

# The dark matter haloes of moderate luminosity X-ray AGN as determined from weak gravitational lensing and host stellar masses

Alexie Leauthaud,<sup>1\*</sup> Andrew J. Benson,<sup>2</sup> Francesca Civano,<sup>3</sup> Alison L. Coil,<sup>4</sup>  
Kevin Bundy,<sup>1</sup> Richard Massey,<sup>5</sup> Malte Schramm,<sup>1</sup> Andreas Schulze,<sup>1</sup> Peter Capak,<sup>6</sup>  
Martin Elvis,<sup>7</sup> Andrea Kulier<sup>8</sup> and Jason Rhodes<sup>9,10</sup>

<sup>1</sup>Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), The University of Tokyo, Kashiwa, Chiba 277-8582, Japan

<sup>2</sup>Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>3</sup>Yale Center for Astronomy and Astrophysics, 260 Whitney ave, New Haven, CT 06520, USA

<sup>4</sup>Department of Physics, Center for Astrophysics and Space Sciences, University of California at San Diego, 9500 Gilman Dr., La Jolla, San Diego, CA 92093, USA

<sup>5</sup>Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK

<sup>6</sup>Spitzer Science Center, 314-6 Caltech, Pasadena, CA 91125, USA

<sup>7</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA

<sup>8</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

<sup>9</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

<sup>10</sup>California Institute of Technology, Pasadena, CA 91125, USA

Accepted 2014 October 21. Received 2014 September 30; in original form 2014 August 10

## ABSTRACT

Understanding the relationship between galaxies hosting active galactic nuclei (AGN) and the dark matter haloes in which they reside is key to constraining how black hole fuelling is triggered and regulated. Previous efforts have relied on simple halo mass estimates inferred from clustering, weak gravitational lensing, or halo occupation distribution modelling. In practice, these approaches remain uncertain because AGN, no matter how they are identified, potentially live a wide range of halo masses with an occupation function whose general shape and normalization are poorly known. In this work, we show that better constraints can be achieved through a rigorous comparison of the clustering, lensing, and cross-correlation signals of AGN hosts to the fiducial stellar-to-halo mass relation (SHMR) derived for all galaxies, irrespective of nuclear activity. Our technique exploits the fact that the global SHMR can be measured with much higher accuracy than any statistic derived from AGN samples alone. Using 382 moderate luminosity X-ray AGN at  $z < 1$  from the COSMOS field, we report the first measurements of weak gravitational lensing from an X-ray-selected sample. Comparing this signal to predictions from the global SHMR, we find that, contrary to previous results, most X-ray AGN do not live in medium size groups – nearly half reside in relatively low mass haloes with  $M_{200b} \sim 10^{12.5} M_{\odot}$ . The AGN occupation function is well described by the same form derived for all galaxies but with a lower normalization – the fraction of haloes with AGN in our sample is a few per cent. The number of AGN satellite galaxies scales as a power law with host halo mass with a power-law index  $\alpha = 1$ . By highlighting the relatively ‘normal’ way in which moderate luminosity X-ray AGN hosts occupy haloes, our results suggest that the environmental signature of distinct fuelling modes for luminous quasars compared to moderate luminosity X-ray AGN is less obvious than previously claimed.

**Key words:** galaxies: abundances – galaxies: active – galaxies: haloes – galaxies: Seyfert – galaxies: stellar content.

## 1 INTRODUCTION

Strong observational evidence suggests a tight coupling between the growth of supermassive black holes (BHs) and the build-up of galaxy bulges (Ferrarese & Merritt 2000; Gebhardt et al. 2000).

\* E-mail: alexie.leauthaud@ipmu.jp

In contrast, we only have a limited understanding of how BH activity relates to dark matter *halo mass* because both halo masses and BH masses are challenging to probe observationally. Improved measurements of the BH–dark matter relation are, however, of great theoretical interest and are key in order to facilitate a more direct comparison between observations and theoretical models of active galactic nuclei (AGN) activity (e.g. Chatterjee et al. 2012; DeGraf et al. 2012; Fanidakis et al. 2013; Hütsi, Gilfanov & Sunyaev 2014).

The AGN–halo mass relation is typically probed by measuring the clustering (e.g. Li et al. 2006; Coil et al. 2009; Gilli et al. 2009; Allevalo et al. 2011; Miyaji et al. 2011; Krumpel et al. 2012; Koutoulidis et al. 2013; Mountrichas et al. 2013; Shen et al. 2013) or the weak gravitational lensing of AGN host galaxies (Mandelbaum et al. 2009). Halo masses (hereafter  $M_h$ ) are typically inferred from these types of approaches by measuring the mean large-scale bias of a given sample. Bias values are then translated into an effective halo mass via the halo mass–bias relation (e.g. Tinker et al. 2010). However, there are several important caveats to this approach. First, large-scale bias is not a sensitive probe of halo mass at lower mass scales (the halo mass–bias relation flattens). Secondly, converting large-scale bias to halo mass requires assumptions about satellite fractions. Thirdly, the effective halo mass measured in this fashion corresponds to a bias-weighted average of the underlying halo mass distribution. For samples which span a wide *range* of halo masses, there is no simple way to relate this effective halo mass to more useful averages such as the mode, mean, or median value of the halo mass distribution.

In principle, these issues can be resolved by adopting an halo occupation distribution (HOD)-type approach which assumes a parametric model to describe the probability distribution  $P(N|M_h)$  that a halo of mass  $M_h$  is host to  $N$  galaxies in a given sample (for a review, see Cooray & Sheth 2002). While an HOD-type approach may work well for galaxy samples defined by simple luminosity thresholds, it is less clear which parametric form should be adopted for occupation functions when considering AGN-type samples (e.g. Allevalo et al. 2011; Miyaji et al. 2011; Kayo & Oguri 2012; Richardson et al. 2013). The AGN duty cycle relative to haloes is unknown, which leads to large uncertainties in both the *shape and normalization* of the AGN occupation functions. Recently, Shen et al. (2013) measured the cross-correlation between quasars (QSOs) and luminous red galaxies (LRGs) from the Sloan Digital Sky Survey (SDSS) at  $z = 0.5$ . Despite the high signal to noise of their cross-correlation measurement, Shen et al. (2013) find that substantially different HODs provide equally good fits to their data. The conclusions from this work suggest that clustering data alone is insufficient to fully constrain the QSO HOD – underlining the difficulty of modelling AGN-type populations.

For samples of less luminous AGN, such as those selected via deep X-ray imaging, these issues are even more pronounced. Typical sample sizes are small, commonly ranging from a few hundred to a few thousand AGN which means that clustering measurements are noisy. To compensate for small sample sizes, many studies measure AGN clustering over a broad range in redshift ( $0 < z < 3$  is not uncommon), X-ray luminosity (hereafter  $L_X$ ), and host galaxy properties (e.g. Coil et al. 2009; Allevalo et al. 2011; Koutoulidis et al. 2013). Even greater caution is required when interpreting HODs or bias measurements in this context.

For moderate luminosity obscured (type-2) AGN samples, however, information about the properties of the host galaxy contains key additional information which has yet to be fully exploited for these types of studies. For obscured systems, the host galaxy light is the dominant component in the optical/near-infrared spec-

tral energy distribution (SED), meaning that the stellar mass of the host galaxy (hereafter  $M_*$ ) can be measured with relatively little contamination from the AGN component.

In this paper, we propose an alternative approach to analysing clustering and/or lensing measurements of moderate luminosity obscured AGN samples that can be employed even with small samples. Our approach relies on using a complete galaxy sample to *first* constrain the overall connection between galaxy mass and halo mass. This model then serves as a fiducial baseline with which to explore the AGN–halo mass relation.

From a global perspective that includes all galaxies, tremendous progress has had been made in recent years in terms of understanding and modelling the connection between galaxy stellar mass and dark matter halo mass out to  $z = 1$  and beyond (Mandelbaum et al. 2006; More et al. 2009; Yang, Mo & van den Bosch 2009; Behroozi, Conroy & Wechsler 2010; Moster et al. 2010; Leauthaud et al. 2011, 2012a). At the core of these models is the stellar-to-halo mass relation (SHMR) for central galaxies. This may be constrained from measurements of either the galaxy stellar mass function (SMF), galaxy clustering, galaxy–galaxy weak lensing, satellite kinematics, or some combination of these four probes. In detail, methods vary between different groups, but all results yield the same global picture:  $M_h(M_*)$  is well described by a power law at low  $M_*$  and then transitions to a more sharply rising function above a characteristic mass scale of  $M_* \sim 10^{10.8} M_\odot$ . The logarithmic scatter in stellar mass at fixed halo mass is also constrained at  $\sigma_{\log M_*} \sim 0.18$  with good agreement between different studies. In addition to the SHMR, these methods also constrain how satellite galaxies populate dark matter haloes as a function of galaxy mass. Finally, the SHMR may also be constrained as a function of other properties beyond stellar mass, such as galaxy colour or star formation activity (Mandelbaum et al. 2006; More et al. 2009; Hearin et al. 2014; Tinker et al. 2013).

