

# A Halo Occupation Interpretation Of Quasars At $z \sim 1.55$ Using Very Small Scale Clustering Information

S. Eftekharzadeh<sup>1,2</sup>, A. D. Myers<sup>1</sup>, E. Kourkchi<sup>3</sup>, et al.

<sup>1</sup> Department of Physics and Astronomy, University of Wyoming, 1000 University Ave., Laramie, WY 82071

<sup>2</sup> Department of Physics, Southern Methodist University, 3215 Daniel Ave., Dallas, TX 75225

<sup>3</sup> Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

4 October 2018

## ABSTRACT

We couple the most accurate small scale ( $< 100 \text{ h}^{-1}\text{kpc}$ ) quasar clustering to date with its most recent counterpart at large scale ( $> 1 \text{ h}^{-1}\text{Mpc}$ ) from the extended Baryon Oscillation Spectroscopic survey (eBOSS) in an attempt to better constrain the satellite fraction of quasars at  $z \sim 1.55$  as one of the more controversial parameters of structure formation hierarchy in the halo occupation framework. We built our Halo Occupation Distribution (HOD) framework based on commonly used analytic forms for the one and two halo terms with the three free parameters of minimum halo mass that host a central quasar, satellite fraction, and the width of the distribution for the dark matter halo masses. Inspired by the latest studies that proposed a steeper number density profile for the dark matter halos to recover the observed excess clustering at kpc scales, we explored the change in the model and the best fit parameters for a range of concentrations. We found that an HOD model with a satellite fraction of  $f_{\text{sat}} = 0.016^{+0.004}_{-0.005}$  and minimum mass of  $M_m = 1.45^{+0.32}_{-0.25} \times 10^{12} \text{ h}^{-1}\text{M}_\odot$  for the host dark matter halos would best describe the *full* quasar clustering at  $z \sim 1.55$ . Our smaller satellite fraction is consistent with recent reported value for quasars at  $z \sim 1.4$  that uses the same HOD formalism. Our findings could be interpreted as an evidence that the commonly used parameterization for the mean number of quasars in each halo is unable to further characterize the halo occupation distribution of quasars.

**Key words:** cosmology: observations, small-scale clustering, halo occupation; quasars: general, surveys, close pairs

## 1 INTRODUCTION

The advent of the first large, homogeneous surveys of the extragalactic sky prompted cosmologists to begin to think about galaxies as discrete points embedded in statistical structures (e.g. Neyman, Scott & Zonn 1962; Neyman & Scott 1974). These structures could be characterized in terms of their size, the distribution of the discrete points that occupied them, and their distribution across the wider Universe (see, e.g. Cooray & Sheth 2002, for a review). As it became increasingly clear that dark matter dominated the mass budget in galaxies (e.g. Rood et al. 1972; Ostriker, Peebles & Yahil 1974; Rubin, Thonnard & Ford 1978), it became more natural to think of occupation statistics in terms of the virialized halos that host luminous tracers (e.g. White & Rees 1978). This line of reasoning ultimately led to empirical approaches that describe how cosmological tracers occupy underlying dark matter structures, such as the Halo Occupation Distribution framework (HOD; e.g., Berlind & Weinberg 2002; Zheng, Coil & Zehavi 2007, and references therein). Parameterizations of the HOD typically consist of a two-halo term, which characterizes how halos of a certain mass cluster around each other, and a one-halo term that relates

galaxy and dark matter statistics through the probability that a halo of a given mass contains a number of galaxies of a given type.

The key observables that are used for constraining HOD descriptions in the common formalism are the number density and the clustering of a given tracer population. As such, the HOD framework has now been successfully used to model galaxy clustering measurements for a wide range of redshifts and galaxy types (e.g., Folkes et al. 1999; Dawson et al. 2016) Quasars, the most luminous of the Active Galactic Nuclei, are driven by accreting supermassive black holes at the centers of galaxies. It is now well-established that the centers of most galaxies contain a supermassive black hole (e.g. Kormendy & Richstone 1995), and that the evolution of active quasars and inactive galaxies is interrelated (e.g. Kauffmann & Haehnelt 2000). It is therefore reasonable to think of quasars simply as a biased tracer of certain types of galaxies that should also, in theory, be empirically describable using HOD statistics.

In the wake of large spectroscopic surveys of quasars such as the 2dF (Folkes et al. 1999) and SDSS (e.g., Fukugita et al. 1996; Gunn et al. 1998; York et al. 2000; Vanden Berk et al. 2001; Stoughton et al. 2002; Strauss et al. 2002; Tegmark et al. 2004;

West et al. 2004; Yip et al. 2004a,b; Bolton et al. 2004; Rojas et al. 2005; Wilhite et al. 2005), the clustering of quasars has been measured at a range of redshifts (e.g., Croom et al. 2004; Porciani, Magliocchetti & Norberg 2004; Myers et al. 2006, 2007a,b; Shen et al. 2007, 2009; Ross et al. 2009; Richardson et al. 2012; White et al. 2012; Richardson et al. 2013; Ross et al. 2013; Eftekhazadeh et al. 2015; Laurent et al. 2017). Typically, these studies focus on the large-scale clustering of quasars via the two-point correlation function, which constrains the “two-halo term” that describes how “central” dark matter halos cluster around each other. The consensus is that, at most redshifts, quasars occupy central halos of masses a few times  $10^{12} h^{-1} M_\odot$ .

Probing how quasars are distributed within halos—the so-called “one-halo” term—is trickier, however. As quasars occupy massive halos, they are rare in general. Further, the halos that host quasars, particularly at high redshift (e.g., White et al. 2012; Eftekhazadeh et al. 2015), are on the steeply falling part of the halo mass function. This implies that instances of two quasars occupying a single halo at high redshift may be very rare indeed. Finding close pairs of quasars is complicated further by the fact that most large spectroscopic surveys use fiber-fed multi-object spectrographs. Such surveys can have restrictions on how closely fibers can be placed together on the sky, which has prompted follow-up surveys of close quasar pairs using long-slit spectrographs (e.g., Hennawi et al. 2006; Myers et al. 2008; Hennawi et al. 2010; Kayo & Oguri 2012; Eftekhazadeh et al. 2017). These long-slit surveys, in combination with the large-scale two-point correlation function, have been used to constrain the one-halo term for quasars via clustering measurements over a wide redshift range of redshift ( $z \sim 0.5\text{--}3$ ) and scale (Richardson et al. 2012; Kayo & Oguri 2012; Shen et al. 2013).

Recent measurements of quasar clustering have used a number of different assumptions for the overall form of the quasar HOD. For instance, Kayo & Oguri (2012) model both the distribution of satellite and central quasars using Gaussians, ultimately expressing the HOD using two-to-three fitting parameters. On the other hand, Zheng, Coil & Zehavi (2007); Richardson et al. (2012, 2013); Shen et al. (2013) use a model that combines a power-law with a Gaussian, which requires five-to-six fitting parameters. These choices, however, seem to have little power to constrain the one-halo term of the quasar HOD as (Kayo & Oguri 2012) and (Shen et al. 2013) derive consistent satellite fractions using their two different HOD parameterizations. Much of this degeneracy regarding the form of the quasar HOD is driven by sizable uncertainties in the parameters that are fit to model quasar clustering on small scales. It is likely, then, that much larger samples of quasar pairs with small separations, or alternative approaches to deriving the Mean Occupation Function of quasars (e.g. Chatterjee et al. 2013), will be needed to probe the overall statistics of how quasars occupy individual halos.

