametrization. When choosing a prior of $[-0.5, 0.5]$, the posterior prediction is $-0.068^{+0.072}_{-0.036}$. We find that models where the red fraction does reach a non-unity value above a halo mass of $10^{13} \ h^{-1} \text{M}_\odot$ are slightly

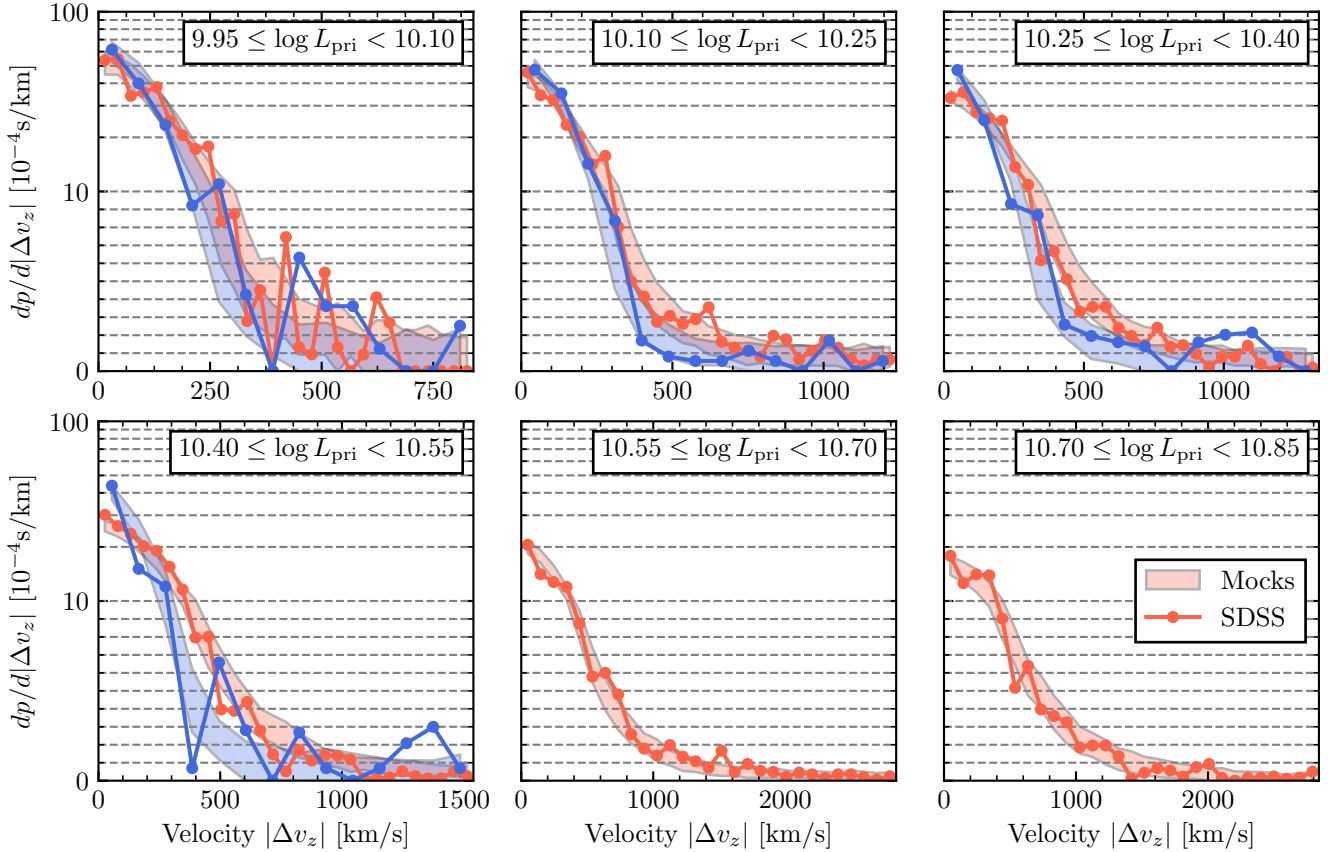


Figure 6. The velocity distribution of secondaries around primaries for SDSS (data points) and the best-fitting bNFW model (bands). Each panel represents a different primary luminosity. We show the results for red and blue primaries separately. Bands show the 68% containment prediction from the best-fitting model for 1000 mock catalogues. Note that the y -scale in each panel is logarithmic in the upper half and linear in the lower in order to show the full dynamic range of the data. Overall, the best-fitting model is able to reproduce all qualitative trends in the data.

favoured by the data. However, the overall evidence ratio tells us that this model is not better in explaining the data than the default model.

5.4.3 Satellite Occupation

In principle, we might expect that at fixed halo mass the satellite occupation depends on the colour of the central galaxy. The reasoning for this is that galaxy colour might correlate with halo formation time, which in turn correlates with subhalo or satellite occupation. [Johannes: Frank, add your preferred citations here.] We test this by allowing the satellite CLF to vary depending on central colour, such that

$$\Phi_{\text{sat}}(L|M, \text{blue central}) = \zeta \times \Phi_{\text{sat}}(L|M, \text{red central}). \quad (29)$$

We choose $[0.5, 2.0]$ as a prior for this boost factor ζ and find $0.80^{+0.15}_{-0.12}$ as the posterior probability. Thus, the posterior does not strongly exclude $\zeta = 1$ and the overall evidence ratio does not significantly favour this model over the default parametrization.

6 DISCUSSION

One of the main goals of this work and Paper I is to improve the analysis of satellite kinematics compared to previous efforts. Specifically, we want to investigate previously reported tensions between results from satellite kinematics and other methods (see e.g. Leauthaud et al. 2012; Mandelbaum et al. 2016). To this end, we compare our results obtained to those obtained from projected galaxy clustering (Zehavi et al. 2011; Vakili & Hahn 2016; Zentner et al. 2016), redshift space clustering (Guo et al. 2015b), group multiplicity functions (Sinha et al. 2017) and combinations thereof. We note that such a juxtaposition is only superficial and a detailed comparison would involve analysing all those observables in our mock catalogues and compare them to SDSS. Nevertheless, it this effort should allow us uncover strong discrepancies, if present.

