Mon. Not. R. Astron. Soc. **000**, ??–?? () Printed 4 June 2016 (MN LATEX style file v2.2)

Constraints on Assembly Bias from Galaxy Clustering

Andrew R. Zentner[1], Andrew P. Hearin[2], Frank C. van den Bosch[3], Antonio Villareal[1], Johannes Lange[3], P. Rogers Nelson[4], Probably some others of unspecifi

[1] *Department of Physics and Astronomy & Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, Pittsburgh, PA 15260*
[2] *Yale Center for Astronomy & Astrophysics, Yale University, New Haven, CT*
[3] *Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT*
[4] *Paisley Park, Chanhassen, MN*

Today

ABSTRACT

We fit SDSS DR7 data with models that include assembly bias.

Important points: -supersedes previous analyses even for standard HOD. -cannot rule out assembly bias using galaxy clustering. -some samples favor assembly bias.

1 INTRODUCTION

For more than a decade, halo occupation modeling has been used to interpret large-scale structure measurements and exploit these measurements to constrain galaxy formation models and cosmology (e.g., Yang et al. 2003; Tinker et al. 2005; Zehavi et al. 2005; Porciani & Norberg 2006; van den Bosch et al. 2007; Zheng et al. 2007; Conroy & Wechsler 2009; Yang et al. 2009; Zehavi et al. 2011; Guo et al. 2011; Wake et al. 2011; Yang et al. 2011, 2012; Leauthaud et al. 2012; Tinker et al. 2013; Cacciato et al. 2013; More et al. 2013; Guo et al. 2014; Zu & Mandelbaum 2015). The key assumptions underlying halo occupation modeling are: (1) all galaxies reside in dark matter halos that are biased tracers of the density field; and (2) galaxies occupy halos as a function of halo masses only. It is now well known that halo bias depends upon halo properties other than mass (e.g. Gao et al. 2005; Wechsler et al. 2006; Gao & White 2007; Zentner 2007; Dalal et al. 2008; Lacerna & Padilla 2011), an effect called halo assembly bias. If galaxies occupy halos as a function of properties other than halo mass, then standard halo occupation methods will be subject to a systematic error due to galaxy assembly bias. Several of us have previously shown that this error can be significant in an analysis of galaxy clustering and can bias inferences about many aspects of galaxy evolution (Zentner et al. 2014). Consequently, we have also developed halo occupation models that enable galaxies to occupy halos in a manner that depends upon several halo properties (Hearin et al. 2016). In this paper, we revisit the interpretation of luminosity-dependent galaxy clustering in the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7) data, analyzed previously by Zehavi et al. (2011), in the context both standard halo occupation models and these new models.

Some history here, blah blah blah, Frank loves this stuff. Therefore, I'm sure this part of the introduction will get quite lengthy.

Our work is important for several reasons. Our work is a re-analysis of the SDSS DR7 data the overcomes shortcomings of previous analyses. In particular, we use direct population of galaxies in a cosmological simulation so that delicate issues present in analytic modeling, such as scale-dependent halo bias, are treated exactly. The simulation we use is based on the latest Planck cosmological parameters, updating previous work. Furthermore, many important differences between this work and the previous work of Zehavi et al. (2011) were caused by insufficiently large Monte Carlo Markov Chain (MCMC) sampling by Zehavi et al. (2011, Z. Zheng & I. Zehavi, Private Communication). Therefore, our analysis *supersedes previous analyses even in the case of standard halo occupation models.*

Furthermore, our work demonstrates explicitly that significant assembly bias in M_r-selected samples from SDSS DR7 cannot be ruled out based on a standard analysis of galaxy clustering only. In fact, we find that several samples *favor galaxy assembly bias to a degree that is statistically significant.* As demonstrated by Zentner et al. (2014), this conclusion could have important consequences for the interpretation of both extant and forthcoming data.

Our paper is organized as follows. In Section 2, we discuss our implementation of halo occupation models and our assumptions in our parameter inference analysis. We present results for both standard halo occupation analysis and analysis in the context of models with galaxy

assembly bias in Section 3. We summarize our results and draw conclusions in Section 4.

2 METHODS

2.1 Halotools Implementation of HOD Models

To generate predictions for galaxy clustering, we directly populate dark matter halos with mock galaxies using `Halotools`. In this section we review the "standard" HOD-style model used in the present work, and in the following section we use the Decorated HOD model described in Hearin et al. (2016). In both cases, we will only briefly review the salient features of our methodology here; interested readers can always refer to `halotools.readthedocs.io` for further details.

2.1.1 Simulation

All of our analyses are based on the Bolshioi-Planck simulation (Riebe et al. 2011). Bolshoi-Planck was run with cosmological parameters based on Planck Collaboration et al. (2014), $\Omega_\Lambda = 0.693, \Omega_\mathrm{m} = 0.307, \Omega_\mathrm{b} = 0.048, h = 0.7, \mathrm{n_s} = 0.96$ and $\sigma_8 = 0.82$, and with a box size of 250 Mpc/h and a particle mass of $m_\mathrm{p} = 1.35\times10^8 M_\odot/h$. Further information about the Bolshoi-Planck simulation is available at `https://www.cosmosim.org`.

We use publicly available[1] dark matter halo catalogs based on the `ROCKSTAR` halo-finder (Behroozi et al. 2011) and `CONSISTENT TREES` algorithm (Behroozi et al. 2013). In particular, we use the `halotools_alpha_version2` version of the $z = 0$ snapshot of the `'bolplanck'` catalog included with `Halotools`. Halos in these catalogs are based on the virial radius density contrast given in Bryan & Norman (1998). When populating this catalog with mock galaxies, we only use present-day host halos with a value of M_peak that exceeds 300 particles.

2.1.2 Occupation statistics

In standard HOD models, there are two types of galaxies: centrals and satellites. The galaxy-halo connection is specified in terms of $P(N_\mathrm{cen}|M_\mathrm{vir})$ and $P(N_\mathrm{sat}|M_\mathrm{vir})$, the probability that a halo of mass M_vir hosts N_cen central and N_sat satellite galaxies, respectively. $P(N_\mathrm{cen}|M_\mathrm{vir})$ is typically a nearest-integer distribution, as a host halo has only either zero or one central galaxy. So occupation statistics of central galaxies are specified by the first moment of $P(N_\mathrm{cen}|M_\mathrm{vir})$, which we model as

$$\langle N_\mathrm{cen}\rangle(M_\mathrm{vir}) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\log_{10} M_\mathrm{vir} - \log_{10} M_\mathrm{min}}{\sigma_{\log_{10} M}}\right)\right) \quad (1)$$

For every host halo in the catalog we draw a random number from a uniform distribution $\mathcal{U}(0,1)$; for a host halo of present-day virial mass M_vir, a central galaxy is assigned to the halo if the associated random number is less than $\langle N_\mathrm{cen}\rangle(M_\mathrm{vir})$; halos with random values exceeding $\langle N_\mathrm{cen}\rangle(M_\mathrm{vir})$ are left devoid of centrals.