Despite the range of possible forms for the Mean Occupation Function of quasars, it remains important to provide empirical constraints on the quasar HOD. Large surveys such as the *extended Baryon Oscillation Spectroscopic Survey* (Myers et al. 2015; Dawson et al. 2016) and surveys with the *Dark Energy Spectroscopic Instrument* (DESI Collaboration et al. 2016a,b) are beginning to use quasars to constrain the cosmological world model at moderate redshift via redshift-space distortions and the Baryon Acoustic Oscillation scale. Sophisticated simulations are required to model these cosmological constraints, which require an assumed form for the quasar HOD on small scales (e.g. Rodríguez-Torres et al. 2017). Recently, we assembled by far the largest sample of quasar pairs that can be used to probe quasar clustering on scales of a few dozen

kiloparsecs, well into the one-halo regime (Eftekhazadeh et al. 2017). In this paper, we combine the sample of Eftekhazadeh et al. (2017) with other clustering results on larger scales (e.g. Kayo & Oguri 2012; Laurent et al. 2017) to provide the best current constraints on the quasar HOD.

This paper is structured as follows: §2 summarizes the properties of the samples that are used in small and large scale clustering measurements. The measurements themselves are detailed in §3, the modeling approach and the chosen parameterization is described in §4, and

We adopt a  $\Lambda$ CDM cosmological model with matter and dark energy and baryon density densities of  $\Omega_m = 0.308$ ,  $\Omega_\Lambda = 0.693$ , and  $\Omega_b = 0.045$  the Hubble parameter  $h = 0.678$ , amplitude of matter fluctuations  $\sigma_8 = 0.814$ , and the slope of the initial power spectrum  $n_s = 0.968$  consistent with Planck Collaboration et al. (2015). All distances quoted throughout the paper are in comoving coordinates unless noted otherwise, denoting the proper coordinates with addition of “p” to the distance units (i.e.,  $h^{-1}$ pkpc or  $h^{-1}$ pMpc).

## 2 DATA

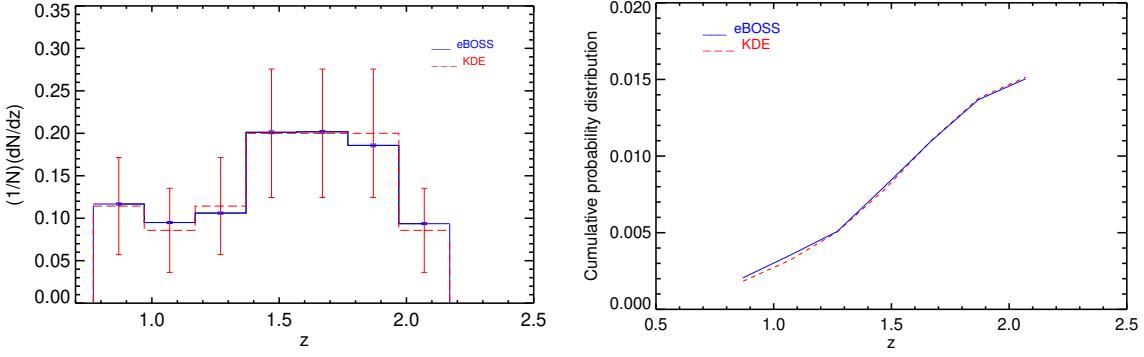
We used two independently compiled samples of confirmed quasars for clustering measurements at kpc and Mpc scales. The small scale clustering measurement used in this work is drawn from a complete sample of  $g < 20.85$  confirmed quasars at  $z \sim 1.5$  and the large scale clustering measurement is made using a sample of  $g < 20.85$  quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) at similar average redshift. We applied a number of adjustments prior to measuring the two-point correlation function and combining these two measurements into one that spans over comoving distances of  $\sim 0.01 - 100 h^{-1}$ Mpc to be used in our halo occupation interpretation. In the following, we summarize the compilation process and properties of the two samples.

### 2.1 KDE-complete sample of close pairs and the associated random catalog

The basis for the follow up spectroscopy that led to the complete sample of 47 spectroscopically confirmed “binaries” with the angular separation of  $2.9'' \leq \theta \leq 7.7''$  at  $0.43 \leq z \leq 2.2$  through a carefully designed long-slit spectroscopy campaign, is a sample of 1,172,157 high probability quasar candidate drawn via a Kernel Density Estimation technique (KDE; Richards et al. 2004) on all the point sources in SDSS-DR6 (Adelman-McCarthy et al. 2008) imaging data down to  $i = 21.3$ . A total of 290,694 quasar candidates at  $0.43 \leq z \leq 2.2$  were collected via a non-parametric Bayesian classifier with  $\sim 92.7\%$  efficiency (see Eftekhazadeh et al. 2017, and references therein for detail).

Out of the 47 pairs in the KDE-complete sample (presented in Figure 4 and Table 5 of Eftekhazadeh et al. 2017), 44 reside in the selected redshift range for this study ( $0.8 \lesssim z \lesssim 2.2$  with  $\bar{z} \sim 1.5$ ). The comoving separations of quasars in these pairs span over  $43.3 \lesssim r_p \lesssim 92.3 h^{-1}$ kpc. This sample constitutes the basis for our sub-Mpc clustering measurement. We later discuss combining our small-scale measurement with a previous measurement at a similar average redshift (see §3) by Kayo & Oguri (2012) to fill-in the gap between the last bin of small-scale (at  $\sim 85 h^{-1}$ kpc) of the measured correlation function at small and large scales.

Following the same procedure for calculating the number of



**Figure 1.** Right: Normalized redshift distribution of the KDE-complete sample of ( $< 7''$ ) pairs in dashed red and the “downsampled” redshift distribution of the eBOSS quasars in solid blue. The “down sampling” process was performed due to the need to match the redshift distribution of the eBOSS quasars to be able to use their co-joint clustering measurements for a halo occupation interpretation (see §2). Left: the cumulative distribution function (CDF) of KDE and eBOSS samples. The Kolmogorov-Smirnov test shows that the probability that these two samples are drawn from the same distribution is 86%.

$z_{min}$	$z_{mid}$	$z_{max}$	$(1/N_{eBOSS}) dN/dz$	Error	$(1/N_{KDE}) dN/dz$	Error
0.77	0.87	0.97	0.117	0.001	0.114	0.057
0.97	1.07	1.17	0.095	0.001	0.086	0.049
1.17	1.27	1.37	0.106	0.001	0.114	0.057
1.37	1.47	1.57	0.201	0.002	0.200	0.076
1.58	1.67	1.77	0.202	0.002	0.200	0.076
1.77	1.87	1.97	0.186	0.002	0.200	0.076
1.97	2.07	2.17	0.094	0.001	0.086	0.049

**Table 1.** Normalized distribution of the spectroscopic redshifts of quasars in eBOSS (4th column) and the KDE-complete sample of close pairs (6th column) in the bins of redshifts depicted in Fig. 1.