6.1 Radial Profile of Satellites

Our findings shown in Figure 4 favour a radial distribution of satellites that is biased with respect to dark matter. While our results are consistent with satellites following an NFW profile, their concentration parameter is likely lower than

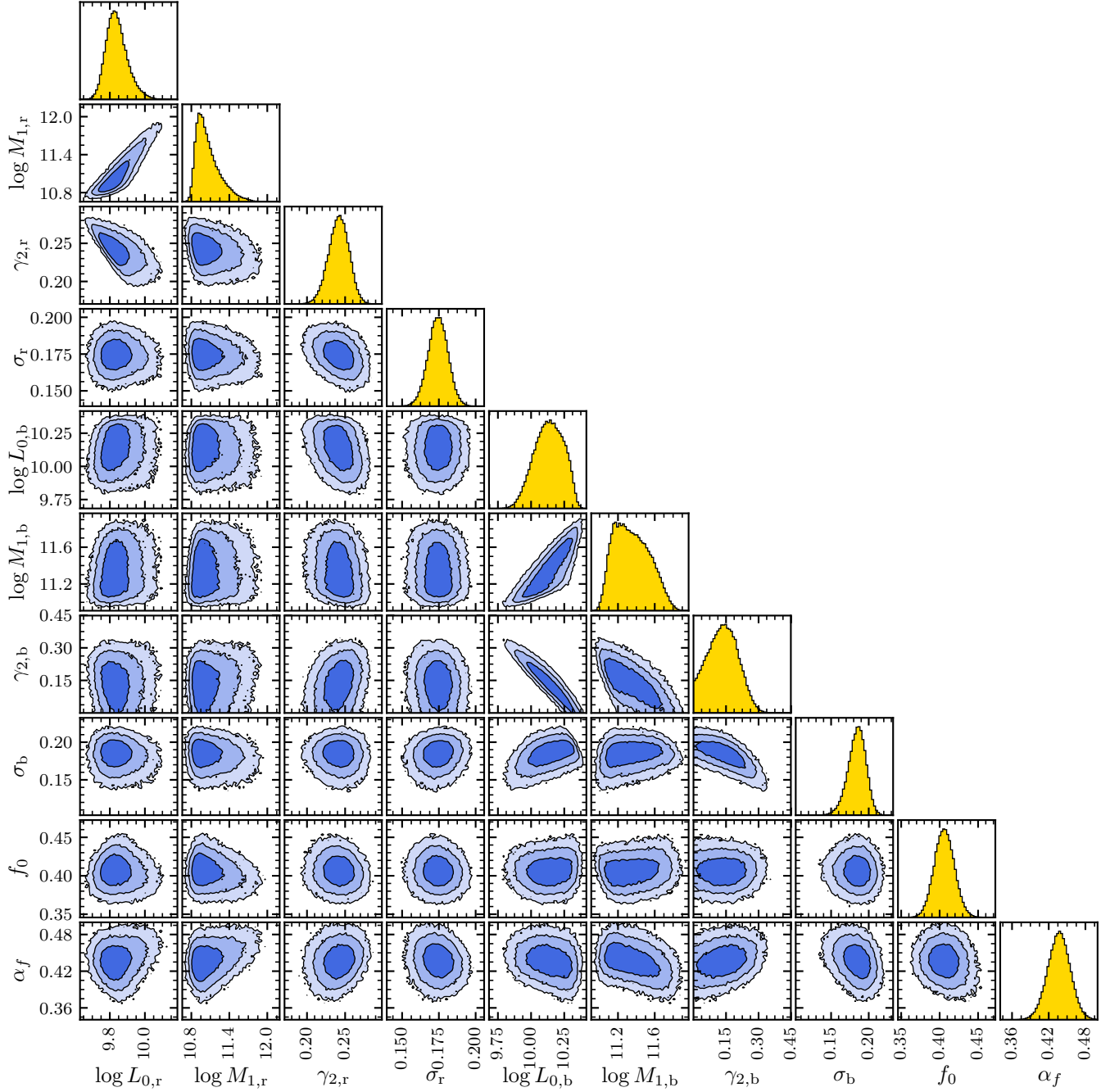


Figure 7. Marginalized posteriors for the galaxy–halo connection for the bNFW radial profile. We do not show parameters related to the satellite occupation, α_s , b_0 , b_1 and b_2 , and the low-mass slopes of the mass–luminosity relation, $\gamma_{1,r}$ and $\gamma_{1,b}$, as they are poorly constrained. The diagonal shows marginalized 1D posteriors and off diagonal panels the 2D posteriors. In the latter case, lines demarcate the 68%, 95% and 99% containment of the posterior.

that of dark matter in the same halo by a factor of around two. At the same time, the cored profile that fits the distribution of M_{peak} selected subhaloes, as described in Paper I, does not provide an adequate description of the SDSS data. Thus, satellites seem to follow a steeper profile than subhaloes.

