Probing Galaxy assembly bias in BOSS galaxies using void probabilities

Kilian Walsh and Jeremy Tinker

Center for Cosmology and Particle Physics, Department of Physics, New York University, New York

ABSTRACT

We measure the void probability function (VPF) of galaxies in the Baryon Oscillation Spectroscopic Survey (BOSS). The VPF provides complementary information to standard two-point statistics in that it is sensitive to galaxy bias in the most extreme underdensities in the cosmic web. Thus, the VPF is ideal for testing whether halo occupation of galaxies depends on large-scale density, an effect known as galaxy assembly bias. We find that standard HOD model – one parameterized by halo mass only – fit only to the two-point function, accurately predicts the VPF. Additionally, for HOD models where density dependence is explicitly incorporated, the best-fitting models fit to the combination of the correlation function and the VPF have zero density dependence. Thus, galaxy assembly bias is not a strong source of systematic uncertainty when modelling the clustering of massive galaxies.

Key words: galaxies: haloes – galaxies: statistics – cosmology: observations.

1 INTRODUCTION

The growth of structure in the dark matter content of the universe leads to the formation of regions of high relative density over time and, contrastingly, also regions of significant sparsity in the matter field. Galaxies act as a biased tracer of the underlying dark matter density field (see Desjacques, Jeong & Schmidt 2018 for a thorough review of large-scale galaxy bias), so these regions of sparsity lead to the existence of large voids within the galaxy distribution (see the review of van de Weygaert & Platen 2011). Since the discovery of voids that were larger in extent than typical galaxy clusters (Gregory & Thompson 1978), they have been considered a probe of interest for the study of cosmology and large-scale structure formation (see e.g. Betancort-Rijo et al. 2009, Einasto et al. 2011, Lavaux & Wandelt 2012, Clampitt, Jain & Sánchez 2016, Kitaura et al. 2016, and Mao et al. 2017 for recent examples).

Voids also present a useful probe to understand the nature of galaxy formation. It is known that galaxies found in voids are typically spiral/irregularly-shaped and found to be H I rich, having low luminosity and stellar mass, and small size relative to the general population of galaxies (Tavasoli et al. 2015, Beygu et al. 2017, and Pustilnik, Tepliakova & Makarov 2019). This suggests that galaxy formation and evolution may have a dependence on the large-scale dark matter density. Controlling for factors such as galaxy mass, or the halo abundances found in underdense regions of simulations, there are also studies done in the context of voids which claim to find no special relationship between galaxies and their environment (Patiri et al. 2006; Croton & Farrar 2008; Tinker et al. 2008a; Tinker & Conroy 2009).

It is well established, theoretically and in high-resolution dark matter simulations, that the properties of dark matter haloes in ΛCDM cosmologies – such as halo mass (Press & Schechter 1974; Mo & White 1996) and formation time (Sheth & Tormen 2004; Wechsler et al. 2006; Zentner 2007; Dalal et al. 2008) – are also strongly dependent on their large-scale dark matter over density. As such, the observed correlations of galaxy properties with environment may be explained with a model that connects galaxies with haloes.

One of the chief models of the galaxy halo connection is the Halo Occupation Distribution (HOD; see the review of the galaxy connection by Wechsler & Tinker 2018) The HOD parameterizes the bias between galaxies and dark matter by the statistical relationship between dark matter haloes and the number of galaxies within them (e.g. Scoccimarro et al. 2001; Berlind & Weinberg 2002; Cooray & Sheth 2002). The HOD has been remarkably successful in explaining a host of observational phenomena (e.g. Zehavi et al. 2011, Tinker et al. 2012, Reid et al. 2014, and Coupon et al. 2015). In its most basic form, the HOD parameterizes the occupation of a halo by galaxies based only on the mass of the halo. However, recent studies have suggested that there may be ‘assembly bias’ in galaxy occupation, whereby the occupation of dark matter haloes may depend on properties other than mass (Reddick et al. 2013; Zentner et al. 2016; Lehmann et al. 2017). Physically, this would manifest as a correlation between galaxy formation efficiency and that large-scale environment of a galaxy at fixed halo mass. Tinker, Weinberg & Warren (2006) proposed using void statistics, in combination with two-point statistics, to determine if such a correlation exists. The two-point correlation function is sensitive to haloes in mean and high-density environments, while most pairs are found, while voids are defined by haloes in the most extreme low-density regions of the universe. Wang et al (2019) recently extended this approach to
counts-in-cells. If galaxy formation efficiency changes from high to low densities, one could not simultaneously fit both statistics using an HOD model that parameterizes occupation solely on halo mass. Zentner, Hearin & van den Bosch (2014) found similar results using abundance matching models of galaxy bias to construct colour-defined galaxy samples. Using this technique, Tinker et al. (2008b) found that the galaxies in the Sloan Digital Sky Survey (York et al. 2000) could be fit with this mass-only approach. Any correlation of the HOD with large-scale environment could only be minimal, or else the models could no longer fit both statistics.

