Mon. Not. R. Astron. Soc. **000**, ??–?? () Printed 27 May 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.

1 INTRODUCTION

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`. Our HOD model is identical to the model described in **?**, and so we will only briefly review the salient features of our methodology here. Interested readers can always refer to `halotools.readthedocs.io` for comprehensive documentation on all aspects of the `Halotools` framework.

$$\langle N_{\rm sat}(M)\rangle = \left(\frac{M - M_0}{M_1}\right)^{\alpha} \qquad (1)$$

2.2 HOD with Assembly Bias: The Decorated HOD

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 **?** as implemented in the `emcee` software package (**?**). 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 priors 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 $A_{\rm cen}$ 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.

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.

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 **?** 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 **?** 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 **?**. 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. **?** 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 $\sigma_8 = 0.82$. 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 **?** 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 **?**. As is evident in Figure 1, $\log(M_{\rm min})$ and $\sigma_{\log M}$ share a relatively narrow degeneracy. 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 $M_{\rm min}$.

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 2.3. This is in stark contrast to several of the results of **?**. For example, for the threshold sample with $M_r < -19.5$ ($M_r < -20.5$), **?** 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. (1), 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 **?** must be an error or must be driven by some prior that is not clearly specified in their manuscript.

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 existing constraints in the literature in two 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 simulations. 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.

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.

The one-dimensional marginalized constraints on all parameters form 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.

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. Visually, these differences appear to be small; however, Table 3.1 shows that they are statistically important.