We model the distribution $P(N_\mathrm{sat}|M_\mathrm{vir})$ to be a Poisson distribution with first moment given by a power law,

$$\langle N_\mathrm{sat}\rangle(M_\mathrm{vir}) = \left(\frac{M_\mathrm{vir} - M_0}{M_1}\right)^\alpha. \quad (2)$$

2.1.3 Galaxy profiles

Central galaxies in the standard HOD models reside at the halo center, moving with the same velocity as the host halo peculiar velocity. We model the intra-halo spatial distribution of satellite galaxies to be located within R_vir of the halo center, with a spherically symmetric NFW profile (Navarro et al. 1997). The concentration c of each halo's satellite galaxy profile is taken to be the same as the concentration of the dark matter particles in the halo.[2]

We model the velocities of satellite galaxies to be a Gaussian whose first moment is the host halo velocity and whose second moment is the solution to the isotropic Jeans equation for an NFW profile (More et al. 2009),

$$\sigma_r^2(\tilde{r}|c) = V_\mathrm{vir}^2 \frac{c^2\tilde{r}(1+c\tilde{r})^2}{g(c)}\int_{c\tilde{r}}^\infty \mathrm{d}y \frac{g(y)}{y^3(1+y)^2}, \quad (3)$$

where $\tilde{r} = r/R_\mathrm{vir}$, $g(x) = \ln(1+\mathrm{x}) - \mathrm{x}/(1+\mathrm{x})$, and $V_\mathrm{vir}^2 = GM_\mathrm{vir}/R_\mathrm{vir}$.

2.2 HOD with Assembly Bias: The Decorated HOD

In addition to the standard occupation statistics described in the previous section, in this paper we also use the decorated HOD formalism to connect galaxies to dark matter halos in a manner that has simultaneous dependence on both M_vir and halo concentration. Briefly, we use Equations 1 and 2 as our "baseline" first occupation moments. At fixed M_vir, halos are divided into one of two categories, those of high- and low-concentration, depending on whether the concentration of the halo places it above or below the rank-order percentile f_split, which we keep fixed to $f_\mathrm{split} = 0.5$ throughout the paper. High-concentration halos have a different first occupation moment relative to low-concentration halos of the same mass, $\langle N_\mathrm{gal}|M_\mathrm{vir}, c_\mathrm{high}\rangle \neq \langle N_\mathrm{gal}|M_\mathrm{vir}, c_\mathrm{low}\rangle$. The difference between the first moment of high- and low-concentration halos is modulated by $\mathcal{A}_\mathrm{bias}$, the novel parameter of the decorated HOD governing assembly bias. Values of $\mathcal{A}_\mathrm{bias} = \pm1$ correspond to the maximum strength of assembly bias allowable by the constraint that the model preserves the marginalized first moment, $\langle N_\mathrm{gal}|M_\mathrm{vir}\rangle$; thus regardless of the value of $\mathcal{A}_\mathrm{bias}$, in the decorated HOD the marginalized first moment of centrals and satellites are unchanged from the baseline value defined by Equations 1 and 2. We refer the reader to Hearin et al. (2016) for further details about the decorated HOD.

[1] `http://www.slac.stanford.edu/~behroozi/BPlanck_Hlists`

[2] We set a maximum value of $c = 25$ to the NFW concentration, because halos with very large values for the concentration tend to be poorly described by an NFW profile, for example due to a recent merger.

Parameter	Prior Interval
$\log(M_{\rm min})$	[9.0,14.0]
$\sigma_{\log M}$	[0.01,1.5]
$\log(M_0)$	[9.0,14.0]
$\log(M_1)$	[10.7,15.0]
α	[0.0,2.0]
$A_{\rm cen}$	[-1.0,1.0]
$A_{\rm sat}$	[-1.0,1.0]

Table 1. Ranges for the priors used in the parameter inference. All prior distributions are uniform over the specified ranges.

2.3 Parameter Inference

To infer parameters for the HOD and Decorated HOD models described in the previous subsections, we performed a Markov Chain Monte Carlo (MCMC) sampling of the posteriors using the affine-invariant ensemble sampler of Goodman & Weare (2010) as implemented in the `emcee` software package (Foreman-Mackey et al. 2013). For most cases, we find that $\sim 3-10\times10^6$ samples are necessary in order for our chains to converge.

The most important detail of this analysis is the priors on the parameters. In all analyses discussed in this paper, we adopt priors that are uniform distributions over the intervals specified in Table 3.2. In the case of the assembly bias parameters $A_{\rm cen}$ and $A_{\rm sat}$, the priors represent physical boundaries. These parameters must satisfy $-1 \leq A_{\rm cen,sat} \leq 1$. Physical considerations require parameter $\sigma_{\log M} > 0$. All other priors have a negligible influence on the posterior aside from $\log M_0$. We find that $\log M_0$ is often very poorly constrained by clustering data.

3 RESULTS

We have performed parameter inference analyses in order to infer the underlying HOD of galaxies from the projected galaxy two-point function $w_{\rm p}(r_{\rm p})$ as described in the preceding section. In this section, we describe the primary results of these analyses. Our marginalized one-dimensional parameter constraints are given in Table 3.1.

3.1 Standard Analysis

Prior to discussing our results using models that include assembly bias, we present results of standard HOD analyses that include no model for assembly bias. The results of the standard HOD analyses and all other analyses are shown in the form of marginalized constraints on individual parameters in Table 3.1. We compare our parameter constraints to the standard HOD analysis performed by Zehavi et al. (2011) in Table 3.1 as well. An example of the inferred posteriors for the HOD parameters is shown in Figure 1. The left-hand panels of Figure 2, Figure 3, and Figure 4 show the projected correlation function data along with predictions for $w_{\rm p}(r_{\rm p})$ from 50 randomly-selected models from the MCMC chains. Note that the significant covariance in the data makes it difficult to determine the quality of fit from visual inspection of these figures.

The inferred parameters from our standard analyses differ in several ways from the Zehavi et al. (2011) analysis. Firstly, in our re-analysis of the projected clustering data, we generally find all mass scales to be slightly higher than in the work of Zehavi et al. (2011). This difference is largely due to the slightly different cosmologies adopted in this work. The most important difference are in the values of $\Omega_{\rm M}$, and σ_8. Zehavi et al. (2011) assumed $\Omega_{\rm M} = 0.25$ and $\sigma_8 = 0.8$, whereas in the present work, we use the BolshoiP simulation in which $\Omega_{\rm M} = 0.307$ and $\sigma_8 = 0.82$. Slightly larger mass scales are necessary in an analysis with higher $\Omega_{\rm M}$ and σ_8 in order to maintain galaxy number densities with larger halo number densities.