In this work, we extend the investigation of the HOD for a far larger sample of galaxies, at higher redshifts, by performing this test using data from the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013). Unlike the SDSS, the BOSS survey is designed for cosmological inference. Attempts to use BOSS (BOSS; Dawson et al. 2013) is one of the spectroscopic surveys of SDSS-III (Eisenstein et al. 2011), comprising 1.5 million galaxies roughly 85 per cent of the full footprint. Additionally, we restrict the survey geometry, a significant fraction of target BOSS galaxies were subject to fibre collisions where only one galaxy in a collided pair was observed (more details in e.g. Anderson et al. 2014, Ross et al. 2017). In our measurements, it was necessary to correct for the missing galaxy pairs that lie within the fibre collision angular radius of 62 arcsec, which corresponds to a projected separation of around \( r_p \sim 0.4 \, h^{-1} \text{Mpc} \) at the CMASS mean redshift and thus has a big impact on small-scale clustering. To estimate the true clustering, we

2 DATA AND MEASUREMENTS

2.1 BOSS data

BOSS (Dawson et al. 2013) is one of the spectroscopic surveys of SDSS-III (Eisenstein et al. 2011), comprising 1.5 million galaxies and intended to measure the Baryon Acoustic Oscillation (BAO) feature (see Weinberg et al. 2013 for a review). The full survey spans 10 000 square degrees and probes up to redshift \( z = 0.8 \). Target selection for the CMASS sample used here is described in Reid et al. (2016). All results here use Data Release 11 of the SDSS, the penultimate data release of BOSS, which comprises roughly 85 per cent of the full footprint. Additionally, we restrict our analyses to the North Galactic Cap (NGC) region, which is the bulk of the BOSS footprint, yielding a total sample covering roughly 6000 deg².

2.2 Data preparation

Our probe of void statistics is the void probability function (VPF) which we will define presently. The VPF is highly sensitive to sample number density, thus we restrict the BOSS data to a range of redshift for which the number density is above some threshold. We then randomly subsampled the data within this redshift interval such that the remaining galaxies would have a constant number density at comoving volume in redshift bins in that range. The redshift range we chose for the Northern Galactic Cap (NGC) CMASS (the high-redshift portion of BOSS targets) galaxies was \( z = [0.4575, 0.5725] \) such that the number densities of the remaining galaxies was 65 per cent of the peak number density of the sample. These values were chosen as a compromise between maintaining a high number density while also probing a large volume of the sample by not overly restricting the redshift range. The resulting galaxy sample has just under 500 000 galaxies. We randomly subsample the galaxies at each location in redshift, rather than a more complicated selection, due to fact that the clustering of CMASS galaxies is roughly constant with redshift (White et al. 2011). Fig. 1 shows the number density of the downsamples galaxies as a function of redshift, compared with the original sample and with the corrected number density, which is discussed in the following section.

2.3 Correlation function measurements

To quantify two-point clustering we use the projected correlation function \( w_p(r_p) \) (Davis & Peebles 1983). This function is defined as the integral of the two-dimensional clustering measurement \( \xi(r_p, \pi) \) along the line-of-sight direction \( \pi \) and is insensitive to redshift space distortions, while offering good constraints on HOD parameters through its signals on small projected scales \( r_p \):

\[
w_p(r_p) = 2 \int_0^{\pi_{\text{max}}} \xi(r_p, \pi) d\pi, \tag{1}
\]

with \( \pi_{\text{max}} = 80 \, h^{-1} \text{Mpc} \). We use the Landy–Szalay estimator (Landy & Szalay 1993) for two-point statistics. We also choose 24 logarithmic spaced bins in \( r_p \) in the range \([0.3, 60.0]\) \( h^{-1} \text{Mpc} \).

