# Constraints on Assembly Bias from Galaxy Clustering

Andrew R. Zentner[1], Andrew Hearin[2], Frank C. van den Bosch[3],
Johannes U. Lange[3], and P. Rogers Nelson[4]

[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 constrain the newly-introduced decorated Halo Occupation Distribution (HOD) model using SDSS DR7 measurements of projected galaxy clustering, $w_{\mathrm{p}}(r_{\mathrm{p}})$ of galaxies in r-band luminosity-threshold samples. The decorated HOD is a model for the galaxy–halo connection that augments the traditional HOD by allowing for the possibility of *assembly bias*: galaxy luminosity may be correlated with dark matter halo properties besides mass, $M_{\mathrm{vir}}$. We demonstrate that it is not possible to rule out galaxy assembly bias using DR7 measurements of galaxy clustering alone. Moreover, galaxy samples $M_r < -20, -20.5$ favor strong levels of central galaxy assembly bias. Our analyses favor models in which high-concentration halos are more likely to host a central galaxy relative to low-concentration halos of the same $M_{\mathrm{vir}}$. We rule out zero assembly bias with high significance for these samples. Satellite galaxy assembly bias is significant for the faintest sample we study, $M_r < -19$. We find no evidence for assembly bias in the $M_r < -21$ sample. In addition to presenting the constraints on HOD models that accommodate assembly bias, our analysis includes numerous improvements over previous analyses of this data set and supersedes existing constraints even on standard HOD parameters.

## 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; Rodríguez-Puebla 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 mass $M_{\mathrm{vir}}$ only. In this paper, we present the first analysis of galaxy clustering using a halo occupation model that violates the second of these assumptions and permits galaxies to occupy halos as a function of multiple halo properties.

It is now well known that the strength of halo clustering, the halo bias, depends upon halo properties other than $M_{\mathrm{vir}}$ (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 halo properties other than $M_{\mathrm{vir}}$, 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).

There is increasing observational evidence that galaxy assembly bias is present in the real Universe. Using an approach based on subhalo abundance matching, Lehmann et al. (2015) showed that the clustering of low-redshift galaxies favors models in which stellar mass depends upon a combination of $M_{\mathrm{vir}}$ and halo concentration. Miyatake et al. (2016) and More et al. (2016) have presented evidence for the presence of assembly bias in massive clusters; in particular, they find that clusters with a more centrally concentrated distribution of satellite galaxies cluster more weakly than clusters with more diffuse satellite galaxy distributions at fixed cluster mass. Indeed, More et al. (2016) suggest that the observed

galaxy bias is considerably stronger than had been expected based on $N$-body simulations coupled with simple models for the galaxy–halo connection, such as subhalo abundance matching. Additional support for the presence of assembly bias has come from a variety of studies that have shown, or suggested, that the large-scale environment of dark matter halos of fixed mass is correlated with the star formation rate of their central galaxies (Yang et al. 2006; Blanton & Berlind 2007; Wang et al. 2008, 2013; Hearin et al. 2014).

Despite this growing evidence, the empirical modeling community has not yet reached a consensus on the question of galaxy assembly bias. For example, Lin et al. (2015) claim that some of the above evidence for assembly bias can be explained by differences in halo mass between the galaxy samples under consideration. Paranjape et al. (2015) urges caution that contamination by satellite galaxies can masquerade as a false signal of assembly bias. Tinker et al. (2008) argue that measurements of void statistics place strong bounds on the possible strength of the signal, though Zentner et al. (2014) constructed an explicit example refuting this claim.

In much of the literature on this topic, the character of the supporting evidence suffers from a severe shortcoming. Until very recently, the only models that have been fit to observational clustering measurements have been traditional models in which assembly bias is assumed from the outset to be zero. Models that include assembly bias such as semi-analytic models are typically too computationally expensive to compare to data in a manner that enables statistical constraints on model parameters. Such models are usually compared to data on a case-by-case basis (e.g., Croton et al. 2007), prohibiting any conclusive statement to be made about the strength of any assembly bias signal.

Motivated by this shortcoming, we have developed a new class of empirical models that enable galaxies to occupy halos in a manner that has *simultaneous* dependence upon multiple halo properties (Hearin et al. 2016), *including continuously variable levels of assembly bias.* Crucially, our implementation is of sufficient computational efficiency to permit a proper likelihood analysis of the model parameter space. Armed with this new methodology, 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 of both standard halo occupation models and these new models that permit parameterized galaxy assembly bias.

This work presents a re-analysis of the SDSS DR7 data that overcomes numerous shortcomings of previous studies. The primary improvements of our approach stem from the fact that we directly populate mock galaxy catalogs within a cosmological simulation, whereas previous analyses relied upon analytical fitting functions with parameters calibrated against a suite of simulations. In direct mock population, delicate issues present in analytic modeling, such as scale-dependent halo bias and halo exclusion, are treated exactly. Our approach also provides a natural framework for studying models of assembly bias, because we can use the exact clustering properties of sim-

ulated halos as a function of any arbitrary halo properties in the catalog. It is not necessary to develop new phenomenological models or fitting functions to deal with this additional complexity. With direct mock population, systematic uncertainty in the model is limited to the sample variance of the simulated box, numerical inaccuracies of halo-finding, and errors related to insufficient numerical resolution (e.g., overmerging; Klypin et al. 1999; Behroozi et al. 2013).

In addition to improvements resulting from direct mock population, the simulation we use is based on the latest Planck cosmological parameters (Planck Collaboration et al. 2014), updating previous work. Furthermore, many important differences between this work and the previous analysis by Zehavi et al. (2011) result from the fact that the Monte Carlo Markov Chains used in Zehavi et al. (2011) did not sufficiently sample the posterior distribution (Z. Zheng, private communication). The result is that the constraints in Zehavi et al. (2011) are overly restrictive, considerably so in some cases, and potentially biased. Therefore, *our analysis supersedes previous work even in the case of standard halo occupation models.*

Most importantly, 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, in agreement with Lehmann et al. (2015), 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 has 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 the parameter inference methodology. We present results from both traditional and assembly-biased halo occupation analyses 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 populate dark matter halos with mock galaxies using the publicly-available, open source, `Halotools` software (Hearin et al. 2016). We explore halo occupation distribution (HOD) models in this work (e.g. Seljak 2000; Ma & Fry 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002), though other techniques that can be used to interpret such data, such as the conditional luminosity function (CLF, e.g., Yang et al. 2003; van den Bosch et al. 2013), exist. In this subsection, we review the "standard" HOD model used in the present work, which assumes that there is no galaxy assembly bias. We refer to such a model as "standard" because all HOD analyses of galaxy clustering to date have assumed no galaxy assembly bias. In the following subsection, we describe the Decorated HOD model described in Hearin et al. (2016). In both cases, we will only review the salient features of our methodology briefly; interested readers can always refer

to `http://halotools.readthedocs.io` and Hearin et al. (2016) for further details.

### 2.1.1 Simulation

All of our analyses are based on the Bolshoi-Planck (BolshoiP) simulation (Riebe et al. 2011). BolshoiP was run with cosmological parameters based on Planck Collaboration et al. (2014): $\Omega_\Lambda = 0.693; \Omega_m = 1 - \Omega_\Lambda = 0.307; \Omega_b = 0.048; h = 0.7; n_s = 0.96;$ and $\sigma_8 = 0.82$. BolshoiP simulated the formation of structure within a cubic box 250 Mpc/$h$ on a side, requiring a particle mass of $m_p = 1.35 \times 10^8 M_\odot/h$. Further information about the BolshoiP 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 $M_{\rm peak}$ that exceeds 300 particles. The minimum peak halo mass considered in our analysis has $M_{\rm peak} \geq 4.05 \times 10^5 \, h^{-1} M_\odot$ and this prevents us from analyzing samples with $M_r > -19$.

