On the small scale clustering of quasars: constraints from the MassiveBlackII simulation

Aklant K. Bhowmick¹, Tiziana DiMatteo¹, Sarah Eftekharzadeh²,³, Adam D. Myers⁴

¹McWilliams Center for Cosmology, Dept. of Physics, Carnegie Mellon University, Pittsburgh PA 15213, USA
²Department of Physics, Southern Methodist University, 3215 Daniel Ave., Dallas, TX 75225
³Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84112
⁴Department of Physics & Astronomy, University of Wyoming, 1000 University Ave., Laramie, WY, 82071

30 November 2018

ABSTRACT
We examine recent high-precision measurements of small-scale quasar clustering (at $z \sim 0.5-2$ on scales of $\sim 25 \text{ kpc}/h$) from the SDSS in the context of the MassiveBlackII (MBII) cosmological hydrodynamic simulation and conditional luminosity function (CLF) modeling. At these high luminosities ($g < 20.85$ quasars), the MBII simulation volume (100 cMpc/h comoving boxsize) has only 3 quasar pairs at distances of 1–4 Mpc. The black-hole masses for the pairs range between $M_{bh} \sim 1-3 \times 10^9 M_\odot/h$ and the quasar hosts are haloes of $M_h \sim 1-3 \times 10^{13} M_\odot/h$. Such pairs show signs of recent major mergers in the MBII simulation. By modeling the central and satellite AGN CLFs as log-normal and Schechter distributions respectively (as seen in MBII AGNs), we arrive at CLF models which fit the simulation predictions and observed measurements of the luminosity function and the small-scale clustering measured for the SDSS sample. The small-scale clustering of our mock quasars is well-explained by central-satellite quasar pairs that reside in $M_h > 10^{14} M_\odot/h$ dark matter haloes. For these pairs, the satellite quasar luminosity is similar to that of the central quasars. Our CLF models imply a relatively steep increase in the maximum satellite luminosity, $L_{sat}^*$, in haloes of $M_h > 10^{14} M_\odot/h$ with associated larger values of $L_{sat}^*/L_*$ at higher redshift. This leads to an increase in the satellite fraction that manifests itself in an enhanced clustering signal at $\lesssim 1$ Mpc/h. For the ongoing eBOSS-CORE sample, we predict $\sim 200-500$ quasar pairs at $z \sim 1.5$ (with $M_h \gtrsim 10^{14} M_\odot/h$ and $M_{bh} \gtrsim 10^8 M_\odot/h$) at $\sim 25 \text{ kpc}$ scales. Such a sample would be $\gtrsim 10$ times larger than current pair samples.

1 INTRODUCTION
Small scale clustering measurements for AGNs/quasars have been of significant interest over the last two decades as they may constrain signatures of the physical processes that trigger AGN activity, such as galaxy mergers (Di Matteo et al. 2005; Mortlock et al. 1999; Kochanek et al. 1999). Several works (Schneider et al. 2000; Hennawi et al. 2006; Myers et al. 2007, 2008; Kayo & Oguri 2012; McGreer et al. 2016; Eftekharzadeh et al. 2017) over the last 20 years have measured the small-scale clustering of quasars, mainly from the SDSS and 2dF-QSO surveys, at scales ranging from $\sim 10$ kpc to $\sim 1$ Mpc. These works typically measure strong clustering at small scales ($\lesssim 200$ kpc), which could be attributed to galaxy mergers as possible triggers of quasar activity.

Cosmological hydrodynamic simulations are valuable tools to study AGN clustering. Theoretical predictions for AGN clustering and luminosity evolution have been made using ILLUSTRIS (DeGraf & Sijacki 2017) and MassiveBlackII (Khandai et al. 2015, hereafter MBII) and they report broad agreement with observational measurements (Croom et al. 2005; Porciani & Norberg 2006; Myers et al. 2006; Shen et al. 2009; White et al. 2012; Eftekharzadeh et al. 2015). However, simulated AGNs are significantly fainter (with bolometric luminosities $L_{bol} < L \lesssim 10^{44}$ ergs/sec) than the observed quasar samples ($L > 10^{44}$ ergs/sec). These bright quasars are too rare for their clustering properties to be directly probed by hydrodynamic simulations (Khandai et al. 2015; Schaye et al. 2015; Nelson et al. 2015; Kaviraj et al. 2017). In general, this is because the volumes of such simulations (boxsizes are 100 cMpc/h for MBII and Horizon AGN, and $\sim 75$ cMpc/h for ILLUSTRIS and EAGLE) are limited due to the computational demand of implementing ‘full physics’. However, analytical models such as Halo Occupation Distribution (HOD) or Conditional Luminosity Function (CLF) modeling can be used to supplement simulations. This mitigates volume limitations and allows simulations to probe the implications of clustering measurements at small scales.

HOD and CLF modeling offer predictions for the properties of dark matter haloes based on the observed clustering of the targets that haloes host. This has been applied extensively to galaxies (Zheng et al. 2007; Guo et al. 2016; Harikane et al. 2018, and references therein) as well as AGNs.
2 Methods

2.1 MassiveBlackII Simulation

MBII is a high-resolution cosmological hydrodynamic simulation which runs from \( z = 159 \) to \( z = 0.06 \). The simulation has a boxsize of 100 cMpc/\( h \) and 2 \( \times 1792^3 \) particles. The simulation used the cosmological parameters inferred from WMAP7 (Komatsu et al. 2011) i.e. \( \Omega_m = 0.275, \Omega_{\Lambda} = 0.725, \Omega_b = 0.046, \sigma_8 = 0.816, h = 0.701, n_s = 0.968 \). The dark matter and gas particle masses are \( 1.1 \times 10^7 M_\odot/h \) and \( 2.2 \times 10^5 M_\odot/h \) respectively. The simulation was run using P-GADGET, an upgraded version of GADGET (Springel 2005). In addition to the N-body gravity solver for the dark matter component and Smooth Particle Hydrodynamics (SPH) solver for the gas component, MBII incorporates subgrid physics modeling such as star formation (Springel & Hernquist 2003), black hole growth and associated feedback.

Haloes and subhaloes were identified using the Friends-of-Friends (FOF) group finder (Davis et al. 1985) and SUBFIND (Springel 2005) respectively. For more details, we refer the reader to Khandai et al. (2015).

2.1.1 Black Hole Growth and Associated Feedback

The prescription for black hole growth used in the simulation is adopted from Di Matteo et al. (2005) and Springel et al. (2005). A seed black hole of mass \( 5 \times 10^5 M_\odot/h \) is placed in a halo of mass \( \geq 5 \times 10^{13} M_\odot/h \) (if the halo does not already contain a black hole). Once seeded, the black hole grows at a rate given by \( M_{bh} = \frac{4 \pi G^2 M^2 c_s^2}{c^4} \rho \) and \( c_s \) are the density and sound speed of the cold phase of the ISM gas, and \( v_{\text{rel}} \) is the relative velocity of the black hole w.r.t. gas. The bolometric luminosity of the black hole is given by \( \epsilon \), where \( \epsilon \) is the radiative efficiency taken to be 0.1. 5% of the energy released is thermodynamically (and isotropically) coupled to the surrounding gas (Di Matteo et al. 2005). Black holes can also grow via merging; two black holes are considered to be merged if they come within the spatial resolution of the simulation (the SPH smoothening length) with a relative speed smaller than the local sound speed of the medium.

2.1.2 Identifying AGNs: Centrals and Satellites

Simulated AGNs are identified to be individual black holes accreting gas from the surrounding medium. We consider objects with bolometric luminosity \( L_{\text{bol}} \geq 3 \times 10^{43} \text{ergs/sec} \) and \( M_{bh} \geq 10^5 M_\odot/h \) (5 times the seed mass of the black hole). Within a host halo (FOF group), the most massive black hole is defined to be the central AGN. Any other AGN within the same halo is defined to be a satellite AGN.