Due to the size of spectroscopic fibres used to take the galaxy redshift data, and the limited number of overlapping regions in the survey geometry, a significant fraction of target BOSS galaxies were subject to fibre collisions where only one galaxy in a collided pair was observed (more details in e.g. Anderson et al. 2014, Ross et al. 2017). In our measurements, it was necessary to correct for the missing galaxy pairs that lie within the fibre collision angular radius of 62 arcsec, which corresponds to a projected separation of around \( r_p \sim 0.4 \, h^{-1} \text{Mpc} \) at the CMASS mean redshift and thus has a big impact on small-scale clustering. To estimate the true clustering, we
use the method of White et al. (2011) and upweight the pair counts of any pair found within that separation to account for the missing pairs. The extra weighting factor was determined by finding the ratio of angular clustering measurements $w(\theta)$ between the BOSS galaxies and the photometric SDSS data from the same field. We find we should use a factor of 2.63 to upweight pairs whose galaxies are within 62 arcsec of one another. The corrected number density of the CMASS galaxies, after accounting for the missing galaxies due to fibre collisions and other systematic factors, is shown in Fig. 1.

To make a direct comparison of the galaxies and underlying dark matter distribution with both the $w_p$ and VPF measurements, we restricted the galaxies to the same redshift range as is used for the VPF measurement before counting the pairs. We don’t randomly subsample the galaxies as is done for the VPF sample. A random sample of galaxies will still have the same clustering properties as the full original sample and therefore this extra step would only increase uncertainty in the $w_p$ measurement. The results of this measurement are shown in Fig. 3, along with the measurements of the VPF, described in the next section.

2.4 VPF measurements

The void probability function measurement $P_\theta(r)$ is made by randomly placing spheres of some radius $r$ in many locations of the survey and counting the fraction of empty spheres. Assuming shot noise error in the number of empty spheres found, an appropriate number of spheres must be placed to make a reasonable determination of $P_\theta$. As sphere scale $r$ increases, empty spheres become increasingly rare and more spheres are required to find any signal. Based on the rarity of void spheres above $35\ h^{-1}\ Mpc$ within our survey volume, and the computational cost of searching sufficient spheres at that scale, we have made a VPF measurement for voids in the range $[5, 35]\ h^{-1}\ Mpc$ in steps of $5\ h^{-1}\ Mpc$, using random samples of up to $10^7$ spheres at each radius.

Due to the nature of the survey geometry, which has irregular borders along with missing points in its interior due to stars and other features, not all possible spheres with centres lying within the survey mask will be fully contained within the survey. As such, care must be taken to ensure that a given sphere can be said to truly be a void sphere. Thus, before assessing whether spheres are devoid of galaxies or not, we must first assess whether they are valid survey spheres. In order to be considered as such, we require any sphere that is accounted for in the VPF measurement to meet a completeness threshold, whereby a minimum volume of the sphere is found to be within the survey. We tested the fraction of conserved spheres at each radius for a range of completeness thresholds by randomly proposing spheres within the survey mask and estimating the percentage of the sphere volume that lay outside the mask by using random ‘veto points’ placed at the survey edges and in the interior points that were compromised. The results after averaging the volume over $10^7$ spheres at each radius are shown in Fig. 2. We decide that spheres up to 95 per cent completeness result in a sufficiently large fraction being conserved for the VPF measurement, where over two-thirds of spheres are valid even up to $40\ h^{-1}\ Mpc$. We will describe our method of making proper comparisons between the mocks and data in the next section.

The results for the CMASS northern sample are shown in Fig. 3 along with the $w_p$ measurement for the same galaxies. Errors are described in the following subsection.