### 2.1.2 Occupation statistics

In standard HOD models, central galaxies and satellite galaxies are treated separately, so the model is specified by two probability distributions, one for each type of galaxy. The galaxy-halo connection is specified in terms of $P(N_{\rm cen}|M_{\rm vir})$ and $P(N_{\rm sat}|M_{\rm vir})$, the probability that a halo of mass $M_{\rm vir}$ hosts $N_{\rm cen}$ central and $N_{\rm sat}$ satellite galaxies, respectively. $P(N_{\rm cen}|M_{\rm vir})$ is typically a nearest-integer distribution, as a host halo has only either zero or one central galaxy. Consequently, the occupation statistics of central galaxies are specified by the first moment of $P(N_{\rm cen}|M_{\rm vir})$, which we model as

$$\langle N_{\rm cen}|M_{\rm vir}\rangle = \frac{1}{2}\left(1 + \mathrm{erf}\left[\frac{\log(M_{\rm vir}) - \log(M_{\rm min})}{\sigma_{\log M}}\right]\right) \tag{1}$$

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 $M_{\rm vir}$, a central galaxy is assigned to the halo if the associated random number is less than $\langle N_{\rm cen}|M_{\rm vir}\rangle$; halos with random values exceeding $\langle N_{\rm cen}|M_{\rm vir}\rangle$ are left devoid of centrals. The parameter $\log(M_{\rm min})$ specifies the halo mass at which the halo has a 50% probability of hosting a central galaxy, while the parameter $\sigma_{\log M}$ specifies the rate at which $\langle N_{\rm cen}|M_{\rm vir}\rangle$ transitions from zero to unity, with smaller values of $\sigma_{\log M}$ corresponding to a more rapid transition.

We model the distribution $P(N_{\rm sat}|M_{\rm vir})$ as a Poisson

distribution with first moment given by a power law,

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

The parameter $M_0$ allows the power-law to be truncated more rapidly at low masses and we set $\langle N_{\rm sat}|M_{\rm vir}\rangle = 0$ for $M_{\rm vir} < M_0$.

The five parameters of HOD models that are varied in standard analyses are $\log(M_{\rm min})$, $\sigma_{\log M}$, $\alpha$, $\log(M_1)$, and $\log(M_0)$, though, as we show below, central galaxies usually outnumber satellite galaxies by a factor of several, so $\log(M_{\rm min})$ and $\sigma_{\log M}$ usually vary along a narrow degeneracy that fixes the total galaxy number density to the observed value. There are many particular choices that can be made for the functional forms of $\langle N_{\rm cen}|M_{\rm vir}\rangle$ and $\langle N_{\rm sat}|M_{\rm vir}\rangle$. We have made choices that mimic the standard SDSS DR7 analysis of Zehavi et al. (2011), to expedite comparisons with their results. However, our choice does deviate from the choice of Zehavi et al. (2011) in one respect. The mean satellite occupation of Zehavi et al. (2011) is that of Eq. (2) multiplied by an overall factor of $\langle N_{\rm cen}|M_{\rm vir}\rangle$. We have chosen not to use this as our default because it can introduce difficulties in some analyses (see the discussion of blue galaxy samples in Zentner et al. 2014) and because, as we have verified explicitly, the extra factor introduces only small quantitative changes to our results and no qualitative changes.

### 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 $R_{\rm vir}$ 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 radial velocity distribution of satellite galaxies as a Gaussian with first moment equal to the host halo velocity and second moment equal to the solution of the isotropic Jeans equation for an NFW profile (More et al. 2009),

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

where $\tilde{r} = r/R_{\rm vir}$, $g(x) = \ln(1+x) - x/(1+x)$, and $V_{\rm vir}^2 = GM_{\rm vir}/R_{\rm vir}$. We assume that velocities are isotropic, setting the peculiar velocities in each Cartesian direction according to random draws from the above radial velocity distribution.

### 2.1.4 Predictions for observables

After populating a halo catalog with mock galaxies, we calculate the comoving number density of our mock

---

[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.

galaxy sample as $n_{\mathrm{g}} = N_{\mathrm{gal}}/L_{\mathrm{box}}^3$, where $N_{\mathrm{gal}}$ is the to-tal number of galaxies in the mock sample. We apply the distant-observer approximation and use the simu-lation z-axis as the line-of-sight direction, and the dis-tance between points in the $xy$-plane to define the pro-jected distance $r_{\mathrm{p}}$. We place mock galaxies into redshift-space by replacing each galaxy's $z$-coordinate with $z_{\mathrm{RS}} = z + V_{\mathrm{z}}/H_0$. We count pairs of points in each of our $r_{\mathrm{p}}$ bins, rejecting pairs with line-of-sight distances in red-shift space, $\Delta z_{\mathrm{RS}}$, exceeding $\pi_{\mathrm{max}} = 60$ Mpc/h, which is the same projection depth chosen by Zehavi et al. (2011), both in their analysis of the SDSS DR7 data and in their modeling. Finally, having populated mocks, we per-form the computation of the projected two-point galaxy correlation function, $w_{\mathrm{p}}(r_{\mathrm{p}})$ using the publicly-available `CorrFunc` package (Sinha 2016) which has been exten-sively optimized for computational speed.

## 2.2 HOD with Assembly Bias: The Decorated HOD

In addition to the standard occupation statistics de-scribed 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 depen-dence on both $M_{\mathrm{vir}}$ and *halo concentration*. Briefly, we use Equations 1 and 2 as our "baseline" first occupation moments. At fixed $M_{\mathrm{vir}}$, halos are divided into one of two categories, those of high- and low-concentration, de-pending on whether the concentration of the halo places it above or below the rank-order percentile $f_{\mathrm{split}}$, which we keep fixed to $f_{\mathrm{split}} = 0.5$ throughout the paper for simplicity. High-concentration halos have a different first occupation moment relative to low-concentration halos of the same mass,

$$\begin{aligned} \langle N_{\mathrm{gal}}|M_{\mathrm{vir}}, c_{\mathrm{high}}\rangle &= \langle N_{\mathrm{gal}}|M_{\mathrm{vir}}\rangle + \delta N_{\mathrm{gal}} \\ \langle N_{\mathrm{gal}}|M_{\mathrm{vir}}, c_{\mathrm{low}}\rangle &= \langle N_{\mathrm{gal}}|M_{\mathrm{vir}}\rangle - \delta N_{\mathrm{gal}} \end{aligned} \tag{4}$$

The difference, $\delta N_{\mathrm{gal}}$, 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 assem-bly bias. Values of $\mathcal{A}_{\mathrm{bias}} = \pm 1$ correspond to the max-imum strength of assembly bias allowable by the con-straint 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 mo-ment of centrals and satellites are *unchanged from the baseline value defined by Equations 1 and 2*. In other words, *decorated HOD models all have the same HODs, when averaged over all halos at fixed mass, as standard HOD models with the same five standard HOD param-eters*. The only change in Decorated HOD models is whether or not an additional property also modulates halo occupation at fixed halo mass. A value of $\mathcal{A}_{\mathrm{bias}} = 0$ indicates no galaxy assembly bias whatsoever. We refer the reader to Hearin et al. (2016) for further details about the decorated HOD.