Many previous studies reported similar findings, particularly that satellite distributions follow an NFW profile (see e.g. van den Bosch et al. 2005a; Yang et al. 2005; Guo

et al. 2012). This is in contrast to the results of More et al. (2009a), who argue that satellites follow a cored profile. As described in Paper I, this is likely due to More et al. (2009a) not correcting for fibre collisions in SDSS. The finding that satellites have a steeper radial profile than subhaloes but less centrally concentrated than dark matter is also in good agreement with independent studies (Yang et al. 2005; Chen 2008) and hydrodynamical simulations (Nagai & Kravtsov 2005; Vogelsberger et al. 2014). Finally, we note for such

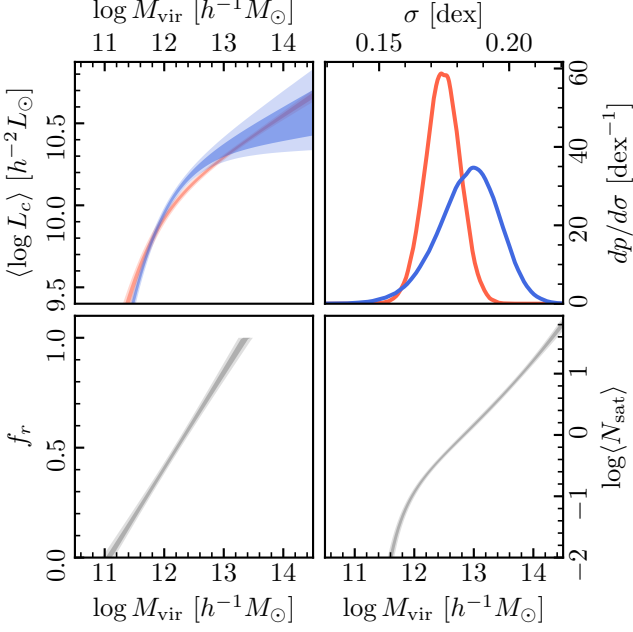


Figure 8. The marginalized posteriors on the galaxy–halo connection. We show the mass–luminosity relation for red and blue centrals (upper, left), the scatter in central luminosity at fixed halo mass for red and blue centrals (upper, right), the red fraction of centrals as a function of halo mass (lower, left) and the average occupation with satellites above $10^{9.5} h^{-2} L_{\odot}$ (lower, right). Bands denote 68% and 95% containment ranges of the posterior.

comparisons, one should keep in mind that the exact radial profile of satellites will likely depend on the halo mass range and the properties of satellites due to colour and luminosity segregation (Chen 2008; Guo et al. 2013).

6.2 Galaxy–Halo Occupation

A detailed comparison with results from clustering and other methods is hampered by two complications. First, all observables like satellite kinematics or clustering will have a dependence on cosmology at fixed models for the galaxy–halo connection. Furthermore, there will be a dependence on the halo definition and the halo finder used. It is thus prudent to only compare results obtained under similar cosmological parameters and halo definitions. We compare our results to the studies of Guo et al. (2015b), Zentner et al. (2016), Vakili & Hahn (2016) and Sinha et al. (2017). Coincidentally, all works assume cosmological parameters compatible with the recent results of Planck Collaboration et al. (2014). Even more, all studies used halo catalogues derived with the **ROCKSTAR** halo finder (Behroozi et al. 2013) and spherical over-density M_{vir} haloes. A second complication is that all those studies use colour-independent halo occupation distribution (HOD) models. Such models only describe the number of galaxies brighter than a luminosity threshold living in a halo of mass M . They do not inherently describe the entire luminosity distribution of galaxy, e.g. the mass–luminosity relation, or their dependence on the colour of the central galaxy. Thus, to allow a comparison we need to downgrade our CLF results to such an HOD framework.

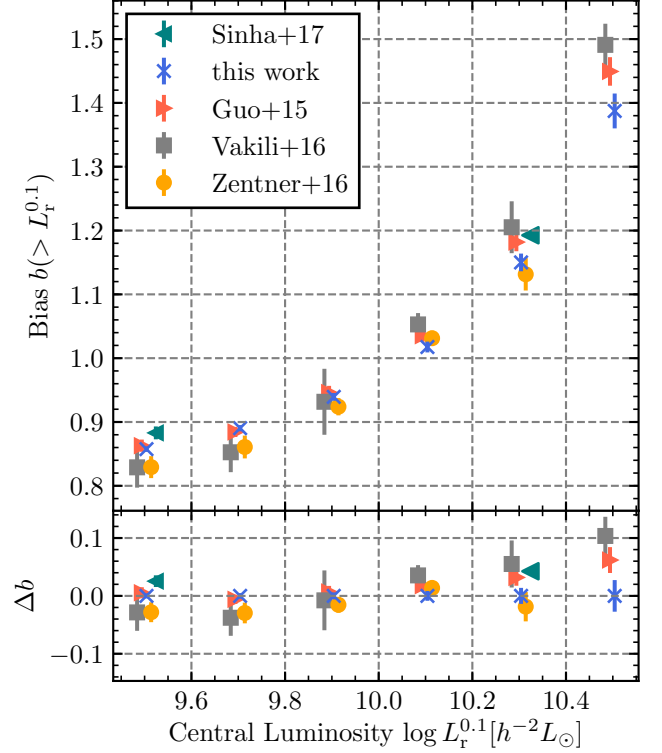


Figure 9. Prediction on the average clustering bias of central galaxies above a given luminosity (upper panel). We compare predictions from Guo et al. (2015b), Vakili & Hahn (2016), Sinha et al. (2017) and our results. All results include 68% error bars. Predictions have been made assuming the Tinker et al. (2008) halo mass function and Tinker et al. (2010) model for the halo bias. We also directly compare the differences between this work and other studies at fixed luminosity (lower panel). The points are slightly offset in the horizontal direction to improve clarity. They all correspond to magnitude limits of $M_r^{0.1} = -19.0, -19.5, -20.0, -20.5, -21.0$ and -21.5 .