A second noteworthy difference between the present work and that of Zehavi et al. (2011) is that we find many parameters to be notably more poorly constrained. At the lower luminosity thresholds, for example, we constrain $\log(M_{\rm min})$ and $\sigma_{\log M}$ with several times lower precision than Zehavi et al. (2011). We do not show our constraints on $\log(M_0)$ as they are very poor in all cases, with 1-sigma constraints on the order of $\pm \geq 1$ dex for nearly all samples. In several cases, the constraint on $\log(M_0)$ is determined by the prior given in Table 3.2. This is in stark contrast to several of the results of Zehavi et al. (2011). For example, for the threshold sample with $M_r < -19.5$ ($M_r < -20.5$), Zehavi et al. (2011) quote $\log(M_0) = 12.23 \pm 0.17$ (12.35 ± 0.24), whereas we infer $\log(M_0) = 11.38^{+0.95}_{-1.57}$ ($11.19^{+0.89}_{-1.39}$). Examining the form of Eq. (2), it is sensible that the parameter $\log(M_0)$ should be unconstrained at the lower end, because the value of M_0 does not alter the predicted satellite number once $M_0 \ll M_1$. Therefore, it seems likely that the tighter constraints quoted by Zehavi et al. (2011) must be an error.

We have confirmed with a subset of the authors of Zehavi et al. (2011) that the number of MCMC samples they included in their analysis was insufficient in a number of cases and that this can lead to a significant underestimation of the uncertainties on the inferred parameters, especially $\log(M_{\rm min})$ and $\sigma_{\log M}$ (Z. Zheng & I. Zehavi, Private Communication). The analysis of Zehavi et al. (2011) used 10^4 samples, whereas we find several $\times10^6$ samples are often necessary for convergence. Additionally, we have recreated qualitatively similar behavior by considering only small subsets of our full MCMC chains. Consequently, insufficient sampling of the posterior seems to be the likely resolution of the discrepancies between our work and that of Zehavi et al. (2011).

Two degeneracies are manifest in Fig. 1 that are common to all of our analyses. The parameters $\log(M_1)$ and α are degenerate with each other and positively correlated. The parameter M_1 is the mass scale at which a halo has one satellite on average and α is the power-law index describing the dependence of average satellite number on halo mass. Increasing M_1 decreases the number of satellites in large halos by placing satellites preferentially in

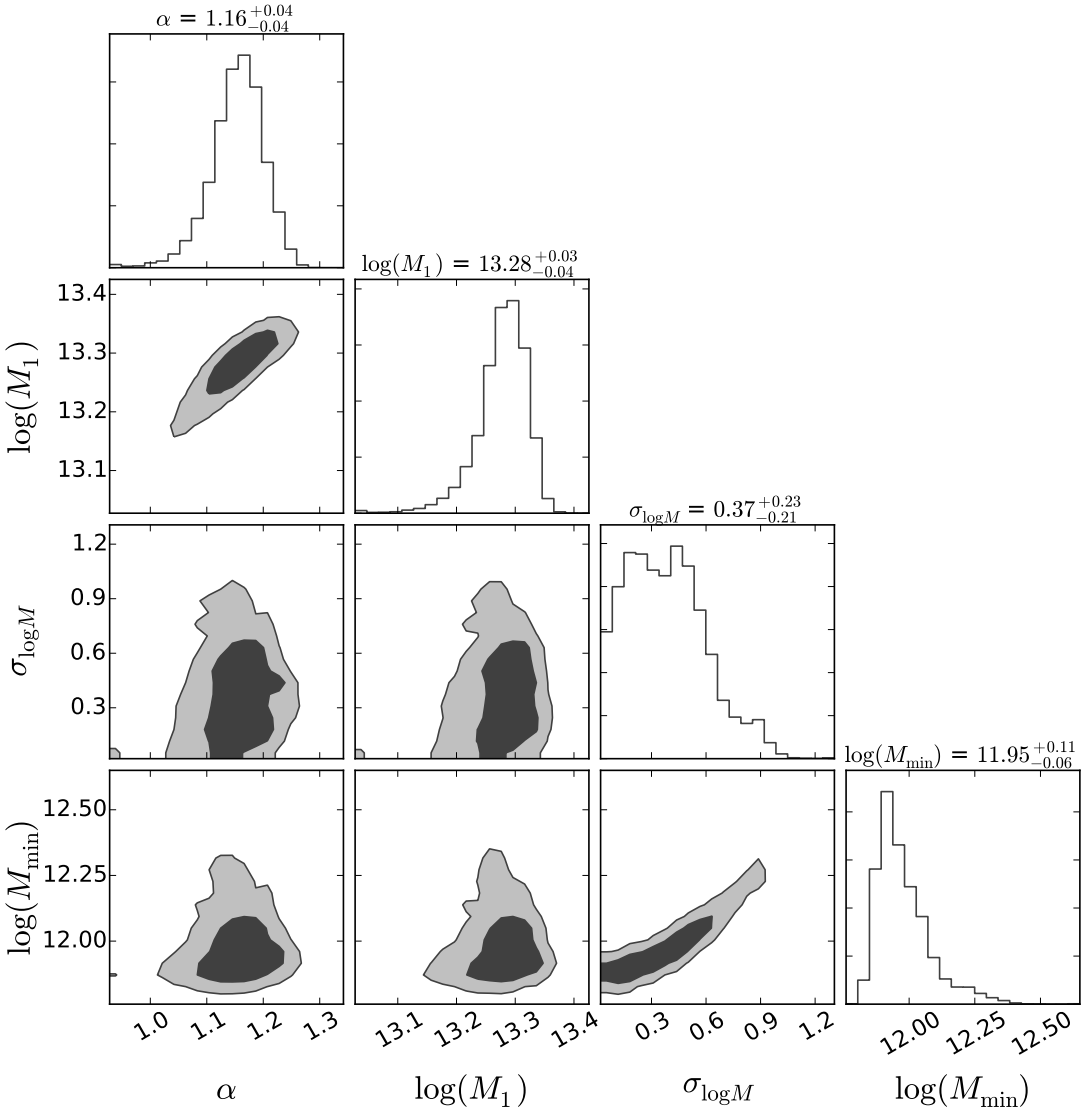

Figure 1. Two-dimensional marginalized constraints on HOD parameters inferred from standard HOD fits to $w_{\rm p}(r_{\rm p})$ data for the $M_r < -20$ sample. The HOD parameter $\log(M_0)$ is extremely poorly constrained by the $w_{\rm p}(r_{\rm p})$ data and has been omitted. The inner contours contain 68% of the posterior probability while the outer contours contain 95% of the probability. The panels along the diagonal show the one-dimensional, marginalized posteriors on each of these parameters. The values above each panel on the diagonal show the median value for the parameter in our chains along with the $16^{\rm th}$ and $84^{\rm th}$ percentiles.

halos of larger mass. An increase in α can partly compensate for an increase in M_1 by increasing the rate at which average satellite number grows with halo mass, leading to a larger number of satellites in high-mass halos.