In this paper, we suggest that whenever information about host mass is available, the AGN–dark matter relation can be probed most effectively by *first* constraining a fiducial SHMR. Once the SHMR is constrained, the distribution of AGN host stellar masses is all that is required to make predictions about AGN occupation statistics. The observed clustering, lensing, or cross-correlations between AGN and stellar mass limited samples may then be interpreted in light of predictions from the fiducial SHMR. The key advantage of this approach is that by using large samples of galaxies that are complete in terms of stellar mass, the SHMR can be built with much higher accuracy than by using any statistic measured from AGN samples alone. Statistics measured from AGN samples (which are necessarily noisy because of small sample sizes) are only used to constrain *deviations* from the fiducial model. Any observed deviations would be of great interest and would provide clues about the mechanisms that fuel AGN. Our method is similar in many respects to the one adopted by Li et al. (2006) and Mandelbaum et al. (2006) for analysing optically selected and radio-loud AGN.

The approach used here alleviates the difficulties raised by Shen et al. (2013) associated with HOD modelling of AGN clustering. However, it can only be employed for samples with host stellar mass measurements and therefore cannot be applied in the context of bright QSO type samples. For these, however, an alternative and closely related approach has been recently developed by Conroy & White (2013) by combining the SHMR with a BH mass–stellar mass relation.

We apply our methodology to a sample of X-ray-selected moderate luminosity obscured AGN at  $z < 1$  from the COSMOS field (Scoville et al. 2007). Despite the small sample size (several hundred AGN), we are able to place robust constraints on AGN halo

occupation statistics. Our choice of the COSMOS field is motivated by the fact that the galaxy SHMR has been previously constrained for this field by Leauthaud et al. (2012a, hereafter L12). The L12 SHMR is determined from measurements of the galaxy mass function, galaxy clustering, and galaxy–galaxy lensing to  $z = 1$ . Here, for the first time, we measure the galaxy–galaxy lensing signal of X-ray-selected obscured AGN. We use this signal to test for differences between the dark matter environment of obscured AGN compared to the overall galaxy population.

The layout of this paper is as follows. The data are described in Section 2 followed by the presentation of our methodology in Section 3. Our main results are presented in Section 4. Finally, we discuss the results and draw up our conclusions in Sections 5 and 6.

We assume a  $\Lambda$  cold dark matter cosmology with  $\Omega_m = 0.258$ ,  $\Omega_\Lambda = 0.742$ ,  $\sigma_8 = 0.796$ ,  $H_0 = 72 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . All distances are expressed in physical Mpc units. The letter  $M_h$  denotes halo mass in general, whereas  $M_{200b}$  is explicitly defined as  $M_{200b} \equiv M(< r_{200b}) = 200\bar{\rho} \frac{4}{3}\pi r_{200b}^3$ , where  $r_{200b}$  is the radius at which the mean interior density is equal to 200 times the mean matter density ( $\bar{\rho}$ ). Stellar mass is noted  $M_*$  and has been derived using a Chabrier initial mass function (IMF). Stellar mass scales as  $1/H_0^2$ . Halo mass scales as  $1/H_0$ . All magnitudes are given on the AB system.

## 2 DATA AND MOCK CATALOGUES

### 2.1 COSMOS X-ray AGN sample

The AGN sample used for this work is selected by combining the COSMOS *XMM–Newton* (*XMM–COSMOS*; Cappelluti et al. 2009) and *Chandra* (C-COSMOS; Elvis et al. 2009) X-ray catalogues. The *XMM–COSMOS* survey covers  $2 \text{ deg}^2$  to a limiting depth of  $5 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$  in the soft (0.5–2 keV) band and  $3 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$  in the hard (2–10 keV) band. The C-COSMOS survey covers  $0.9 \text{ deg}^2$  to a limiting depth of  $1.9 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$  in the soft band and  $7.3 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$  in the hard band. The combined catalogue of X-ray sources contains  $\sim 1800$  objects from *XMM–COSMOS* and  $\sim 950$  objects from C-COSMOS. Details concerning the X-ray catalogues, the spectroscopic observing programme, and the spectroscopic/photometric classifications can be found in Brusa et al. (2010), Civano et al. (2012), and Salvato et al. (2009, 2011).

Full band (0.5–10 keV) fluxes are provided in the C-COSMOS catalogue but are not available in the *XMM–COSMOS* catalogue. We compute the full band flux for *XMM–COSMOS* sources by summing fluxes in the soft and hard bands. If a source is not detected in one of the bands, only the detected flux is included. Rest-frame X-ray luminosities are homogeneously derived for both catalogues assuming a power-law spectral model with a slope of  $\Gamma = 2$  and absorption from a Galactic column density of  $N_{\text{H,Gal}} = 2.6 \times 10^{20} \text{ cm}^{-2}$  (Kalberla et al. 2005). Given that a flat slope has been assumed, no *K*-correction is needed.

The aim of this work is to consider moderately obscured and moderate luminosity AGN for which the host galaxy light is the dominant component in the optical/near-infrared SED. We select AGN in the redshift range  $0.2 < z < 1$ . All spectroscopically identified broad line AGN are removed from the sample. A photometric classification (Salvato et al. 2009, 2011) is used to identify obscured AGN when a spectroscopic classification is not available. Spectroscopic redshifts are available for 71 per cent of our sample (272/382). We also impose a lower limit on host mass of  $\log_{10}(M_*) > 10.5$  (see Section 2.2). This cut is designed to (only very) roughly match samples used in previous studies of the clus-

tering of X-ray-selected AGN (see Section 4). This mass cut is well above the COSMOS stellar mass completeness limit at  $z = 1$  ensuring that our sample is complete in terms of galaxy mass.

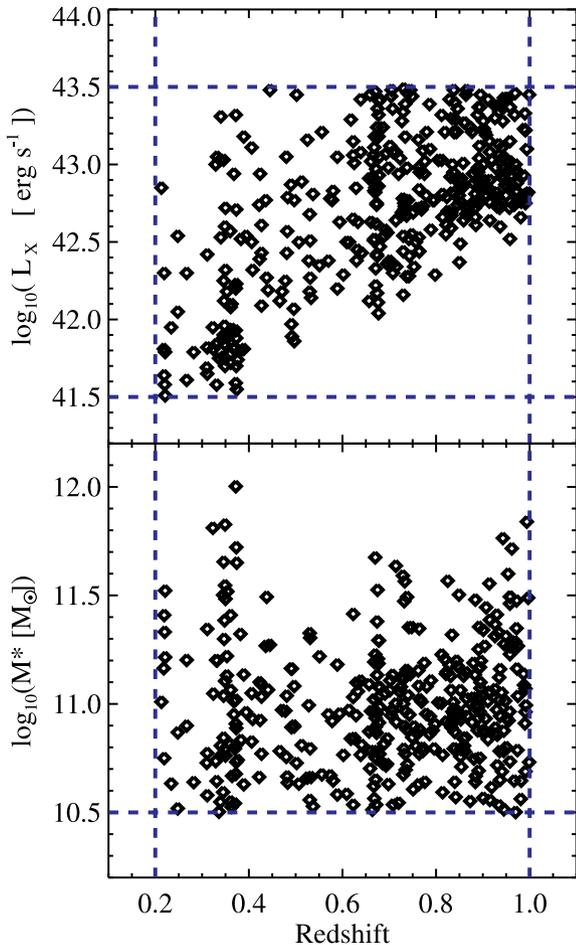
In addition, we also limit our sample to AGN with a rest-frame 0.5–10 keV band luminosity in the range  $10^{41.5} < L_X < 10^{43.5} \text{ erg s}^{-1}$ . The upper limit on  $L_X$  is set to avoid bright AGN which might contaminate the host galaxy light. The lower boundary on  $L_X$  is set to limit contamination from star-forming sources and early-type galaxies with pure hot gas X-ray emission (e.g. Civano et al. 2014). Our results are reasonably robust to contamination from galaxies outside our sample, provided these span a similar stellar mass range as our AGN sample. In this case, a 5 per cent contamination will have no impact on our mean/median halo mass estimates, and will simply modify the amplitude of our inferred HOD by 5 per cent. Our mean/median halo mass estimates are more sensitive to contamination from galaxies with preferentially low or high  $M_*$  values compared to our AGN sample. As an extreme example, if all the most massive (least massive) galaxies in our sample are contaminants (at the 5 per cent level), our mean halo masses will be biased by 8 per cent (2 per cent) and our median halo masses by 7 per cent (6 per cent).

The conclusions drawn in this paper are specific to the AGN sample described above. In particular, we do not probe all AGN down to our mass limit of  $\log_{10}(M_*) > 10.5$ . Because of the lower limit imposed on  $L_X$ , our sample will miss AGN with low Eddington ratios (Aird et al. 2012). In total, our sample contains 382 AGN with a mean redshift of  $\langle z \rangle = 0.7$ , a mean-log X-ray luminosity of  $\langle \log_{10}(L_X) \rangle = 42.7$ , and a mean stellar mass of  $\langle M_* \rangle = 1.3 \times 10^{11} M_\odot$ . Fig. 1 shows the  $L_X$  and  $M_*$  distributions for our sample.