2.2 AGN Clustering

In this work, we focus on the two-point (pairwise) clustering statistics quantified by the two-point correlation function. The one-component spatial correlation function is defined to be

\[
dN = n_{\text{AGN}}(1 + \xi(r))4\pi r^2 dr
\]

where \( r \) is the inter-particle distance; assuming a particle located at \( r = 0 \), \( dN \) is the total number of paired particles within a spherical shell at a distance \( r \), thickness \( dr \); \( n_{\text{AGN}} \) is the volume density of AGNs. Similarly, the two-component redshift space correlation function is defined as

\[
dN = n_{\text{AGN}}(1 + \xi(r, s))4\pi r_p ds
dr
\]

where \( r_p \) and \( s \) are the distance parallel and perpendicular (in redshift space) to the observer’s line of sight, respectively; assuming a particle located at the origin i.e. \( r_p = 0 \), \( s = 0 \), \( dN \) is the total number of paired particles within a cylindrical shell at projected distance (inner radius) \( r_p \), thickness \( dr_p \) and height \( ds \).

The projected correlation function \( w_p(r_p) \) is defined as

\[
w_p(r_p) = \int_0^\infty \xi(r_p, s) ds = 2 \int_0^\infty \frac{r_p^2(r)}{\sqrt{r^2 - r_p^2}} dr.
\]
Recent works (Hennawi et al. 2006; Kayo & Oguri 2012; Eftekhazadeh et al. 2017) on measurements of quasar clustering at small (≲ kpc) scales have used the “Volume averaged projected correlation function” \( W_p \). For quasar pairs within transverse separations of \( r_{\text{min}} < r_p < r_{\text{max}} \) and a line of sight velocity separation of \( v_{\text{max}} \), this statistic is defined as

\[
W_p(z; r_{\text{min}}, r_{\text{max}}) = \frac{1}{V_{\text{shell}}} \int_{v_{\text{max}}/aH(z)}^{v_{\text{max}}/aH(z)} \int_{r_{\text{min}}}^{r_{\text{max}}} \frac{\xi(r_p, s, z)}{2\pi r_p dr_p ds} \tag{4}
\]

where \( V_{\text{shell}} = \pi(r_{\text{max}}^2 - r_{\text{min}}^2)v_{\text{max}}/aH(z) \). Unlike \( w_p \) which has a dimension of length, \( W_p \) is dimensionless and is essentially the two component correlation function \( \xi(r_p, s) \) averaged over the volume \( V_{\text{shell}} \). In practice, if the velocity separation is large enough (true in our case with \( v_{\text{max}} = 2000 \) km/sec), the redshift space distortions can be effectively removed by the integration over \( s \) in Eq. (4), and the following approximation can be used to convert between \( w_p \) and \( W_p \)

\[
w_p \approx 2v_{\text{max}}/aH(z) W_p; \tag{5}
\]

(see also Eq. A2 of Richardson et al. 2012).

From the simulated sample of objects, we determine \( W_p \) using

\[
W_p = QQ/(QR) - 1 \tag{6}
\]

where \( QQ \) is the number of AGN-AGN pairs within a projected distance of \( r_{\text{min}} \) to \( r_{\text{max}} \) and a line of sight separation of \( v_{\text{max}}/aH(z) \) in redshift space; \( (QR) \) is the expected number of AGN-random pairs inside the same cylindrical shell (the random objects are uniformly distributed over the volume). Alternatively, \( W_p \) can also be determined from \( w_p \) by using Eq. (5), where \( w_p \) can be determined from \( \xi(r_p, s) \) or \( \xi(r) \) using Eq. (3).

### 2.3 Conditional luminosity function modeling

The Conditional Luminosity Function (CLF) approach has been widely used for clustering analyses of galaxies (Cooray 2006; Trevisan & Mamon 2017, and references therein). In this work, we shall be constructing a CLF model for the one-halo clustering of AGNs. The CLF, denoted by \( \Phi(L, M_h) \), measures the distribution of AGN bolometric luminosities \( L \) within a halo of mass \( M_h \). Its definitive properties are

1) \( \int \Phi(L, M_h) dM_h = \Phi(L) \) where \( \Phi(L) \) is the overall luminosity function, and

2) \( \int_{L_{\text{min}}}^{\infty} \Phi(L', M_h) dL' = N(L > L_{\text{min}}, M_h) \) \( \tag{8} \)

where \( N(L > L_{\text{min}}, M_h) \) is the mean halo occupation number of AGNs with \( L > L_{\text{min}} \). We separate our CLF into contributions from central and satellite AGNs (according to Section 2.1.2)

\[
\Phi(L, M_h) = \Phi_{\text{cen}}(L, M_h) + \Phi_{\text{sat}}(L, M_h). \tag{9}
\]

As we shall see in Section 4, CLFs for central AGNs can be modelled as a log-normal distribution

\[
\Phi_{\text{cen}}(L, M_h) = \frac{1}{2\pi\sigma_{\text{cen}}} \exp \left( - \frac{(L - L^*_{\text{cen}})^2}{\sigma_{\text{cen}}^2} \right), \tag{10}
\]

where \( L^*_{\text{cen}} \) is the probability distribution of luminosities of central AGNs within a halo of mass \( M_h \), \( L^*_{\text{cen}} \) is the average central AGN luminosity, and \( \sigma_{\text{cen}} \) is the width of the log-normal distribution. CLFs for satellite AGNs can be modelled as a Schechter distribution

\[
\Phi_{\text{sat}}(L, M_h) = \langle N_{\text{sat}} \rangle \frac{dN}{dL} = Q_{\text{sat}} \left( \frac{L}{L^*_{\text{sat}}} \right)^{\alpha_{\text{sat}}} \exp \left( - \frac{L}{L^*_{\text{sat}}} \right) \tag{11}
\]

where \( \langle N_{\text{sat}} \rangle \) and \( P_{\text{sat}} \) are, respectively, the average number and the probability distribution of luminosities of satellite AGNs within a halo of mass \( M_h \). \( L^*_{\text{sat}} \) is a measure of the most luminous satellite for a given halo of mass \( M_h \), \( \alpha_{\text{sat}} \) determines the slope of the distribution for \( L < L^*_{\text{sat}} \), and \( Q_{\text{sat}} \) is the overall normalization.

\( \xi(r) \) can be decomposed as

\[
\xi(r) = \xi_{1h}(r) + \xi_{2h}(r) \tag{12}
\]

where \( \xi_{1h} \) and \( \xi_{2h} \) are ‘one-halo’ and ‘two-halo’ contributions respectively. The one-halo contribution for the power spectrum \( P_{1h} \) is given by

\[
P_{1h}(k) = \frac{1}{n_{\text{AGN}}} \int ((\langle N_{\text{cen}} \rangle + \langle N_{\text{sat}}(N_{\text{sat}} - 1) \rangle) \frac{dn}{dM_h} u(k, M_h)^2 dM_h \tag{13}
\]

where \( \langle N_{\text{cen}} \rangle \) and \( \langle N_{\text{sat}}(N_{\text{sat}} - 1) \rangle \) are the expected number of central-satellite and satellite-satellite pairs, \( u(k, M_h) \) is the Fourier transform of the normalized satellite AGN density profile, and \( \frac{dn}{dM_h} \) is the halo mass function. We assume that central and satellite occupations are independent i.e. \( \langle N_{\text{cen}} \rangle \langle N_{\text{sat}} \rangle = \langle N_{\text{cen}} \rangle \langle N_{\text{sat}} \rangle \), and satellite distributions are Poisson i.e. \( \langle N_{\text{sat}}(N_{\text{sat}} - 1) \rangle = \langle N_{\text{sat}} \rangle^2 \). \( \langle N_{\text{cen}} \rangle \) and \( \langle N_{\text{sat}} \rangle \) can be obtained from \( \Phi_{\text{cen}}(L, M_h) \) and \( \Phi_{\text{sat}}(L, M_h) \) respectively using Eq. (8). \( \xi_{1h}(r) \) can then be determined using

\[
\xi_{1h}(r) = \frac{1}{2\pi^2} \int dk \frac{k^2 P_{1h}(k) \sin kr}{kr}. \tag{14}
\]

The two-halo contribution is given by

\[
\xi_{2h}(r) = \xi_{\text{matter}}(r) b_{\text{eff}}^2 \tag{15}
\]

where \( \xi_{\text{matter}}(r) \) is the matter power spectrum. Here, \( b_{\text{eff}} \) is the effective bias of AGNs given by

\[
b_{\text{eff}} = \frac{1}{n_{\text{AGN}}} \int dM_h (\langle N_{\text{cen}} \rangle + \langle N_{\text{sat}} \rangle) \frac{dn}{dM_h} b(M_h) \tag{16}
\]

where \( b(M_h) \) is the halo bias.