### 2.5 Estimating covariances with mock samples

We determine an estimate of the covariance in our measurements by using a set of mock galaxy samples that were made for the purpose of constructing covariance matrices for the CMASS survey data. The mocks used were the quick particle mesh (QPM) mocks (White, Tinker & McBride 2014) created for BOSS. We made the measurements to match the methods described in the previous sections exactly, using the same estimators and bins and radii, and found the full covariance of these measurements over a set of 400 mocks. The covariance matrix we calculate using these mocks is used for making parameter inference of the HOD model, as described below. The diagonal variances of the matrix were used to determine the error bars in the measurements shown in Fig. 3. For the VPF, we show the fractional error bars in log-space, as the range is very broad. At small scales, the VPF does not contain significant information about large-scale structure and the errors are consistent with simple Poisson errors. However, at scales larger than the mean galaxy spacing (roughly $16\ h^{-1}\ Mpc$) the amplitude of the VPF is reflective of the structure, but the fractional error bar is significantly larger. This is both due to the rarity of empty spheres at these scales and sample variance in the clustering of galaxies.

### 3 Galaxy clustering and voids from a halo occupation model

#### 3.1 Dark matter simulation and halo catalogue

To compare the measurements with models of halo occupation, we use a halo catalogue from a high-resolution dark matter simulation. For this work, use the Big MultiDark Planck (BigMDPL) simulation (Prada et al. 2012; Riebe et al. 2013). The simulation has $2.5 h^{-1}\ Gpc$ comoving volume and $3840^3$ dark matter particles, giving it a mass resolution of $2.359 \times 10^{10}\ h^{-1}\ M_\odot$. It uses a Planck cosmology (Planck Collaboration VI 2018) with $h = 0.6777$, $\Omega_m = 0.693$, $\Omega_{\Lambda} = 0.307$, and $n = 0.96$. The simulation was performed using GADGET-2 code (Springel 2005) and further details can be found in Klypin et al. (2016). The Planck cosmological parameters are used throughout this work to remain consistent with the simulation.

The MultiDark Database also provides halo catalogues for each simulation, which we use for our analysis. The catalogues were made using the Rockstar Halo Finder (Behroozi, Wechsler & Wu 2013). We use the host haloes from this catalogue – those not
Figure 3. This figure shows the projected correlation function \( w_p \) with fractional errors and void probability function with fractional errors. The errors of the VPF are shown in log space since they increase effectively exponentially with void radius.

3.2 Halo occupation distribution

To construct model galaxy catalogues from the haloes, we use the Halo Occupation Distribution (HOD) described in Zheng, Coil & Zehavi (2007). This is a statistical prescription for the number of central and satellite galaxies that live in a given halo, depending on the mass of that halo. For centrals, there can be either zero or one galaxy, with the mean central occupation at any mass given by

\[
\langle N_{\text{cen}} \rangle_M = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right],
\]

(2)

where \( M_{\text{min}} \) is the halo mass at which the probability for a galaxy existing at its centre goes from less than a half to greater than a half, and \( \sigma_{\log M} \) effectively determines the scatter in halo mass between haloes with no central galaxies and haloes which all have a central galaxy. If a halo has a central galaxy, drawn from a Bernoulli distribution with probability \( \langle N_{\text{cen}} \rangle \) the galaxy is given the position and velocity of the halo centre of mass.

The average number of satellite galaxies in a host halo of mass \( M \) is given by

\[
\langle N_{\text{sat}} \rangle_M = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right] \left[ \left( \frac{M}{M_1} \right)^\alpha \exp \left( -\frac{M}{M_{\text{cut}}} \right) \right],
\]

(3)

which is a slight modification of the original Zheng formulation. It consists of a power law in halo mass, scaled by \( M_1 \) with slope \( \alpha \) modulated by the central occupation relationship and with an extra exponential cut-off \( M_{\text{cut}} \). The satellite number for each individual halo is drawn from a poisson distribution with mean \( \langle N_{\text{sat}} \rangle \) (only if the halo has a central galaxy) and these galaxies are placed around the halo centre according to an NFW profile with the halo’s concentration. The satellites are given velocities assuming an isotropic distribution of virialized objects in the halo’s gravitational potential.