### 2.2 Stellar masses

In this paper, we use the stellar mass-dependent SHMR models and mock catalogues from L12. For consistency, we adopt the same galaxy stellar mass estimates as derived in L12. Contamination of the optical light by emission from the AGN is a potential issue since our masses were derived using galaxy templates without an AGN component. However, this effect should be small – our sample of AGN have moderate luminosities ( $L_{\text{x,full}} < 10^{43.5} \text{ erg s}^{-1}$ ) and are not powerful enough to significantly affect the optical light of the host galaxy (Nandra et al. 2007; Bundy et al. 2008). To test for contamination effects, we compare our mass estimates with those from Bongiorno et al. (2012) which were derived using both galaxy and AGN templates. We find an overall offset of 0.18 dex between our mass estimates which is within the expected range of systematic uncertainties (Behroozi et al. 2010). More importantly, however, this mass offset does not exhibit any trends with  $L_X$  suggesting that our mass estimates are robust at these moderate luminosities.

Here, we only give a brief description of the stellar mass estimates and refer the reader to L12 and Bundy et al. (2010) for further details. Stellar mass estimates are based on point spread function (PSF)-matched 3.0 arcsec diameter aperture photometry from the ground-based COSMOS catalogues (filters  $u^*$ ,  $B_J$ ,  $V_J$ ,  $g^+$ ,  $r^+$ ,  $i^+$ ,  $z^+$ ,  $K_s$ ; Capak et al. 2007; Ilbert et al. 2009; McCracken et al. 2010). The depth in all bands reaches at least 25th magnitude (AB) with the  $K_s$ -band limited to  $K_s < 24$ . Stellar masses are derived using the Bayesian code described in Bundy et al. (2006) assuming a Chabrier IMF and a Charlot & Fall (2000) dust model. An observed galaxy’s SED and redshift is referenced to a grid of models constructed using the Bruzual & Charlot (2003) synthesis code. The grid includes models that vary in age, star formation history, dust content, and metallicity. At each grid point, the probability that the



**Figure 1.** Sample of 382 AGN host galaxies as a function of  $L_X$ , stellar mass, and redshift. Our selection cuts are shown by the blue dashed lines. The sample is not complete in terms of  $L_X$  but it is complete in terms of host stellar mass. The redshift-dependent structures that can be seen in this figure are due to the relatively small size of COSMOS. Our work accounts for sample variance using a suite of mock catalogues.

observed SED fits the model is calculated, and the corresponding stellar mass to  $K$ -band luminosity ratio and stellar mass is stored. By marginalizing over all parameters in the grid, the stellar mass probability distribution is obtained. The median of this distribution is taken as the stellar mass estimate.

### 2.3 Weak lensing catalogue

The COSMOS program has imaged the largest contiguous area ( $1.64 \text{ deg}^2$ ) to date with the *Hubble Space Telescope* (*HST*) using the Advanced Camera for Surveys (ACS) Wide Field Channel (Koekemoer et al. 2007). The imaging quality of ACS and the stability of the *HST* PSF makes this a prime data set with which to perform weak lensing measurements at  $z < 1$ . The details of the COSMOS weak lensing catalogue are already described in detail elsewhere (Leauthaud et al. 2007; Massey et al. 2007; Rhodes et al. 2007; L12). The COSMOS weak lensing catalogue contains  $3.9 \times 10^5$  galaxies with accurate shape measurements which represents a number density of 66 source galaxies per arcminute<sup>2</sup>.

The galaxy–galaxy lensing signals presented in Section 4.1 are measured following the same methodology as L12. The only minor difference compared to L12 is that here we use an updated version of

the COSMOS photo- $z$  catalogue (v1.8) of the photometric redshifts (hereafter photo- $z$ 's) presented in Ilbert et al. (2009) which have been computed with over 30 bands of multiwavelength data. This update to the photo- $z$  catalogue does not affect any of the lensing results.

### 2.4 Mock catalogues

The COSMOS ACS survey covers a relatively small volume. To estimate sample variance, we use a series of mock catalogues described in L11 and L12. These mocks are extracted from a  $1400^3$  particle,  $420 h^{-1} \text{ Mpc}$   $N$ -body simulation (‘Consuelo’ from the Las Damas suite) with a particle mass of  $1.87 \times 10^9 h^{-1} M_\odot$ <sup>12</sup> (McBride et al., in preparation). In this paper, we use 100 mock catalogues that are created from random lines of sight through the simulation volume for three redshift intervals:  $z_1 = [0.22, 0.48]$ ,  $z_2 = [0.48, 0.74]$ , and  $z_3 = [0.74, 1]$ . Mocks are populated with galaxies using the SHMR HOD model of L12. By design, this suite of mock catalogues matches the stellar-mass-dependent clustering and galaxy–galaxy weak lensing of COSMOS galaxies from  $0.2 < z < 1.0$ . The mocks are largely complete in terms of stellar mass for the  $\log_{10}(M_*) > 10.5$  sample considered in this paper. Mock galaxies have stellar masses, redshifts, halo masses, and a central/satellite identification flag.

## 3 METHODOLOGY

We begin with an outline of the rationale underlying our investigation. Our goal is to clearly sketch out the steps in our proposed methodology so that they may be easily followed by future studies. Although we focus here on a sample of moderate luminosity AGN, our methodology can be applied to any subpopulation with stellar mass measurements.

Our approach begins with the assumption that AGN can be described by the same SHMR as the overall galaxy population. Here, we use an SHMR parametrized as a function of stellar mass, but one could consider additional parameters, such as galaxy colour (e.g. Tinker et al. 2013). The details of the SHMR-based model<sup>3</sup> that we use and how it is constrained from COSMOS data are described in Leauthaud et al. (2011, 2012b). Other models based on the conditional SMF or abundance matching techniques would also our purpose (e.g. Yang, Mo & van den Bosch 2008; Behroozi, Conroy & Wechsler 2010; Moster et al. 2010; Hearin et al. 2014).

The first step in our methodology is to choose a statistic (or multiple statistics) to test the assumption that AGN can be described by the same SHMR as the overall galaxy population. In this paper, we use galaxy–galaxy lensing but our method can be applied to other statistics such as the AGN autocorrelation function, or cross-correlations between AGN and galaxies (ideally binned by stellar mass).

After computing the statistic of interest, the second step is to compare the results of this measurement with the prediction from the fiducial SHMR-based model. The goal of this step is to perform

<sup>1</sup> In this paragraph, numbers are quoted for  $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

<sup>2</sup> <http://lss.phy.vanderbilt.edu/lasdams/simulations.html>

<sup>3</sup> The L12 model uses an SHMR for central galaxies and an HOD-based prescription for satellite galaxies. For convenience, throughout this paper, we refer to the combined model (for centrals and satellites) as our ‘fiducial SHMR model’, even though technically speaking, the SHMR only refers to central galaxies.

a null-test of whether AGN populate dark matter haloes in the same fashion as the overall galaxy sample. Predictions from this model can be computed both analytically or from mock catalogues. Here, we mainly rely on mock catalogues to generate our predictions – these have the added advantage of providing an estimate of the sample variance.

An important point to stress here is that when performing this null-test, ideally the AGN sample should be independent from the sample used to derive the fiducial SHMR. However, the L12 SHMR model was derived using all galaxies in the COSMOS field, including the subset of AGN hosts considered here. A better approach would be to use one half of the COSMOS survey to derive the SHMR and the second half to compute AGN host statistics. Certainly, this type of approach can be easily adopted in future large-area surveys which will have more than ample statistical constraining power. In our case, however, the AGN sample only represents  $\sim 3$  per cent of the galaxy population with  $\log_{10}(M_*) > 10.5$  and should only have a minor impact on the overall SHMR.

A negative null-test would be highly interesting and would indicate that AGN (or more generally, the subpopulation in question) ‘know’ something about the dark matter haloes in which they reside. In this case, step three is to vary a subset of parameters (those we expect might be different for active populations). This choice can be informed by predictions from semi-analytic models (SAM) of galaxy formation or from direct hydrodynamic simulations (e.g. Chatterjee et al. 2012; DeGraf et al. 2012). As discussed in more detail in Section 5.4, one parameter to consider is the AGN satellite fraction  $f_{\text{sat}}$ . Another parameter of interest might be  $c_{\text{sat}}$ , the halo concentration of satellite AGN (e.g. Chatterjee et al. 2012). Step three is to vary a small set of parameters (e.g.  $f_{\text{sat}}$  and/or  $c_{\text{sat}}$ ) to fit the statistic of choice while marginalizing over other parameters in the SHMR-based model. In this paper, however, step three is unnecessary because the null-test is positive (see Section 4.2).

The final step in our methodology is to use the fiducial SHMR (or the modified version from step three) to study halo distributions, satellite fractions, and halo occupation statistics. Again, this step can be achieved both analytically or by using mock catalogues. This final step combines two key sets of information. These are (a) the fiducial (or modified) SHMR-based model and (b) the AGN fraction as a function of stellar mass and redshift.

## 4 RESULTS

### 4.1 Galaxy–galaxy lensing of X-ray AGN

To obtain high signal-to-noise measurements, we stack the weak lensing signal around our sample of 382 AGN hosts as a function of radial transverse separation  $r$ . All of our stacks are performed in physical coordinates. The galaxy–galaxy lensing signal that we measure yields an estimate of the mean *surface mass density contrast* profile for our AGN host sample:

$$\Delta\Sigma(r) \equiv \bar{\Sigma}(< r) - \bar{\Sigma}(r). \quad (1)$$

Here,  $\bar{\Sigma}(r)$  is the azimuthally averaged and projected surface mass density at radius  $r$  and  $\bar{\Sigma}(< r)$  is the mean projected surface mass density within radius  $r$  (e.g. Miralda-Escudé 1991; Wilson et al. 2001). For the radial ranges that we probe in this study ( $< 2$  Mpc), our lensing signals are mainly due to the dark matter haloes associated with the stacked galaxy sample (the ‘one-halo’ term).