4 CONCLUSIONS

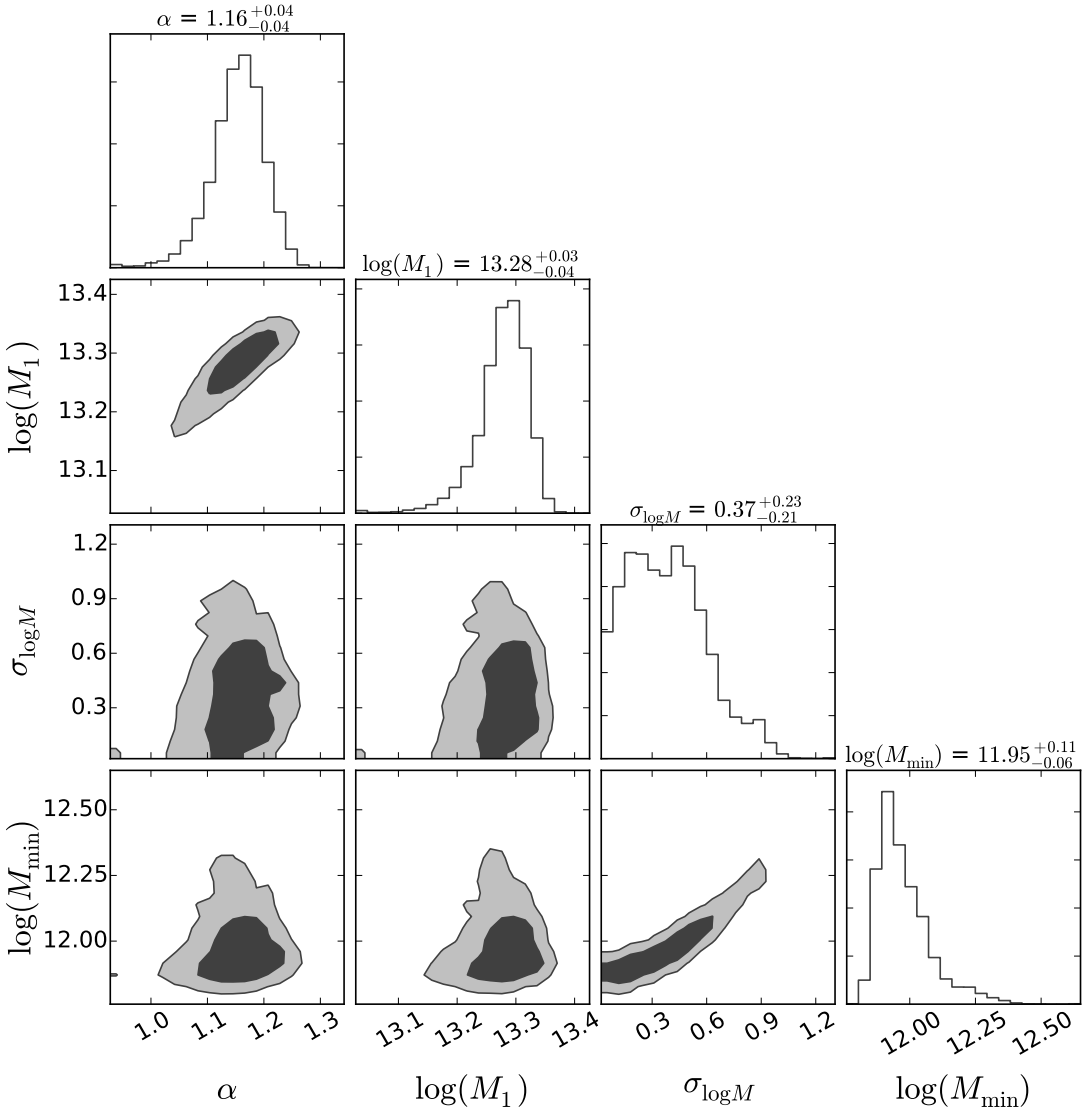

Figure 1. Two-dimensional marginalized constraints on HOD parameters inferred from standard HOD fits to $w_{\rm p}(r_{\rm p})$ data for the $M_r < -20$ sample. The HOD parameter $\log(M_0)$ is extremely poorly constrained by the $w_{\rm p}(r_{\rm p})$ data and has been omitted. The inner contours contain 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.

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.

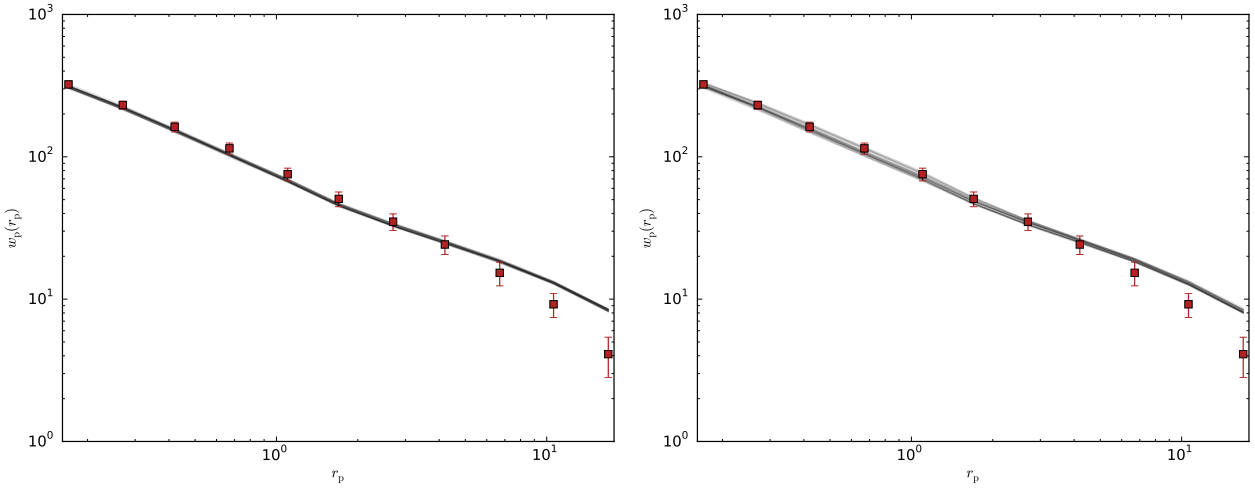

Figure 2. Left: The $M_r < -19$ threshold sample projected correlation function with diagonal elements of covariance (points with errorbars). The grey lines are 25 randomly-selected HOD models that yield $\Delta\chi^2 < 1$ compared to the best-fitting model. **Right:** Same as the left panel but using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