We make the following additional assumptions in our modeling:

- We use the halo mass function from Tinker et al. (2008). MBII mass functions are consistent with Tinker et al. (2008) as shown in Khandai et al. (2015).
- For the two-halo term, we use the linear halo bias model from Tinker et al. (2010) and the non-linear matter power spectrum predicted by MBII.
For the one-halo term, we use a power-law satellite profile with exponent ‘−2’, which is consistent with the radial distribution of satellite AGNs in MBII, as discussed in Appendix B.

Our objective is to use the methods summarized in Section 2.2 and 2.3 to obtain theoretical predictions for $W_p$ and compare with observed measurements. For the sample of simulated AGNs (assuming it is large enough), one can use Eq. (6) to obtain $\xi(r)$ of simulated AGNs and compare with observed measurements. For the sample of SDSS quasars (Shen 2013) do find evidence of such a correlation. Also, for a given black hole mass, bolometric luminosity increases with increasing redshift for both central and satellite AGNs. $M_{bh}$ vs. $M_h$ relations are shown in the middle panels; there is no significant redshift evolution in the $M_{bol}$ vs. $M_h$ relation. The $M_{bh}-L_{bol}$ and $M_{bh}-M_h$ relations then inevitably produce the positive correlation between $L_{bol}$ vs. $M_h$ that is shown in the rightmost panels of Figure 1. Note that for satellite black holes the slope of $L_{bol}$ vs. $M_h$ is significantly smaller than for central AGNs. In a nutshell, we find that the relation between AGN luminosity and halo mass is primarily governed by the associated black hole mass - halo mass relation. As a result, more luminous AGNs also consistently have slightly higher (by factors of 1.5-2.5) stellar masses ranging from $M_*/h = 10^5$ to $10^6 M_\odot/h$. The morphology and luminosities of these satellite AGNs are comparable to those of central AGNs, making them an extreme and rare subset of the host population. This hints at the possibility of these pairs originating from recent halo mergers, such that both members of the pairs were two central AGNs prior to the merger.

In order to further investigate whether these pairs are activated by mergers, we look into the properties of their host galaxies and haloes, and their progenitors. Figure 2 (left and middle panels) shows the projected positions of 3 pairs (at $z = 0.6$) with $g < 20.85$ along with their host galaxies and haloes. The host haloes are shown as grey histograms and have masses $\sim 1 - 3 \times 10^{14} M_\odot/h$. The morphology of the host haloes in both cases (left and middle panels) suggest that they are involved in a major merger. The colored histograms show the host galaxies of these individual quasars. The green histograms show the host galaxy of the central AGN; red and blue histograms show the host galaxy of the satellite AGNs. Note that host galaxies of the central AGNs also consistently have slightly higher (by factors of 1.5-2.5) stellar masses compared to satellites, making them consistent with the usual definition of “central galaxies”. As we can see, these simulated quasar pairs reside in separate (central and satellite) galaxies within the same halo. They all have roughly comparable (within a half order of magnitude) stellar masses ranging from $M_\star \sim 1 - 3 \times 10^{12} M_\odot/h$ and are both among the most massive galaxies in the simulation. This suggests that they were initially two centrals in different haloes, which then merged to form a central-satellite pair (according to our definition in Section 2.1.2) with similar stellar masses. We now look at the growth history of the three $g < 20.85$ quasar pairs and their host haloes. The panels in Figure 3 show the evolution of $M_{bh}$, $L_{bol}$ and $M_{bh}$ at $z = 0.6$ are denoted by red, blue and green lines. The host halo masses of their progenitor AGNs (top panels) are different at $z = 1$, implying that a merger happened between $z = 1$ and $z = 0.6$. If we look at the evolution of $L_{bol}$ and $M_{bh}$ in the middle and bottom panels, we see that the activity of one or both members is significantly enhanced after the merger. Particularly for the solid green and dashed red lines, $L_{bol}$ increases from $\sim 10^{45.5}$ ergs/sec at $z = 1$ to $\sim 10^{48}$ ergs/sec; consequently, $M_{bh}$ rises from $\sim 10^6 M_\odot/h$ at $z = 1$ to $\sim 10^9 M_\odot/h$ at $z = 0.6$. This confirms that the AGN activity required for the formation of these bright quasar pairs was indeed triggered by recent halo mergers.

The cyan lines correspond to the pair at $z = 1$ (shown as red circles in Figure 1), and is also an interesting example demonstrating the extent to which halo mergers trigger AGN activity (in this pair, only the satellite AGN is bright enough to be observable) in the simulation. At $z = 2$, the mass of the central AGN (dashed cyan line in the bottom panel) progenitor is 2 orders of magnitude less than that of the satellite (solid cyan line in the bottom panel); but the halo merger around $z \sim 2$ triggered AGN activity in the central such that it surpasses the black hole mass of the satellite AGN at $z = 0.6$. Similar to $g < 20.85$ pairs, we also report

3 AGN PROPERTIES IN MBII

3.1 Scaling relations

There are $\gtrsim 70000$ black holes in MBII at $0.6 \lesssim z \lesssim 2$. We first present the scaling relations between various properties of the full AGN population (drawn from this sample of black holes) in the MBII simulation. Figure 1 presents the relations between the bolometric luminosity $L_{bol}$ black hole mass $M_{bh}$, and host halo mass $M_h$ of the quasars in MBII for central BHs (residing in the most massive (central) galaxies) and satellite BH/AGN within the host halo (which are typically associated with a satellite galaxy). The solid lines from blue to orange show the redshift evolution (the mean at each redshift) from $z = 0.6$ to $z = 2$. As expected, we find that the black hole mass is correlated with the bolometric luminosity for both central and satellite AGNs (see leftmost panels). Observational estimates of black hole masses of SDSS quasars (Shen 2013) do find evidence of such a correlation. Also, for a given black hole mass, bolometric luminosity increases with increasing redshift for both central and satellite AGNs. $M_{bh}$ vs. $M_h$ relations are shown in the middle panels; there is no significant redshift evolution in the $M_{bol}$ vs. $M_h$ relation. The $M_{bh}-L_{bol}$ and $M_{bh}-M_h$ relations then inevitably produce the positive correlation between $L_{bol}$ vs. $M_h$ that is shown in the rightmost panels of Figure 1. Note that for satellite black holes the slope of $L_{bol}$ vs. $M_h$ is significantly smaller than for central AGNs. In a nutshell, we find that the relation between AGN luminosity and halo mass is primarily governed by the associated black hole mass - halo mass relation. As a result, more luminous AGNs live in more massive haloes. We also find that the correlation is stronger for central AGNs compared to satellite AGNs. Furthermore, AGNs for fixed halo mass become more luminous with increasing $z$.

3.2 Quasar pairs in MassiveBlackII

The primary objective of our work is to interpret the small-scale clustering of quasar pairs measured in E17. These pairs are limited to a magnitude of $g < 20.85$. Within the MBII simulation volume, we find only 3 pairs with $g < 20.85$ at $z = 0.6$ at distances of $\sim 1 - 4$ Mpc.; there are no such pairs at $z = 1, 1.5, 2$. These numbers are reasonably consistent with the observed luminosity functions of quasars; as discussed later in Section 5. The properties of these quasar pairs are marked as filled blue circles in Figure 1, which shows that the central quasars lie reasonably close to the mean trends (solid lines). However, for the satellite quasars, we find (see Figure 1) that both black hole masses ($M_{bh} \sim 10^8 M_\odot/h$ and luminosities ($L_{bol} \sim 10^{45}$ ergs/sec) are $\sim 10^8 - 10^9$ times higher than the typical values ($M_{bh} \sim 10^7 M_\odot/h$ and $L_{bol} \sim 3 \times 10^{41}$ ergs/sec). In fact, the black hole masses and luminosities of these satellite AGNs are comparable to those of central AGNs, making them an extreme and rare subset of the host population. This hints at the possibility of these pairs originating from recent halo mergers, such that both members of the pairs were two central AGNs prior to the merger.