Uncertainties on the lensing signal are derived using two different methods. The first, most naive estimate assumes that the data bins are independent, and that measurement error and shape noise

are the dominant sources of error. The uncertainty on  $\Delta\Sigma$  is then simply  $\sigma_w = \sqrt{1/\sum w_i}$  where the sum is performed over all lens–source pairs and where  $w_i$  is an estimate of the shear variance for each source (see L12). However, at larger radii, bins may become correlated due to the fact that the same source galaxy may be associated with multiple lens galaxies (‘correlated shape noise’). To test for the magnitude of this effect, we also derive jackknife uncertainties on  $\Delta\Sigma$ , noted hereafter as  $\sigma_{\text{jk}}$ . The two uncertainty estimates are in good agreement with the jackknife errors being somewhat larger for the outer radial bins suggesting small levels of correlated shape noise. Jackknife estimates of covariance for the outer radial bins suggest that the correlation coefficient between the three last radial bins is at most  $|\rho| < 0.3$ . Throughout this paper, we quote values using both of these uncertainty estimates and we neglect the small amount of covariance for the outermost radial bins. Finally, we use 100 mock catalogues (described in the previous section) to estimate the sampling variance for our lensing signal. These include both shot noise due to the small number of lens galaxies in our sample, as well as sample variance in the underlying dark matter realization for a field the size of COSMOS. These errors are subdominant (less than 10 per cent) compared to shape noise.

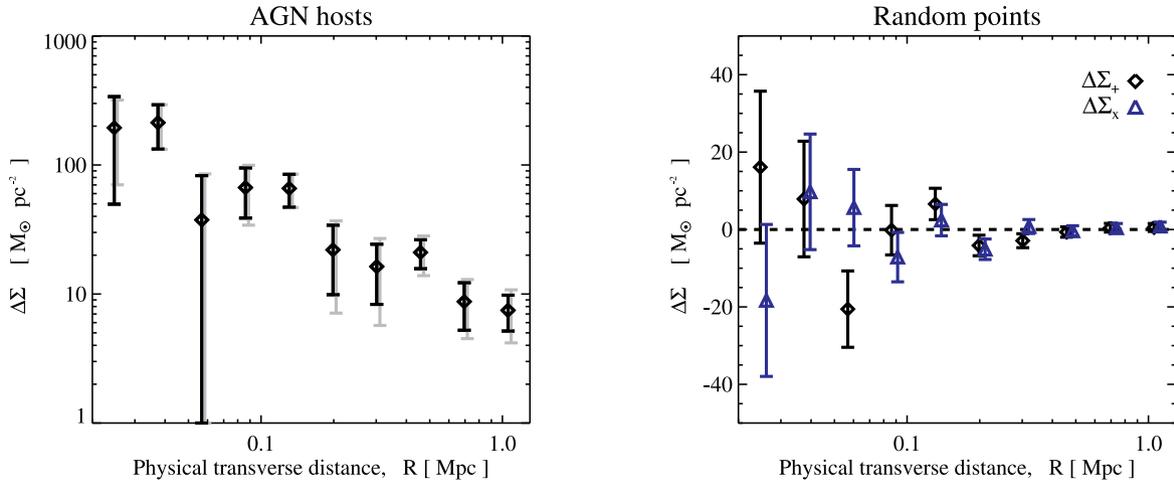
Our weak lensing signal for the AGN sample is shown in Fig. 2. For this measurement, we have used 10 logarithmically spaced bins from  $r = 20$  kpc to 1.3 Mpc. The weak lensing signal is clearly detected out to the largest scales with a mean signal to noise per data point of  $S/N \sim 2.4$  using shape-noise uncertainties and  $S/N \sim 2.1$  using jackknife uncertainties. As a test for systematic effects, we also compute the lensing signal around 7000 random points that are drawn from the same redshift distribution as our AGN lens sample. The result is shown in the right-hand panel of Fig. 2. No evidence for systematic shear patterns is detected around random points.

### 4.2 Weak lensing signal of AGN hosts compared to fiducial SHMR model

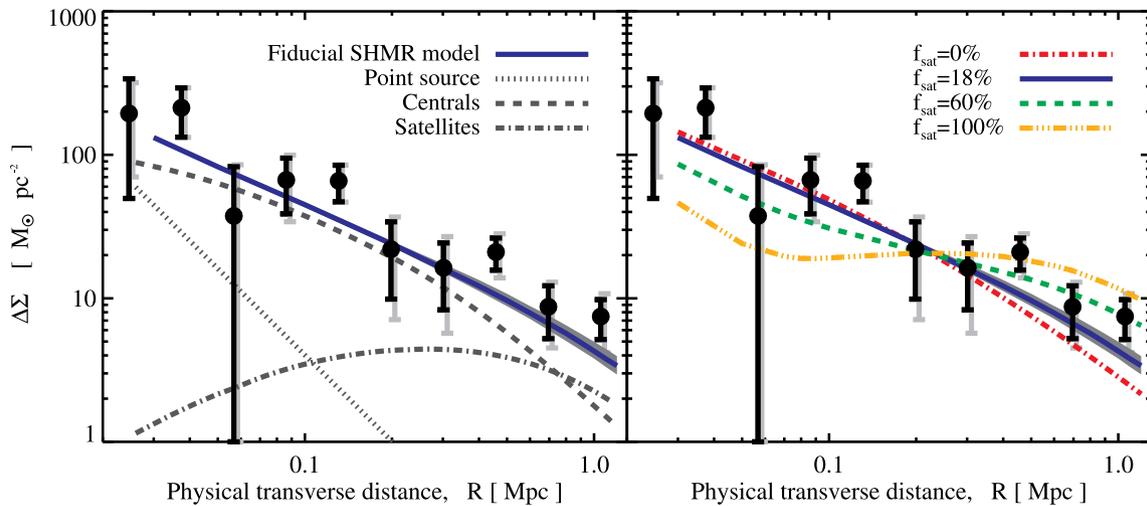
Given the host mass and redshift of each AGN in our sample, we use the SHMR model of L12 to compute the predicted AGN galaxy–galaxy lensing signal. The predicted galaxy–galaxy lensing signal is shown in Fig. 3 and is composed of three terms: (1) a contribution from the stellar mass of the AGN host galaxy, (2) a contribution from the dark matter haloes associated with central galaxies that follows the standard Navarro–Frenk–White profile (Navarro, Frenk & White 1997), and (3) a contribution from the dark matter haloes associated with satellite galaxies. The total weak lensing signal is the sum of these three terms. Contributions from subhaloes associated with satellites are neglected. The ‘two-halo’ term is negligible on these small radial scales.

The grey shaded region in Fig. 3 shows the field-to-field variance derived from mock catalogues; this is subdominant compared to the measurement errors on the lensing signal. Overall, we find that our fiducial SHMR model does an excellent job at matching the weak lensing signal of AGN. The  $\chi^2$  between the measured lensing signal and the SHMR prediction is  $\chi^2/\text{d.o.f} = 8.5/10$  ( $\chi^2 = 11.6/10$  for shape-noise errors). Since there are no free parameters in this model, the number of degrees of freedom is simply the number of data points,  $\text{d.o.f} = 10$ . As mentioned in Section 3, however, the AGN sample is not strictly independent from the data used to infer the SHMR which does place a caveat on this comparison.

Our null-test is positive. Hence, we confirm the null hypothesis that the AGN host occupation is no different than that defined by galaxies with the same  $M_*$ , regardless of nuclear activity.



**Figure 2.** Left: first reported weak lensing signal of X-ray AGN as measured from 382 X-ray-selected hosts with  $(\log(L_X) = 42.7)$  from the COSMOS field. Black error bars show shape-noise uncertainties, grey error bars indicate jackknife uncertainties. Right: as a test for systematics, we also compute the lensing signal measured around random points.