In the present work, we fix our model to the simplest class of galaxy assembly bias models, though a recipe for generalizing to more complicated models can be found in

Hearin et al. (2016). In particular, we split halos into two populations as specified in the previous paragraph. We then populate halos with satellite galaxies specified by an assembly bias parameter $-1 \leq A_{\mathrm{sat}} \leq 1$ and central galaxies with a distinct assembly bias parameter $-1 \leq A_{\mathrm{cen}} \leq 1$. For the value of $f_{\mathrm{split}} = 0.5$ adopted here, we then have that

$$\begin{aligned} \delta N_{\mathrm{cen}} &= A_{\mathrm{cen}} \, \mathrm{MIN}\left\{\langle N_{\mathrm{cen}}|M_{\mathrm{vir}}\rangle, 1 - \langle N_{\mathrm{cen}}|M_{\mathrm{vir}}\rangle\right\} \\ \delta N_{\mathrm{sat}} &= A_{\mathrm{sat}}\langle N_{\mathrm{sat}}|M_{\mathrm{vir}}\rangle \end{aligned} \tag{5}$$

In our analyses that include assembly bias, we vary these two additional parameters, bringing the total number of parameters that vary in these analyses to seven. As we show below, these additional parameters are in many in-stances poorly constrained by clustering data of the qual-ity of SDSS DR7 alone, so exploring more complex mod-els of assembly bias does not yet seem justified in such analyses.

## 2.3 Parameter Inference

We constrain HOD parameters based on SDSS DR7 measurements of the projected galaxy two-point func-tions, $w_{\mathrm{p}}(r_{\mathrm{p}})$, and galaxy number densities, $n_{\mathrm{g}}$, for luminosity threshold samples published in Ze-havi et al. (2011). We use the full covariance ma-trix of the projected correlation function available at `astroweb.cwru.edu/izehavi/dr7_covar/table8`. We assume a likelihood of the form $\mathcal{L} \propto e^{-\chi^2/2}$, where

$$\chi^2 = \sum_{i,j} \Delta w_{\mathrm{p},i} \, [C^{-1}]_{ij} \, \Delta w_{\mathrm{p},j} + \frac{(n_{\mathrm{g}}^{\mathrm{mock}} - n_{\mathrm{g}}^{\mathrm{meas}})^2}{\sigma_{\mathrm{n}}^2}, \quad (6)$$

$\Delta w_{\mathrm{p},i} = w_{\mathrm{p}}^{\mathrm{mock}}(r_{\mathrm{p},i}) - w_{\mathrm{p}}^{\mathrm{meas}}(r_{\mathrm{p},i})$ is the difference be-tween the projected two-point function predicted by the mock catalog and the measured value in the $i^{\mathrm{th}}$ separation bin (of 12 bins total), $C^{-1}$ is the inverse of the covariance matrix of the measurements, and the last term in Eq. (6) is the contribution from the difference between the pre-dicted and measured galaxy number density. The error on the galaxy number density assumes Poisson statistics for both the measured and predicted galaxy number den-sities.

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

The most important detail of this analysis is the pri-ors on the parameters. In all analyses discussed in this paper, we adopt priors that are uniform distributions over the intervals specified in Table 2.3. In the case of the assembly bias parameters $A_{\mathrm{cen}}$ and $A_{\mathrm{sat}}$, the priors have strict boundaries. Mathematically, these parameters must satisfy $-1 \leq A_{\mathrm{cen,sat}} \leq 1$. Physical considerations require parameter $\sigma_{\log M} > 0$. All other priors have a neg-ligible influence on the posterior aside from $\log M_0$. We find that $\log M_0$ is often very poorly constrained by clus-

| Parameter | Prior Interval |
|-----------|----------------|
| $\log(M_{\min})$ | [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] |
| $\alpha$ | [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.

tering data and priors on $\log M_0$ can have a non-negligible influence on inferred parameters.

## 3   RESULTS

We have performed parameter inference analyses in order to infer the underlying HODs 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 treatment of assembly bias. In our standard HOD analyses, the parameters $\log(M_{\min})$, $\sigma_{\log M}$, $\alpha$, $\log(M_1)$, and $\log(M_0)$ are permitted to vary. 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. Though the quality of our fits, as measured by the minimum of $\chi^2$ per Degree of Freedom (DoF), varies from case to case, all fits have a probability $\gtrsim 1\%$ of obtaining a higher value of $\chi^2$ by chance. An example of the inferred posteriors for the HOD parameters is shown in Figure 1 for the sample with $M_r < -20$. We do not show the full posteriors for all five threshold samples in the interest of brevity. 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 within $\Delta\chi^2 \leq 5.89$ of the best-fit model. Under the assumption of a Gaussian posterior distribution, $\Delta\chi^2 = 5.89$ contains 68% of the posterior probability for a five-parameter model. Note that the significant covariance in the data makes it difficult to determine the quality of fits from visual inspection of these figures.