“expected” quasar-random pairs ( $\langle QR \rangle = N_Q/N_R \times QR$ ) we construct the random catalog by populating a circular polygon with angular radius of  $7.7''$  around each of the 243,110 of quasar candidates (see section 3.1 of Eftekharzadeh et al. 2017, for a detailed description of the taken steps).

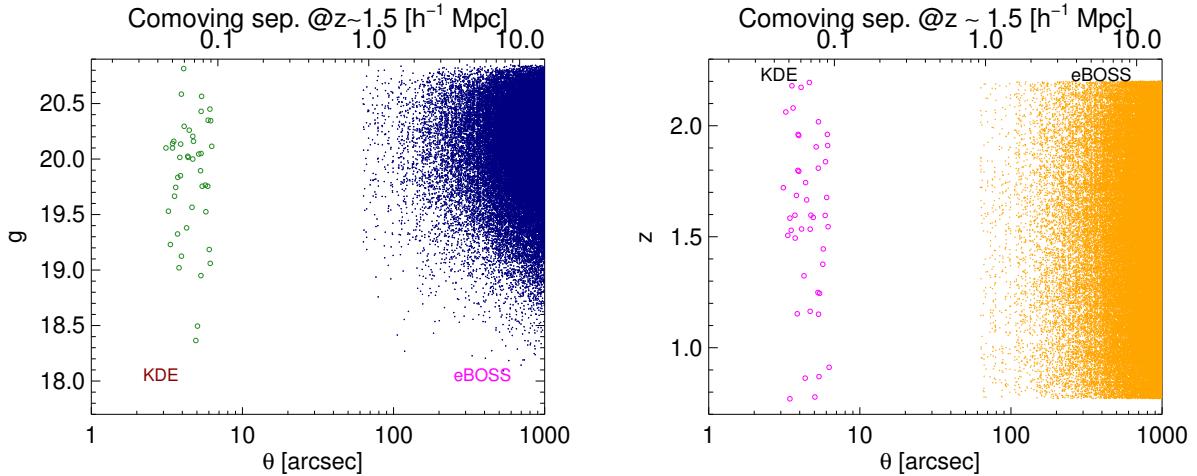
## 2.2 eBOSS quasars with $g < 20.85$ and the associated random catalog

For Mpc-scale clustering measurement, we used a large sample of spectroscopically confirmed quasars from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) that aims to extend BAO studies to higher redshifts by measuring the redshift for about 0.5 million quasars at  $0.8 \lesssim z \lesssim 2.2$  as well as Ly $\alpha$  forest quasars at  $z > 2.1$ , luminous red galaxies at  $0.6 < z < 1.0$ , emission line galaxies at  $z > 0.6$  (see Myers et al. 2015, for target selection criteria and process). In order to reduce the risk of missing targets by the target selection algorithm, eBOSS corrects for variance of the  $5 - \sigma$  detection limit for the photometrically identified point sources across the survey footprint and for different bands. The correction includes applying a depth-dependence systematic weight,  $w_{sys}$  : WEIGHT\_SYSTOT, to each quasar as well as considering redshift failure ( $w_{zf}$  : WEIGHT\_NOZ) and fiber collision ( $w_{cp}$  : WEIGHT\_CP) weights (see, Laurent et al. 2017; Rodríguez-Torres et al. 2017; Bianchi & Percival 2017; Anderson et al. 2012; Ross et al. 2012, for detail of the weight determination). The total weight for each quasar is therefore defined as

$w_{qso} = w_{fkp}w_{sys}(w_{zf} + w_{cp} - 1)$  where  $w_{fkp}$  : (WEIGHT\_FKP) is the applied weight for optimum estimation of the two point correlation function (see, Feldman, Kaiser & Peacock 1994).

As eBOSS utilizes the same fiber-fed optical spectrograph as BOSS, the fibers can not get closer than  $62''$  in the spectroscopic plates except for the overlapping regions (Blanton et al. 2003). This angular separation limit is equivalent to the comoving separation of  $0.37 h^{-1}\text{Mpc}$  at  $z \sim 1.5$ . For clustering measurement at small scales, it is more important to be complete as the interplay between the fiber collision radius and the *unknown* small-scale clustering of the targets is impossible to perfectly reconstruct. Although quasar samples are sparse enough that the number of collided pairs is very small (e.g., Rodríguez-Torres et al. 2017), we set the first bin of comoving separation at which we started to measure the two-point correlation function (2PCF), to be at  $1.1 h^{-1}\text{Mpc}$  (see, e. g., Guo, Zehavi & Zheng 2012; Hahn et al. 2017, for studies on the efficiency of fiber collision corrections at small scales).

We started with a set of 116,866 eBOSS quasars (now publicly available at DR14 release Abolfathi et al. (2018)) and its associated  $\sim 44\times$  larger random catalog that considers the variation of the completeness pattern across the survey footprint and mocks the number density and redshift distribution of the targets in building units of the Swanson et al. (see 2008); Laurent et al. (see 2017). A total of 88,764 of these quasars fall in the redshift limits of  $0.8 \lesssim z \lesssim 2.2$  from which 40,821 also have  $g < 20.85$ . This is the brightness lower limit for the quasar pairs in the KDE-complete sample at small-scale. As implied earlier, maintaining similar red-



**Figure 2.** Angular separation of quasars in each of the pairs that are used for the small and large-scale clustering measurement as a function of their g-magnitudes (left) and their redshifts (right). Blue dots and green circles in the left and orange dots and pink circles in right each represent one pair of two confirmed quasars. The open circles in both panels represent the 40 ( $< 7''$ ) pairs drawn from a KDE-complete ‘parent sample’ of 243,110 quasar candidates (see Eftekhari et al. (2017) for detail). The blue and orange dots in left and right panels are the quasar pairs in a sample of 39,083  $g \leq 20.85$  quasars from eBOSS. For consistency, we applied the same magnitude and redshift limits to both samples. As is shown, both set of pairs from the small and large angular separations exhibit similar distributions in magnitude and redshift. The angular separations of the pair members are translated into the comoving transverse separation at the average redshift of both samples ( $z \sim 1.55$ ) and shown on the top axis of both panels. Fiber collision of multi-fiber spectrographs in eBOSS (and BOSS), imposes a lower limit of  $62''$  for the angular separation of the quasar pairs, causing a  $\sim 55''$  gap between the close pairs from the KDE-complete sample at  $7''$  and the eBOSS pairs.

shift and luminosity range for quasars in both samples would guarantee that our clustering measurement is measured within one class of quasars (see, e.g., Shen et al. 2007; Allevato et al. 2011; Shen et al. 2013; Allevato et al. 2014; Eftekhari et al. 2015; McGreer et al. 2016, for discussions on whether quasars from different luminosity classes posses different clustering properties).

An important factor in combining the clustering measurement from two independently compiled samples is to ascertain that they have similar redshift distributions and span over the same range of luminosities. We match the distributions by *downsampling* the much larger eBOSS quasar sample in over populated redshift bins to that of the KDE-complete sample of close pairs by first determining the over population at each bin of redshift and then randomly removing the *extra* number of targets to be removed from that redshift bin. Fig. 1 shows the normalized redshift distribution of the KDE-complete sample of ( $< 7''$ ) pairs and the ‘downsampled’ and matched redshift distribution of the eBOSS quasars. The left-hand-side of this figure illustrates the cumulative probability distribution function (CDF) of KDE and eBOSS samples. The two-sample Kolmogorov-Smirnov test shows that the probability that these two samples are drawn from the same distribution is 86%. The errors on the redshift distribution of the KDE sample are Poisson error only to show that the match is well within the one-sigma uncertainty.