Guo et al. (2015b) model the redshift space clustering of galaxies in SDSS using the framework of Zheng & Guo (2016) to accurately and efficiently calculate correlation functions. Specifically, they model both the projected correlation function and the monopole, dipole and quadrupole moments of the redshift space clustering. Furthermore, in addition to HOD parameters, they fit parameters describing the central and satellite velocity bias. Zentner et al. (2016) and Vakili & Hahn (2016) use **halotools** (Hearin et al. 2016) to fit the projected correlation function of galaxies in SDSS DR7. We note that the observational data used in Zentner et al. (2016) comes from Zehavi et al. (2011), whereas Vakili & Hahn (2016) use data from Guo et al. (2015b). Compared to Guo et al. (2015b), they also specifically allow for the presence of galaxy assembly bias (Zentner et al. 2014). Finally, Sinha et al. (2017) also fit the friends-of-friends group multiplicity function in addition to the projected correlation function. Their focus lies on the realistic simulation of observational effects by comparing observational results to a large number of mock catalogues.

To make a most “fair” comparison of our results with clustering studies, we analyse the average spatial bias of all

centrals above a given luminosity threshold. This is most closely related to observables in galaxy clustering, specifically the clustering strength in the 2-halo regime, and should test agreement with clustering predictions despite the different parametrizations of the galaxy–halo relationship. Due to the tight relation between halo mass and bias, it is also indicative of the average halo mass. In Figure 9 we show one such comparison between the results of this work (blue crosses) and those of Guo et al. (2015b) (red triangles), Zentner et al. (2016) (orange circles), Vakili & Hahn (2016) (grey squares) and Sinha et al. (2017) (turquoise triangles). In all cases, we made predictions by applying the HOD model to the Tinker et al. (2008) mass function and the bias description of Tinker et al. (2010). Slight differences in the cosmological parameters and redshifts of the halo catalogues have been accounted for in the calculation of the halo mass function and the bias. For example, Sinha et al. (2017) use slightly different cosmological parameters and halo catalogues derived for redshifts $z = 0.054$ and $z = 0.132$ to model the $M_r^{0.1} = -19$ and $M_r^{0.1} = -21$ sample, respectively. On the other hand, Guo et al. (2015b), Zentner et al. (2016) and Vakili & Hahn (2016) all use $z = 0$ catalogues. All results include 68% posterior uncertainties. The predictions of Zentner et al. (2016) and Vakili & Hahn (2016) are from models without the possibility for assembly bias. The posterior from Sinha et al. (2017) is the one using the M_{vir} halo definition and without principal component analysis (PCA).

Overall, we see an excellent qualitative agreement between all studies compared here. In all works, the bias of central galaxies increases strongly with their luminosity. Given the tight correlation of halo mass with bias, this indicates a positive correlation between halo mass and galaxy luminosity. At larger luminosities, our results predict a slightly lower central galaxy bias than those of Guo et al. (2015b), Vakili & Hahn (2016) and Sinha et al. (2017). The disagreement is generally at the level of around 2σ . However, Zentner et al. (2016) predict values similar to our results. Note that, for example, for the $L > 10.5 h^{-2} L_\odot$ ($M_r = -21.5$) sample, all clustering studies use a similar sample of galaxies with redshifts of up to ~ 0.2 , whereas our results are derived with galaxies at lower redshifts. Thus, all other studies are intrinsically correlated at fixed luminosity and their findings cannot be regarded as independent. Thus, any apparent agreement should not be seen as evidence that true average spatial bias is indeed higher than inferred from our study.

In our work, we also constrain the scatter in luminosity at fixed halo mass. Unfortunately, this is not directly constrained by results of simple HOD fitting, so we cannot compare our results with the studies previously discussed. **[Johannes: In principle, $\sigma_{\log M}$ in HOD models is related to the scatter in luminosity. However, these constraints are poor to begin with. So a comparison would not be useful. The comparison is also very difficult to make since it depends on the slope of the mass-luminosity relation and is not constrained in HOD models. Should I mention that here?]** Yang et al. (2008) have estimated the scatter in luminosity as inferred from their group catalogues. To estimate group masses, they assumed a one-to-one relationship between total group luminosity of galaxies above $M_r^{0.1} < -19.5$ and halo mass. They find a scatter in BHG luminosity of around $\sigma \sim 0.15$ dex

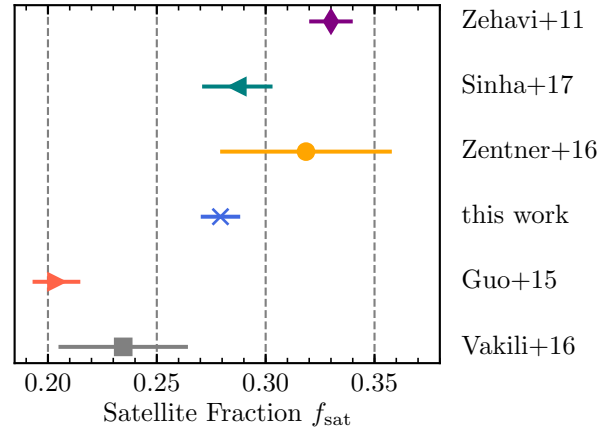


Figure 10. Predictions on the satellite fraction of galaxies brighter than $10^{9.5} h^{-2} L_\odot$ for various studies. Our results, as well as those of Guo et al. (2015b), Vakili & Hahn (2016) and Sinha et al. (2017), have been derived assuming the Tinker et al. (2008) halo mass function. Results for Zehavi et al. (2011) are given directly in their respective studies. Error bars show the 1σ uncertainty.