In order to further investigate whether these pairs are activated by mergers, we look into the properties of their host galaxies and haloes, and their progenitors. Figure 2 (left and middle panels) shows the projected positions of 3 pairs (at $z = 0.6$) with $g < 20.85$ along with their host galaxies and haloes. The host haloes are shown as grey histograms and have masses $\sim 1 - 3 \times 10^{14} M_\odot/h$. The morphology of the host haloes in both cases (left and middle panels) suggest that they are involved in a major merger. The colored histograms show the host galaxies of these individual quasars. The green histograms show the host galaxy of the central AGN; red and blue histograms show the host galaxy of the satellite AGNs. Note that host galaxies of the central AGNs also consistently have slightly higher (by factors of 1.5-2.5) stellar masses compared to satellites, making them consistent with the usual definition of “central galaxies”. As we can see, these simulated quasar pairs reside in separate (central and satellite) galaxies within the same halo. They all have roughly comparable (within a half order of magnitude) stellar masses ranging from $M_\star \sim 1 - 3 \times 10^{12} M_\odot/h$ and are both among the most massive galaxies in the simulation. This suggests that they were initially two centrals in different haloes, which then merged to form a central-satellite pair (according to our definition in Section 2.1.2) with similar stellar masses. We now look at the growth history of the three $g < 20.85$ quasar pairs and their host haloes. The panels in Figure 3 show the evolution of $M_{bh}$, $L_{bol}$ and $M_{bh}$. The pairs at $z = 0.6$ are denoted by red, blue and green lines. The host halo masses of their progenitor AGNs (top panels) are different at $z = 1$, implying that a merger happened between $z = 1$ and $z = 0.6$. If we look at the evolution of $L_{bol}$ and $M_{bh}$ in the middle and bottom panels, we see that the activity of one or both members is significantly enhanced after the merger. Particularly for the solid green and dashed red lines, $L_{bol}$ increases from $\sim 10^{45}$ ergs/sec at $z = 1$ to $\sim 10^{48.5}$ ergs/sec; consequently, $M_{bh}$ rises from $\sim 10^6 M_\odot/h$ at $z = 1$ to $\sim 10^9 M_\odot/h$ at $z = 0.6$. This confirms that the AGN activity required for the formation of these bright quasar pairs was indeed triggered by recent halo mergers.

The cyan lines correspond to the pair at $z = 1$ (shown as red circles in Figure 1), and is also an interesting example demonstrating the extent to which halo mergers trigger AGN activity (in this pair, only the satellite AGN is bright enough to be observable) in the simulation. At $z = 2$, the mass of the central AGN (dashed cyan line in the bottom panel) progenitor is 2 orders of magnitude less than that of the satellite (solid cyan line in the bottom panel); but the halo merger around $z \sim 2$ triggered AGN activity in the central such that it surpasses the black hole mass of the satellite AGN at $z = 0.6$. Similar to $g < 20.85$ pairs, we also report
halo merger-driven AGN activity for $z \sim 1.5$ satellites above the magnitude limit of eBOSS-CORE quasars ($g \lesssim 22$; Myers et al. 2015), which are shown as green squares in Figure 1 (bottom panels).

It is not possible to directly calculate the (statistical) small-scale clustering of quasars with so few simulated objects. In the next section, we therefore investigate the conditional luminosity functions (CLFs) of MBII AGNs; we shall then build a CLF model to study the small-scale clustering of $g < 20.85$ quasars.

4 CONDITIONAL LUMINOSITY FUNCTIONS (CLFS) OF MASSIVEBLACKII AGNS

Filled circles in the top and bottom panels of Figure 4 show the MBII predictions for the Conditional Luminosity Functions (CLFs) for central and satellite AGNs respectively. For the central AGNs, the CLFs can be well described by a log-normal distribution given by Eq. (10) (Note: we use the scipy.optimize.curve_fit package to perform our fits). The key parameter of interest is $L_{cen}$, which measures the characteristic mean luminosity of central AGNs as a function of halo mass. Filled blue circles in the top panel of Figure 5 show the best fit values of $L_{cen}$ for different halo mass bins. We can see that the maximum halo masses probed by MBII at different redshifts are $M_h \sim 10^{14}, 10^{15.5}, 10^{16}, 10^{16.5}$ for $z \sim 0.6, 1.0, 1.5, 2.0$ respectively. Within this range, $L_{cen}$ increases with halo mass $M_h$ as a power law; this is expected from the scaling relations shown in the top-right panel of Figure 1.

For the satellite AGNs, we find that CLFs can be well fit by a Schechter distribution which is given by Eq. (11). Here, the key parameter of interest is $L_{sat}$, which corresponds to the “edge” of the satellite CLF; i.e. the most luminous satellite quasars within haloes of a given mass $M_h$. $L_{sat}$ therefore would be sensitive to halo mergers if mergers indeed triggered the formation of satellite quasars that are $10^4$ times more luminous than the average values (as seen with the simulated quasar pairs in Section 3.2). Filled green circles in the bottom panel of Figure 5 show the best fit values of $L_{sat}$.

Figure 1. Scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 2. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 3. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 4. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 5. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 6. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 7. Scaling relations of satellite quasars

scaling relations of central (top panels) and satellite (bottom panels) quasars. The blue histogram shows the relations for $z = 0.6$. Solid lines show the mean relations at various redshifts and the error bars show 1σ scatter. Filled circles in all the panels correspond to the central-satellite pairs with $g < 20.85$ (the E17 magnitude limit). There are 3 such pairs at $z = 0.6$ (blue circles) and 1 pair at $z = 1$ (red circle) (Note that for the blue circles, two of the satellites are paired with the same central, as shown in the leftmost panel of Figure 2). The green squares show the central-satellite pairs above the eBOSS limit ($g < 22$) at a target redshift of $z = 1.5$. 

Figure 8. Scaling relations of satellite quasars
for different halo mass bins. As in the case of centrals, the maximum halo mass probed by MBII at different redshifts are $M_h \lesssim 10^{14}, 10^{13.5}, 10^{13}, 10^{12}$ for $z \sim 0.6, 1.0, 1.5, 2.0$ respectively. At $z = 0.6, 1.0, 2.0$, we do not find a significant dependence of $L^*_\text{sat}$ on $M_h$ for $M_h \lesssim 10^{14}, 10^{13.5}, M_\odot/h$. However, at $z = 1.5, 2.0$ $L^*_\text{sat}$ tends to increase with $M_h$ for $M_h \lesssim 10^{13} M_\odot/h$.

We shall use the trends we have noted in this section in the next few sections, where we build CLF models to populate AGNs in $M_h \gtrsim 10^{14}, 10^{13.5}, 10^{13}, 10^{12} M_\odot/h$ haloes at $z \sim 0.6, 1.0, 1.5, 2.0$ respectively.

5 AGN-LUMINOSITY FUNCTION

5.1 Simulations vs. observations

Filled blue circles and green circles in the top panels of Figure 6 show the MBII predictions for the luminosity functions of central and satellite AGNs respectively. We see that the central AGNs are the dominant contribution to the luminosity functions. We can compare the blue circles to the observational measurements from Croom et al. (2009) and Richards et al. (2006) shown as open and closed black squares respectively. This shows the range of magnitudes probed by simulations and observations at various redshifts. At $z = 0.6, 1.0, 2.0$, we see that there is a significant (albeit partial) overlap between the simulations and the magnitude range of the observations over $-20 \gtrsim M_g \gtrsim -24$ and there is a reasonable agreement between the two. However at $z \sim 1.5, 2.0$, the magnitude range of simulated and observed AGNs do not overlap because the number of AGNs in the simulation rapidly declines.