Table 1 lists the matched redshift distributions of KDE-complete sample of 40 pairs and 33,245 eBOSS  $g < 20.85$  quasars.

The left and right panels of Figure 2 show the angular separation of quasars in each of the pairs that are used for the small and large-scale clustering measurement as a function of their g-magnitudes and their redshifts respectively. Blue dots and green circles in the left and orange dots and pink circles in right each represent one pair of two confirmed quasars. The open circles in both

panels represent the 40 ( $< 7''$ ) pairs drawn from a KDE-complete ‘parent sample’ of 243,110 quasar candidates (see Eftekhari et al. (2017) for detail). The blue and orange dots in left and right panels are the quasar pairs in a sample of 39,083  $g \leq 20.85$  quasars from eBOSS. All the targets used in this sample are publicly available as part of the 14th data release of the Sloan Digital Sky Survey (DR14; Abolfathi et al. 2018). For consistency, we applied the same magnitude and redshift limits to both samples. As is shown, both set of pairs from the small and large angular separations exhibit similar distributions in magnitude and redshift. The angular separations of the pair members are translated into the comoving transverse separation at the average redshift of both samples ( $z \sim 1.5$ ) and shown on the top axis of both panels. Fiber collision of multi-fiber spectrographs in eBOSS (and BOSS), imposes a lower limit of  $62''$  for the angular separation of the quasar pairs, causing a  $\sim 55''$  gap between the close pairs from the KDE-complete sample at  $7''$  and the eBOSS pairs.

### 3 CLUSTERING MEASUREMENT

Given the considerations discussed in §2, the final samples became suitable for covering the full range of comoving separations for measuring the real-space correlation function of quasars at  $z \sim 1.55$ . Here we summarize the steps taken in clustering measurements at kpc and Mpc-scales.

#### 3.1 kpc-scale clustering

The adapted method for measuring the volume-averaged correlation function in real space with sparse samples of close pairs is discussed in Eftekhari et al. (2017) and Hennawi et al. (2006).

Fig. 2 of Eftekharzadeh et al. (2017) shows how the sample of 47 pairs with angular separations of  $< 7''$  constitute the complete sample and how they are positioned in the redshift-physical separation plane. We apply the redshift cut of  $0.8 \lesssim z \lesssim 2.2$  that limits the eBOSS sample and count the remaining pairs in the smaller box and still in the same four bins of separation but now in comoving coordinates. This constitutes the origin of four bins of separation distance for the 44 pairs in the measured correlation function shown in Fig. 3. Multiple studies have argued the independency of the pair counts in small scales. We thus adopt a Poisson error statistics from ? for the uncertainty of the measured  $\bar{W}_p$  and later translate them to errors for  $w_p(r_p)$ . The volume averaged correlation function was measured for the 6, 13, 12 and 13 pairs at four bins of separation at the comoving average separation of 48, 58, 70 and 85  $\text{h}^{-1}\text{kpc}$  respectively. Similarly to the measurement presented in Eftekharzadeh et al. (2017), we also measure  $\bar{W}_p$  for the full bin of 43.4 to 92.3  $\text{h}^{-1}\text{kpc}$  (shown as single bin at  $\bar{r}_p = 67.8 \text{ h}^{-1}\text{kpc}$  in the right-hand-side panel of Fig. 3). The process through which the *expected* Quasar-Random pairs,  $\langle QR \rangle$ , is calculated has been discussed in detail in Eftekharzadeh et al. (2017).

While the reason to divert from conventional definition of the projected correlation function (i.e.,  $w_p(r_p) = 2 \int_0^\infty d\pi \xi(\pi, r_p)$ ), to a volume-averaged correlation function,  $\bar{W}_p$ , at small scales is the scarcity of the samples of complete and confirmed close pairs, converting  $\bar{W}_p$  to their equivalent  $w_p(r_p)$  makes it easier to compare the measurement to the HOD model. We convert  $\bar{W}_p$  to their equivalent  $w_p(r_p)$  using the approximation:

$$\bar{W}_p(r_p) \sim \frac{1}{N_{qso}} \int_{z_{min}}^{z_{max}} dz \frac{dV_c}{dz} n(z) \frac{1}{v_z} \int_0^{v_z} d\pi \xi(\sqrt{(r_p^2 + \pi^2)}, z), \quad (1)$$

where  $n(z)$  is the comoving number density of quasars at bins of redshift,  $v_z \equiv v_{max}(1+z)/H(z)$ ,  $v_{max} = 2000 \text{ s}^{-1}\text{km}$ ,  $H(z)$  is the expansion rate at redshift  $z$ ,  $N_{qso} \sim \int_{z_{min}}^{z_{max}} dz \frac{dV_c}{dz} n(z)$  and  $\int_0^{v_z} d\pi \xi(\sqrt{(r_p^2 + \pi^2)}, z)$  is essentially  $w_p(r_p, v_z \rightarrow \infty)$ . Described by numerous clustering works, the correlation function  $\xi(r)$  is best described by a two-parameter power-law  $(\frac{r}{r_0})^{-\gamma}$  with  $\gamma \sim -2.0$  and  $r_0 = 5.0 \text{ h}^{-1}\text{Mpc}$  from Eftekharzadeh et al. (2017). Similar approximation with different numerical approach has been used in Kayo & Oguri (2012) for  $\bar{W}_p(r_p) \rightarrow w_p(r_p)$  conversion.

### 3.2 Mpc-scale clustering

We calculate the Mpc-scale section of the correlation function using the downsampled eBOSS quasars that matches the number density of the parent sample with which the kpc-scale correlation function is measured. The corresponding random catalog for the “downsampled” quasar catalog is created by keeping track of the fraction of objects that are brighter than  $g=20.85$  in each bin of redshift ( $f$ ) and then for each object in the original random catalog and its assigned redshift, we pick a random number between 0 and 1 and if that random number is less than  $f$ , then that random object retain in the random catalog otherwise, discarded. We calculate  $w_p(r_p)$  via its conventional definition by Landy & Szalay (1993). As discussed in §2.2, the quasars participating in the Quasar-Random and Quasar-Quasar pair counts, will have the associated weight  $w_{qso}$ . the measured correlation function in bins of comoving separation across  $1 < r_p < 100 \text{ h}^{-1}\text{Mpc}$  are shown with filled black circles in Fig. 3 and Fig. 5. The error bars are calculated through a jackknife resampling (see, e.g., eqn. 4 of Eftekharzadeh et al. 2015).