3.3 Modelling assembly bias

To determine whether there is any environment-dependent bias in the galaxy occupation of dark matter haloes which is independent of the halo mass, we use the dark matter overdensity of host haloes.
We measure the relative density $\rho$ of a halo by counting the number of dark matter particles in a sphere of $15\, h^{-1}\, \text{Mpc}$ located at the halo’s centre of mass and dividing by the average number expected over the whole simulation, for a random subsample of dark matter particles for the simulation. Tinker et al. (2006, 2008b) used a simple model where halo occupation changed by a set amount at densities below some critical density. The results of Tinker et al. (2006) demonstrated that a model of this kind could produce changes to both the $w_p(r_p)$ and VPF, but that the VPF was far more sensitive to these changes and could be detected to high significance even when the change to $w_p(r_p)$ was negligible within the statistical precision of the measurements.

Here, we extend the method of Tinker et al. (2006) to increase the flexibility of the model. We modify the HOD described above by changing the value of $\log M_{\text{min}}$ as a function of $\rho$. For each halo, we add an offset to $\log M_{\text{min}}$ of

$$\Delta M_{\text{min}} = \frac{f_p}{2} \left[ 1 + \text{erf} \left( \frac{\log \rho - \log \rho_0}{\sigma_\rho} \right) \right].$$

(4)

This leads to a change in the minimum halo mass for which a halo will be occupied by central galaxies and also changes the modulation of the satellite galaxy occupation by the same factor. The value for $M_{\text{min}}$ is effectively changed to have two regimes of halo occupation – one value at higher densities, and another value at lower densities, with a difference in $\log M_{\text{min}}$ that is separated by a threshold density $\rho_0$, with a soft transition dictated by $\sigma_\rho$. This model allows us to assess whether the data allow for a mechanism of galaxy formation that prefers galaxies to be formed in haloes of the same mass when they occur in different environments. The choice of parameterization in equation (4) is meant to give maximum flexibility to the model. It allows galaxy formation efficiency to both increase or decrease in voids, it changes the density scale at which this change can happen, and the rapidity with which the change happens along the edge of the void. Fig. 4 shows several examples of how shifts in the mass scale impact clustering. We also compare these results to a model of assembly bias where galaxies are assigned to haloes using the abundance matching approach, with galaxy luminosity matched to the peak maximum circular velocity over the history of each halo (see Wechsler & Tinker 2018 for a review of abundance matching methods). Peak circular velocity is also correlated with large-scale density at fixed halo mass, thus imprints an assembly bias signal on the galaxy population without explicitly using the large-scale density. However, the change to the clustering created by this abundance matching model is easily within the parameter space of equation (4). But we do not wish to restrict ourselves to only the types of assembly bias created in this way – thus, the increased flexibility of equation (4) is more appropriate for the purposes of this study.

3.4 Modelling galaxy clustering

3.4.1 Projected correlation function

Using the halo catalogue and HOD described above, we can simulate a galaxy catalogue which we can use to make model measurements to compare with the measurements made in the data. As described in Section 2.3, the projected correlation function measurement made in the CMASS sample has been corrected for any geometric effects and fibre collisions through the use of a carefully constructed random point distribution over the survey mask and with collision pair-upweighting. Therefore, we need not make any special considerations in our model measurement. We make a distant observer approximation for the simulation box and take one dimension to be the line-of-sight axis. Along this dimension $x_3$, we displace each galaxy according to its parallel velocity
3.4.2 Void probability function

In order to make an honest comparison of the simulated galaxies to the observed galaxies with respect to the VPF, we must first match the galaxy number density. We downsample the simulated galaxies randomly until we reach the same constant number density as we had for the redshift range for the VPF measurement discussed in Section 2.2 and shown in Fig. 1.

Since the measurement of the VPF made in the data was made using a sample of galaxies with missing observations due to fibre collisions, it was necessary to also simulate such fibre collisions in our mock galaxies. Otherwise, downsampling the simulated galaxy sample to our chosen number density would lead to more voids, since there would be no preferential removal of collided pairs, which are typically in denser regions. Removing ‘fibre collisions’ first in our simulated mock leads to a smaller fraction of voids at increasingly larger radii.