5.2 CLF modeling of the AGN luminosity function

Because the simulated and observed luminosity functions do not overlap, we use CLF modeling to extend the range of our predictions to higher luminosities. We consider only central galaxies, thereby assuming that the centrals continue to dominate over satellites at magnitudes brighter than those probed by the simulation ($M_g \lesssim -24$).

In section 4, we noted that MBII constrains the dependence of $L^*_\text{cen}$ on $M_h$ for halo masses that can be effectively probed by the simulation ($M_h \lesssim 10^{14}, 10^{13.5}, 10^{13}, 10^{12} M_\odot/h$ for $z \sim 0.6, 1.0, 1.5, 2.0$ respectively). We find that within the range probed by simulations, $L^*_\text{cen} \propto M_h$. In order to reach the observed range of magnitudes, we need to extend the $L^*_\text{cen}$ vs. $M_h$ relation to haloes more massive than MBII can effectively probe. We therefore define a scaling

$$L^*_\text{cen} \propto M_h^{\beta_{\text{cen}}}$$

for $M_h \gtrsim 10^{14}, 10^{13.5}, 10^{13}, 10^{12} M_\odot/h$ for $z \sim 0.6, 1.0, 1.5, 2.0$ respectively. $\beta_{\text{cen}}$ is a power-law exponent which essentially determines the shape of the luminosity function at $M_g \lesssim -24$.

We determine the values of $\beta_{\text{cen}}$ required to produce a model luminosity function consistent with the observed luminosity functions. The red and green dashed lines (enclosing the shaded blue regions) in the middle panels of Figure 6 show the CLF model predictions of the luminosity functions, and we see that they are reasonably consistent with the observed measurements. The values of $\beta_{\text{cen}}$ are listed in the legends; the corresponding $L^*_\text{cen}$ vs. $M_h$ relations are shown as red and green dashed lines in the top panels of Figure 5. Note that for $z = 0.6$ (the top-left panel), $\beta_{\text{cen}} \sim 1$ and is an extrapolation of the best fit line in the simulated regime ($M_h \lesssim 10^{14} M_\odot/h$). This is expected because at

Figure 2. Illustrations of quasar pairs with $g < 20.85$ at $z = 0.6$ snapshot (white crosses) in MBII. The stars in their host galaxies are shown in colors. The green histogram shows the host galaxy of the central (most massive black hole) quasar, and the red and blue histograms show the host galaxies of the satellite quasars. The grey histogram shows the host dark matter halo (FOF group), for the pairs indicated by filled circles in Figure 1. The stellar masses of the host galaxies are $2 \times 10^{12}, 1 \times 10^{12}, 1 \times 10^{12} M_\odot$ for the crosses within the green, blue and red histograms respectively. Properties of the quasars and their hosts are shown in Figure 3.

MNRAS 000, 000–000 (0000)
can now focus on our main objective, which is to construct a CLF model for satellite AGNs in order to probe the one-halo clustering.

5.3 Modeling the Satellite AGNs

For the CLF modeling of satellite AGNs, we adopt a similar approach to centrals i.e. we define a scaling relation

\[ L_{\text{sat}}^* \propto M_h^{\beta_{\text{sat}}} \]

for \( M_h \gtrsim 10^{13}, 10^{13.5}, 10^{14} \, M_\odot/h \) for \( z \sim 0.6, 1.0, 1.5, 2.0 \) respectively; \( \beta_{\text{sat}} \) is a power law exponent which determines the satellite LF at \( M_\phi \lesssim -22 \). However, unlike the centrals, there are no observational measurements of the satellite AGN luminosity functions with which to constrain \( \beta_{\text{sat}} \). Therefore, we consider a family of CLF models with various possible values \( \beta_{\text{sat}} \), which are shown as dashed lines in the bottom panels of Figure 5. The dashed lines in the lowermost panels of Figure 6 show the corresponding values of the satellite LFs; as we can see, we consider models wherein the satellites do not overshoot the central LFs.

In the next section, we shall look at the one-halo clustering predicted by the CLF models for \( g < 20.85 \) quasars, and compare to the observational measurements of E17.

6 SMALL-SCALE CLUSTERING: COMPARISON WITH OBSERVATIONAL CONSTRAINTS

E17 measured the volume-averaged projected correlation function \( \langle \hat{W}_p \rangle \) over scales of 17.7–36.6 kpc/h for quasar pairs with velocity differences of \(< 2000 \, \text{km/sec} \). We compute the \( \hat{W}_p(17.7–36.6 \, \text{kpc/h}) \) predicted by our CLF model, which we present as dashed lines for various \( \beta_{\text{sat}} \) values in Figure 7. At fainter magnitudes (23 \( \lesssim g \lesssim 26 \)), these dashed lines converge and are reasonably consistent with the simulation predictions (shown as filled green circles). For \( g < 20.85 \), we see that the different models can predict a range of clustering amplitudes. The black squares show the measurements of E17 for their sample of quasar pairs with \( g < 20.85 \). We find that in all redshift bins, there is a set of \( \beta_{\text{sat}} \) values (shown as red and green dashed lines that enclose the shaded green region) which predict a clustering amplitude consistent with the measurements of E17.

7 IMPLICATIONS OF CURRENT OBSERVATIONAL CONSTRAINTS

We now select the CLF models which are consistent with observed small-scale clustering constraints and investigate the implications of these models for the AGN population.

The green circles in Figure 8 show \( \beta_{\text{sat}} \) as a function of redshift for our final CLF model for satellite AGNs. The observed clustering measurements imply \( \beta_{\text{sat}} > 0 \), which means that \( L_{\text{sat}}^* \) is positively correlated with halo mass for massive \( (M_h \gtrsim 10^{14} M_\odot/h) \) haloes. A possible physical origin of the positive correlation is the merger of gas-rich massive haloes as triggers of AGN activity (as was seen for the host haloes of the simulated pairs). Higher \( M_h \) implies more incoming gas available to fuel the AGNs; as a consequence, the AGNs
can reach higher luminosities (higher $L_{\text{sat}}$). In addition, the scaling between $L_{\text{sat}}$ and $M_h$ evolves from $L_{\text{sat}} \propto M_h^{1/2}$ at $z \sim 0.6$ to $L_{\text{sat}} \propto M_h$ at $z \sim 2$, thereby becoming steeper at higher redshifts for $M_h \gtrsim 10^{14}$ $\text{M}_\odot/h$ haloes. This is expected if the halo merger rate increases with redshift, which is indeed the case (Fakhouri et al. 2010). Thus, while we model the one-halo clustering as originating primarily from central-satellite quasar pairs, it is likely that these are actually central AGNs that recently became "satellites" after their host haloes underwent a merger.

We calculate the AGN satellite fraction as defined by $f_{\text{sat}} \equiv \Phi_{\text{sat}}/(\Phi_{\text{sat}} + \Phi_{\text{cen}})$. Figure 9 shows the redshift evolution of $f_{\text{sat}}$ from $z \sim 0.6$ to $z \sim 2$. For $M_g > -22$ satellites, the satellite fraction is $\sim 20$–30% at all redshifts. For $M_g < -22$ satellites, the satellite fraction is 10% for $z \sim 1.5, 2.0$ but drops to 0.1% at $z \sim 0.6, 1.0$, implying that satellite quasars are significantly more abundant (by factors of $10^2$–$10^3$) at $z \sim 2.0$ compared to at $z \sim 0.6$. In order to explain the significant increase in satellite quasars at $z \sim 1.5, 2$, we show the evolution of the $L_{\text{sat}}$-$M_h$ relation from $z \sim 0.6$ to $z \sim 2$ in the inset panel of Figure 9 (the shaded regions are from CLF modeling and the filled circles are for simulations). We see that for $M_h \gtrsim 10^{13}$ $\text{M}_\odot/h$ haloes, $L_{\text{sat}}$ values are significantly higher (by factors of 10–100) at $z \sim 1.5, 2$ compared to at $z \sim 0.6, 1$. This implies that satellite fractions of $M_g \lesssim -22$ quasars increase significantly from $z \sim 0.6$ to $z \sim 2$.

Finally, it is worth noting that our CLF models, when compared to the observed measurements, imply halo masses of $\gtrsim 10^{14}$ $\text{M}_\odot/h$ for the SDSS $g < 20.85$ quasar pairs, which is consistent with our simulated quasar pairs.