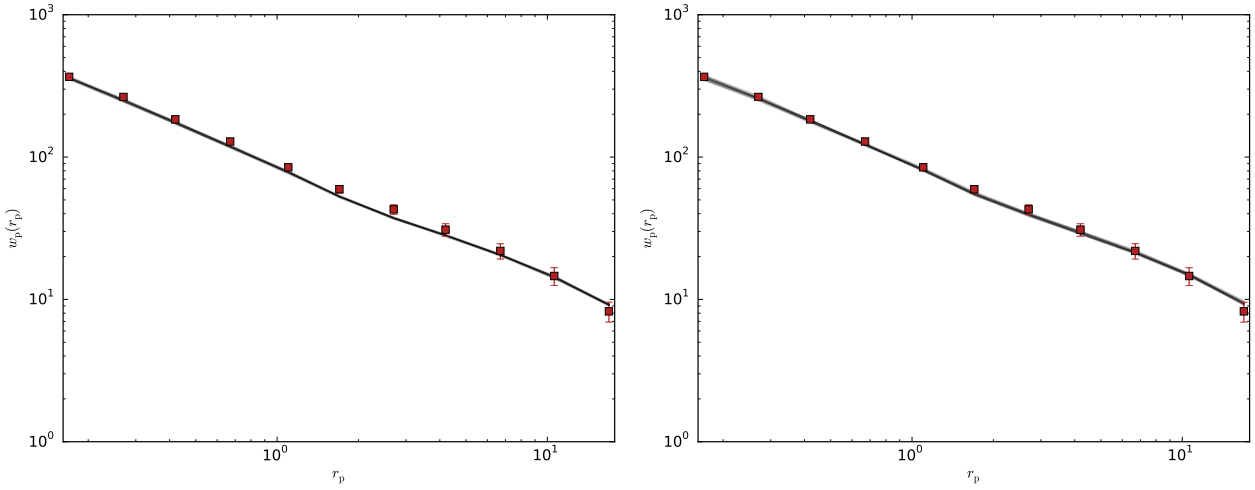

Figure 3. The same as Figure 2, but for the $M_r < -20$ threshold sample.

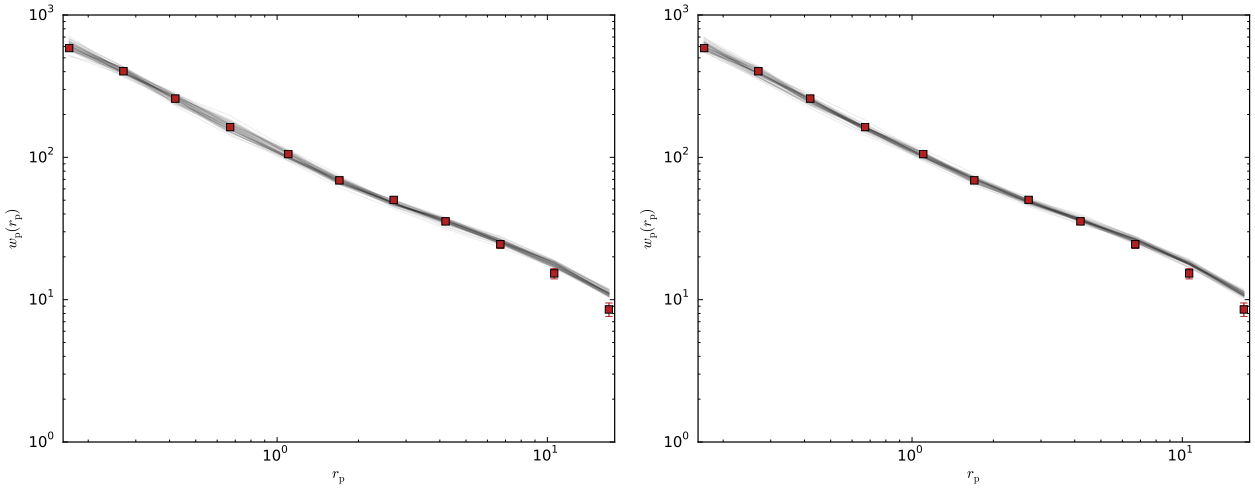

Figure 4. The same as Figure 2, but for the $M_r < -21$ threshold sample.