### 4 HOD MODELLING

One important application of measuring the real-space correlation function of a given population is to provide the statistics of how those objects populate individual dark matter haloes. Halo Occupation Distribution framework, provides the number of residing objects within one halo as well as an analytical form for their distribution in each halo under the assumption that the number of objects that reside in individual haloes is a function of the halo mass ( $\langle N(M) \rangle$ ) (Peacock & Smith 2000; Seljak 2000; Scoccimarro et al. 2001). Different incarnations of this modelling approach have been used for interpreting the measured correlation function of AGNs and quasars in recent years (see, e.g., Porciani, Magliocchetti & Norberg 2004; Coil et al. 2004; Abazajian et al. 2005; Coil et al. 2006, 2007, 2009; Miyaji et al. 2011; Richardson et al. 2012; Kayo & Oguri 2012; Krumpe et al. 2012; Richardson et al. 2013; Shen et al. 2013; Coil et al. 2016, 2017).

Although a number of studies have attempted to constrain the HOD parameters both in large and small scales, only a few had access to observed correlation function in halo scales ( $< 1 \text{ h}^{-1}\text{Mpc}$ ). We chose a modelling approach similar to a recent study with an observed correlation function over a similar scale and redshift regime (Kayo & Oguri 2012).

The projected two-point correlation function can be modeled using the matter power spectrum (see, e.g., Cooray & Sheth 2002):

$$w_p(r_p) = \int_0^\infty \frac{k dk}{2\pi} P(k) J_0(kr_p) \quad (2)$$

where  $J_0$  is the zeroth order of the Bessel function of the first kind. The power spectrum  $P(k)$  can be separated into two independent *one* and *two-halo* terms:

$$P(k) = P_{1h}(k) + P_{2h}(k) \quad (3)$$

The two halo term can be simply obtained by

$$P_{2h}(k) = b^2 P_{lin}(k) \quad (4)$$

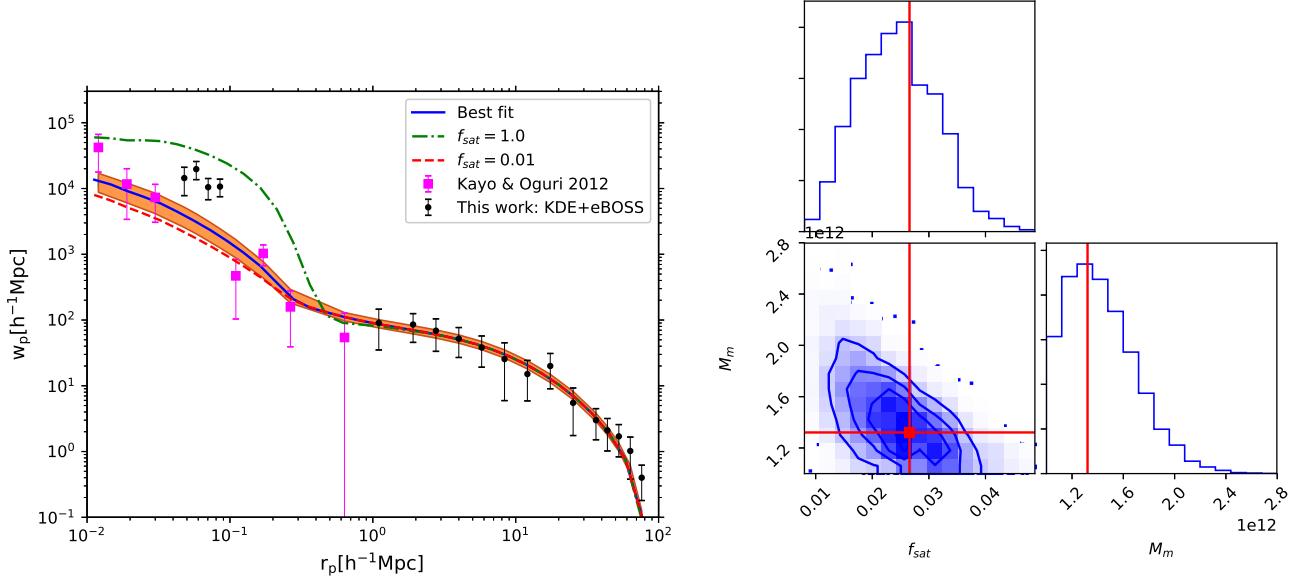
Where the linear matter power spectrum,  $P_{lin}$  is computed using Eisenstein & Hu (1999)’s fitting form with our chosen cosmological parameters and  $b$  is the bias parameter that can be modeled as:

$$b = \frac{\int b_h(M) \frac{dn}{dM} dM}{\int \frac{dn}{dM} dM} \quad (5)$$

Modeling the Halo Occupation Distribution is essentially estimating the mean number of quasars in individual haloes by modeling their distribution as a function of a number physical parameters that characterize the interplay between virialized dark matter overdensities and the embedded baryonic matter within them. Following a commonly assumed mass-dependence, we model the mean number of occupying quasars in individual haloes simply by a three parameter Gaussian form:

$$\langle N(M) \rangle = f_N \times \frac{1}{\sqrt{(2\pi)\Delta_m}} \exp\left[-\frac{\ln^2(M/M_m)}{2\Delta_m^2}\right], \quad (6)$$

where  $M_m$  is the minimum mass for a dark matter halo that can host a quasar and  $\Delta_m$  is the width of the initial mass distribution.  $f_N$  is the normalization factor that matches the observed number density of the quasars in our sample ( $\sim 6.462 \times 10^{-6} \text{ h}^3\text{Mpc}^{-3}$ ) with its expected value from the halo mass function and the assumed distribution of haloes. In the one-halo term of the power spectrum where the correlation function is measured with the pairs of haloes within boundaries of individual



**Figure 3.** Left: Our HOD prediction of the projected correlation function that best fits the observational data. The solid blue curve is the best fit and the red dashed and green dotted-dashed lines are the model with best fit  $M_m$  but  $f_{sat} = 0.01$  and  $f_{sat} = 1.0$ . We verified that relaxing  $\Delta_m$ , as the third parameter, has no significant effect on the shape of the best-fit model. The best fit value for  $\Delta_m$  stays consistent with 0.75 that was derived independently by Kayo & Oguri (2012). The shaded envelope around the best fit model is the extent of the 2-sigma certainty for the shape of the HOD model at each bin of the projected distance ( $r_p$ ). All the data points depicted here (filled pink squares from Kayo & Oguri (2012) and filled black circles) have participated in the fit. See §2 for a description of the clustering measurement from two separate sets of observations at  $z \sim 1.55$  at small (kpc) and large (Mpc) scales. Best fit model has  $M_m = 1.43^{+0.31}_{-0.23} \times 10^{12} h^{-1} M_\odot$ , and  $f_{sat} = 0.0244^{+0.0075}_{-0.0070}$  with reduced  $\chi^2 = 1.59$ . Right: The result of the 2-parameter Monte-Carlo fit to the projected two point correlation function over  $\sim 0.01 - 100 h^{-1}\text{Mpc}$  for the model shown in the left. Diagonal panels show the probability distribution functions (PDFs) of the fitting parameters (i.e., satellite fraction ( $f_{sat}$ ), minimum halo mass that hosts a quasar ( $M_{min}$ ), and the range of the halo masses that quasars occupy ( $\Delta_m$ )). The off-diagonal panels show the density contours of the chosen sets of HOD parameters within one to three-sigma certainties. As the Markov Chain explores the parameter space for each of the fitting parameters, each step of the chain takes a random walk with a randomly chosen step size and calculates the likelihood of the new set of parameters based on the difference between their deduced chi-squared and that of the previous set of chosen parameters.