for red and blue BHGs, in good agreement with our result of $\sigma_r = 0.169^{+0.006}_{-0.006}$ and $\sigma_b = 0.178^{+0.008}_{-0.008}$. Only at the lower mass end, they find a scatter of ~ 0.10 dex. However, for such low-mass groups, the central dominates the total group luminosity, and thereby the halo mass estimate. Thus, the inferred scatter is likely to be artificially suppressed. Cacciato et al. (2013) have analysed the luminosity function, galaxy–galaxy lensing and clustering in SDSS in order to constrain cosmological parameters. Nevertheless, the scatter in luminosity at fixed halo mass was a by-product of this analysis. They find $\sigma_c = 0.157^{+0.007}_{-0.007}$, without making a distinction between red and blue galaxies. It is straightforward to show that this scatter for all galaxies is related to the individual scatter of red and blue galaxies via

$$\sigma_c^2 = f_r \sigma_r^2 + f_b \sigma_b^2 + 2f_r f_b (\langle \log L_r \rangle - \langle \log L_b \rangle)^2, \quad (30)$$

where f_r and f_b are the red and blue fraction of centrals at any given halo mass. As expected, the scatter is slightly increased if red and blue centrals have different average luminosities. For example, for $\sigma_r = \sigma_b = 0.17$, $f_r = f_b = 0.1$ and $|\langle \log L_r \rangle - \langle \log L_b \rangle| = 0.1$, we get $\sigma_c = 0.184$. Ultimately, bearing this in mind, our results imply a slightly larger scatter than what is found by Cacciato et al. (2013), but still in reasonable agreement.

In Figure 10 we compare predictions for the satellite fraction of galaxies above $L = 10^{9.504} h^{-2} L_\odot$ ($M_r^{0.1} = -19$). We used the Tinker et al. (2008) mass function and accounted for slightly different cosmologies and redshifts to calculate the values for our work and those of Guo et al. (2015b), Zentner et al. (2016), Vakili & Hahn (2016) and Sinha et al. (2017). We also include the results reported in Zehavi et al. (2011), which are, however, derived under significantly different cosmological parameters. In this comparison, we note a stronger disagreement between different predictions than in the previous one. The results of Guo et al. (2015b) imply a satellite fraction of around $\sim 21 \pm 1\%$, whereas our finding is $\sim 28 \pm 1\%$. The results of

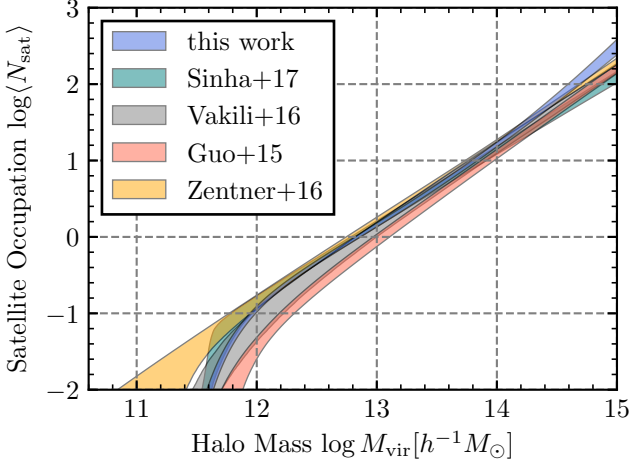


Figure 11. Posterior predictions on the satellite occupation above $10^{9.504} h^{-2} L_{\odot}$ ($M_r^{0.1} = -19$) from several studies. Bands denote the 68% posterior uncertainty. We compare the results of our work, Guo et al. (2015b), Vakili & Hahn (2016) and Sinha et al. (2017).

Vakili & Hahn (2016) fall in between those two values, but with roughly three times larger uncertainties. Sinha et al. (2017) find a satellite fraction that is in good agreement with our findings and incompatible with those of Guo et al. (2015b). Zentner et al. (2016) predict a similarly high fraction $\sim 32 \pm 4\%$ but with the largest uncertainty. Finally, Zehavi et al. (2011) find $\sim 33 \pm 1\%$, but for a different cosmology and halo definition. It is unclear at this point what the reason for the disagreement between the different studies is. Given that Zentner et al. (2016) and Vakili & Hahn (2016) only used the projected correlation function and find the satellite fraction to be loosely constrained, it seems that this observation is compatible with a large range of satellite fractions. On the other hand, taken at face value, the redshift space clustering Guo et al. (2015b) seems to imply a lower fraction than the group multiplicity function Sinha et al. (2017) and satellite kinematics.

We further explore this point in Figure 11, where we show the satellite occupation as a function of halo mass. The threshold is $10^{9.504} h^{-2} L_{\odot}$ ($M_r^{0.1} = -19$) in all cases. Qualitatively, the results all agree very well with each other. However, at lower masses, $\log M_{\text{vir}}/(h^{-1} M_{\odot}) \sim 12 - 13$, the results of Guo et al. (2015b) imply a lower satellite occupation than our findings and those of Zentner et al. (2016) and Sinha et al. (2017). On the other hand, at the high mass end, $\log M_{\text{vir}}/(h^{-1} M_{\odot}) \gtrsim 14.5$, we find a higher satellite occupation than all other studies.

Altogether, we find a satisfactory agreement of our results with a variety of different studies. In all the comparisons here, we do not find our predictions to be strong outliers. Although certain tensions, especially for the satellite fraction, exist, they affect all studies together, not just the results of satellite kinematics. If there were truly a discrepancy in those results and satellite kinematics, it could be due to a variety of reasons. In addition to systematic uncertainties affecting satellite kinematics, which we will discuss in the next section, there are also uncertainties affecting the

modelling of clustering and group multiplicity functions. A prime factor could be assembly bias, the fact that the spatial clustering of dark matter haloes depends on secondary halo properties other than halo mass (Gao et al. 2005; Wechsler et al. 2006). If the galaxy occupation is correlated with those secondary properties at fixed halo mass, it will affect the clustering at fixed values for the galaxy occupation as a function of halo mass (see e.g. Zentner et al. 2014; Hearin et al. 2016). Another uncertainty is the radial profile of satellites. For example, all other studies assumed that satellites follow the dark matter radial profile in an unbiased fashion. Our results, on the other hand, imply that satellites are anti-biased with respect to dark matter. Another uncertainty is whether satellite numbers follow a Poisson distribution. Since clustering is a 2-point statistic, it is very sensitive to the second moment of the distribution. Furthermore, Guo et al. (2015b), Vakili & Hahn (2016) and Sinha et al. (2017) all assumed the 5-parameter model of Zheng et al. (2007) to model observational data. However, as argued by Sinha et al. (2017), this model might not be flexible enough to fit a variety of observations. Finally, we note that the agreement between satellite kinematics and clustering likely depends on the cosmological parameters, as it does for clustering and lensing (Cacciato et al. 2009, 2013). Thus, we might expect an agreement between all those studies only under the right cosmological parameters.