As is evident in Figure 1, $\log(M_{\rm min})$ and $\sigma_{\log M}$ share a relatively narrow degeneracy as well. This degeneracy is largely induced by the measured number density of the sample. Increasing $\log(M_{\rm min})$ decreases galaxy number density, but this can be compensated by an increase in $\sigma_{\log M}$, which places galaxies in a fraction of the considerably more numerous halos with masses less than $M_{\rm min}$. The consequence is that $\log(M_{\rm min})$ and $\sigma_{\log M}$ are degenerate with each other such that most of the posterior probability lies in a narrow band along which $\log(M_{\rm min})$ and $\sigma_{\log M}$ are positively correlated, as shown in Fig. 1. In the following plots, we suppress the parameter $\sigma_{\log M}$, in order to increase the clarity of the plots, because the viable range of $\sigma_{\log M}$ is determined by this simple degeneracy with $\log(M_{\rm min})$.

The results of this subsection demonstrate that we achieve reasonable fits to projected galaxy clustering data using direct HOD population of a high-resolution numerical simulation of structure formation. These results also update and supersede existing constraints in the literature in at least three respects. First, we work within the best-fit Planck cosmology. Second, we perform our parameter inference analysis using direct population of halos identified in a numerical simulation of cosmological structure formation (BolshioP). This greatly mitigates modeling uncertainties associated with nonlinear density

Sample M_r	Authors	$\log(M_{\rm min})$	$\sigma_{\log M}$	$\log(M_1)$	α	$A_{\rm cen}$	$A_{\rm sat}$	χ^2/DoF
-21	Zehavi+11	12.78 ± 0.10	0.68 ± 0.15	13.80 ± 0.03	1.15 ± 0.06	$--$	$--$	3.1
-21	Zentner+16	$12.92^{+0.07}_{-0.11}$	$0.74^{+0.09}_{-0.16}$	$13.93^{+0.04}_{-0.05}$	$1.23^{+0.10}_{-0.12}$	$--$	$--$	1.59
-21	Zentner+16	$12.83^{+0.11}_{-0.09}$	$0.60^{+0.15}_{-0.17}$	$13.93^{+0.05}_{-0.08}$	$1.16^{+0.12}_{-0.14}$	$0.29^{+0.44}_{-0.35}$	$0.08^{+0.49}_{-0.36}$	1.34
-20.5	Zehavi+11	12.14 ± 0.03	0.17 ± 0.15	13.44 ± 0.03	1.15 ± 0.03	$--$	$--$	2.7
-20.5	Zentner+16	$12.25^{+0.07}_{-0.03}$	$0.23^{+0.17}_{-0.15}$	$13.59^{+0.02}_{-0.02}$	$1.20^{+0.04}_{-0.04}$	$--$	$--$	1.90
-20.5	Zentner+16	$12.32^{+0.13}_{-0.08}$	$0.45^{+0.21}_{-0.25}$	$13.59^{+0.04}_{-0.04}$	$1.14^{+0.05}_{-0.06}$	$> 0.08(90\%)$	$0.22^{+0.40}_{-0.31}$	1.40
-20	Zehavi+11	11.83 ± 0.03	0.25 ± 0.11	13.08 ± 0.03	1.00 ± 0.05	$--$	$--$	2.1
-20	Zentner+16	$11.95^{+0.11}_{-0.6}$	$0.37^{+0.23}_{-0.21}$	$13.28^{+0.03}_{-0.04}$	$1.16^{+0.04}_{-0.04}$	$--$	$--$	2.19
-20	Zentner+16	$12.23^{+0.33}_{-0.21}$	$0.84^{+0.37}_{-0.31}$	$13.20^{+0.06}_{-0.08}$	$1.05^{+0.06}_{-0.08}$	$> 0.28(99\%)$	$0.01^{+0.32}_{-0.26}$	1.16
-19.5	Zehavi+11	11.57 ± 0.04	0.17 ± 0.13	12.87 ± 0.03	0.99 ± 0.04	$--$	$--$	1.00
-19.5	Zentner+16	$11.76^{+0.33}_{-0.11}$	$0.51^{+0.51}_{-0.29}$	$13.05^{+0.04}_{-0.08}$	$1.12^{+0.04}_{-0.07}$	$--$	$--$	1.24
-19.5	Zentner+16	$11.80^{+0.36}_{-0.16}$	$0.63^{+0.53}_{-0.37}$	$13.04^{+0.09}_{-0.12}$	$1.06^{+0.07}_{-0.10}$	$> -0.01(84\%)$	$> -0.16(84\%)$	0.69
-19	Zehavi+11	11.45 ± 0.04	0.19 ± 0.13	12.64 ± 0.04	1.02 ± 0.02	$--$	$--$	1.8
-19	Zentner+16	$11.72^{+0.33}_{-0.19}$	$0.69^{+0.52}_{-0.46}$	$12.78^{+0.04}_{-0.04}$	$1.03^{+0.04}_{-0.04}$	$--$	$--$	2.77
-19	Zentner+16	$11.62^{+0.33}_{-0.13}$	$0.53^{+0.57}_{-0.35}$	$12.83^{+0.06}_{-0.07}$	$1.02^{+0.04}_{-0.04}$	$0.35^{+0.45}_{-0.66}$	$> 0.02(84\%)$	2.01

Table 2. Results of standard HOD fits to SDSS DR7 $w_{\rm p}(r_{\rm p})$ as well as fits using a parameterized model of assembly bias. Assembly bias is quantified by the parameters $A_{\rm cen}$ ($A_{\rm sat}$) for central (satellite) galaxies. The secondary property that we assume to determine the galaxy HOD is halo formation time. $A_{\rm cen,sat} = 0$ means that there is no assembly bias while $A_{\rm cen,sat} = 1$ ($A_{\rm cen,sat} = -1$) means that halo formation time is maximally correlated (anticorrelated) with halo formation time. Thus the parameters range over $-1 \leq A_{\rm cen,sat} \leq 1$. If the constraints on $A_{\rm cen}$ and $A_{\rm sat}$ are unspecified, then the model used to interpret the data does not include assembly bias. In our analyses, quoted parameter values with errors correspond to the median value of the parameter and the $16^{\rm th}$ and $84^{\rm th}$ percentiles. In cases for which the posterior is monotonic within the physical parameter space, we show one-sided percentiles.