Our parameter constraints can be compared to the standard HOD analysis performed by Zehavi et al. (2011) by examining the top two rows in each luminosity threshold grouping in Table 3.1. 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 primarily due to the slightly different cosmological model used in this work as compared to that of Zehavi et al. (2011). The most important differences are in the values of $\Omega_{\rm M}$, and $\sigma_8$. 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 $\sigma_8$ in order to maintain galaxy number densities fixed 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 $\log(M_{\min})$ 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, with 1-sigma constraints $\gtrsim 1$ dex for all samples. In several cases, the constraint on $\log(M_0)$ is determined by the prior given in Table 2.3. 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 $\log(M_0) = 12.23 \pm 0.17$ ($12.35 \pm 0.24$), 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 one of the authors of Zehavi et al. (2011) that the MCMC chains used in a number of their analyses may not have been properly converged and that this may have led to a significant underestimation of the uncertainties on the inferred parameters, especially $\log(M_0)$, $\log(M_{\min})$, and $\sigma_{\log M}$. The analysis of Zehavi et al. (2011) used $10^4$ samples, whereas we find that several $\times 10^5$ to $10^6$ samples are often necessary for convergence. In all cases, our final results are from $> 10^6$ samples of the posterior. The cause of the problem is likely a restrictive proposal distribution that causes the chain to diffuse through the posterior distribution only extremely slowly (Z. Zheng, Private Communication). We have recreated qualitatively similar behavior 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 $\alpha$ 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 $\alpha$ is the power-law index describing the dependence of average satellite number on halo mass. Increasing $M_1$ *decreases* the number of satellites in massive halos by increasing the mass scale where the power law abundance becomes operative. An increase in $\alpha$ can partly compensate for an increase in $M_1$
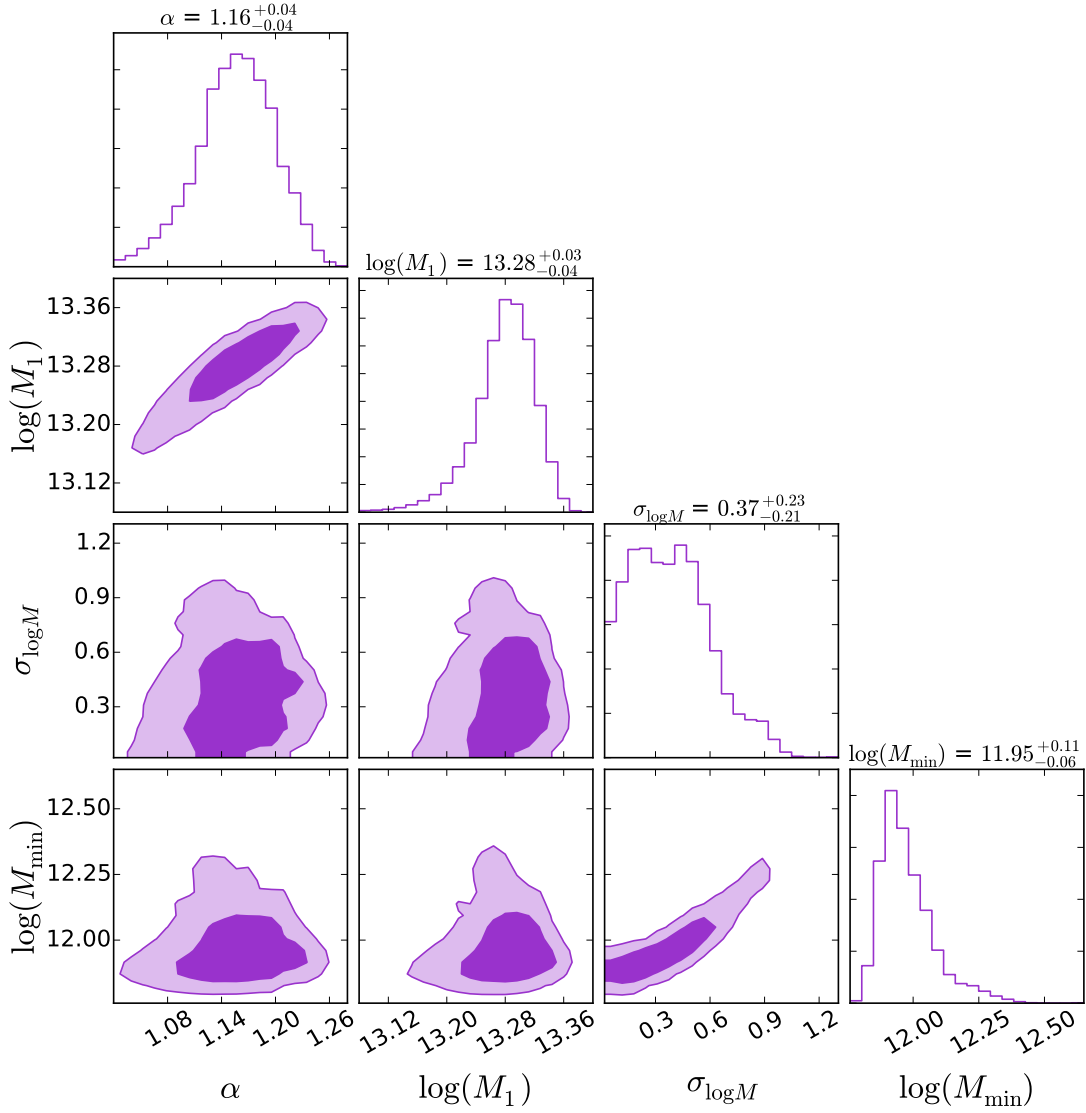
**Figure 1.** Two-dimensional marginalized constraints on HOD parameters inferred from standard HOD fits to $w_{\mathrm{p}}(r_{\mathrm{p}})$ data for the $M_r < -20$ sample. The HOD parameter $\log(M_0)$ is extremely poorly constrained by the $w_{\mathrm{p}}(r_{\mathrm{p}})$ data and has been omitted. 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 $16^{\mathrm{th}}$ and $84^{\mathrm{th}}$ percentiles.

by increasing the rate at which average satellite number grows with halo mass.

Look for the purple banana[3] in the third panel from the left in the bottom row of Figure 1. This panel illustrates that the parameters $\log(M_{\mathrm{min}})$ 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 $\log(M_{\mathrm{min}})$ 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 $M_{\mathrm{min}}$. The consequence is that $\log(M_{\mathrm{min}})$ and $\sigma_{\log M}$

are degenerate with each other such that most of the posterior probability lies in a narrow band along which $\log(M_{\mathrm{min}})$ and $\sigma_{\log M}$ are positively correlated. In the following plots, we exclude 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 $\log(M_{\mathrm{min}})$.

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 ha-

---

[3] 'Til they put us in the truck.

| Sample $M_r$ | Authors | $\log(M_{min})$ | $\sigma_{\log M}$ | $\log(M_1)$ | $\alpha$ | $A_{cen}$ | $A_{sat}$ | $\chi^2/\text{DoF}$ |
|---|---|---|---|---|---|---|---|---|
| $-21$ | Zehavi+11 | $12.78 \pm 0.10$ | $0.68 \pm 0.15$ | $13.80 \pm 0.03$ | $1.15 \pm 0.06$ | $--$ | $--$ | 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 | $12.14 \pm 0.03$ | $0.17 \pm 0.15$ | $13.44 \pm 0.03$ | $1.15 \pm 0.03$ | $--$ | $--$ | 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 | $11.83 \pm 0.03$ | $0.25 \pm 0.11$ | $13.08 \pm 0.03$ | $1.00 \pm 0.05$ | $--$ | $--$ | 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 | $11.57 \pm 0.04$ | $0.17 \pm 0.13$ | $12.87 \pm 0.03$ | $0.99 \pm 0.04$ | $--$ | $--$ | 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 | $11.45 \pm 0.04$ | $0.19 \pm 0.13$ | $12.64 \pm 0.04$ | $1.02 \pm 0.02$ | $--$ | $--$ | 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 $w_p(r_p)$ as well as fits using a parameterized model of assembly bias. Assembly bias is quantified by the parameters $A_{cen}$ ($A_{sat}$) for central (satellite) galaxies. The secondary property that we assume to determine the galaxy HOD is halo concentration. $A_{cen,sat} = 0$ means that there is no assembly bias while $A_{cen,sat} = 1$ ($A_{cen,sat} = -1$) means that galaxy abundance is maximally correlated (anticorrelated) with halo concentration at fixed $M_{vir}$. Thus the $A_{cen,sat}$ parameters span the range $[-1, 1]$. If the constraints on $A_{cen}$ and $A_{sat}$ are unspecified in the table, 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 $16^{th}$ and $84^{th}$ percentiles. In cases for which the posterior is monotonic within the physical parameter space, we show one-sided percentiles.
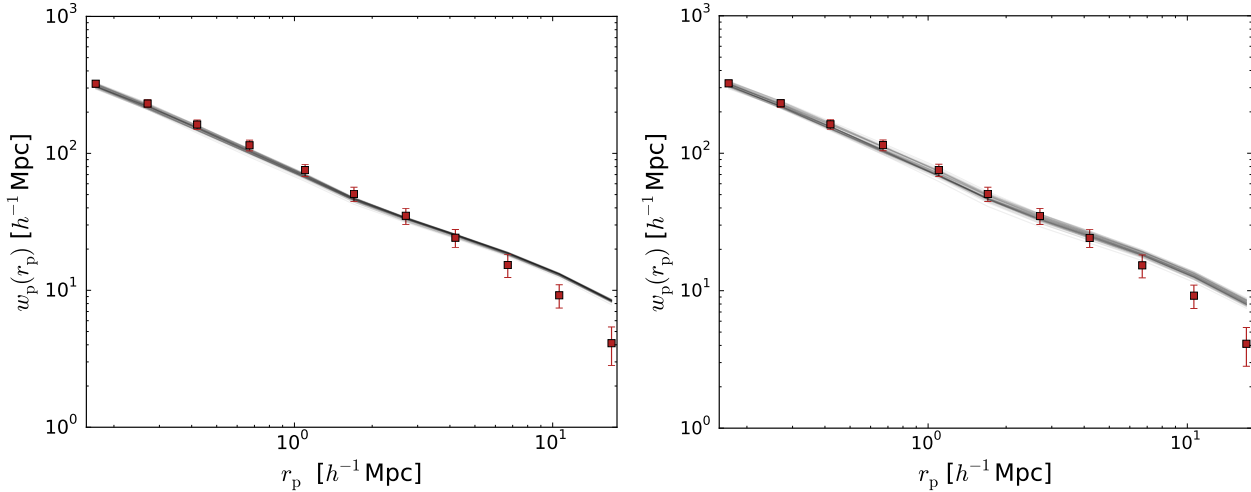