6.3 Caveats

It is important to consider that several simplified assumptions have been made in this work that might potentially bias our inferences. A prime concern is that we have assumed that satellite galaxies obey the spherically symmetric, time-independent Jeans equation without anisotropy. However, it is known that satellite distributions are not generally spherically symmetric, do have some radial anisotropy (Diemand et al. 2004; Wojtak & Mamon 2013) and are not fully relaxed, i.e. time-independent (Ye et al. 2017). Observationally, Guo et al. (2015a,b) argue that satellite galaxies have a lower velocity dispersion than dark matter particles, if both follow the same radial distribution. Thus, both populations cannot simultaneously obey the Jeans equation in the form we assumed. Additionally, it is also well known that the halo core and the central galaxy are unlikely to be at rest with the entire halo (van den Bosch et al. 2005b; Behroozi et al. 2013; Guo et al. 2015a,b, 2016; Ye et al. 2017), something we have implicitly assumed.

In Paper I, we have addressed such concerns by analysing mocks where satellites are placed on subhaloes and centrals in the dark matter halo core. Evidently, all the simplifications are violated in such mock catalogues. Nevertheless, our inferences were not strongly biased with respect to the input parameters, indicating that those simplifications do not have a strong impact on observables. We perform a similar exercise here. We take the best-fit parameters of the Cored profile and create 100 mocks where satellites are placed on subhaloes and centrals in the dark matter halo core. We then calculate the average of all observables and compare it to the case where satellites follow our analytic prescription. As we show in Figure 12, differences between the two predictions are generally below the 1σ level, especially for $\log \sigma_{\text{hw}}$ and $\sigma_{\text{hw}}^2/\sigma_{\text{sw}}^2$, for all luminosities and in-

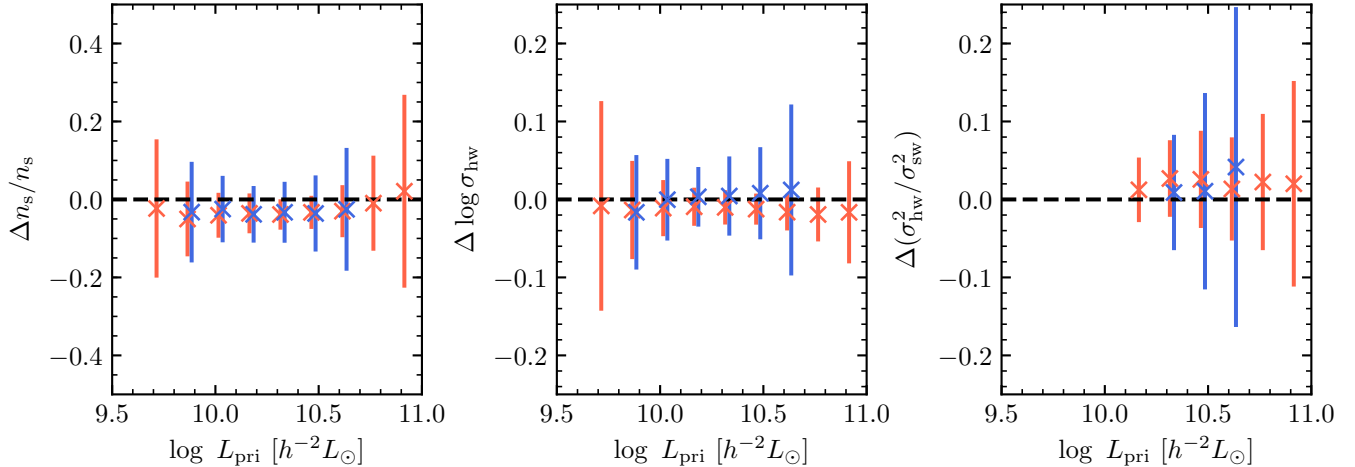


Figure 12. Difference between the predicted kinematic observables when satellites are placed on subhaloes, as opposed to following the simplified analytical prescription. In all panels, we show the difference of the subhalo based mocks with respect to the analytical phase-space results. In each panel, the different colours denote the colour of the primary and the points are slightly offset horizontally for clarity. Error bars show the 68% scatter between different mock realizations, as inferred from the analytical mocks.

dependent of whether the primaries are red or blue. Thus, if satellites would follow M_{peak} -selected subhaloes, our simplifications would only have a mild impact on observables. Note however that such subhaloes have a different radial profiles than satellites in SDSS. Thus, these results might not be indicative of how well our simplifications fare with respect to satellites in SDSS. Finally, we note that for future work on satellite kinematics, one could use phase-space coordinates extracted from hydrodynamical simulations to model satellite populations.

Another potential concern is the fact that we do not have 100% completeness in BHGs. Particularly, our selection criterion will be biased against low-luminosity BHGs in high-density regions. If the mass-luminosity relation had a dependence on environment, a scenario we do not model, then, similar to the effects of assembly bias in clustering, our inferences would be biased.

We also assumed that centrals are always BHGs in their respective halo. However, we know that this is not the case (Skibba et al. 2011; Hikage et al. 2013; Wang et al. 2014; Hoshino et al. 2015; Lange et al. 2018). However, Lange et al. (2018) have shown that this simplification only has a very small effect on the measured velocity dispersion. Nevertheless, generally this should lead our analysis to slightly overestimate halo masses.