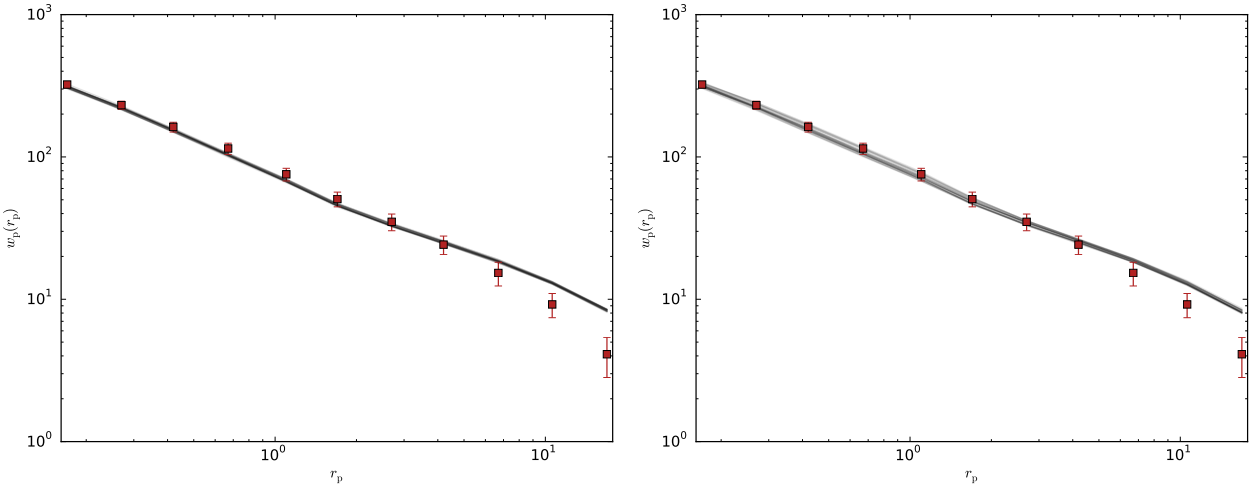

Figure 2. Left: The $M_r < -19$ threshold sample projected correlation function with diagonal elements of covariance (points with errorbars). The grey lines are 25 randomly-selected HOD models that yield $\Delta\chi^2 < 1$ compared to the best-fitting model. **Right:** Same as the left panel but using a fit to a Decorated HOD model that contain parameters to describe the strength of assembly bias.

field evolution, scale-dependent halo bias, halo exclusion, or other effects that have been difficult to incorporate into analytical halo models with high precision. Third, we have explored the posteriors of the parameters with significantly more samples (roughly two orders of magnitude), thereby mitigating errors on inferred parameters and their errors induced by insufficient sampling of the posterior.

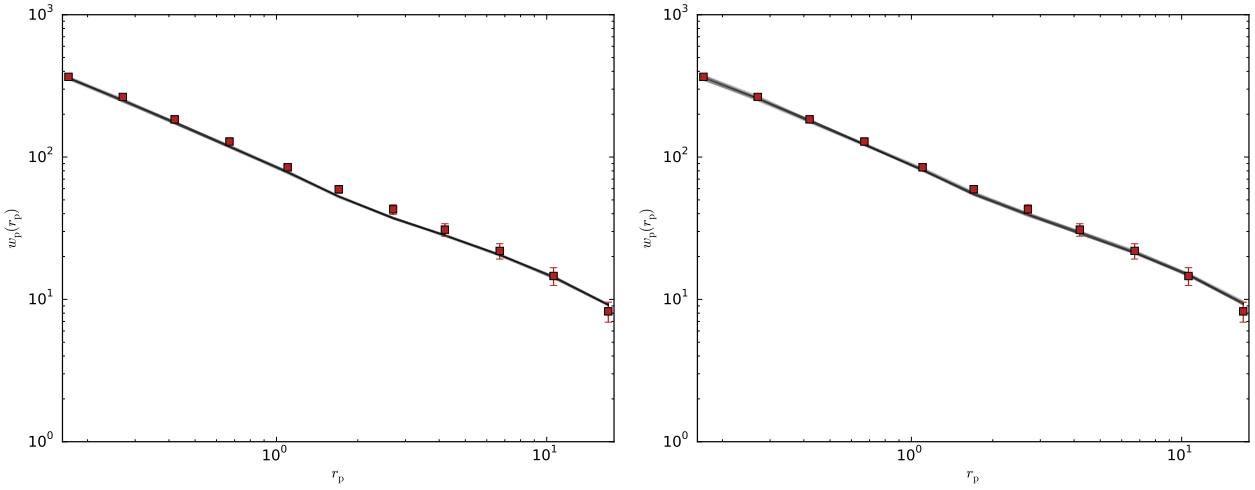

Figure 3. The same as Figure 2, but for the $M_r < -20$ threshold sample.

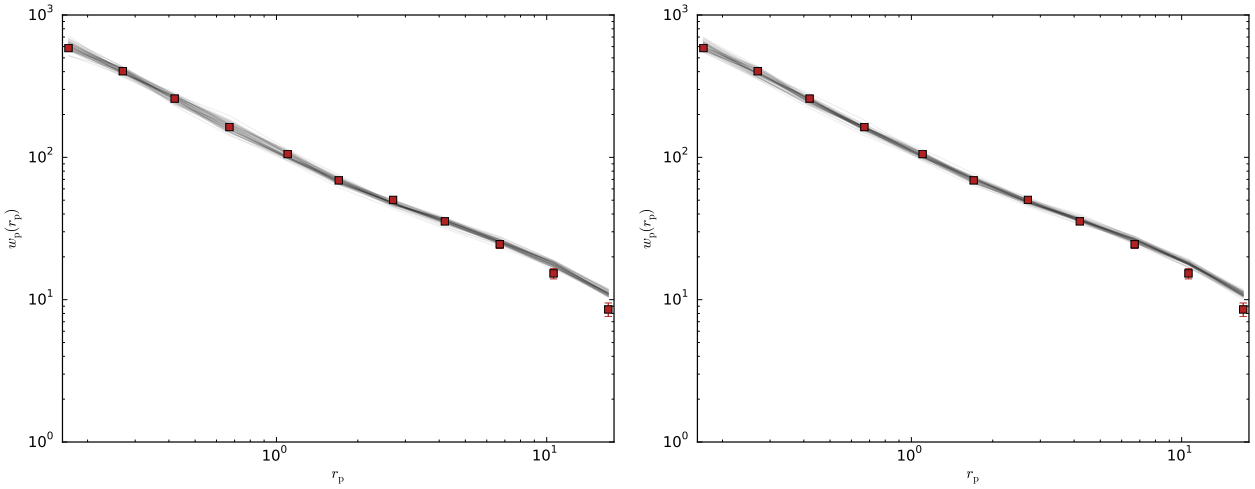

Figure 4. The same as Figure 2, but for the $M_r < -21$ threshold sample.