**Figure 2. Left:** The $M_r < -19$ threshold sample projected correlation function with diagonal elements of covariance (points with errorbars). The grey lines are 50 randomly-selected HOD models that yield $\Delta\chi^2 < 5.89$ compared to the best-fitting model. For a Gaussian posterior distribution, $\Delta\chi^2 = 5.89$ would contain 68% of the probability in the full five-dimensional parameter space. **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.
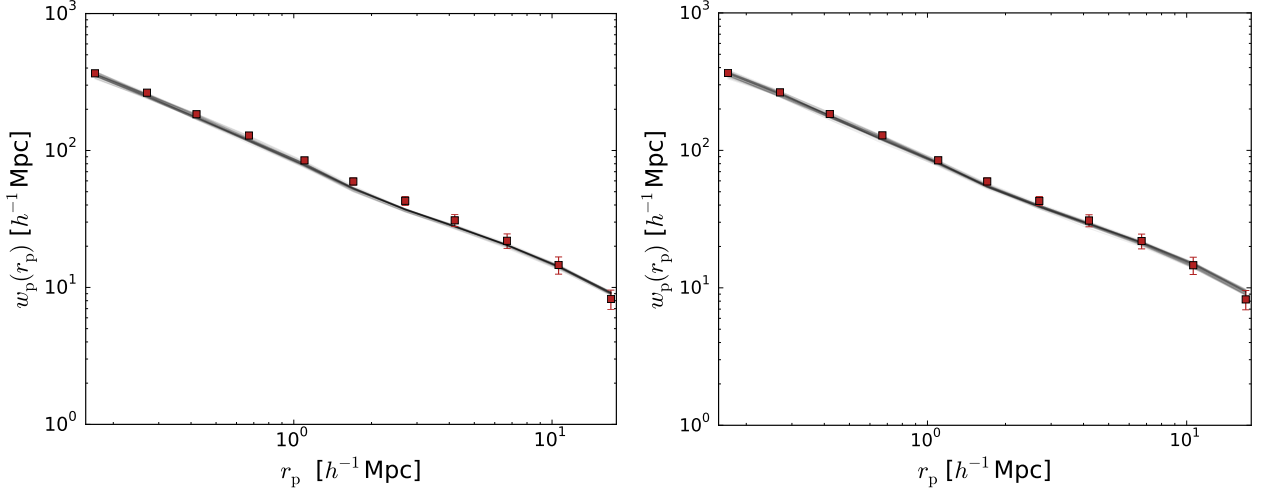
**Figure 3.** The same as Figure 2, but for the $M_r < -20$ threshold sample.
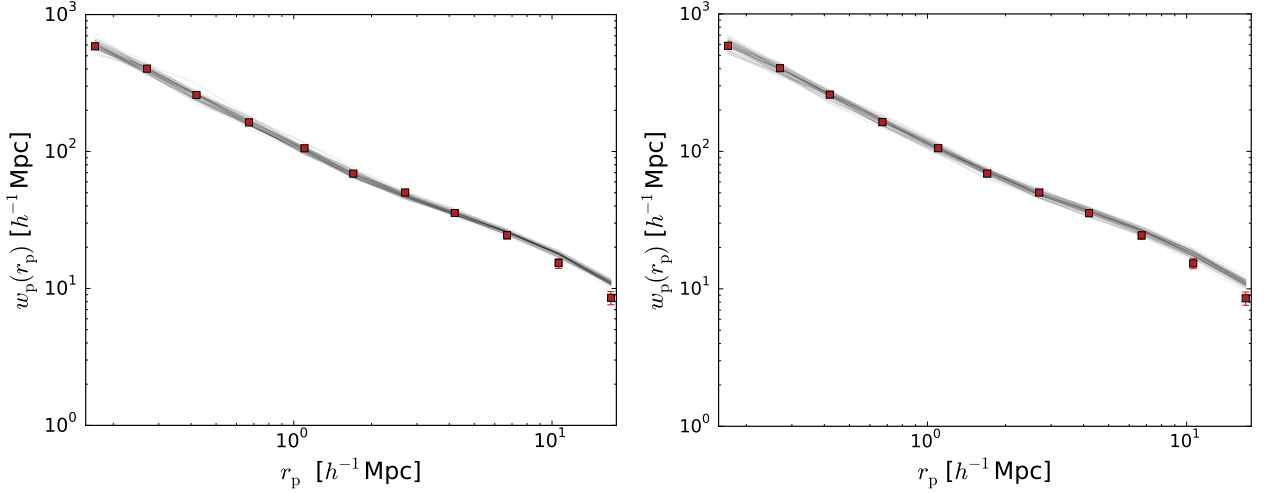


**Figure 4.** The same as Figure 2, but for the $M_r < -21$ threshold sample.

los identified in a numerical simulation of cosmological structure formation (BolshoiP). This greatly mitigates modeling uncertainties associated with nonlinear density 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 caused by insufficient sampling of the posterior by Zehavi et al. (2011).

### 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, $A_{cen}$ and $A_{sat}$, that describe the strength of central galaxy and satellite galaxy assembly bias respectively. These parameters are limited to values of $-1 \leq A_{cen,sat} \leq 1$, and $A_{cen,sat} = 0$ when there is no galaxy assembly bias. In this work, we use halo concentration as our secondary halo property, so $A_{cen,sat} = 1$ ($A_{cen,sat} = -1$) means that the mean number of galaxies per halo is maximally correlated (anti-correlated) with halo concentration. The model and its implementation in `Halotools` Hearin et al. (2016) is discussed further in Section 2.2 above and in more detail 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 from the one-halo (highly nonlinear) to two-halo (nearly linear) regimes

near $r_{\rm p} \sim 2\,h^{-1}{\rm Mpc}$, an effect that could be anticipated by the known scale-dependence of galaxy assembly bias (Sunayama et al. 2016). 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 distribution of either $A_{\rm cen}$ or $A_{\rm sat}$ is a monotonic within the physical parameter range, we quote an upper or lower limit on the parameter (e.g., see the posterior distribution for $A_{\rm cen}$ in Fig. 6). Examples of one- and two-dimensional visualizations of the posteriors from our analyses are shown in Figure 5, Figure 6, and Figure 7 for the $M_r < -19$, $M_r < -20$, and $M_r < -21$ samples respectively.