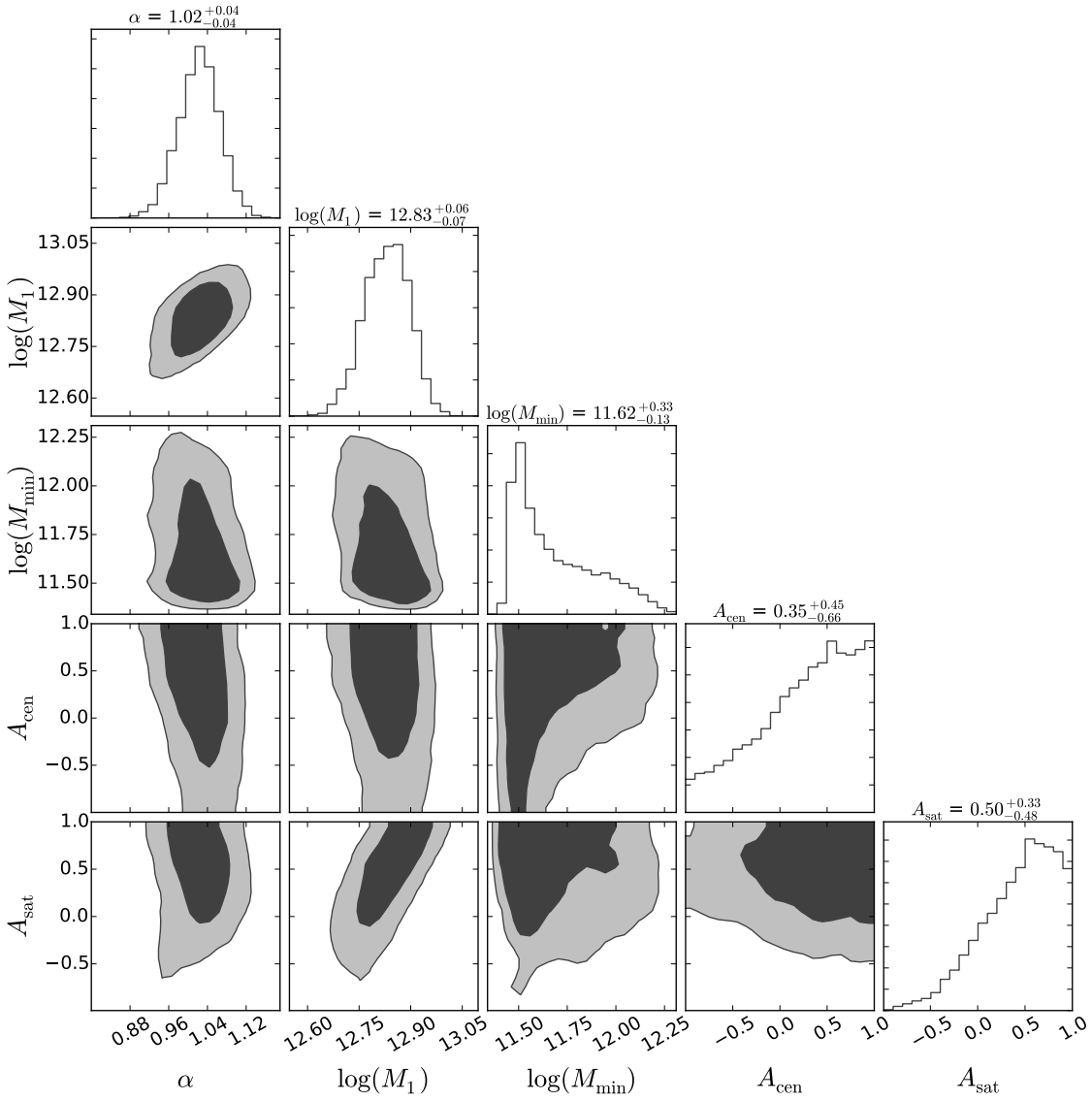

Figure 5. Two-dimensional marginalized constraints on decorated HOD parameters inferred from fits to $w_{\rm p}(r_{\rm p})$ data for the $M_r < -19$ sample. The contours and histograms along the diagonal panels are as in Fig. 1. The decorated HOD models include a two-parameter model for assembly bias. 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.