3.2 Analysis with Decorated HOD

We turn now to a discussion of our parameter inference analysis of projected galaxy clustering in Decorated HOD models that include a treatment of galaxy assembly bias. In this work, we consider only the simplest model of galaxy assembly bias, introducing only two new parameters, $A_{\rm cen}$ and $A_{\rm sat}$, that describe the strength of central galaxy and satellite galaxy assembly bias respectively. These parameters are limited to values of $-1 \leq A_{\rm cen,sat} \leq 1$, and $A_{\rm cen,sat} = 0$ when there is no galaxy assembly bias. In this work, we use halo concentration as our secondary halo property, so $A_{\rm cen,sat} = 1$ ($A_{\rm cen,sat} = -1$) means that the mean number of galaxies per halo is maximally correlated (anti-correlated) with halo concentration. The model and its implementation in `halotools` is discussed further in Section 2.2 above and in Hearin et al. (2016).

Examples of our fits are given in the right-hand panels of Figure 2, Figure 3, and Figure 4. The general trend that can be gleaned from these figures is that introducing assembly bias improves the ability of the predicted two-point functions to match the measured two-point functions across the transition form the one-halo (highly nonlinear) to two-halo (nearly linear) regimes near $r_{\rm p} \sim 2\,h^{-1}$Mpc. This is most apparent for the $M_r < -20$ threshold sample shown in Fig. 3. Visually, these differences appear to be small; however, Table 3.1 shows that they are statistically important.

The one-dimensional marginalized constraints on all parameters from these analyses are given in the lowest row of each luminosity threshold grouping in Table 3.1. In cases where the posterior on a parameter is monotonic within the physical parameter range, we quote an upper or lower limit on the parameter. One- and two-dimensional visualizations of the posteriors from our analysis are shown in Figure 5, Figure 6, and Figure 7.

Table 3.1 and Figures 5-**??** all make several simple, generic points. Introducing additional parameter freedom associated with galaxy assembly bias generally increases

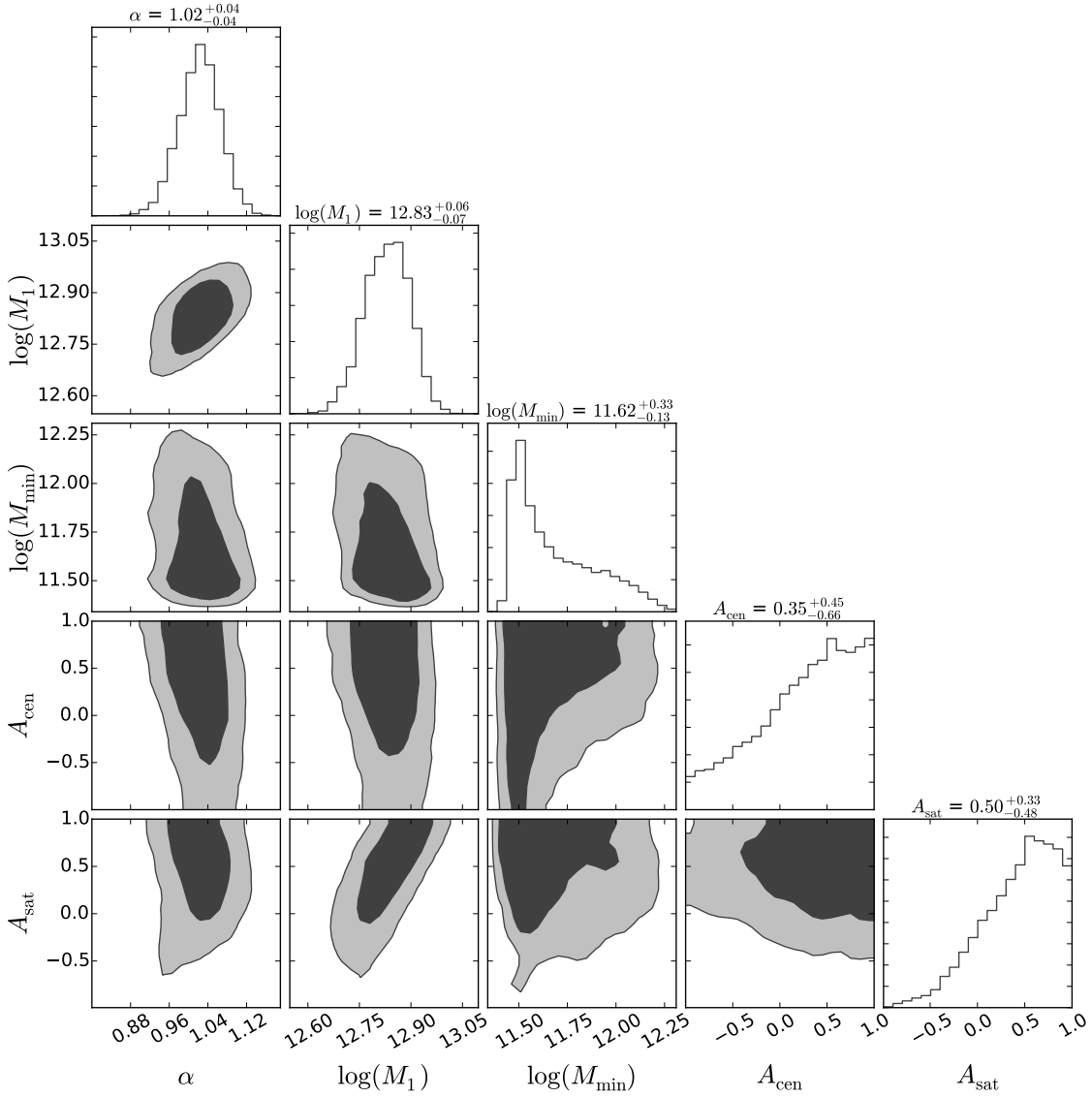

Figure 5. Two-dimensional marginalized constraints on decorated HOD parameters inferred from fits to $w_{\rm p}(r_{\rm p})$ data for the $M_r < -19$ sample. The contours and histograms along the diagonal panels are as in Fig. 1. The decorated HOD models include a two-parameter model for assembly bias. The HOD parameter $\log(M_0)$ is extremely poorly constrained by the data and has been suppressed for clarity. Likewise, as in Fig. 1, $\sigma_{\log M}$ and $\log(M_{\rm min})$ share a narrow degeneracy, so we have suppressed $\sigma_{\log M}$ in order to make constraints on other parameters more easily visible.

the viable parameter space, even for the subset of standard HOD parameters. Constraints on the standard HOD parameters are generally less restrictive. This is exactly what is expected from the introduction of additional parameter freedom.