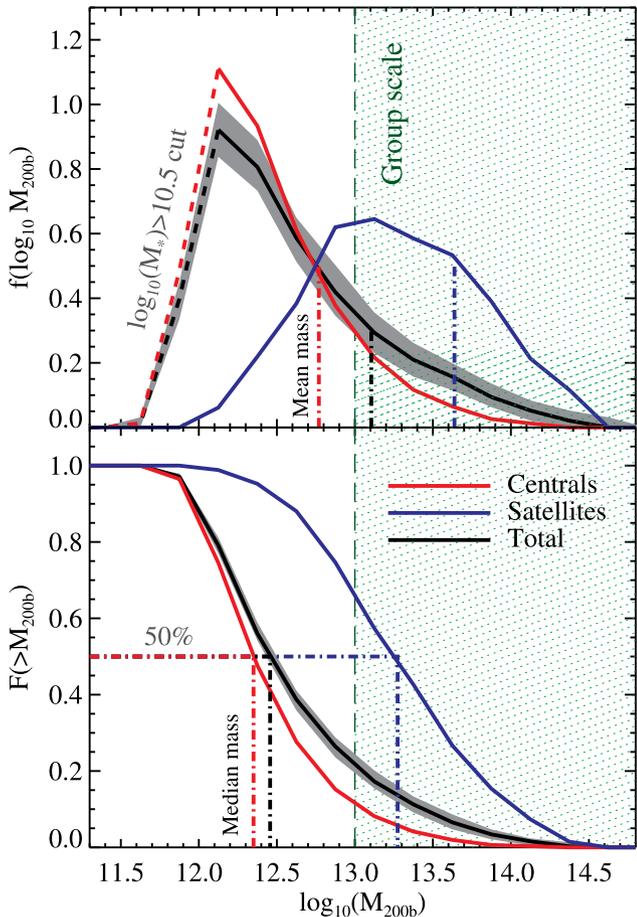
**Figure 3.** Left: lensing signal of AGN sample compared to the prediction from our fiducial SHMR model (blue line). The grey shaded region shows the field-to-field variance derived from 100 mock catalogues. These uncertainties are insignificant compared to shape-noise uncertainties. The lensing signal has contributions from the host stellar mass (dotted line) and from the dark matter haloes associated with both central (dashed line) and satellite (dash-dotted line) galaxies. We confirm the null hypothesis that the AGN host occupation is no different than that defined by galaxies with the same  $M_*$ , regardless of nuclear activity. Right: predicted lensing signal for varying satellite fractions. The satellite fraction of the fiducial SHMR model is  $f_{\text{sat}} = 18$  per cent.

Step three in our methodology is unnecessary for this sample (we do not need to vary any parameters to describe the lensing signal). As an example, however, of how step three might proceed – the right-hand panel of Fig. 3 shows how the predicted lensing signal of AGN hosts varies with  $f_{\text{sat}}$  (keeping all other parameters fixed). We find that reducing the satellite fraction to  $f_{\text{sat}} = 0$  only has a relatively small impact on the overall lensing signal. The predicted lensing signal is mostly unchanged on small scales and decreases slightly on 1 Mpc scales. If we increase the satellite fraction to 100 per cent, then the predicted signal increases on large scales but decreases on small scales creating a clear scale-dependent signature which should be easily detectable with the next generation of lensing surveys. Small values of  $f_{\text{sat}}$  may be difficult to detect with lensing alone, but joint measurements of lensing and clustering will be able to pin down  $f_{\text{sat}}$  with greater accuracy.

#### 4.3 Dark matter environment of AGN sample as inferred from host mass

In the previous section, we compared the weak lensing signal of AGN hosts with the prediction from our fiducial SHMR model. The fact that they are statistically indistinguishable suggests that AGN in our sample populate haloes in the same fashion as the overall galaxy population. One caveat, however, is that our AGN lensing signal is relatively noisy. Upcoming lensing surveys with better signal to noise may find differences that we have been unable to detect. In the meantime, given that we have no evidence to suggest otherwise, in the remainder of this paper we proceed under the assumption that host stellar mass and redshift are sufficient to predict the mean dark matter environment for this AGN sample.

We now use our mock catalogues to investigate the predicted halo mass distribution for this AGN sample. A mock AGN



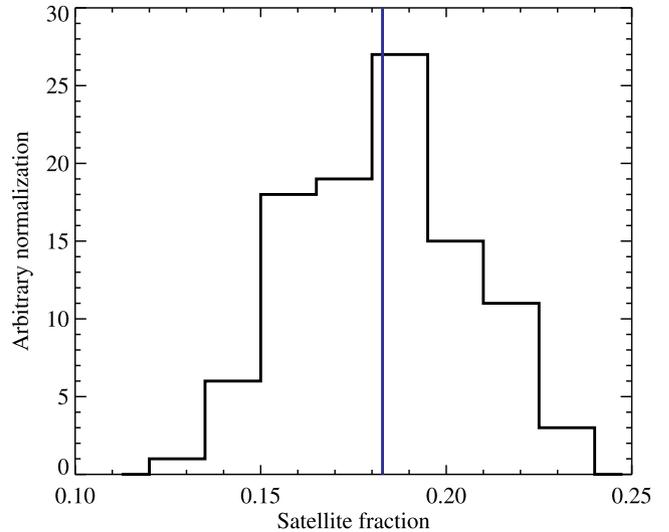
**Figure 4.** Top: probability per  $\log_{10}(M_{200b})$  that an AGN in our sample is hosted by a halo of mass  $M_{200b}$ . Note that this is a probability density function and may take on values greater than 1. Distributions are shown separately for central (red line) and satellite AGN (blue line). For satellites,  $M_{200b}$  represents the mass of the parent halo (not the subhalo). The black solid line is the full sample (centrals plus satellites) with grey shaded regions representing the variance from mock catalogues. Vertical lines (dash-dotted) indicate a mean halo mass. The dashed green vertical line shows a typical mass limit for galaxy groups. Note that the sharp drop-off at  $\log_{10}(M_{200b}) \sim 12$  is simply due to the fact that we select AGN with hosts mass  $\log_{10}(M_*) > 10.5$ . This cut was motivated to select a sample that is roughly similar to previous work on the clustering of X-ray AGN. The black and red curves would continue to rise had we included lower mass AGN hosts in our sample. Bottom: complementary cumulative distribution function. Vertical lines (dash-dotted) indicate a median halo mass. Only  $\sim 60$  per cent of AGN satellites are contained in haloes with  $\log_{10}(M_{200b}) > 13$ .

population is extracted from each mock catalogue (see Section 2.4) by matching mock galaxies and AGN hosts in terms of stellar mass and redshift. There are 100 mock catalogues in total, each mock has the same volume as COSMOS. Fig. 4 shows the halo mass probability density function as well as the complementary cumulative distribution function for mock AGN samples. Errors in Fig. 4 represent the field-to-field variance between mock catalogues. Table 1 summarizes the mean and median halo masses for centrals, satellites, and for the combined sample (centrals and satellites). For the combined sample, we find that the mean halo mass,  $\langle M_{200b} \rangle = 1.3 \times 10^{13} M_{\odot}$ , is a factor of 4.5 times larger than the median halo mass,  $M_{200b}^{\text{med}} = 2.9 \times 10^{12} M_{\odot}$ . We underscore the fact that the mean and the median halo masses may be markedly different because of the skewed tail in the halo mass distribution.

**Table 1.** Mean and median halo masses.

Halo mass	Mean $\langle M_{200b} \rangle$ ( $10^{13} M_{\odot}$ )	Median $M_{200b}^{\text{med}}$ ( $10^{13} M_{\odot}$ )
Centrals	$0.59 \pm 0.08$	$0.22 \pm 0.01$
Satellites	$4.3 \pm 1.3$	$1.9 \pm 0.5$
Cen + Sat	$1.3 \pm 0.3$	$0.29 \pm 0.02$

*Notes:* errors represent the field-to-field variance derived from mock catalogues.

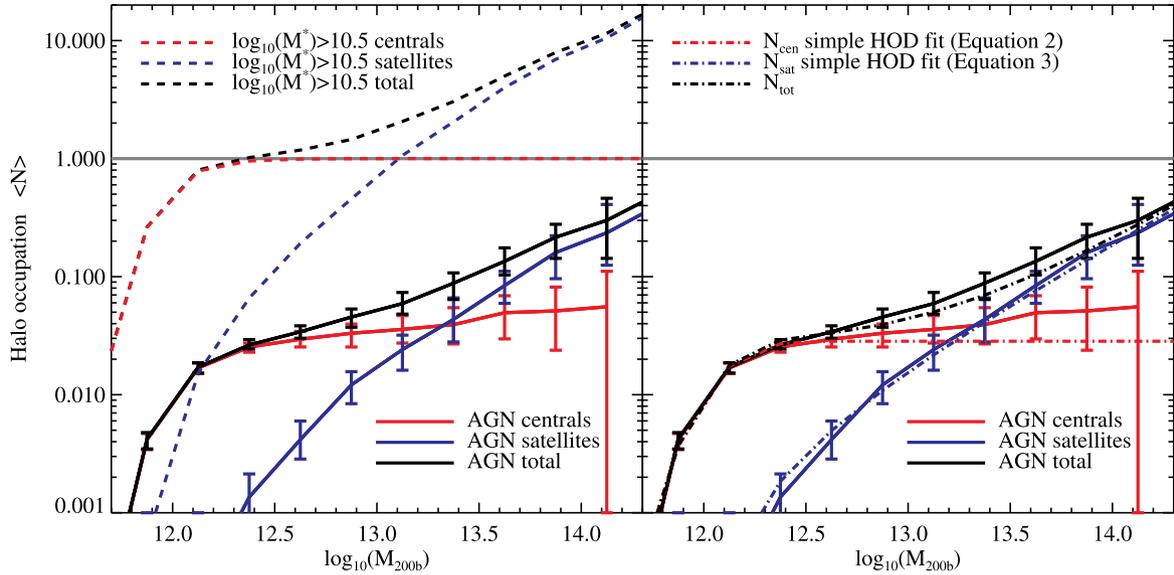


**Figure 5.** AGN satellite fraction distribution from mock catalogues. The mean satellite fraction is  $\langle f_{\text{sat}} \rangle = 18$  per cent with an rms dispersion between mock catalogues of 2 per cent.