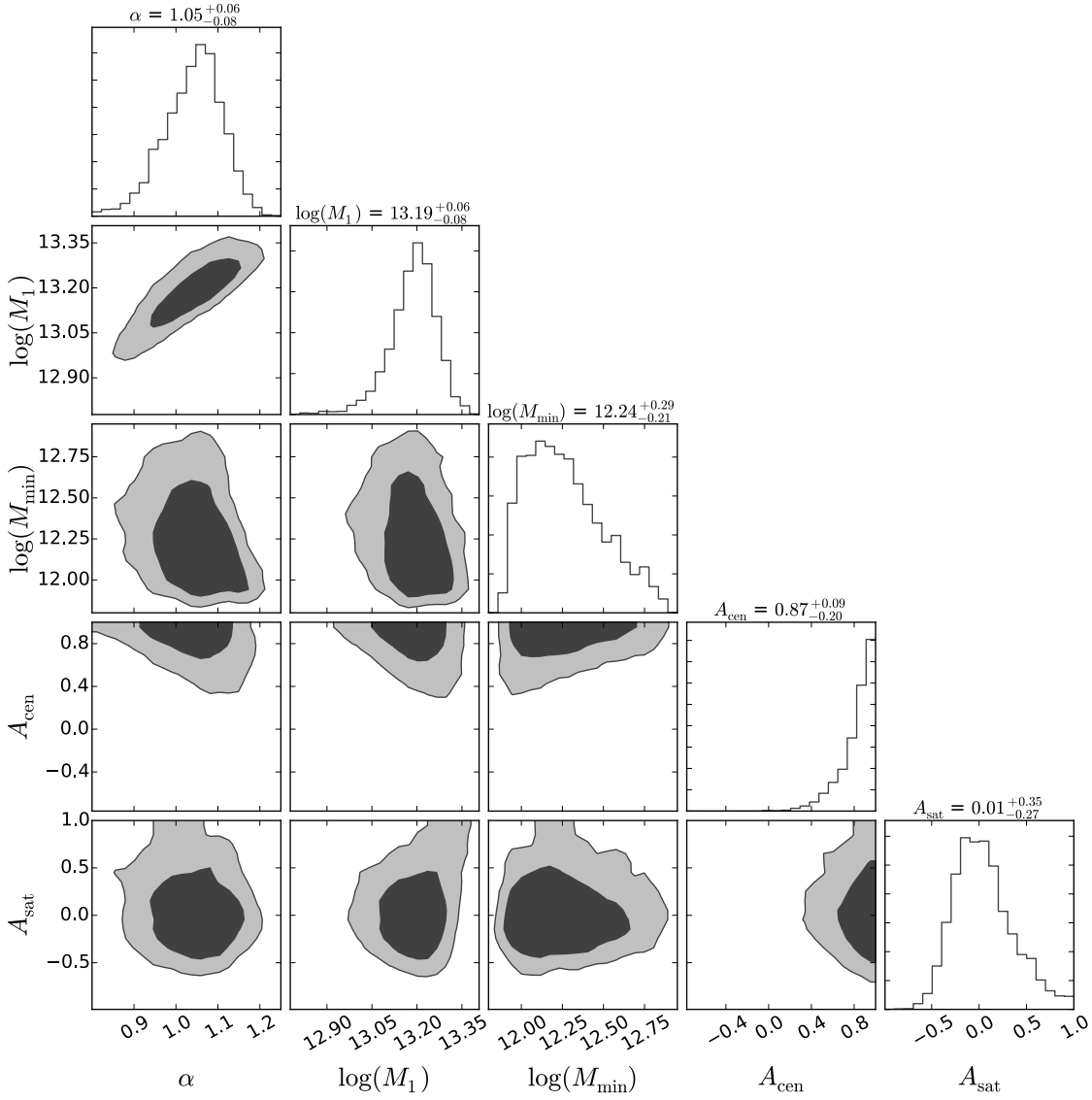

Figure 6. The same as Figure 5, but for the $M_r < -20$ sample.