8 FORECASTS FOR EBOSS QUASARS

We now use our CLF model to make predictions for spatial clustering of quasars in the ongoing eBOSS survey (Dawson et al. 2016). eBOSS is expected to detect $\gtrsim 500,000$ quasars with $g < 22$ between $z \sim 0.9$ to 2.2 (Myers et al. 2015). Figure 10 presents the one-halo and two-halo contributions of $W_p$ at a target redshift of $z = 1.5$. We note that the one-halo term ($r_p \lesssim 1$ Mpc/$h$) is somewhat enhanced and steeper.
compared to the two-halo term ($r_p \gtrsim 1$ Mpc$/h$), which is expected given the high satellite fractions at $z \sim 1.5$ discussed in the previous section. Strong clustering at small scales, consistent with a high satellite fraction, has been reported in multiple measurements of small-scale quasar clustering (e.g. Hennawi et al. 2006; Myers et al. 2007, 2008; Kayo & Oguri 2012; Eftekharzadeh et al. 2017).

We focus on scales targeted by E17 i.e. $17.7 \lesssim r_p \lesssim 36$ kpc$/h$ (the shaded region in Figure 10). At these scales the clustering amplitude $W_p$ of eBOSS quasar pairs is predicted to be $\sim 100$–200. We can use the clustering amplitude to calculate the expected number of quasar pairs within the survey area of eBOSS i.e. $\sim 7500$ deg$^2$ (Dawson et al. 2016). For a redshift bin-width of 0.46 (the same as the sample of E17) centered at $z = 1.5$, we expect 200–500 quasar pairs at scales of $17.7 \lesssim r_p \lesssim 36$ kpc$/h$. This is $\gtrsim 10$ times larger than the E17 sample. Additionally, our CLF model predicts that binary quasars in the eBOSS-CORE sample are expected to have host halo masses of $\gtrsim 10^{13} M_\odot/h$, which is consistent with the host halo masses of our simulated $g < 22$ satellite AGNs.

9 SUMMARY AND CONCLUSION

In this work, we have analysed recent high-precision small-scale clustering measurements of SDSS quasar pairs ($g < 20.85$) using the MBII simulation and CLF modeling. Within the MBII volume, there are 3 quasar pairs at these high luminosities at distances of $\sim 1$–4 Mpc. Within $M_h \sim 1$–$3 \times 10^{14} M_\odot/h$ haloes, they reside in separate central and satellite galaxies both of which have stellar masses $M_\star \gtrsim 1$–$3 \times 10^{12} M_\odot$ and black halo masses $1$–$3 \sim 10^9 M_\odot/h$. These galaxies and black holes are among the most massive in the entire simulation. The growth history of these quasars revealed that for all three of the pairs, their progenitor black holes lived in massive holes (the shaded region in Figure 7) centered at $z = 0.6$ (the same as the sample of E17). The growth history of these quasars is consistent with the host halo masses of our simulated $g < 22$ satellite AGNs.
by log-normal and Schechter distributions respectively. Assuming that the central AGNs dominate over satellites in the luminosity function, we built a CLF model for central AGNs which predicts an AGN luminosity function consistent with that of the simulated MBII AGNs as well as observed quasars. For the satellite AGNs, based on the trends exhibited by the CLFs of MBII AGNs in $M_h \lesssim 10^{14} M_\odot/h$ haloes, we considered a family of CLF models to extrapolate to $M_h \gtrsim 10^{14} M_\odot/h$ haloes and make predictions for the small-scale (one-halo) clustering; in particular, the central-satellite term. We compare these predictions to small-scale clustering measurements from observations at $\sim 25$ kpc scales, and arrive at a final model which is consistent with the measurements. Our key parameter of interest is $L_{\text{sat}}^*$, which is the maximum luminosity of a satellite AGN in a halo of a given mass $M_h$. Constraining our model with the observed measurements leads to three interesting findings about the $L_{\text{sat}} - M_h$ relation:

(i) $L_{\text{sat}}^*$ has significant positive correlation with $M_h$ for $M_h \gtrsim 10^{14} M_\odot/h$ haloes.

(ii) The correlation gets stronger with redshift evolving from $L_{\text{sat}}^* \sim M_h^{1/2}$ at $z = 0.6$ to $L_{\text{sat}}^* \sim M_h$ at $z = 2$.

(iii) For fixed halo mass, $L_{\text{sat}}^*$ steeply increases (by 2–3 orders of magnitude) from $z \sim 0.6$ to $z \sim 2$ for $M_h \gtrsim 10^{14} M_\odot/h$ haloes. This leads to a significant increase in

---

Figure 6. **Top Panels:** Blue and green circles show the luminosity functions (LFs) of central and satellite AGNs respectively; error bars are 1σ Poisson errors. Filled and open black squares show the observed measurements. **Middle Panels:** The red and green dashed lines (enclosing the shaded blue region) show the CLF model predictions of central AGN LFs for values of $\beta_{\text{cen}}$ shown in Figure 5 (top panels); these models produce LFs consistent with observations. **Bottom Panels:** Dashed lines show the CLF model predictions of satellite AGN LFs for the values of $\beta_{\text{sat}}$ shown in the bottom panels of Figure 5. The red and green dashed lines (enclosing the shaded green region) correspond to the range of models which produce a small-scale clustering prediction consistent with current constraints (as shown in Figure 7).
Figure 7. $W_p$ is the volume-averaged projected correlation function averaged over $17.0 < r_p < 36.5$ kpc/h for quasars brighter than a given magnitude threshold, which we denote by ‘$g(<)$’. The dashed lines correspond to predictions from the different CLF models (corresponding to the different values of $\beta_{\text{sat}}$) that are considered in the bottom panels of Figures 5 and 6. The black squares correspond to the observational constraints from E17 at $g < 20.85$. The red and green dashed lines (enclosing the green shaded region) correspond to the range of models that is consistent with the observed small-scale clustering measurements. The filled circles are predictions from MBII with error bars showing 1$\sigma$ Poisson errors.

Figure 8. Redshift evolution of $\beta_{\text{cen}}$ and $\beta_{\text{sat}}$ from $z = 0.6$ to $z = 2$. Error bars correspond to the range of values obtained from the current observational constraints on the luminosity functions (for $\beta_{\text{cen}}$) and small-scale clustering (for $\beta_{\text{sat}}$).

The AGN satellite fraction from $f_{\text{sat}} \sim 10^{-3}$ at $z = 0.6$ to $f_{\text{sat}} \sim 10^{-1}$ at $z = 2$ for $M_g \lesssim -22$ quasars.

These findings are consistent with a scenario where binary quasars are triggered by mergers of massive $\gtrsim 10^{14} M_\odot/h$ haloes (as seen for our simulated pairs). In this scenario, a merger of two such haloes can funnel a significant amount of gas to a black hole in a relatively short time, thereby increasing the activity in satellite AGNs to $\sim 10^{44.5-45}$ ergs/sec, making them $\sim 10^2 - 10^3$ times more luminous compared to a typical satellite AGN (with mean luminosity of $\sim 10^{41.5-42}$ ergs/sec, see Figure 1). Therefore, these mergers potentially affect the most luminous “edge” of the satellite CLF, quantified by $L_{\text{sat}}$. To explain point i) in our summary, above, mergers of increasingly massive haloes are accompanied by increasing amounts of incoming gas now available to feed the black holes. Thus, quasars can reach increasingly higher luminosities with increasing halo mass, thereby leading to a positive correlation between $L_{\text{sat}}$.
and $M_h$. Points ii) and iii) can then be explained by the fact that the merger rate increases with redshift (Fakhouri et al. 2010). Higher numbers of satellites at $z \sim 1.5$ leads to enhanced one-halo clustering (at $r_p \lesssim 1$ Mpc/h).