Focusing attention on the parameters describing galaxy assembly bias, it is evident that these parameters are often quite poorly constrained by galaxy clustering data. This is important as it implies that galaxy clustering of the precision of SDSS DR7 measurements cannot rule out, or strongly restrict galaxy assembly bias in many cases. Nonetheless, it is apparent that the presence of assembly bias can alter the inferred HOD, or more generally, the inferred relationship between galaxies and halo mass. This is most evident for the $M_r < -20$ threshold sample, for which there are significant differences in the inferred values of all baseline HOD parameters between models with and without galaxy assembly bias. Other threshold samples exhibit significant changes particularly for α, and to a lesser degree for $\log(M_{\rm min})$ and $\sigma_{\log M}$.

Beyond those generic conclusions, a few specific cases are worthy of further examination. Consider the $M_r < -20$ sample. The inferred value of $A_{\rm cen} > 0.28$ at 99% confidence. In this case the data strongly prefer $A_{\rm cen} > 0$ and thus strongly prefer galaxies to reside in halos of larger concentration at fixed halo mass. This particular threshold sample is the most significant outlier in this regard. Nevertheless, there are hints of assembly bias in other samples. Satellite galaxies show a marginal prefer-

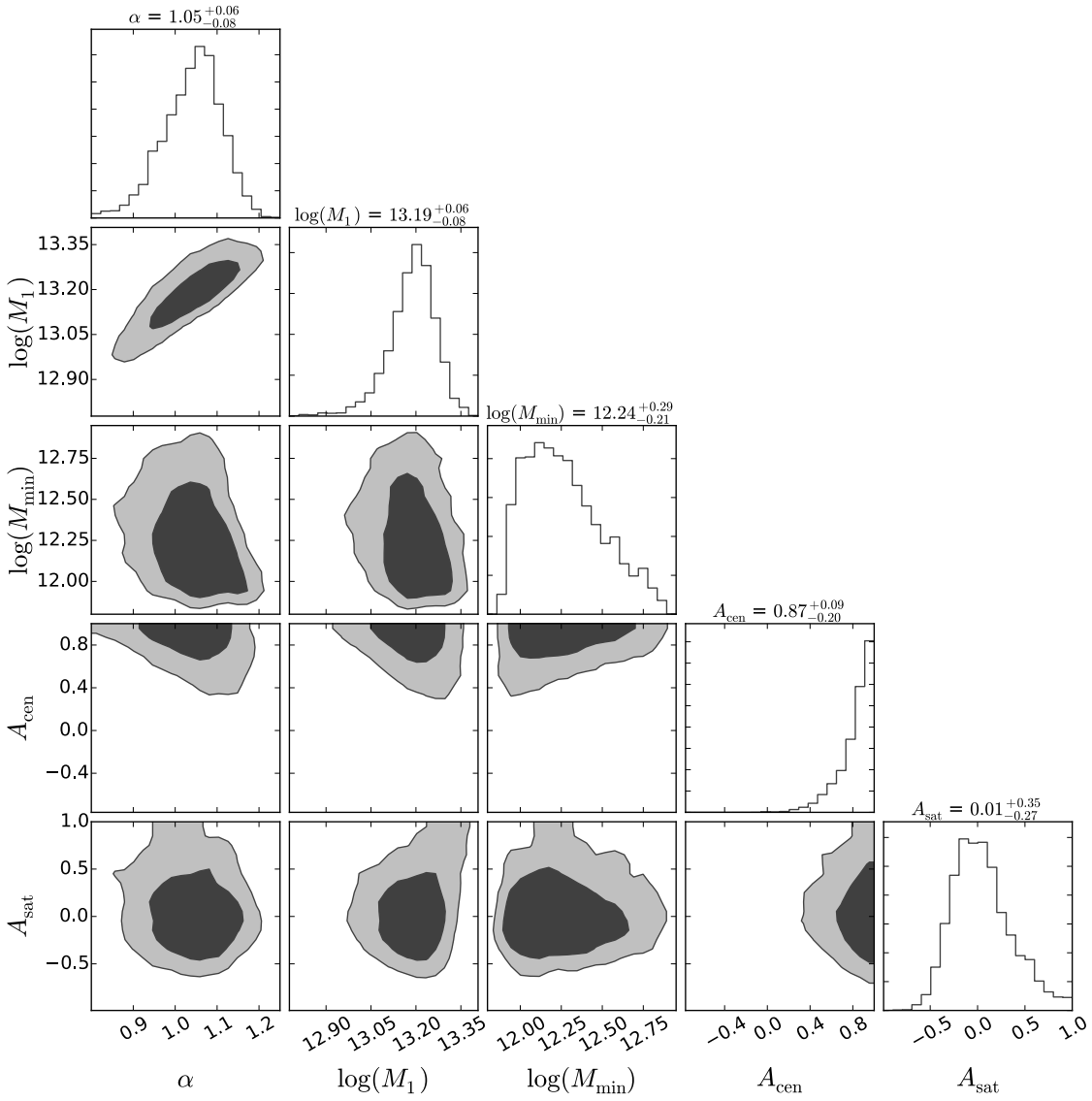

Figure 6. The same as Figure 5, but for the $M_r < -20$ sample.

ence for occupying halos of higher concentration in the $M_r < -19$ threshold sample. The $M_r < -19$ sample exhibits weak preference for a positive correlation of galaxy occupation with halo concentration at fixed mass for both satellite galaxies and central galaxies. Continuing upward with luminosity, the $M_r < -20.5$ sample exhibits a significant preference for central galaxy assembly bias. Lastly, there is no preference for either central galaxy or satellite galaxy assembly bias for the $M_r < -21$ threshold sample, for which both $A_{\rm cen}$ and $A_{\rm sat}$ are consistent with zero will within 1σ. These data suggest that assembly bias may be present in the real universe.

M_r Threshold	ΔBIC
-21	-0.54
-20.5	1.33
-20	4.56
-19.5	0.26
-19	4.37

Table 3. Change to the Bayesian Information Criterion, ΔBIC, after introducing additional parametric freedom to accommodate galaxy assembly bias. Sign convention is such that positive values favor models including assembly bias, negative values favor standard HOD models with no assembly bias parameters. Changes in the Bayesian Information Criterion $|\Delta{\rm BIC}| \geq 5$ strongly favor one model over another.