We stress that these values are specific to our particular AGN sample selection. In our case, the most important factor in determining the exact halo mass distribution is the  $\log_{10}(M_*) > 10.5$  cut that we applied to the AGN host masses. This cut drives the sharp drop-off at  $\log_{10}(M_{200b}) \sim 12$  in Fig. 4. However, in practice, COSMOS AGN catalogues do contain X-ray AGN in galaxies with  $\log_{10}(M_*) < 10.5$ . According to the SHMR, on average, these are expected to live in even lower mass haloes.

Let us now turn our attention to the halo mass distributions of satellite AGN.<sup>4</sup> Our predictions are based on our fiducial SHMR model where AGN hosts have the same satellite fractions as inactive galaxies (see Section 4.2 and Fig. 3). We find that 50 per cent of satellite AGN in our sample live in haloes less massive than  $\log_{10}(M_{200b}) = 13.2$ . Fig. 5 shows the predicted satellite fractions for our sample. We find a mean satellite fraction of  $\langle f_{\text{sat}} \rangle = 18$  per cent with an rms dispersion between mock catalogues of 2 per cent. How does this compare with previous results derived from clustering studies of X-ray AGN? Reliable constraints on satellite fractions derived from HOD modelling are limited by modelling uncertainties (e.g. Miyaji et al. 2011; Shen et al. 2013). A perhaps more robust estimate of satellite fractions may be obtained by measuring the effects of satellite peculiar velocities on the 2D redshift space correlation function. Using this technique, Starikova et al. (2011) report a 90 per cent confidence level upper limit on the satellite fraction of  $f_{\text{sat}} < 8$  per cent. Their sample, however, is truncated

<sup>4</sup> For satellites, halo mass refers to the mass of the parent halo, not subhalo masses.



**Figure 6.** Left: central and satellite halo occupation functions for our AGN sample derived from mock catalogues based on our fiducial SHMR model (solid lines). Errors represent field-to-field variance for a COSMOS-like survey. Dashed lines indicate the occupation functions for all galaxies with  $\log_{10}(M_*) > 10.5$ . The turnover in the HOD at  $\log(M_{200b}) \sim 12$  is set by the  $\log_{10}(M_*) > 10.5$  cut that defines our AGN sample. The amplitude of the AGN HOD ( $\langle N_{\text{cen}} \rangle \sim 0.01\text{--}0.06$ ) indicates that X-ray obscured AGN from our sample only represent a few per cent of all central galaxies in any given halo mass bin. Right: comparison between the occupation function inferred from our analysis and a fit using a simple HOD given by equations (2) and (3).

to brighter hosts than ours for which we do indeed expect lower satellite fractions. As discussed in more detail in Section 5, it is unclear how much this difference might be of genuine interest as opposed to simply due to sample selection effects.

#### 4.4 Halo occupation functions

The occupation functions for AGN in our sample are shown in Fig. 6. For reference, we also show the occupation functions for a galaxy sample with  $\log_{10}(M_*) > 10.5$ . We find that  $\log_{10}(M_{200b}) \sim 12$  haloes will host on average  $\langle N_{\text{cen}} \rangle \sim 0.01$  central AGN in our sample and at  $\log_{10}(M_{200b}) \sim 14$  this number rises to  $\langle N_{\text{cen}} \rangle \sim 0.06$  (see Table 2). Including both centrals and satellites, we expect that  $\log_{10}(M_{200b}) \sim 14.0$  haloes host on average  $\langle N_{\text{tot}} \rangle \sim 0.3$  AGN in our sample. The errors on the occupation functions due to field-to-field variance are non-negligible for groups with mass  $\log_{10}(M_{200b}) \sim 14.0$  for a survey the size of COSMOS.

Our occupation functions are mostly comparable to those obtained by Allevato et al. (2012) from direct counting of AGN

in groups within the COSMOS field. Allevato et al. (2012) measure  $\langle N_{\text{tot}} \rangle \sim 0.2\text{--}0.6^5$  for haloes with masses above  $10^{13} M_{\odot}$ . Our values are in fair agreement with these estimates, especially given that we apply different selection criteria to the COSMOS AGN samples (we apply host mass and  $L_X$  cuts for example) which can easily translate into factor of 2 differences in the amplitude of the inferred HODs.

We stress that the goal of this paper is not so much the exact values of the HOD presented in Fig. 6 since these will depend sensitively on our particular AGN sample selection (varying the  $L_X$  cuts will affect the amplitude of the HOD for example). Instead, our main point here is to demonstrate that, under the assumption that active and inactive galaxies inhabit similar dark matter environments, the SHMR-based approach advocated here makes firm predictions for the *shape and normalization* of the AGN occupation functions. Miyaji et al. (2011) investigated three different HOD parameterizations to model the cross-correlation function between *ROSAT* All-Sky Survey detected AGN and SDSS LRGs. Among the three models explored by Miyaji et al. (2011), their model B provides the best description of the HODs presented here. This is a model that is similar to those used for threshold galaxy samples but with an additional free parameter  $f_A$  that allows the global normalization of  $\langle N_{\text{cen}} \rangle$  to float.

Fig. 6 shows that our HOD is reasonably well fit by an HOD of the form

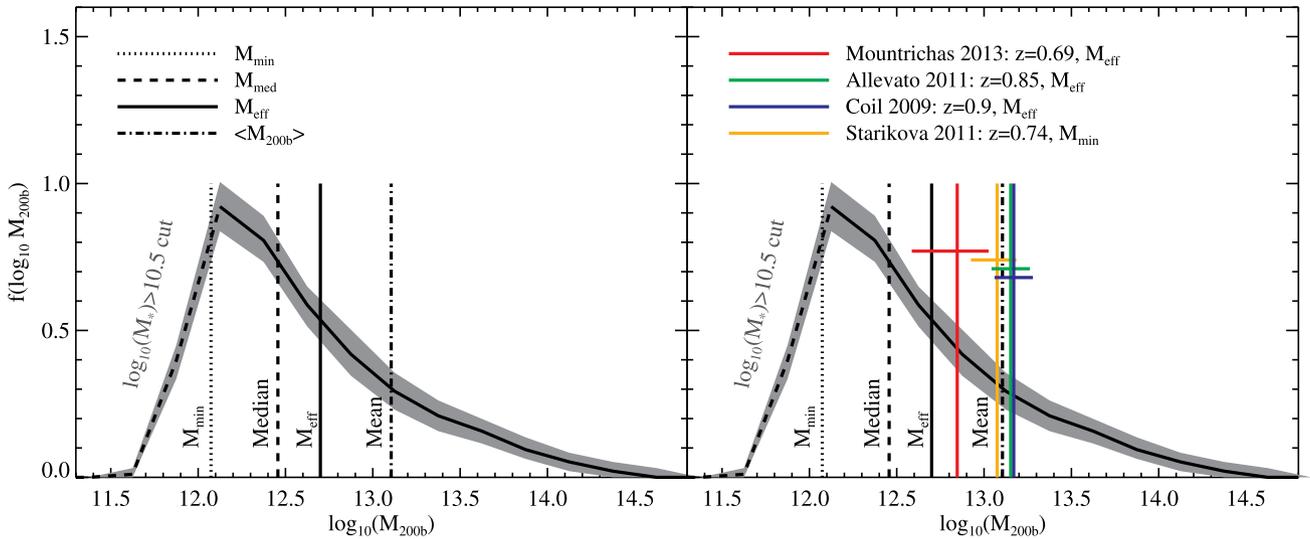
$$\langle N_{\text{cen}} \rangle = \frac{f_A}{2} \left[ 1 + \text{erf} \left( \frac{\log_{10}(M_{200b}/M_{\text{min}})}{\sigma_{\log M}} \right) \right] \quad (2)$$

$$\langle N_{\text{sat}} \rangle = \langle N_{\text{cen}} \rangle \left( \frac{M_{200b}}{M_{\text{sat}}} \right)^{\alpha} \exp \left( \frac{-M_{\text{cut}}}{M_{200b}} \right) \quad (3)$$

<sup>5</sup> Allevato et al. (2012) correct their HODs for incompleteness in  $L_X$ , but we attempt no such corrections here. The values quoted here from Allevato et al. (2012) are taken from their fig. 3 before any luminosity and redshift evolution corrections.