Finally, for the ongoing eBOSS-CORE sample ($g < 22$), we predict a small-scale clustering amplitude of $W_p \sim 100–200$ at $17.7 \lesssim r_p \lesssim 36$ kpc/h. This corresponds to $\sim 200–500$ pairs (with separations of $< 2000$ km/s) expected at these scales at $z \sim 1.5$. Furthermore, these pairs are expected to live in haloes with $M_h \gtrsim 10^{13} M_\odot/h$ and have black-hole masses $M_{bh} \gtrsim 10^8 M_\odot/h$. This predicted sample is $\gtrsim 10$ times larger than the size of current samples of quasar pairs.

10 GENERAL REMARKS AND FUTURE WORK

Our work demonstrates that hydrodynamic simulations are invaluable tools to study properties of AGN and quasar populations. In particular, simulations such as MBII can probe faint AGNs ($M_g \gtrsim -18$) which cannot be accessed by observations because their AGN activity is masked by the luminosity of star formation activity within the host galaxy. CLF modeling was a crucial tool in our work to establish the link between faint ($M_g \gtrsim -18$) AGNs (which are difficult to access by observations) and bright ($M_g \lesssim -22$) quasars (which are difficult to access by simulations). It helped us build a model for the AGN-halo connection across a very wide range of luminosities ($-16 \gtrsim M_g \gtrsim -32$) and study the redshift evolution of quasars from $z = 0.6$ to $z = 2$.

It is however important to recognize that our CLF modeling does rely on some implicit assumptions about the AGN population. They are as follows:

- We assumed that central AGNs are the dominant contributor to the luminosity function across the entire range of AGN luminosities ($-16 \gtrsim M_g \gtrsim -32$). This enabled us to use the observed quasar luminosity function to constrain the CLF model for central AGNs in haloes not well probed by MBII simulation i.e. $M_h \gtrsim 10^{14} M_\odot/h$.
- We assumed that the model CLFs for satellite and central AGNs follow Schechter and log-normal distributions respectively in $M_h \gtrsim 10^{14} M_\odot/h$ haloes (which are not well probed by simulations).
- Among the CLF model parameters, $L^*_\text{cen}$ and $L^*_\text{sat}$ (and their dependence on halo mass at $M_h \gtrsim 10^{14} M_\odot/h$) primarily determine the quasar luminosity function at $M_g < -22$ and the small-scale clustering of $g < 20-85$ quasars; therefore we only allowed 2 parameters ($\beta_{\text{cen}}$ and $\beta_{\text{sat}}$) to vary in our initial family of CLF models. The modeling of other parameters ($\sigma_{\text{cen}}, M_{\alpha_{\text{sat}}}, Q_{\text{sat}}$) were fixed based on the trends seen in the MBII simulation (see Appendix A for details).

Relaxing one or more of the above assumptions greatly increases the complexity of the problem because the parameter space becomes large (5 parameters). Future work should involve the use of more sophisticated techniques (Markov chain Monte Carlo for example) to constrain the CLF models without depending on one or more of these assumptions, and therefore look for potential degeneracies in the modeling.

Efforts will also continue to exploit the rapid progress in computational power to push hydrodynamic simulations to larger volumes. This may ultimately allow simulations to directly probe the statistical properties of close pairs of quasars.

ACKNOWLEDGMENTS

AKB, TD, SE and ADM were supported by the National Science Foundation through grant number 1616168. ADM also acknowledges support through the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0019022. AKB is thankful to Duncan Campbell and Francois Lamusse for useful discussions.

APPENDIX A: MODELING CLFS

A1 Central AGNs:

We modeled the central CLF as a log-normal distribution shown in Eq. (10) with mean luminosity $L^\ast_{\text{cen}}(M_h)$ for a given halo mass $M_h$. Section 4 discusses the modeling of $L^\ast_{\text{cen}}(M_h)$. Here, we complete the discussion of CLF modeling of central AGNs by presenting the other parameter $\sigma_{\text{cen}}(M_h)$, which quantifies the scatter of log-normal distribution.

Filled circles in Figure A1 (left panel) show the best fit values of $\sigma_{\text{cen}}$ in various halo mass bins. We find that the values of $\sigma_{\text{cen}}$ range from 0.4–0.7 with a mild increase with halo mass. We find no obvious trend in the redshift evolution. We model the mass dependence using a linear regression between $\sigma_{\text{cen}}$ and log$_{10} M_h$, which are shown as solid lines in the left-hand panel of Figure A1.

A2 Satellite AGNs

We modeled the satellite CLF as a Schechter distribution shown in Eq. (11) with a maximum luminosity $L^\ast_{\text{sat}}(M_h)$ for
a given halo mass $M_h$. Section 4 discusses the modeling of $L_{sat}(M_h)$. Here, we complete the discussion of CLF modeling of satellite AGNs by presenting the remaining parameters $\alpha_{sat}(M_h)$ and $Q_{sat}(M_h)$.

$\alpha_{sat}$ measures the slope of the Schechter distribution for $L_{sat} < L_{sat}$. Filled circles in Figure A1 (middle panel) show the best fit values of $\alpha_{sat}$ in various halo mass bins. We find that $\alpha_{sat}$ increases with $M_h$ as a power law up to $M_h \sim 10^{14}, 10^{14.5}, 10^{15}, 10^{15.5} M_\odot/h$ for $z = 0.6, 1.0, 1.5, 2.0$ respectively; the relations are therefore modelled as solid lines as shown in Figure A1 (middle panel). For more massive haloes, simulations indicate that the $\alpha_{sat}$ vs. $M_h$ relation flattens; in this regime, we simply model $\alpha_{sat}$ as a constant value independent of $M_h$ shown as dashed lines. This makes sense because otherwise, a naive extrapolation of the solid lines would imply $\alpha_{sat} > 0$ for large enough halo masses; this would invert the slope of the Schechter distribution implying that AGN abundance increases with luminosity, which is somewhat unphysical.

The final parameter to be modelled is $Q_{sat}(M_h)$, which simply determines the overall normalization of the Schechter function, given the shape parameters $\alpha_{sat}$ and $L_{sat}^*$. Interestingly, we find that $Q_{sat}$ has a tight correlation with $\alpha_{sat}$ as shown in the right-hand panel of Figure A1. Furthermore, their relation does not exhibit a significant redshift dependence. We therefore simply model the dependence of $Q_{sat}$ and $\alpha_{sat}$ as a linear regression shown as the solid black line; the relation is given as:

$$\log_{10} Q_{sat} = (3.0 \pm 0.1) \alpha_{sat} + (1.25 \pm 0.1)$$

(A1)

where the uncertainties in the coefficients are 1σ errors in the regression. Eq. (A1) along with Tables A1 and A1 summarize our final CLF model parameters.

**APPENDIX B: RADIAL PROFILES OF SATELLITE AGNS IN MBII**

For a given satellite occupation, the radial satellite distributions determine how the small-scale clustering power is distributed across different scales. Figure B1 shows the radial profiles of satellite AGNs around the central AGN averaged over all haloes of mass $M_h > 10^{11} M_\odot/h$ at $z = 1$. We find that across the entire range of luminosities (shown as different colors), profiles trace a power-law with exponent $\sim -2$. This is consistent with previous works on AGN HODs using smaller volume (33.75 cMpc/h) hydrodynamic simulations (Chatterjee et al. 2012). Therefore, for the modeling of the one-halo term, we use a satellite profile with exponent $-2$.

In order to validate our CLF model and our underlying assumptions, we compare the projected clustering predictions against the direct predictions of our simulations for AGNs from $g < 26$ up to $g < 23$ (the regime that is well probed by our simulations). Figure B2 shows the projected clustering profiles. The filled circles show the MBII predictions and solid lines show the CLF model predictions. We can see that the simulations and CLF model predictions are within reasonable agreement for a wide range of redshifts and magnitude thresholds.
bars are covariance error estimates obtained from the linear fits.