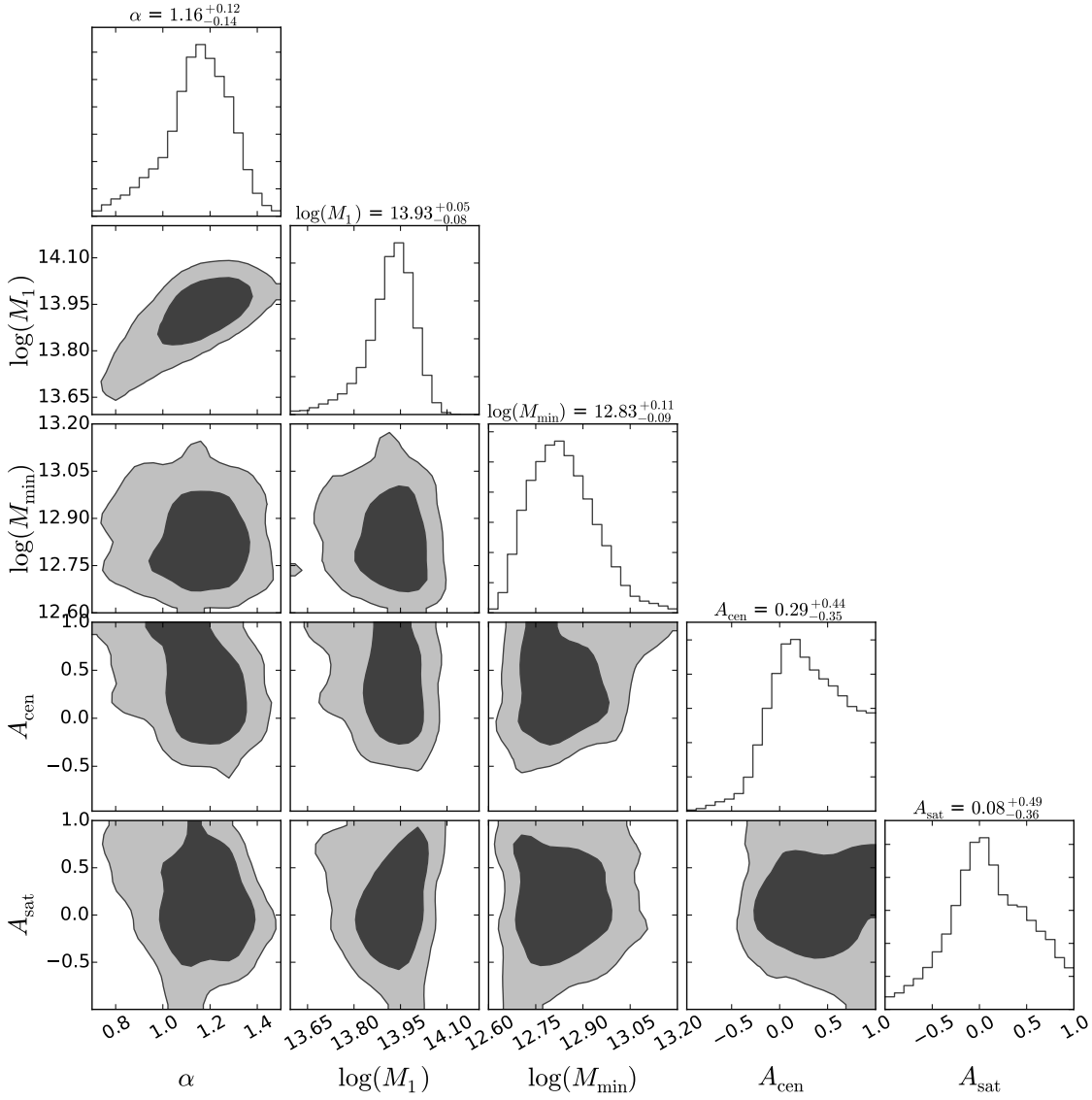

Figure 7. The same as Figure 5, but for the $M_r < -21$ sample.

4 CONCLUSIONS

5 ACKNOWLEDGEMENTS

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 Astro-Computing Center (UC-HiPACC) and were run at the NASA Ames Research Center.

REFERENCES

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2011, ArXiv:1110.4372

Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013, ApJ, 763, 18

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

Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, Mon. Not. R. Astron. Soc., 430, 767

Conroy C., Wechsler R. H., 2009, ApJ, 696, 620

Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Gao L., Springel V., White S. D. M., 2005, Mon. Not. R. Astron. Soc., 363, L66

Gao L., White S. D. M., 2007, Mon. Not. R. Astron. Soc., 377, L5

Goodman J., Weare J., 2010, Comm. App. Math. and Comp. Sci, 5

Guo H., Zheng Z., Zehavi I., Xu H., Eisenstein D. J.,

Weinberg D. H., Bahcall N. A., Berlind A. A., Comparat J., McBride C. K., Ross A. J., Schneider D. P., Skibba R. A., Swanson M. E. C., Tinker J. L., Tojeiro R., Wake D. A., 2014, ArXiv:1401.3009
Guo Q., White S., Boylan-Kolchin M., De Lucia G., Kauffmann G., Lemson G., Li C., Springel V., Weinmann S., 2011, Mon. Not. R. Astron. Soc., 413, 101
Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, Mon. Not. R. Astron. Soc.
Lacerna I., Padilla N., 2011, Mon. Not. R. Astron. Soc., 412, 1283
Leauthaud A., et al., 2012, ApJ, 744, 159
More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, Mon. Not. R. Astron. Soc., 392, 801
More S., van den Bosch F. C., Cacciato M., More A., Mo H., Yang X., 2013, Mon. Not. R. Astron. Soc., 430, 747
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Planck Collaboration Ade P. A. R., Aghanim N., Armitage-Caplan C., Arnaud M., Ashdown M., Atrio-Barandela F., Aumont J., Baccigalupi C., Banday A. J., et al. 2014, Astron. Astrophys., 571, A16
Porciani C., Norberg P., 2006, Mon. Not. R. Astron. Soc., 371, 1824
Riebe K., Partl A. M., Enke H., Forero-Romero J., Gottloeber S., Klypin A., Lemson G., Prada F., Primack J. R., Steinmetz M., Turchaninov V., 2011, ArXiv:1109.0003
Tinker J. L., Leauthaud A., Bundy K., George M. R., Behroozi P., Massey R., Rhodes J., Wechsler R., 2013, ArXiv:1308.2974
Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, ApJ, 631, 41
van den Bosch F. C., Yang X., Mo H. J., Weinmann S. M., Macciò A. V., More S., Cacciato M., Skibba R., Kang X., 2007, Mon. Not. R. Astron. Soc., 376, 841
Wake D. A., Whitaker K. E., Labbé I., van Dokkum P. G., Franx M., Quadri R., Brammer G., Kriek M., Lundgren B. F., Marchesini D., Muzzin A., 2011, ApJ, 728, 46
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
Yang X., Mo H. J., van den Bosch F. C., 2003, Mon. Not. R. Astron. Soc., 339, 1057
Yang X., Mo H. J., van den Bosch F. C., 2009, ApJ, 693, 830
Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752, 41
Yang X., Mo H. J., Zhang Y., van den Bosch F. C., 2011, ApJ, 741, 13
Zehavi I., et al., 2005, ApJ, 630, 1
Zehavi I., et al., 2011, ApJ, 736, 59
Zentner A. R., 2007, International Journal of Modern Physics D, 16, 763
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, Mon. Not. R. Astron. Soc., 443, 3044
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760
Zu Y., Mandelbaum R., 2015, ArXiv:1509.06758