The comoving distance from the nearest galaxies to the furthest in our observed sample is approximately 300 $h^{-1}$ Mpc. The depth of the simulation box, along our line-of-sight dimension, is far greater, at 2500 $h^{-1}$ Mpc. The number of galaxy pairs that will seem to be ‘collided’ by the fibre collision radius of our simulation redshift (which comes to 0.4 $h^{-1}$ Mpc in comoving separation) is thus also far greater in the box than in the survey. It becomes an issue to obtain the desired number density in our mock because there can be enough collided galaxies, that removing the same fraction of them as are unobserved in the data, leads to a lower mock number density than the final data sample. As a workaround, we restrict the simulation to a fraction of its size along the line-of-sight dimension. To be as consistent with the survey data as possible, we choose the fraction such that the ratio of simulation box ‘depth’ to ‘area’, is similar to the ratio of survey depth and area. The full survey volume is approximately $10^9 h^{-3}$ Mpc$^3$, so we choose our effective depth for the simulation box to be 600 $h^{-1}$ Mpc and make the measurement in this restricted portion of the mock volume.

Once we have removed collided mock galaxies appropriately and sampled the mock down to match the observed constant number density sample, we measure the VPF in the mock. We make the measurement with $10^7$ spheres at the largest radii, as in the data, and use random points distributed evenly throughout the mock. The
mock is assumed periodic in the transverse dimensions, but voids are kept one sphere radius from the edges in the restricted line-of-sight dimension.

3.5 Fitting HOD parameters to the measurements

To fit the model values for $w_p$ and $P_0(r)$ to the observed values, we use a standard $\chi^2$ statistic. We compute the statistic for samples in the HOD parameter space of $M_{\text{min}}$, $\sigma_{\log M}$, $M_1$, $M_{\text{cut}}$, and $\alpha$ (and including $f_\rho$, $\rho_{\text{th}}$, and $\sigma_\rho$ for the density dependent model). For each sample, we populate the halo catalogue with galaxies accordingly and then compute the simulated measurements. The samples are drawn from a posterior distribution obtained by multiplying the $\chi^2$ values with prior probabilities, chosen to have flat distributions within physically reasonable ranges.

Any implementation of an HOD requires the number density of the sample to be modelled. For the CMASS sample, which has a variable number density, the choice of which density to use is not entirely clear. Here, we use the mean corrected number density within the redshift range of our sample. The sample variance on this value is very small, at the percent level, but we allow the number density of the HOD model to vary by $\pm10$ per cent relative to the mean. This incorporates the possibility that the mean value is not the most appropriate, and marginalizes over other possibilities. For each mock, however, the VPF is always measured after downsampling the mock to the same value as the CMASS sample on which we
make our VPF measurements. The posteriors are sampled using a Markov Chain Monte Carlo (MCMC) sampler based on an affine-invariant search algorithm (Goodman & Weare 2010).

We perform the fit for multiple configurations of the model and measurements. In the first experiment, we fit only to the \(w_p\) measurement, using the standard, non-density dependent HOD, to get a fiducial set of posteriors and compare the VPF distribution of the mocks from the posterior samples to the observed VPF. We make another fit with the standard HOD, fitting to both the \(w_p\) and VPF \(P_0\). Finally we make a fit using the density-dependent HOD, fitting to both the \(w_p\) and \(P_0\).

4 RESULTS

The comparison of posterior predictions for both \(w_p\) and \(P_0\) for the fiducial fit (constrained using only \(w_p\)) are shown in Fig. 5. The fiducial model provides a good fit to the VPF, despite only using the information of the two-point function in the data to constrain the model. This suggests already that there is no need for the inclusion of density dependence in the galaxy halo model to describe galaxy clustering statistics that are sensitive to a range of densities within the cosmic web.

As a confirmation of this lack of density dependence in halo occupation, we show the posterior parameter distributions of a model including density dependence – as described in Section 3 – in Fig. 6. The density parameters, particularly the offset magnitude \(f_\rho\), are all consistent with having no impact on the model. An analogous comparison of the predictions for \(w_p\) and \(P_0\) as in Fig. 5, but made using the HOD including this density dependence, is shown in Fig. 7. Unsurprisingly, there is similarly good correspondence with the data.