Table 3.1 and Figures 5–7 all make several simple, generic points. Introducing additional parameter freedom associated with galaxy assembly bias generally increases the volume of the viable parameter space, even for the subset of five standard HOD parameters. Constraints on the standard HOD parameters are generally less restrictive in models that include assembly bias. 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. This is the case, for example, for the $M_r < -21$ sample with posteriors shown in Fig. 7. In this case, both $A_{\rm cen}$ and $A_{\rm sat}$ have posteriors that peak near zero, but are very broad. Nonetheless, it is apparent that the presence of galaxy 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 differences particularly for $\alpha$, 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. Figure 8 compares the inferred HODs for the $M_r < -20$ sample in a standard HOD analysis with the inferred HOD from our decorated HOD analysis that includes parameters to describe assembly bias. This is an explicit example of the degree to which the inferred relationship between galaxies and halos can be altered by assembly bias. In this case, the efficiency of forming a sufficiently luminous galaxy is a function that increases slowly over nearly two decades in halo mass in contrast to a standard analysis HOD where this efficiency rises rapidly over roughly one third of a decade in $M_{\rm vir}$. Taking these results at face value, the

| $M_r$ Threshold | $\Delta$BIC |
|:---:|:---:|
| -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, $\Delta$BIC, 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 $|\Delta$BIC$| \geq 5$ strongly favor one model over another.

reason that the HOD of central galaxies can vary slowly over several orders of magnitude in $M_{\rm vir}$ is that larger galaxies form more efficiently in high-concentration halos at fixed mass.

Though the evidence for assembly bias is the most significant for the $M_r < -20$ sample, there are hints of galaxy assembly bias in other samples. Satellite galaxies show a marginal preference for occupying halos of higher concentration in the $M_r < -19$ threshold sample. The $M_r < -19.5$ 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 well within $1\sigma$. The implications of all of these results are too numerous to address in the present manuscript; however, we plan to address them in forthcoming publications.

Beyond the fit results, it is also possible to ask whether or not the fits to the data warrant the use of additional parameters (in this case, to describe assembly bias). This requires a metric that penalizes the use of additional parameters. A commonly used metric is the Bayesian Information Criterion (BIC, Schwarz 1978), defined as

$$\mathrm{BIC} = -2\mathcal{L}_{\rm max} + k\ln N, \qquad (7)$$

where $\mathcal{L}_{\rm max}$ is the maximum of the likelihood, $k$ is the number of paramters in the model and $N$ is the number of data points. We compute the change in the BIC, $\Delta$BIC, by subtracting the BIC for the analysis with assembly bias from the BIC in the standard analysis in which there is no assembly bias. With this convention, $\Delta$BIC $> 0$ favors assembly bias, suggesting that the data warrant the additional parametric freedom. Conversely, $\Delta$BIC $< 0$ favors the standard analysis, suggesting that the additional freedom in the model does not improve the description of the data to a sufficiently great degree to support the added complexity. A common convention is to refer to the preference as *strong* when $|\Delta$BIC$| \geq 5$.
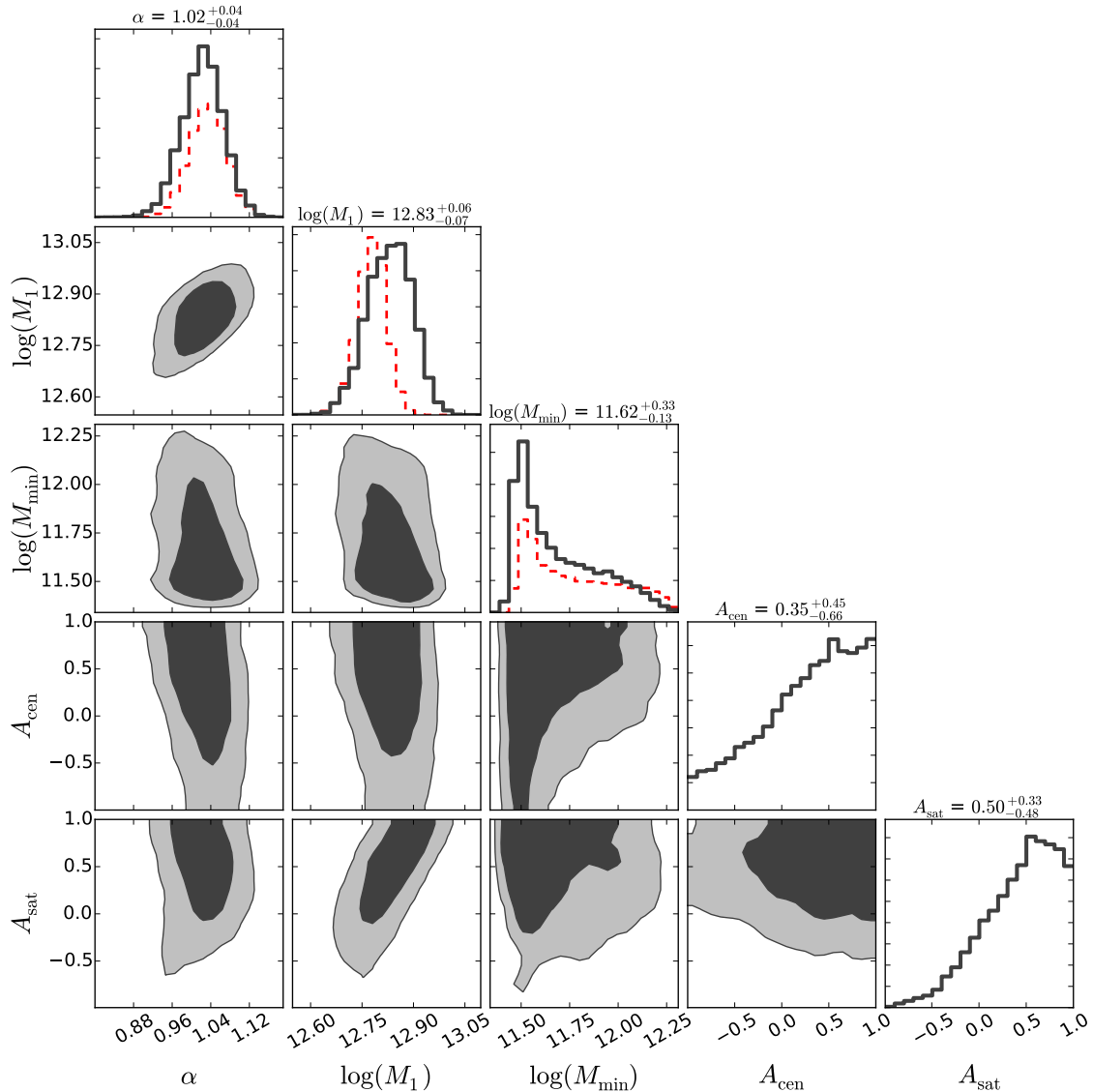
**Figure 5.** Two-dimensional marginalized constraints on decorated HOD parameters inferred from fits to $w_{\mathrm{p}}(r_{\mathrm{p}})$ data for the $M_r < -19$ sample. The contours and histograms along the diagonal panels are as in Fig. 1. The dashed line in the panels along the diagonal show the posteriors on these parameters from the standard analysis without any parameterization of galaxy assembly bias. 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 $\log(M_{\min})$ share a narrow degeneracy, so we have suppressed $\sigma_{\log M}$ in order to make constraints on other parameters more easily visible.

The values of $\Delta$BIC are given in Table 3.2. In the analysis of the $M_r < -21$ sample, $\Delta$BIC $< 0$, indicating that the additional parameters describing assembly bias are not warranted by the data. For all other samples, $\Delta$BIC $> 0$, showing a preference for the models with assembly bias. The cases that most support the additional complexity of the decorated HOD are $M_r < -19$ and $M_r < -20$, for which $\Delta$BIC rises nearly to the level of a strong preference.