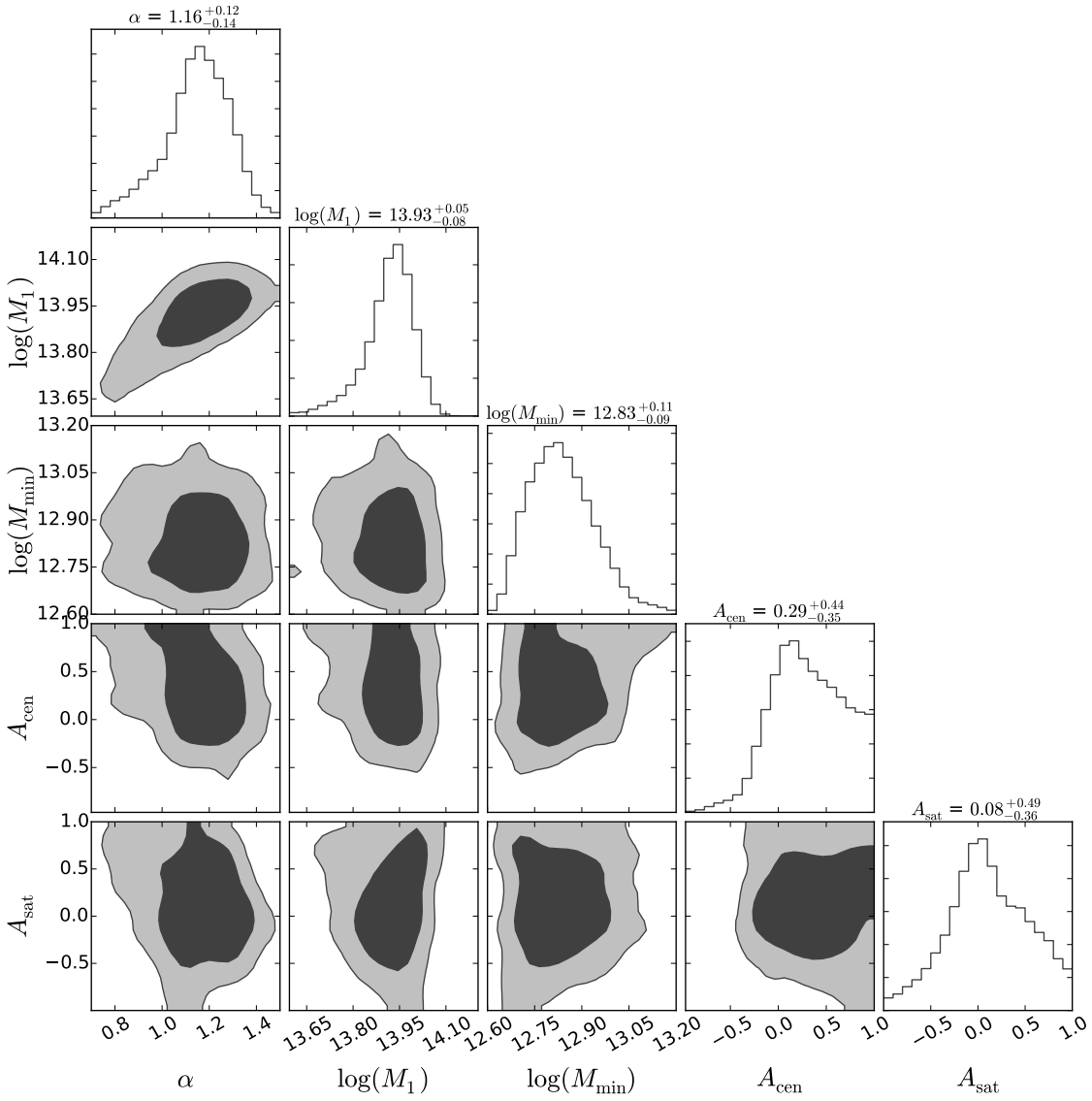

Figure 7. The same as Figure 5, but for the $M_r < -21$ sample.