haloes themselves, there needs to be assumptions for the fraction of quasars that exist as “satellite”s to a central quasar ( $f_{sat}$ ) and the fraction of quasars that reside at the center of dark matter haloes ( $p$ ). The halo occupation model then has “central” and “satellite” components such that  $\langle N_{cen}(M) \rangle = (1 - f_{sat})\langle N(M) \rangle$  and  $\langle N_{sat}(M) \rangle = f_{sat}\langle N(M) \rangle$  (Berlind & Weinberg 2002; Kravtsov et al. 2004; Zheng et al. 2005). Kayo & Oguri (2012) justifies ignoring the mass dependence of the satellite fraction (i.e.,  $f_{sat}(M) \simeq f_{sat}$ ) by choosing a relatively narrow ( $\sim 2$  dex) range for the halo mass distribution around the characterized minimum mass,  $M_m$ . Similar to their approach, we assume a constant satellite fraction and investigate any difference by relaxing the width of the halo mass distribution,  $\Delta_m$ , as an additional fitting parameter in our HOD model.

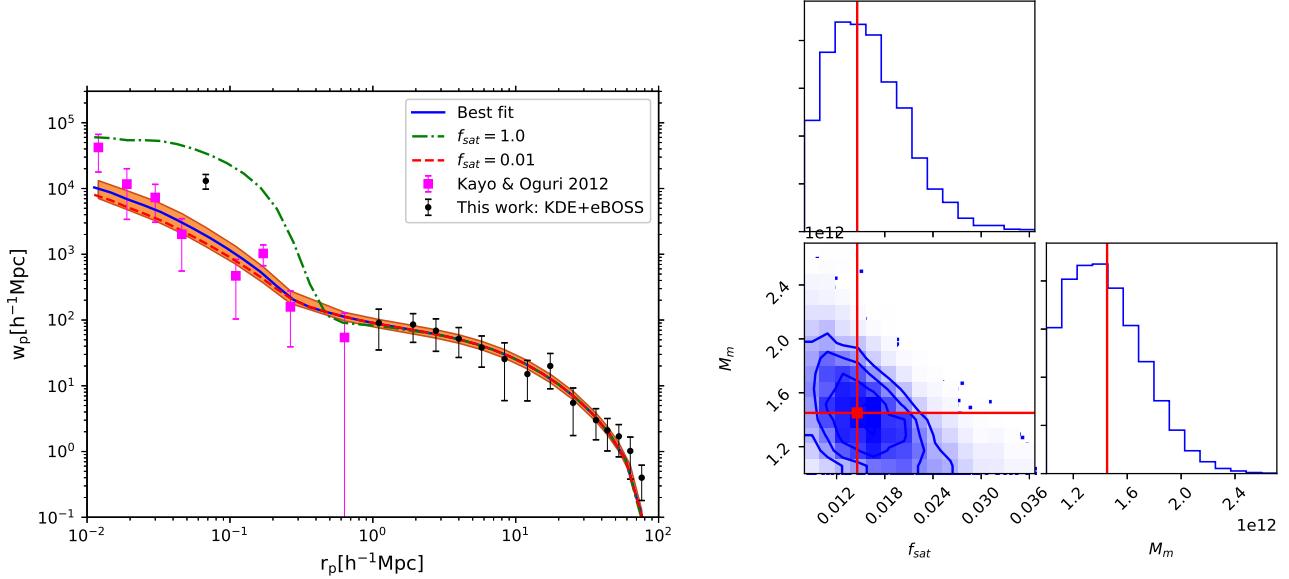
We further adopt the “independent quasar activity” assumption for the central and satellite quasars that facilitates the use of Poisson statistics for the integrator in (Seljak 2000)’s description of the one-halo term of the power spectrum (Kayo & Oguri 2012):

$$\begin{aligned} P_{1h}(k) &= \frac{1}{n_q^2} \int \langle N(N-1) \rangle u(k, M)^p \frac{dn}{dM} dM \\ &\equiv 2\langle N_{cen} N_{sat} \rangle u(k, M) + \langle N_{sat}(N_{sat}-1) \rangle |u(k, M)|^2 \\ &= [2f_{sat}(1-f_{sat})u(k, M) + f_{sat}^2|u(k, M)|^2]\langle N(M) \rangle, \end{aligned} \quad (7)$$

where  $u(k, M)$  is the Fourier transform of the quasar number density profile in a dark matter halo of mass  $M$ . We adopt an NFW profile (Navarro, Frenk & White 1997) with concentration parameter

$$c(M, z) = \frac{c_0}{1+z} \left[ \frac{M}{M^*} \right]^{-0.13} \quad (8)$$

where  $M^*$  is the nonlinear mass scale for collapse at  $z = 0$ ,  $\beta = -0.13$  (Cooray & Sheth 2002). In this work, we initially chose  $c_0 = 9$  and investigated the change in the best-fit parameters for a model with haloes with 10 times higher concentration (see the right panel of Fig. 5). A weak dependency of the HOD model to the concentration parameter is reported by a number of studies (Richardson et al. 2012, 2013; Shen et al. 2013) that used a different HOD formalism where the distribution of the central and satellite haloes are assumed to be a softened three-parameter step function and a two-parameter power-law respectively. We further discuss the differences between our chosen parameters and those previously used in the literature in §5 where we summarize the differences between assumptions and their potential effect on the best-fit parameters.



**Figure 4.** Left: Similar to the right panel of Fig. 3 but for a single bin of the measured correlation function using the KDE-complete sample of close pairs. Best fit model has  $M_m = 1.45_{-0.25}^{+0.32} \times 10^{12} h^{-1} M_\odot$ , and  $f_{\text{sat}} = 0.016_{-0.004}^{+0.005}$  with reduced  $\chi^2 = 1.18$ . The shaded envelope around the best fit model is the extent of the 2-sigma certainty for the shape of the HOD model at each bin of the projected distance ( $r_p$ ). Right: Fitting statistics similar to the one shown for the fitting process in the right-hand panel of Fig. 3.

## 5 RESULTS AND DISCUSSION

We investigated relaxing  $\Delta_m$  as the third parameter for the HOD model (in addition to  $f_{\text{sat}}$  and  $M_m$ ) to have a significant effect in changing the best fit model.

Inspired by the test conducted in Kayo & Oguri (2012), we fit an HOD model with 10 times higher concentration parameter to the full-scale correlation function (with and without inclusion of Kayo & Oguri (2012)'s measurement) and found no significant improvement compared the previous fit although the implication for steeper slope for the one-halo term of the model exist (see Fig. 5).

Recent HOD analysis of the local AGNs at  $z \sim 0.01 - 0.1$  has revealed that AGNs reside in similar dark matter haloes to that of galaxies with similar stellar mass (reference a recent analysis with BASS-2MASS CCF measurement at  $z=0.01-0.1$  using HALOTOOLS package). This could be interpreted as a by product of observing no significant luminosity dependency of quasar and AGN clustering.