# 4  CONCLUSIONS

We have re-analyzed the SDSS DR7 measurements of projected galaxy clustering, $w_{\mathrm{g}}$, and number density, $n_{\mathrm{g}}$, originally published in Zehavi et al. (2011). The novelty of our work is that we have performed these analyses using the decorated Halo Occupation Distribution (HOD) formalism, which allows for galaxy assembly bias and was introduced by Hearin et al. (2016) precisely for this purpose. In this work, we provide the first quantitative constraints on assembly bias derived from the Decorated HOD. We enumerate our most important conclusions below.
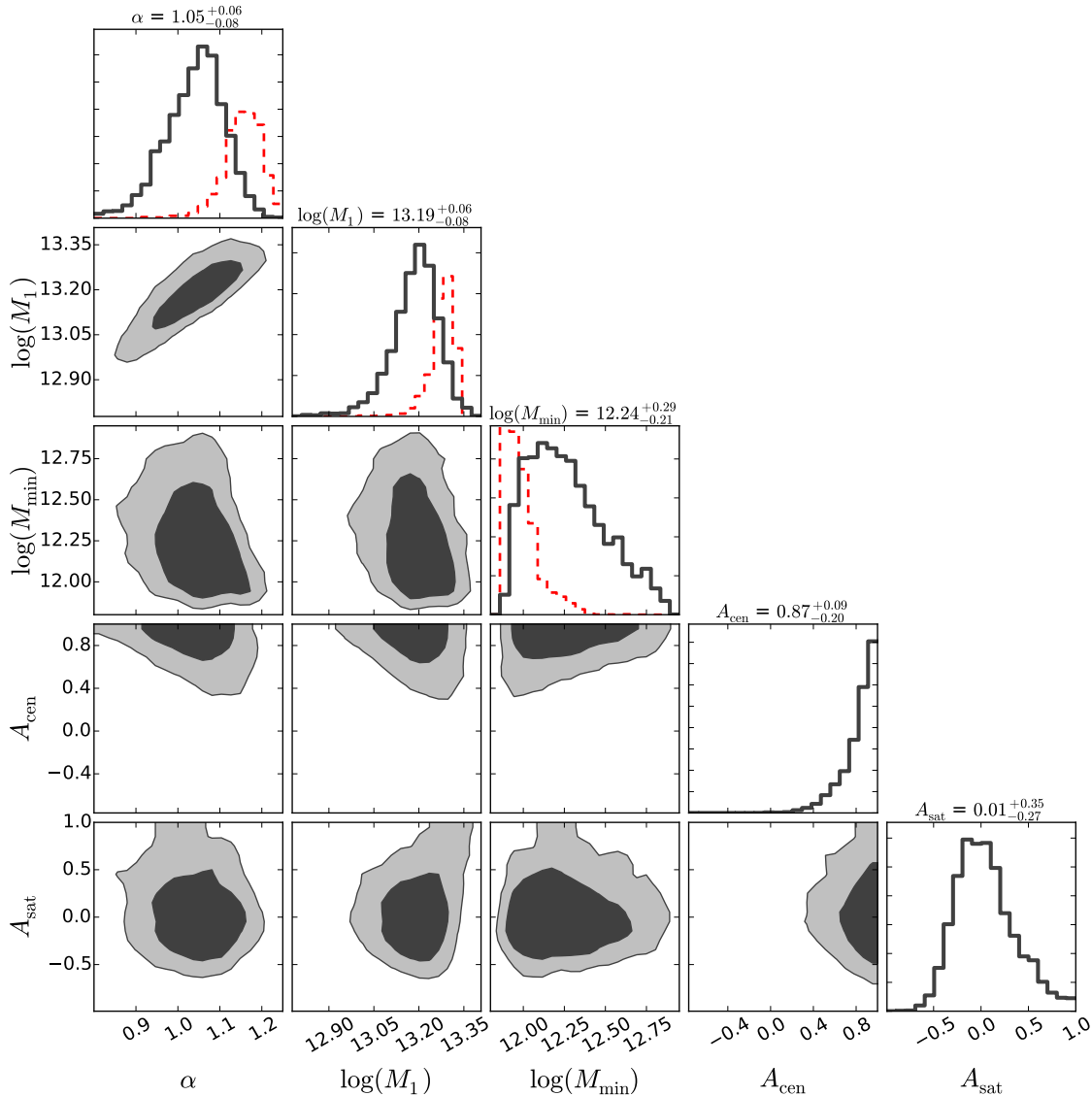
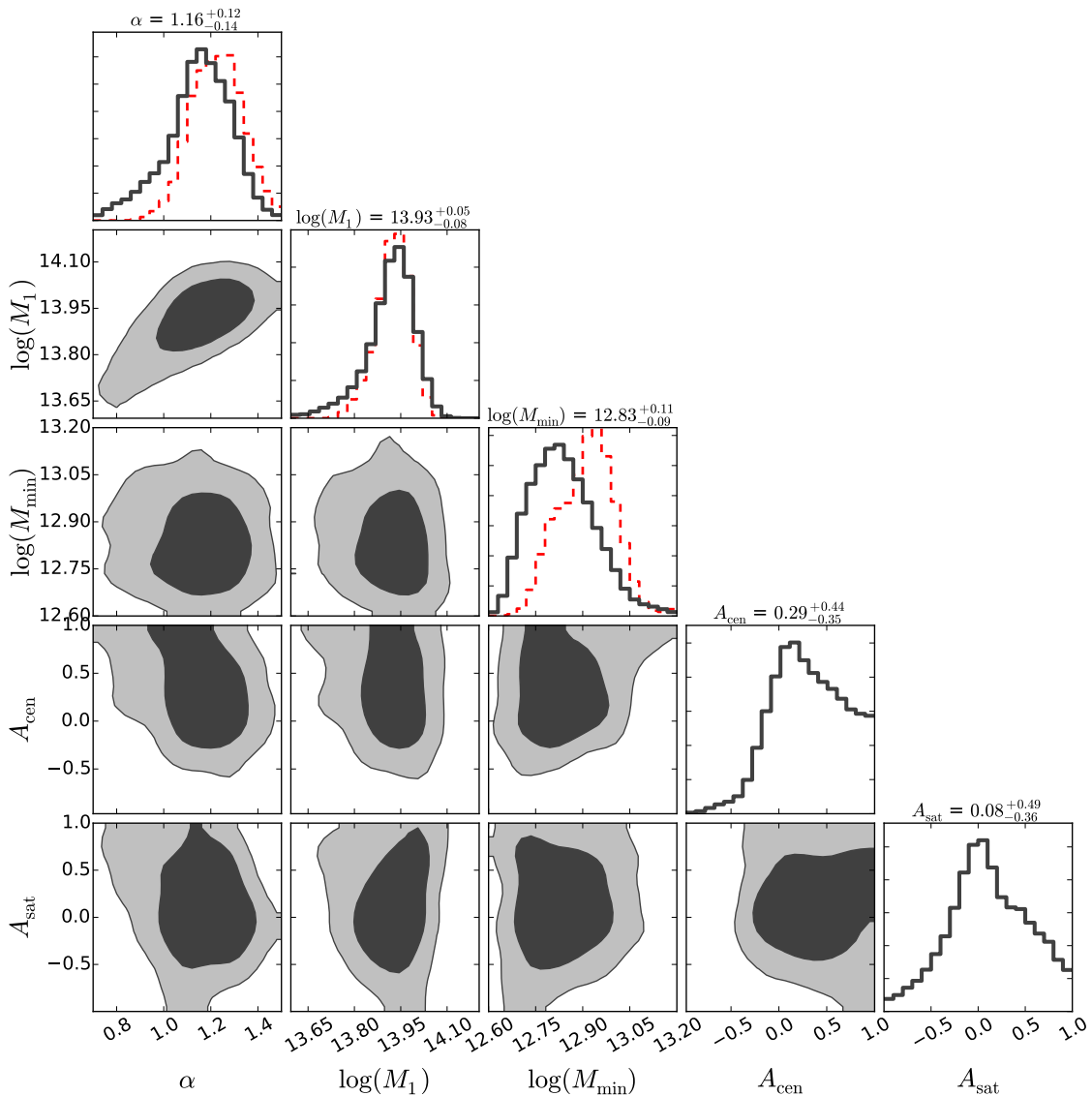**Figure 6.** The same as Figure 5, but for the $M_r < -20$ sample.