Chatterjee S., Nguyen M. L., Myers A. D., Zheng Z., 2013, ApJ,

Chatterjee S., Degraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, MNRAS, 419, 2657

Chatterjee S., Nguyen M. L., Myers A. D., Zheng Z., 2013, ApJ, 779, 147

Cooray A., 2006, MNRAS, 365, 842

Croom S. M., et al., 2005, MNRAS, 356, 415

Croom S. M., et al., 2009, MNRAS, 399, 1755

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dawson K. S., et al., 2016, AJ, 151, 44

DeGraf C., Sijacki D., 2017, MNRAS, 466, 3331

Table A1. Summary of the CLF model for central AGNs. $L_{\text{cen}}^*$ and $M_h$ are written in units of ergs/sec and $M_\odot/h$ respectively. Error bars are covariance error estimates obtained from the linear fits.

| Redshift | $M_\text{H}^*$ ($M_\odot/h$) | Halo Mass Range | CLF model for Central AGNs |
|----------|-------------------------------|----------------|-----------------------------|
| 0.6      | $10^{13.8}$                 | $M_h \leq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{30.0^{+0.2}_{-0.2}}(M_h - M_\text{H}^*)^{1.2^{+0.2}_{-0.2}}; \sigma_{\text{cen}} = (0.07 \pm 0.01) \log_{10} M_h + (0.32 \pm 0.18)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{44.1}(M_h - M_\text{H}^*)^{1.2^{+0.2}_{-0.2}}; \sigma_{\text{cen}} = (0.07 \pm 0.01) \log_{10} M_h + (0.32 \pm 0.18)$ |
| 1.0      | $10^{13.5}$                 | $M_h \leq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{29.9^{+0.7}_{-0.7}}(M_h - M_\text{H}^*)^{1.0^{+0.5}_{-0.5}}; \sigma_{\text{cen}} = (0.05 \pm 0.02) \log_{10} M_h + (-0.07 \pm 0.23)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{44.1}(M_h - M_\text{H}^*)^{1.7^{+0.2}_{-0.2}}; \sigma_{\text{cen}} = (0.05 \pm 0.02) \log_{10} M_h + (-0.07 \pm 0.23)$ |
| 1.5      | $10^{13}$                   | $M_h \leq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{30.5^{+0.4}_{-0.4}}(M_h - M_\text{H}^*)^{2.0^{+0.4}_{-0.4}}; \sigma_{\text{cen}} = (0.02 \pm 0.01) \log_{10} M_h + (0.31 \pm 0.06)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{44.8}(M_h - M_\text{H}^*)^{2.0^{+0.4}_{-0.4}}; \sigma_{\text{cen}} = (0.02 \pm 0.01) \log_{10} M_h + (0.31 \pm 0.06)$ |
| 2.0      | $10^{13}$                   | $M_h \leq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{26.7^{+1.7}_{-1.7}}(M_h - M_\text{H}^*)^{2.0^{+0.4}_{-0.4}}; \sigma_{\text{cen}} = (0.05 \pm 0.04) \log_{10} M_h + (-0.06 \pm 0.50)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{cen}}^* = 10^{44.8}(M_h - M_\text{H}^*)^{2.0^{+0.4}_{-0.4}}; \sigma_{\text{cen}} = (0.05 \pm 0.04) \log_{10} M_h + (-0.06 \pm 0.50)$ |

Table A2. Summary of the CLF model for satellite AGNs. $L_{\text{sat}}^*$ and $M_h$ are written in units of ergs/sec and $M_\odot/h$ respectively. Error bars are covariance error estimates obtained from the linear fits.

| Redshift | $M_\text{H}^*$ ($M_\odot/h$) | Halo Mass Range | CLF model for Satellite AGN |
|----------|-------------------------------|----------------|----------------------------|
| 0.6      | $10^{13.8}$                 | $M_h \leq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{47.5^{+0.1}_{-0.1}}(M_h - M_\text{H}^*)^{0.5^{+0.1}_{-0.1}}; \alpha_{\text{sat}} = (0.56 \pm 0.02) \log_{10} M_h + (-8.2 \pm 0.2)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{44.5}(M_h - M_\text{H}^*)^{0.5^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = -0.29 \pm 0.01$ |
| 1.0      | $10^{13.5}$                 | $M_h \leq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{42.4^{+0.1}_{-0.1}}(M_h - M_\text{H}^*)^{0.6^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = (0.43 \pm 0.07) \log_{10} M_h + (-6.3 \pm 1.0)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{44.3}(M_h - M_\text{H}^*)^{1.0^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = -0.51 \pm 0.01$ |
| 1.5      | $10^{13}$                   | $M_h \leq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{36.2^{+0.2}_{-0.2}}(M_h - M_\text{H}^*)^{0.6^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = (0.40 \pm 0.05) \log_{10} M_h + (-5.8 \pm 0.7)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{44.8}(M_h - M_\text{H}^*)^{1.0^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = -0.57 \pm 0.02$ |
| 2.0      | $10^{13}$                   | $M_h \leq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{37.2^{+0.2}_{-0.2}}(M_h - M_\text{H}^*)^{0.6^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = (0.45 \pm 0.06) \log_{10} M_h + (-6.4 \pm 0.8)$ |
|          |                              | $M_h \geq M_\text{H}^*$ | $L_{\text{sat}}^* = 10^{44.8}(M_h - M_\text{H}^*)^{1.2^{+0.2}_{-0.2}}; \alpha_{\text{sat}} = -0.42 \pm 0.01$ |

Figure B2. Solid lines show the predicted one-halo and two-halo contributions of the projected clustering profiles of our CLF model summarized in Tables A1 and A2. The filled circles are the predictions from MBII.

REFERENCES

Ballantyne D. R., 2017, MNRAS, 464, 613

Chatterjee S., Degraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, MNRAS, 419, 2657

Chatterjee S., Nguyen M. L., Myers A. D., Zheng Z., 2013, ApJ, 779, 147

Cooray A., 2006, MNRAS, 365, 842

Croom S. M., et al., 2005, MNRAS, 356, 415

Croom S. M., et al., 2009, MNRAS, 399, 1755

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dawson K. S., et al., 2016, AJ, 151, 44

DeGraf C., Sijacki D., 2017, MNRAS, 466, 3331
On the small scale clustering of quasars: constraints from the MassiveBlack II simulation

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Di Matteo T., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, ApJ, 745, L29
Eltekharzadeh S., et al., 2015, MNRAS, 453, 2779
Eltekharzadeh S., Myers A. D., Hennawi J. F., Djorgovski S. G., Richards G. T., Mahabal A. A., Graham M. J., 2017, MNRAS, 468, 77
Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267
Guo H., et al., 2016, MNRAS, 459, 3040
Harikane Y., et al., 2018, PASJ, 70, S11
Hennawi J. F., et al., 2006, AJ, 131, 1
Kaviraj S., et al., 2017, MNRAS, 467, 4739
Kayo I., Oguri M., 2012, MNRAS, 424, 1363
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349
Kochanek C. S., Falco E. E., Muñoz J. A., 1999, ApJ, 510, 590
Komatsu E., et al., 2011, ApJS, 192, 18
McGreer I. D., Eltekharzadeh S., Myers A. D., Fan X., 2016, AJ, 151, 61
Mitra K., Chatterjee S., DiPompeo M. A., Myers A. D., Zheng Z., 2018, MNRAS, 477, 45
Mortlock D. J., Webster R. L., Francis P. J., 1999, MNRAS, 309, 836
Myers A. D., et al., 2006, ApJ, 638, 622
Myers A. D., Brunner R. J., Richards G. T., Nichol R. C., Schneider D. P., Bahcall N. A., 2007, ApJ, 658, 99
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
Myers A. D., et al., 2015, ApJS, 221, 27
Nelson D., et al., 2015, Astronomy and Computing, 13, 12
Porciani C., Norberg P., 2006, MNRAS, 371, 1824
Richards G. T., et al., 2006, AJ, 131, 2766
Richardson J., Zheng Z., Chatterjee S., Nagai D., Shen Y., 2012, ApJ, 755, 30
Schaye J., et al., 2015, MNRAS, 446, 521
Schneider D. P., et al., 2000, AJ, 120, 2183
Shen Y., 2013, Bulletin of the Astronomical Society of India, 41, 61
Shen Y., et al., 2009, ApJ, 697, 1656
Springel V., 2005, MNRAS, 364, 1105
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Trevisan M., Mamon G. A., 2017, MNRAS, 471, 2022
White M., et al., 2012, MNRAS, 424, 933
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760