A look at the recent investigations on the effective environmental factors that lead to presence of two AGN types and the contradictory findings on the difference between their bias with respect to dark matter, halo mass, black hole mass, mean number of satellites (see, e.g., DiPompeo et al. 2015, 2017; Krumpe et al. 2018) is a shining evidence of the need for theoretical followup and pairing up the measured correlation function with complementary measurements that enhances the presence of the observationally challenging bins at kpc scales (see, Starikova et al. 2011, for a proposed alternative).

In recent years, an increasing number of reports on the persistent degeneracies between the best-fit parameters of widely used HOD formalism (initially designed for local AGNs) have emerged (Coil et al. 2016, 2017).

These studies differ from this work (as well as (Kayo & Oguri

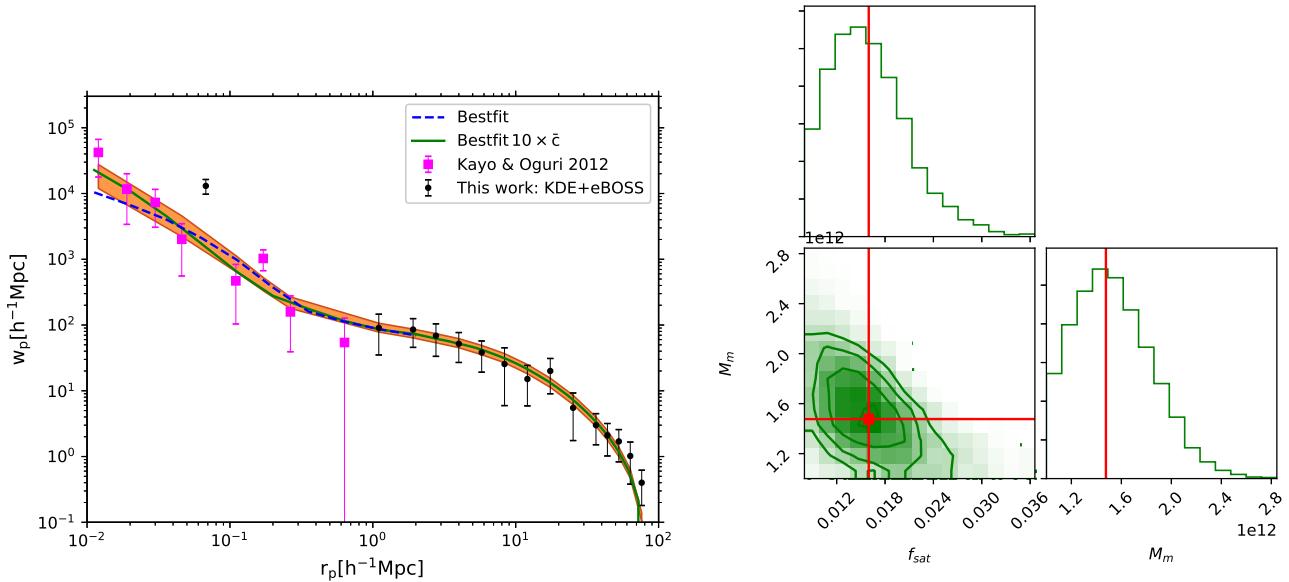
2012)) in two significant ways: (i) The two-point correlation function in those studies fall short of measurements over scales well within individual halo ( $\sim 10 - 100 h^{-1} \text{kpc}$ ) where the one halo term of the power spectrum has an ideal chance of constraining with observation at those scales. (ii) We assumed a Gaussian form for distribution of both central and satellite haloes (add an argument for what this assumption terms of the assembly history). Shen et al. (2013) performed a comprehensive investigation of the formalism introduced by Zheng et al. (2005) and implemented by a number of clustering analysis ever since (Zheng, Coil & Zehavi 2007; Miyaji et al. 2011; Zehavi et al. 2011; Richardson et al. 2012, 2013)(add more refs) by performing a 5 and six parameter fit to their galaxy-quasar cross correlation function at  $\bar{z} \sim 0.5$

## ACKNOWLEDGMENTS

SE and ADM acknowledge support by the National Science Foundation (NSF) through the grant number 10 times.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is <http://www.sdss.org>.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berke-



**Figure 5.** Left: Our HOD prediction of the projected correlation function that best fits the full-scale measurement (all the shown data points participated in the best fit determination) and where the dark matter halos are assumed to have 10 times higher average concentrations than the best fit model shown with the dashed blue line (solid blue curve in Fig.4). The best fit model (solid green line) has  $M_m = 1.55^{+0.35}_{-0.29} \times 10^{12} h^{-1} M_\odot$ , and  $f_{sat} = 0.016^{+0.005}_{-0.004}$  with reduced  $\chi^2 = 1.21$ . The shaded envelope around the best fit model is the extent of the 2-sigma certainty for the shape of the HOD model at each bin of the projected distance ( $r_p$ ). Although the steeper shape of the one halo term seems to better follow the measurements reported in small scales, the fitting results do not show a significantly better constraint on the fitting parameters.

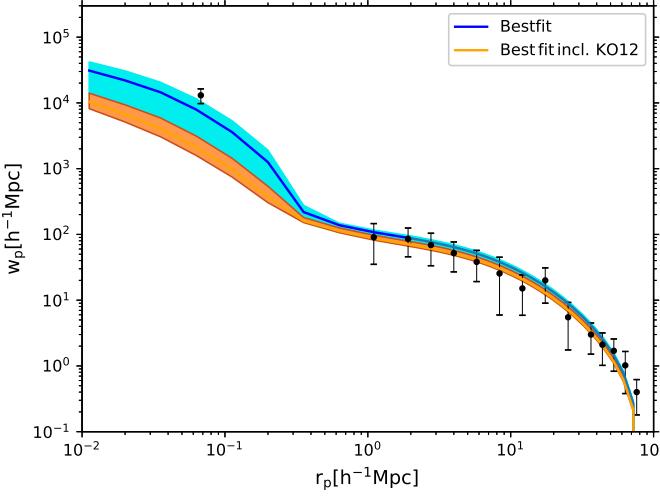
Data	$M_m (\times 10^{12} h^{-1} M_\odot)$	$f_{sat}$	$\chi^2_{red}$ (dof)
4KDE+eBOSS+KO12	$1.43^{+0.31}_{-0.25}$	$0.024^{+0.007}_{-0.007}$	1.46 (22)
1KDE+eBOSS+KO12	$1.45^{+0.32}_{-0.25}$	$0.016^{+0.005}_{-0.004}$	1.180 (20)
1KDE+eBOSS+KO12 (10c)	$1.55^{+0.29}_{-0.35}$	$0.016^{+0.007}_{-0.007}$	1.213 (20)
1KDE+eBOSS	$2.49^{+0.50}_{-0.46}$	$0.052^{+0.016}_{-0.033}$	1.086 (12)

**Table 2.** Summary of the fitting results based on the data that participated in the fit. The second and third columns are the best-fit parameters for the fit and their 1-sigma uncertainties. The last column lists the reduced  $\chi^2_{red}$  (and the degrees of freedom).

ley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatorio Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