(i) It is not possible to rule out galaxy assembly bias using SDSS DR7 measurements of $w_{\rm p}$ and $n_{\rm g}$.

(ii) Decorated HOD fits to $w_{\rm p}$ and $n_{\rm g}$ favor significant levels of assembly bias, particularly in the lower luminosity thresholds we study. Both the $M_r < -20, -20.5$ samples prefer relatively strong central galaxy assembly bias. The $M_r < -19$ sample favors satellite assembly bias. Assembly bias in these samples improves the capability of the model to describe galaxy clustering near the one-halo to two-halo transition at $r_{\rm p} \sim 2\,h^{-1}{\rm Mpc}$.

(iii) The evidence for galaxy assembly bias weakens for brighter galaxy samples: Decorated HOD fits to the $M_r < -21$ sample are consistent with both $A_{\rm cens} = 0$ and $A_{\rm sats} = 0$. This is consistent with the well-established result that *halo assembly bias* weakens with increasing halo mass over the dynamic range relevant to these galaxy samples (see, e.g., Figure 8 of Hearin et al. 2016, and references therein).

(iv) Our posteriors and best-fit parameters summarized in Table 3.1 supersede the values published in Zehavi et al. (2011), as direct-mock-population together with the Decorated HOD allows us to account for highly significant systematics that have heretofore been neglected from all HOD fits to SDSS data. We note that our findings update the original Zehavi et al. (2011) results *even for our fits to standard HODs without assembly bias*, because the original results were derived from MCMC chains that did not sufficiently sample the HOD model posteriors.

We conclude by noting that since $w_{\rm p}$ and $n_{\rm g}$ are already very well-measured in DR7, it is likely that further improvements on assembly bias constraints at low-redshift will require additional observational measurements, e.g., galaxy–galaxy lensing, group statistics, etc. I need to think of a more clever ending.

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

## 5  ACKNOWLEDGEMENTS

## REFERENCES

Behroozi P. S., et al., 2013, ApJ, 762, 109
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
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Blanton M. R., Berlind A. A., 2007, ApJ, 664, 791
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
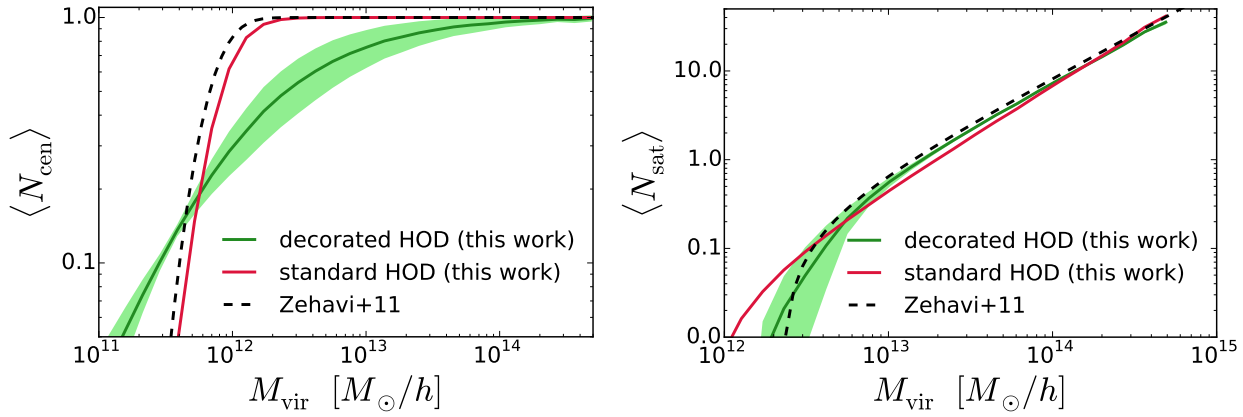
**Figure 8.** We show the first moments of the central and satellite occupation distributions from the best-fit models to the $M_r < -20$ sample in the left and right panels, respectively. The green curve shows the HOD of our best-fit model that includes the effect of assembly bias; the red curve shows the HOD of the best-fit model in which assembly bias is assumed to be zero; the dashed black curve shows the best-fit HODs from Zehavi et al. (2011). The green error band around the decorated HOD results shows the variance in the HOD from randomly sampling one hundred points in parameter space with $\chi^2_{\rm model} - \chi^2_{\rm min} < 1$.

Croton D. J., Gao L., White S. D. M., 2007, Mon. Not. R. Astron. Soc., 374, 1303

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., Campbell D., Tollerud E., Behroozi P., Diemer B., Goldbaum N. J., Jennings E., Leauthaud A., Mao Y.-Y., More S., Parejko J., Sinha M., Sipocz B., Zentner A., 2016, ArXiv:1606.04106

Hearin A. P., Watson D. F., van den Bosch F. C., 2014, ArXiv:1404.6524

Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, Mon. Not. R. Astron. Soc.

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

Lacerna I., Padilla N., 2011, Mon. Not. R. Astron. Soc., 412, 1283

Leauthaud A., et al., 2012, ApJ, 744, 159

Lehmann B. V., Mao Y.-Y., Becker M. R., Skillman S. W., Wechsler R. H., 2015, ArXiv:1510.05651

Lin Y.-T., Mandelbaum R., Huang Y.-H., Huang H.-J., Dalal N., Diemer B., Jian H.-Y., Kravtsov A., 2015, ArXiv:1504.07632

Ma C.-P., Fry J. N., 2000, ApJ, 543, 503

Miyatake H., More S., Takada M., Spergel D. N., Mandelbaum R., Rykoff E. S., Rozo E., 2016, Physical Review Letters, 116, 041301

More S., Miyatake H., Takada M., Diemer B., Kravtsov A. V., Dalal N. K., More A., Murata R., Mandelbaum R., Rozo E., Rykoff E. S., Oguri M., Spergel D. N., 2016, ArXiv e-prints

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

Paranjape A., Kovač K., Hartley W. G., Pahwa I., 2015, Mon. Not. R. Astron. Soc., 454, 3030

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

Rodríguez-Puebla A., Drory N., Avila-Reese V., 2012, ApJ, 756, 2

Schwarz G., 1978, Ann. Statist., 6, 461

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20

Seljak U., 2000, Mon. Not. R. Astron. Soc., 318, 203

Sinha M., 2016, http://dx.doi.org/10.5281/zenodo.55161

Sunayama T., Hearin A. P., Padmanabhan N., Leauthaud A., 2016, Mon. Not. R. Astron. Soc., 458, 1510

Tinker J. L., Conroy C., Norberg P., Patiri S. G., Weinberg D. H., Warren M. S., 2008, ApJ, 686, 53

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., More S., Cacciato M., Mo H., Yang X., 2013, Mon. Not. R. Astron. Soc., 430, 725

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

Wang L., Weinmann S. M., De Lucia G., Yang X., 2013, Mon. Not. R. Astron. Soc., 433, 515

Wang Y., Yang X., Mo H. J., van den Bosch F. C., Weinmann S. M., Chu Y., 2008, ApJ, 687, 919

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., 2006, Astrophys. J. Lett., 638, L55

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