**Table 2.** Halo occupation functions.

$\log_{10}(M_{200b})$	$\langle N_{\text{cen}} \rangle$	$\langle N_{\text{sat}} \rangle$	$\langle N_{\text{tot}} \rangle$
11.875	$0.0043^{+0.0005}_{-0.0007}$	0.0	$0.0042^{+0.0006}_{-0.0008}$
12.125	$0.017^{+0.002}_{-0.002}$	0.0	$0.017^{+0.002}_{-0.002}$
12.375	$0.025^{+0.002}_{-0.002}$	$0.0014^{+0.0008}_{-0.0007}$	$0.026^{+0.003}_{-0.002}$
12.625	$0.028^{+0.005}_{-0.004}$	$0.0041^{+0.002}_{-0.001}$	$0.034^{+0.004}_{-0.004}$
12.875	$0.033^{+0.006}_{-0.008}$	$0.012^{+0.004}_{-0.004}$	$0.046^{+0.008}_{-0.008}$
13.125	$0.036^{+0.01}_{-0.008}$	$0.024^{+0.008}_{-0.008}$	$0.059^{+0.01}_{-0.01}$
13.375	$0.039^{+0.02}_{-0.01}$	$0.04^{+0.02}_{-0.02}$	$0.088^{+0.02}_{-0.02}$
13.625	$0.05^{+0.02}_{-0.02}$	$0.08^{+0.03}_{-0.02}$	$0.13^{+0.04}_{-0.03}$
13.875	$0.05^{+0.03}_{-0.03}$	$0.16^{+0.06}_{-0.06}$	$0.22^{+0.06}_{-0.07}$
14.125	$0.06^{+0.05}_{-0.06}$	$0.2^{+0.2}_{-0.1}$	$0.3^{+0.2}_{-0.2}$



**Figure 7.** Left: probability per  $\log_{10}(M_{200b})$  that an AGN in our sample is hosted by a halo of mass  $M_{200b}$ . The dotted, dashed, solid and dash–dotted vertical lines show  $M_{\min}$  (equation 2),  $M_{\text{med}}$ ,  $M_{\text{eff}}$  and  $\langle M_{200b} \rangle$ , respectively. Right: comparison with the values inferred from studies of X-ray AGN clustering. Previous results from clustering studies tend to lie in between our effective halo mass and our mean halo mass.

with  $f_A = 0.028$ ,  $M_{\min} = 1.2 \times 10^{12} M_{\odot}$ ,  $\sigma_{\log M} = 0.25$ ,  $M_{\text{sat}} = 1.5 \times 10^{13} M_{\odot}$ , and  $M_{\text{cut}} = 2 \times 10^{12} M_{\odot}$ . This model is similar to model B from Miyaji et al. (2011) except that  $\langle N_{\text{cen}} \rangle$  is modelled with an error function instead of a step function, and our satellite occupation includes an exponential cutoff with a scale set by  $M_{\text{cut}}$ .

One interesting feature in Fig. 6 is that our empirically determined HODs display a rise towards higher halo mass that is not well captured by a constant  $f_A$ . This parameter is sometimes interpreted as an AGN duty cycle (e.g. Martini & Weinberg 2001; Shen et al. 2007; Miyaji et al. 2011; Richardson et al. 2013). In this case, the rising nature of  $\langle N_{\text{cen}} \rangle$  could indicate a varying AGN duty cycle with halo mass. We caution however that at least part of this trend will be imposed by sample selection effects introduced by our  $L_X$  cut. AGN show a wide distribution of Eddington ratios (e.g. Heckman et al. 2004; Kauffmann & Heckman 2009; Schulze & Wisotzki 2010; Aird et al. 2012). For a fixed  $L_X$  cut, we sample AGN in massive galaxies over a wide range of Eddington ratios, while for less massive galaxies we only sample AGN with larger Eddington ratios (Schulze & Wisotzki 2010; Aird et al. 2012). This may lead to an apparent increase in  $f_A$  with halo mass for  $L_X$ -selected samples.

## 5 DISCUSSION

### 5.1 Comparison with previous results based on clustering measurements

Before comparing with previous results, let us first briefly review how clustering studies typically infer halo mass. What most studies based on clustering measurements actually derive<sup>6</sup> is the linear effective bias,  $b_{\text{eff}}$ . The effective halo mass is then the mass which satisfies  $b(M_{\text{eff}}) = b_{\text{eff}}$ , where  $b(M_h)$  is the mean bias of haloes of mass  $M_h$  (e.g. Tinker et al. 2010). What exactly does this effective halo mass correspond to when considering samples that span a

wide range in halo mass? The effective bias measured by clustering studies is

$$b_{\text{eff}} = \frac{\int b(M_h) N_{\text{AGN}}(M_h) n(M_h) dM_h}{\int N_{\text{AGN}}(M_h) n(M_h) dM_h}, \quad (4)$$

where  $N_{\text{AGN}}(M_h)$  and  $n(M_h)$  are, respectively, the mean number of AGN and the number density of haloes as a function of  $M_h$  (e.g. Baugh et al. 1999; Fanidakis et al. 2013). For our purpose, it is perhaps more clear to rewrite equation (4) so as to highlight the AGN halo mass probability density function  $f_{\text{AGN}}$ :

$$f_{\text{AGN}} = \frac{N_{\text{AGN}}(M_h) n(M_h)}{\int N_{\text{AGN}}(M_h) n(M_h) dM_h}. \quad (5)$$

Using  $f_{\text{AGN}}$ , equation (4) simply becomes

$$b_{\text{eff}} = \int b(M_h) f_{\text{AGN}}(M_h) dM_h. \quad (6)$$

Written in this fashion, it is clear that  $M_{\text{eff}}$  measured from  $b_{\text{eff}}$  corresponds to a bias-weighted average of  $f_{\text{AGN}}$ . Halo bias  $b(M_h)$  is not a simple linear function of halo mass. Broadly speaking,  $b(M_h)$  is a shallow function at low halo mass and then rises sharply at higher halo mass (e.g. Tinker et al. 2010). Hence,  $M_{\text{eff}}$  may be different than other, perhaps more useful averages such as the mode, median, or mean value of  $f_{\text{AGN}}$ .

We now investigate the difference between  $M_{\text{eff}}$ , the median, and the mean of our halo mass distribution. At our mean redshift of  $\bar{z} = 0.7$ , our SHMR model predicts  $b_{\text{eff}} = 1.8$  and  $M_{200b}^{\text{eff}} = 5.0 \times 10^{12} M_{\odot}$ . Hence,  $M_{200b}^{\text{eff}}$  is roughly mid-way between the median and the mean halo mass of  $f_{\text{AGN}}$ .

Fig. 7 compares our halo mass distribution with  $M_{\text{eff}}$  values derived from clustering studies. We focus on samples that (only very roughly) span the same redshift and luminosity range as ours (Coil et al. 2009; Allevato et al. 2011; Mountrichas et al. 2013).<sup>7</sup> We also compare with the results from Starikova et al. (2011) at  $0.5 < z < 1.0$

<sup>6</sup> Here, we refer specifically to studies that infer halo mass from  $b_{\text{eff}}$  assuming that  $b(M_{\text{eff}}) = b_{\text{eff}}$ . Studies that model clustering with an HOD-type approach may quote a mean, median, or a minimum halo mass instead of an effective halo mass.

<sup>7</sup> Unless stated otherwise, we used a compilation of halo mass values provided in table 2 of Fanidakis et al. (2013). For Allevato et al. (2011), we use their results from table 3 of their paper for obscured X-ray AGN at  $\bar{z} = 0.85$ . All halo masses have been converted to our mass definition,  $M_{200b}$ .

which are inferred from the ratio of the projected autocorrelation function integrated along and across the line of sight. The technique employed in Starikova et al. (2011) is different compared to the other studies mentioned above because it uses additional information from the peculiar velocities of satellites.

Fig. 7 shows that the halo mass values reported by previous studies are typically larger than ours and tend to lie in between our effective halo mass and our mean halo mass. Our results are most different compared to Starikova et al. (2011) who report a *minimum* halo mass,<sup>8</sup> not an effective halo mass. However, our host sample selection is also most different compared to Starikova et al. (2011), who limit their selection to brighter hosts than we do. This brings us to an important point, discussed in the following section, which is that when comparing studies of AGN clustering – any cuts on host galaxy mass/luminosity must be taken into account because host mass/luminosity *correlates* with halo mass.

### 5.2 A simple selection effect that cannot be neglected: brighter hosts live in more massive haloes

Studies of X-ray AGN clustering are typically limited to brighter hosts simply due to the fact that such measurements require spectroscopic redshifts. The samples we compare with in Fig. 7 are typically limited to hosts with  $I_{AB} < 21.5$ – $23$ . More specifically, Coil et al. (2009) make no explicit cut on host luminosity but their sample is roughly limited at  $R_{AB} < 22.6$  due to the availability of optical spectroscopy. The Alleinato et al. (2011) AGN sample is limited to  $I_{AB} < 23$  (with a spectroscopic completeness of 53 per cent). Koutoulidis et al. (2013) make no explicit cut on host luminosity, but spectroscopic requirements drive an implicit cut on host luminosity which varies between the different surveys in their compilation. Starikova et al. (2011) apply a host magnitude cut of  $I < 21.5$  to their sample. There is not a one-to-one relation between stellar mass and optical luminosity, but to give some idea of the characteristic luminosity of our hosts, galaxies in COSMOS with  $\log_{10}(M_*) > 10.5$  have a median magnitude of  $I_{AB} \sim 21$  at  $z \sim 0.5$ . At  $z \sim 1$ , they have a median magnitude of  $I_{AB} \sim 22.6$ .

On average, hosts with brighter luminosities live in larger dark matter haloes. In addition, samples defined by a fixed observed host luminosity threshold probe different stellar mass (hence halo mass) limits as a function of redshift. We stress that these (sometimes implicit) cuts on host properties need to be considered when comparing X-ray-selected samples to one another, when comparing AGN samples selected at different wavelengths (e.g. X-ray AGN versus QSOs), and also when comparing with theoretical predictions from SAMs or hydrodynamical simulations. Differences in AGN luminosities between samples are commonly taken into account, but cuts on host properties must also be considered.