## REFERENCES

- Abazajian K. et al., 2005, ApJ, 625, 613
- Abolfathi B. et al., 2018, ApJS, 235, 42
- Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297
- Allevato V. et al., 2011, ApJ, 736, 99
- Allevato V. et al., 2014, ApJ, 796, 4
- Anderson L. et al., 2012, MNRAS, 427, 3435
- Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
- Bianchi D., Percival W. J., 2017, MNRAS, 472, 1106
- Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, AJ, 125, 2276
- Bolton A. S., Burles S., Schlegel D. J., Eisenstein D. J., Brinkmann J., 2004, AJ, 127, 1860
- Chatterjee S., Nguyen M. L., Myers A. D., Zheng Z., 2013, ApJ, 779, 147
- Coil A. L. et al., 2004, ApJ, 609, 525
- Coil A. L. et al., 2009, ApJ, 701, 1484
- Coil A. L. et al., 2006, ApJ, 638, 668
- Coil A. L., Hennawi J. F., Newman J. A., Cooper M. C., Davis



**Figure 6.** Left: Our HOD prediction of the projected correlation function that best fits the full-scale measurement where only the measurement from KDE-complete sample of close pairs and the data from eBOSS clustering is added (solid blue curve in Fig.4). The best fit model (solid green line) has  $M_m = 2.49_{-0.465}^{+0.498} \times 10^{12} h^{-1} M_\odot$ , and  $f_{\text{sat}} = 0.052_{-0.033}^{+0.016}$  with reduced  $\chi^2 = 1.09$ . The shaded envelope around the best fit model is the extent of the 2-sigma certainty for the shape of the HOD model at each bin of the projected distance ( $r_p$ ).

- M., 2007, ApJ, 654, 115  
 Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2016, ArXiv:1609.09090  
 Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, ApJ, 838, 87  
 Cooray A., Sheth R., 2002, Phys. Rep., 372, 1  
 Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, MNRAS, 349, 1397  
 Dawson K. S. et al., 2016, AJ, 151, 44  
 DESI Collaboration et al., 2016a, ArXiv e-prints  
 DESI Collaboration et al., 2016b, ArXiv e-prints  
 DiPompeo M. A., Hickox R. C., Myers A. D., Geach J. E., 2017, MNRAS, 464, 3526  
 DiPompeo M. A., Myers A. D., Hickox R. C., Geach J. E., Holder G., Hainline K. N., Hall S. W., 2015, MNRAS, 446, 3492  
 Eftekharzadeh S., Myers A. D., Hennawi J. F., Djorgovski S. G., Richards G. T., Mahabal A. A., Graham M. J., 2017, MNRAS, 468, 77  
 Eftekharzadeh S. et al., 2015, MNRAS, 453, 2779  
 Eisenstein D. J., Hu W., 1999, ApJ, 511, 5  
 Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23  
 Folkes S. et al., 1999, MNRAS, 308, 459  
 Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748  
 Gunn J. E. et al., 1998, AJ, 116, 3040  
 Guo H., Zehavi I., Zheng Z., 2012, ApJ, 756, 127  
 Hahn C., Scoccimarro R., Blanton M. R., Tinker J. L., Rodríguez-Torres S. A., 2017, MNRAS, 467, 1940  
 Hennawi J. F. et al., 2010, ApJ, 719, 1672  
 Hennawi J. F. et al., 2006, AJ, 131, 1  
 Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576  
 Kayo I., Oguri M., 2012, MNRAS, 424, 1363  
 Kormendy J., Richstone D., 1995, ARA&A, 33, 581  
 Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35  
 Krumpe M., Miyaji T., Coil A. L., Aceves H., 2012, ApJ, 746, 1  
 Krumpe M., Miyaji T., Coil A. L., Aceves H., 2018, MNRAS, 474, 1773  
 Landy S. D., Szalay A. S., 1993, ApJ, 412, 64  
 Laurent P. et al., 2017, JCAP, 7, 017  
 McGreer I. D., Eftekharzadeh S., Myers A. D., Fan X., 2016, AJ, 151, 61  
 Miyaji T., Krumpe M., Coil A. L., Aceves H., 2011, ApJ, 726, 83  
 Myers A. D., Brunner R. J., Nichol R. C., Richards G. T., Schneider D. P., Bahcall N. A., 2007a, ApJ, 658, 85  
 Myers A. D., Brunner R. J., Richards G. T., Nichol R. C., Schneider D. P., Bahcall N. A., 2007b, ApJ, 658, 99  
 Myers A. D. et al., 2006, ApJ, 638, 622  
 Myers A. D. et al., 2015, ApJS, 221, 27  
 Myers A. D., Richards G. T., Brunner R. J., Schneider D. P., Strand N. E., Hall P. B., Blomquist J. A., York D. G., 2008, ApJ, 678, 635  
 Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493  
 Neyman J., Scott E. L., 1974, in IAU Symposium, Vol. 63, Confrontation of Cosmological Theories with Observational Data, Longair M. S., ed., pp. 129–140  
 Neyman J., Scott E. L., Zonn W., 1962, AJ, 67, 583  
 Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJ, 193, L1  
 Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144  
 Planck Collaboration et al., 2015, ArXiv:1502.01589  
 Porciani C., Magliocchetti M., Norberg P., 2004, MNRAS, 355, 1010  
 Richards G. T. et al., 2004, ApJS, 155, 257  
 Richardson J., Chatterjee S., Zheng Z., Myers A. D., Hickox R., 2013, ApJ, 774, 143  
 Richardson J., Zheng Z., Chatterjee S., Nagai D., Shen Y., 2012, ApJ, 755, 30  
 Rodríguez-Torres S. A. et al., 2017, MNRAS, 468, 728  
 Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2005, ApJ, 624, 571  
 Rood H. J., Page T. L., Kintner E. C., King I. R., 1972, ApJ, 175, 627  
 Ross A. J. et al., 2012, MNRAS, 424, 564  
 Ross N. P. et al., 2013, ApJ, 773, 14  
 Ross N. P. et al., 2009, ApJ, 697, 1634  
 Rubin V. C., Thonnard N., Ford, Jr. W. K., 1978, ApJ, 225, L107  
 Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20  
 Seljak U., 2000, MNRAS, 318, 203  
 Shen Y. et al., 2013, ApJ, 778, 98  
 Shen Y. et al., 2007, AJ, 133, 2222  
 Shen Y. et al., 2009, ApJ, 697, 1656  
 Starikova S. et al., 2011, ApJ, 741, 15  
 Stoughton C. et al., 2002, AJ, 123, 485  
 Strauss M. A. et al., 2002, AJ, 124, 1810  
 Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391  
 Tegmark M. et al., 2004, ApJ, 606, 702  
 Vanden Berk D. E. et al., 2001, AJ, 122, 549  
 West A. A. et al., 2004, AJ, 128, 426  
 White M. et al., 2012, MNRAS, 424, 933  
 White S. D. M., Rees M. J., 1978, MNRAS, 183, 341  
 Wilhite B. C., Vanden Berk D. E., Kron R. G., Schneider D. P., Pereyra N., Brunner R. J., Richards G. T., Brinkmann J. V., 2005, ApJ, 633, 638  
 Yip C. W. et al., 2004a, AJ, 128, 585  
 Yip C. W. et al., 2004b, AJ, 128, 2603

- York D. G. et al., 2000, AJ, 120, 1579  
Zehavi I. et al., 2011, ApJ, 736, 59  
Zheng Z. et al., 2005, ApJ, 633, 791  
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760