7 CONCLUSION

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1125915. 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 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**². 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.

We thank Hong Guo, Manodeep Sinha and Mohammadjavad Vakili for making the posteriors of their analysis available to us.

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
- Astropy Collaboration 2013, *A&A*, 558, A33
- 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
- Blanton M. R., et al., 2005, *AJ*, 129, 2562
- Buchner J., et al., 2014, *A&A*, 564, A125

² <https://github.com/esheldon/pymangle>

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
 Chen J., 2008, A&A, 484, 347
 Diemand J., Moore B., Stadel J., 2004, MNRAS, 352, 535
 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)
 Foreman-Mackey D., 2016, The Journal of Open Source Software, 24
 Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66
 Guo Q., Cole S., Eke V., Frenk C., 2012, MNRAS, 427, 428
 Guo Q., Cole S., Eke V., Frenk C., Helly J., 2013, MNRAS, 434, 1838
 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., et al., 2016, preprint, (arXiv: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
 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
 Mandelbaum R., Wang W., Zu Y., White S., Henriques B., More S., 2016, MNRAS, 457, 3200
 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
 Nagai D., Kravtsov A. V., 2005, ApJ, 618, 557
 Planck Collaboration et al., 2014, A&A, 571, A16
 Sinha M., Berlind A. A., McBride C. K., Scoccimarro R., Piscione J. A., Winking B. D., 2017, preprint, (arXiv:1708.04892)
 Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417
 Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391
 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
 Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
 Vakili M., Hahn C. H., 2016, preprint, (arXiv:1610.01991)
 Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in Science & Engineering, 13, 22
 Vogelsberger M., et al., 2014, Nature, 509, 177
 Wang L., et al., 2014, MNRAS, 439, 611
 Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
 Wojtak R., Mamon G. A., 2013, MNRAS, 428, 2407
 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., 2008, ApJ, 676, 248
 Ye J.-N., Guo H., Zheng Z., Zehavi I., 2017, ApJ, 841, 45
 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)

Zheng Z., Guo H., 2016, MNRAS, 458, 4015
 Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760
 Zu Y., Mandelbaum R., 2015, MNRAS, 454, 1161
 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

APPENDIX A: GALAXY–HALO CONNECTION

Here, we briefly describe our parametrization of the galaxy–halo connection. We refer the reader to Paper I for a more in-depth discussion.

The occupation of dark matter halos with galaxies is governed by a CLF model. Particularly, the CLF specifies the average number of galaxies with a luminosity in the range $[L - dL/2, L + dL/2]$ residing in a halo of mass M . We split the galaxy occupation into a central and satellite component,

$$\Phi_{\text{tot}}(L|M) = \Phi_c(L|M) + \Phi_s(L|M). \quad (31)$$

For the central galaxies, we choose two different parametrizations for red and blue centrals,

$$\Phi_c(L|M) = f_r(M)\Phi_{c,r}(L|M) + f_b(M)\Phi_{c,b}(L|M). \quad (32)$$

Here, $f_r(M) = 1 - f_b(M)$ is the probability that a galaxy residing in a halo of mass M is red. This probability is parametrized by

$$f_r(M) = f_0 + \alpha_f \log M_{12}, \quad (33)$$

where $M_{12} = M/[10^{12} h^{-1} M_\odot]$. We assume that every halo hosts exactly one central whose luminosity is drawn from a log-normal distribution with scatter σ ,

$$\Phi_c(L|M)dL = \frac{\log e}{\sqrt{2\pi}\sigma_c^2} \exp\left[-\frac{(\log L - \log L_c(M))^2}{2\sigma_c^2}\right] \frac{dL}{L}. \quad (34)$$

The median luminosity, which can be different for red and blue centrals, is parametrized by

$$L_c(M) = L_0 \frac{(M/M_1)^{\gamma_1}}{(1 + M/M_1)^{\gamma_1 - \gamma_2}}. \quad (35)$$

Additionally, the scatter in central luminosity is allowed to be different for red and blue centrals. Altogether, $L_{0,r}$, $M_{1,r}$, $\gamma_{1,r}$, $\gamma_{2,r}$, σ_r , $L_{0,b}$, $M_{1,b}$, $\gamma_{1,b}$, $\gamma_{2,b}$, σ_b , f_0 and α_f are the free parameters governing the occupation of dark matter haloes with central galaxies.

The occupation of satellite galaxies is assumed to follow of Poisson distribution with expectation value

$$N_s(M) = \int_{L_{\text{th}}}^{\infty} \Phi_s(L|M)dL, \quad (36)$$

where L_{th} is a luminosity threshold above which we populate dark matter haloes. Following Yang et al. (2008), the CLF of satellite galaxies follows a modified Schechter function,

$$\Phi_s(L|M) = \phi_s^*(M) \left(\frac{L}{L_s^*(M)}\right)^{\alpha_s} \exp\left[-\left(\frac{L}{L_s^*(M)}\right)^2\right] \frac{dL}{L}.$$

(37)

α_s , a free parameter, describes the low-luminosity slope of the satellite CLF. $L_s^*(M)$ describes the position of the exponential cut-off. Yang et al. (2008) found that this is closely related to the average luminosity of centrals, $L_s^*(M) \sim 0.562L_c(M)$. Thus, we choose

$$L_s^*(M) = 0.562L_{c,r}(M). \quad (38)$$

We note that, in principle, this CLF model allows satellites to be brighter than centrals. We explicitly remove satellites that are brighter than their respective centrals from our mock galaxy catalogues. Finally, the overall mass-dependent normalization of the satellite CLF is described by

$$\log[\phi_s^*(M)] = b_0 + b_1 \log M_{12} + b_2 (\log M_{12})^2. \quad (39)$$

Finally, satellites are also randomly assigned colours based on the parametrization in eq. (33) with $f_{0,\text{sat}} = 0.44$ and $\alpha_{f,\text{sat}} = 0.14$ (compare Yang et al. 2008). This colour assignment only matters in the rare cases that satellites are identified as primary candidates. Thus, we have α_s , b_0 , b_1 and b_2 as free parameters for the satellite CLF.