A table summarizing the maximum a posteriori values along with the standard deviation of marginalized samples in each parameter, for each of the model cases, is shown in Table 1. This includes the test where we simultaneously fit both the \(w_p\) and \(P_0\) for the density-independent model. As expected from Fig. 5, the parameter constraints are consistent with the model in which only \(w_p\) is fit. The only significant difference is in \(M_{\text{cut}}\), but the difference is only in the numerical values – at such small values of \(M_{\text{cut}}\), the cut-off in the satellite occupation function has no quantitative impact on the clustering because it is so much smaller than \(M_{\text{min}}\). The agreement on the HOD parameters extends to the density-dependent model as well. The best-fitting value of \(f_\rho\) is 0.02 ± 0.04, clearly consistent with zero – i.e. no density dependence. Indeed, values significantly away from zero are strongly excluded by the data.

5 DISCUSSION

We have shown that the information on galaxy bias encoded in the galaxy two-point correlation function is sufficient to model galaxy bias in extremely low-density environments. Thus, halo occupation does not change from high to low densities. Additionally,
incorporating a flexible model for galaxy assembly bias – a model in which halo occupation depends on both $M_h$ and large-scale density – only fits the combined $w_p$ and VPF data when the density-dependent parameters are consistent with zero. Studies meant to extract cosmological information from non-linear clustering of large-scale galaxies redshift surveys, like BOSS, its successor eBOSS (Dawson et al. 2015) and the near-term DESI survey (DESI Collaboration et al. 2016), can expect minimal-to-no degeneracies between cosmology and assembly bias.

This result appears to be in tension with those of Saito et al. (2016), where measurements of $w_p$ are better fit using an age-matching model for galaxy occupation in a halo-abundance matching context. However, the results of Saito et al. (2016) are mostly inferred from redshift-dependent effects in their model, which we do not encounter, given the relatively narrow redshift range of this work. Another difference is our implementations of assembly bias, which in our model is done explicitly through a dependence on $\rho$, and in theirs done implicitly through abundance matching. However, the results of Saito et al. (2016) show no detection of assembly bias for the brightest samples that they examine: $M_h - 5\log h < -21$ and $-22$, respectively. The CMASS sample is not immediately comparable to a volume-limited, complete SDSS sample, but the number density of the CMASS sample is roughly consistent with the number density of the $M_h - 5\log h < -21.5$ sample, given some groundings of the halo mass scales involved in each sample.

Additionally, the CMASS sample, as well as the brighter SDSS samples, are comprised almost exclusively of red-and-dead, passive galaxies. Many studies of colour-dependent clustering have shown that the quenching process cannot correlate with halo formation history (Tinker et al. 2008b, 2017, 2018; Sin, Lilly & Henriques 2017; Zu & Mandelbaum 2018). Additionally, the small observed scatter between halo mass and stellar mass for massive galaxies is most easily explained by a quenching process that depends exclusively on the galaxy stellar mass (Tinker 2017). Thus, galaxy quenching may erase any correlation between galaxy mass and halo formation history that may exist for star-forming galaxies which gave rise to the positive detections of galaxy assembly bias for lower luminosity SDSS galaxies.

Thus, the luminous red galaxy class of targets may be the best opportunity to obtain cosmological constraints from non-linear galaxy clustering.

### ACKNOWLEDGEMENTS

Funding for SDSS-III has been provided by the Alfred P Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is [https://www.sdss3.org](https://www.sdss3.org).

### REFERENCES

Anderson L. et al., 2014, MNRAS, 441, 24
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Betancort-Rijo J., Patiri S. G., Prada F., Romano A. E., 2009, MNRAS, 400, 1835
Beygu B., Peletier R. F., van der Hulst J. M., Jarrett T. H., Kreckel K., van de Weygaert R., van Gorkom J. H., Aragon-Calvo M. A., 2017, MNRAS, 464, 666
Clampitt J., Jain B., Sánchez C., 2016, MNRAS, 456, 4425
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coupon J. et al., 2015, MNRAS, 449, 1352
Croton D. J., Farrar G. R., 2008, MNRAS, 386, 2285
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2015, AJ, 151, 34
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Einasto J. et al., 2011, A&A, 534, A128