Central galaxies are assumed to be located at the centre of each dark matter halo and to be at rest with the bulk halo velocity. On the other hand, satellite galaxies are assumed to follow spherically symmetric profile whose number density can be described by a generalized NFW profile,

$$n_s(r) \propto \left(\frac{r}{\mathcal{R}r_s}\right)^{-\gamma} \left(1 + \frac{r}{\mathcal{R}r_s}\right)^{\gamma-3}, \quad (40)$$

where r_s is the scale radius of the dark matter profile. γ and \mathcal{R} are free parameters. For $\gamma = \mathcal{R} = 1$, satellites follow the dark matter distribution in an unbiased fashion and become increasingly shallow for smaller γ and larger \mathcal{R} . In this work, we consider three choices for (γ, \mathcal{R}) to describe the SDSS data: (1, 1) which we call “NFW”, (1, 2) referred to as “bNFW” and (0, 2.5) which we call “Cored”. We assume that satellite galaxies obey the spherically symmetric Jeans equation without velocity anisotropy. In this case, the velocity dispersion at any given radius can be computed via

$$\sigma^2(r|V_{\text{vir}}, c_{\text{vir}}) = \frac{c_{\text{vir}} V_{\text{vir}}^2}{\mathcal{R}^2 g(c_{\text{vir}})} \left(\frac{r}{\mathcal{R}r_s}\right)^\gamma \left(1 + \frac{r}{\mathcal{R}r_s}\right)^{3-\gamma} \int_{r/r_s}^{\infty} \frac{g(y)dy}{(y/\mathcal{R})^{\gamma+2}(1+y/\mathcal{R})^{3-\gamma}}, \quad (41)$$

where $V_{\text{vir}} = GM_{\text{vir}}/r_{\text{vir}}$ is the virial velocity and $c_{\text{vir}} = r_{\text{vir}}/r_s$ the virial concentration. Finally,

$$g(y) = \ln(1+y) - \frac{y}{1+y}. \quad (42)$$

Higher order moments of the velocity distribution are ignored because we also do not analyse them in this work.

APPENDIX B: MOCK CATALOGUES

A key advantage of our study over works using satellite kinematics is to use detailed mock catalogues to estimate observables and their associated errors. Here, we outline how mock catalogues are created.

All mock catalogues used in this work are based on the

SMDPL dark matter-only simulation. Specifically, we use the $z = 0$ ROCKSTAR (Behroozi et al. 2013) halo catalogue³. From this, we extract all main haloes with a virial mass of $M_{\text{vir}} \geq 3 \times 10^{10} h^{-1} M_\odot$, corresponding to 300 times the dark matter particle mass in SMDPL. The cosmology used in SMDPL is the same as the one we assume. Furthermore, the halo mass resolution is more than sufficient to resolve galaxies above $10^{9.5} h^{-2} L_\odot$ (compare e.g. Guo et al. 2015b; Sinha et al. 2017) and the volume of SMDPL is more than 10 times bigger than the volume of SDSS we analyse. Thus, SMDPL is the ideal simulation to base our analysis on.

We use `halotools` to populate the dark matter catalogue with galaxies according to the CLF recipe described above. Specifically, for each dark matter halo, we first draw a colour for the central and afterwards a luminosity. For the satellites, we first calculate the expected occupation and then draw a number from a Poisson distribution with the same mean. Afterwards, we assign luminosities and colours based on the CLF parameters. Satellites that are brighter than their respective centrals are removed. Generally, this only eliminates roughly 1% of the satellite population. For both the centrals and satellites we consider all galaxies above a luminosity of $10^9 h^{-2} L_\odot$. This threshold is lower than that of galaxies we analyse in SDSS and is chosen to have a sufficient number of targets for potential fibre collisions. Finally, central galaxies are assigned the phase-space position of the halo core and satellites phase-space positions relative to that based on the assumed radial profile and the spherical Jeans equation.

Now having a real-space galaxy catalogue, we place a virtual observer with a random position and orientation inside the simulation box. We periodically repeat the galaxy catalogue to include all galaxies within $z = 0.15$. Next, we calculate apparent magnitudes for all galaxies, taking into account average k and evolution corrections. We remove all targets with $m_r^{0.1} > 17.6$. Afterwards, redshift space distortions and SDSS spectroscopic redshift errors are simulated by perturbing the cosmological redshift z of all galaxies by

$$\Delta z = (1+z) \times \frac{(v_{\text{los}} + \mathcal{N}(0, (35 \text{ km s}^{-1})^2))}{c}, \quad (43)$$

where v_{los} is the line-of-sight velocity and $\mathcal{N}(0, (35 \text{ km s}^{-1})^2)$ represents a random variable drawn from a Gaussian distribution with 0 mean and 35 km s^{-1} dispersion. Afterwards, we apply the NYU VAGC `bright0` survey window and mask. Finally, we simulate fibre collisions. For each galaxy in the catalogue, we determine all other targets within an angular separation of $55''$. We make a table of all galaxies having such close neighbours. We then remove a random galaxy from that list. Furthermore, with a probability of 65% we assign all close neighbours still on the list the redshift of this galaxy and also remove them. This process is repeated until all galaxies have been removed from the table. In Paper I, we have shown that this fibre collision scheme, although simplified compared to what happens in SDSS (?), is sufficient for our mock catalogues.

³ <http://yun.ucsc.edu/sims/SMDPL/hlists/index.html>