### 5.3 The dark matter environment of moderate luminosity X-ray AGN compared to UV luminous QSOs

The prevailing wisdom from clustering studies of X-ray AGN is that moderate luminosity X-ray selected AGN populate group-sized dark matter haloes with  $M_h \sim 10^{13} M_\odot$  (e.g. Fanidakis et al. 2013; Koutoulidis et al. 2013; Hütsi et al. 2014, to cite a few recent examples). In contrast, UV luminous QSOs in the 2dF and SDSS

surveys are found to reside in lower mass haloes with  $M_h \sim 10^{12} M_\odot$  (Croom et al. 2005; da Ângela et al. 2008; Ross et al. 2009; Shanks et al. 2011). This environmental dependence has led to the suggestion that moderate luminosity X-ray AGN and luminous QSOs may have different fuelling mechanisms (e.g. Fanidakis et al. 2013). In this scenario, QSOs are fuelled from cold-gas reservoirs that are funnelled to galaxy centres by catastrophic events such as mergers or disc instabilities whereas moderate luminosity X-ray AGN may be connected with an additional fuelling channel in which gas is accreted directly from a diffuse state in massive dark matter haloes ( $M_h > 10^{12} M_\odot$ , the ‘radio’ or ‘hot-halo’ mode).

Under closer consideration, however, the difference between the dark matter environment of moderate luminosity X-ray AGN and QSOs may not be so clear. First, selection cuts on host properties (see the previous section) must be taken into account. Spectroscopic requirements impart different selections on host properties for X-ray and QSOs samples (e.g. Hopkins et al. 2009) – this will naturally lead to difference in the underlying dark matter distributions. Secondly, clustering studies often report a single halo mass scale which may be difficult to interpret in the context of samples that span a wide range of halo masses.

The results of this paper favour a different picture for the dark matter haloes of galaxies hosting moderate luminosity X-ray AGN. Fig. 4 suggests that most AGN in our sample do *not* live in group environments ( $M_h > 10^{13} M_\odot$ ) – 50 per cent of the AGN in our sample are found in haloes less massive than  $M_{200b} \sim 3 \times 10^{12} M_\odot$ , and hence live in relatively low mass haloes. Recently, Conroy & White (2013) showed that a simple phenomenological model in which QSOs live in a wide range of halo masses successfully predicts both the QSO luminosity function and the two-point correlation function from  $0.5 < z < 2$ . Taken together, these two results suggest that both QSOs and moderate luminosity X-ray AGN may occupy haloes in a relatively ‘normal’ way compared to galaxies without active nuclei. The notion that they may share similar dark matter environments calls into question the need for different physical mechanisms to explain the fuelling of moderate luminosity X-ray AGN and QSOs.

### 5.4 Insights from Galacticus SAM: halo mass distributions of active galaxies versus all galaxies

The approach adopted in this paper is valid if AGN populations can be described by varying a few simple parameters in the SHMR description for the overall galaxy population. How does this premise compare with theoretical models of AGN activity and which parameters in the SHMR are most likely to differ? To investigate these questions, we turn to the state-of-the-art Galacticus SAM (Benson 2012). We use Galacticus because its modelling of BH physics is relatively realistic compared to other SAMs, comparable to the detailed BH evolution modelling developed by Fanidakis et al. (2011). Specifically, for this work, we use v0.9.1 (revision 1456) of Galacticus and the default set of parameters supplied with that version. A description of the key features of this SAM relevant for this paper is given in the appendix. The full details of the Galacticus model can be found in Benson (2012).

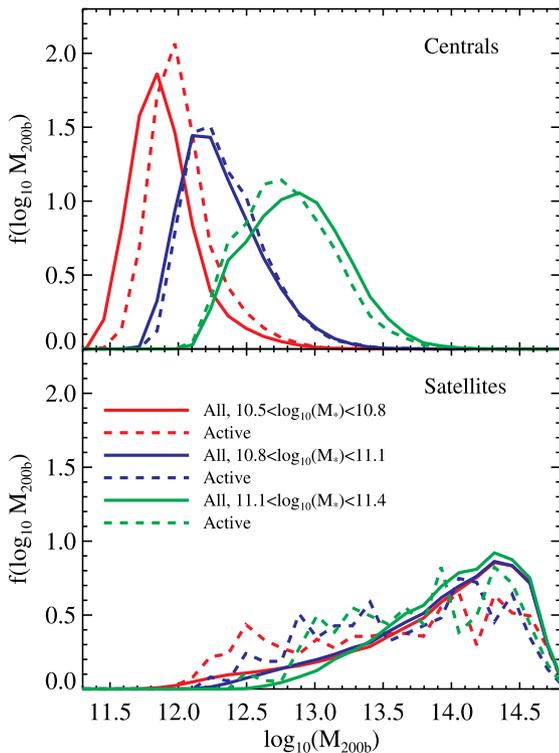
Our goal is not a one-to-one comparison between Galacticus and our COSMOS results. Many aspects of the COSMOS data are not reproduced by this Galacticus model. For example, as discussed in the following section, the overall SHMR is different. Also, at fixed stellar mass, satellites in Galacticus populate more massive haloes than in COSMOS. With these caveats in mind, we use Galacticus

<sup>8</sup> Starikova et al. (2011) quote the halo mass that corresponds to the minimum  $V_{\max}$  (maximum circular velocity) of haloes that host an X-ray AGN in their model.

to investigate qualitative differences between active and inactive galaxies that may be informative in interpreting our observational results.

We select a sample of active galaxies from the Galacticus simulation at  $z = 0.61$  (close to our mean redshift of  $z = 0.7$ ) using the same host mass and  $L_X$  cuts as our COSMOS AGN sample. We do not however mimic the  $L_X$  incompleteness in the COSMOS data. Fig. 8 shows the halo mass distributions of active galaxies in three stellar mass bins compared to halo mass distributions for the overall galaxy population. There are some small differences between these halo mass distributions. For example, the central halo masses of active galaxies are larger by  $\sim 0.15$  dex compared to inactive galaxies in the lowest stellar mass bin ( $\log_{10}(M_*) \sim 10.6$ ). Broadly speaking, however, the halo mass distributions of active galaxies and inactive galaxies are remarkably similar in this SAM. There is however one key difference between the two samples. The AGN satellite fraction ( $\sim 3$  per cent) is an order of magnitude lower than the satellite fraction for all galaxies in the SAM (note that this difference is not obvious from Fig. 8 which shows probability density functions).

The similarity between the halo mass distributions in Fig. 8 supports our approach of using a fiducial SHMR as a starting point to model this AGN population (at least when compared to this particular SAM). The low satellite fraction of the active galaxy sample suggests, however, that in step three of our methodology (see Section 3),  $f_{\text{sat}}$  should be left as a free parameter. The fact that this SAM predicts a much lower satellite fraction for active galaxies compared to inactive galaxies also raises the possibility of an interesting tension between this SAM and our weak lensing results. However,



**Figure 8.** Halo mass probability density functions for centrals (upper panel) and satellites (lower panel) from the Galacticus SAM. Solid lines correspond to all galaxies in three stellar mass bins spanning the range  $\log_{10}(M_*) = 10.5$  to  $\log_{10}(M_*) = 11.4$ . Dashed lines correspond to a sample of active galaxies selected to roughly mimic our COSMOS sample. Broadly speaking, active galaxies have similar halo mass distributions compared to inactive galaxies of similar stellar mass.

there are important differences between our data and this SAM that preclude a direct comparison. Above all, the SAM needs to better match other global properties of the galaxy distribution given the expectation that the same physical processes that regulate galaxy growth may also affect AGN activity. To first order, this requires matching the galaxy mass function and its evolution with redshift – a non-trivial task (recent progress on this topic is described by Vogelsberger et al. 2013; Benson 2014). In addition, the AGN weak lensing signal is more sensitive to high satellite fractions than to low satellite fractions (Fig. 3). An interesting direction for future work would be to use a joint analysis of lensing and clustering to pin down  $f_{\text{sat}}$  with greater accuracy.

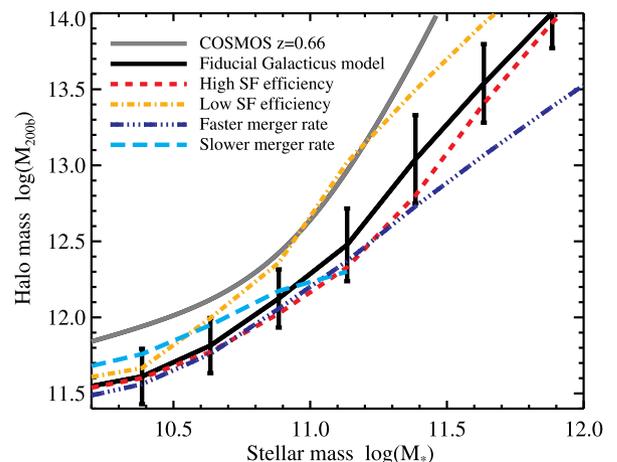
### 5.5 Insights from galacticus SAM: physical parameters that regulate the AGN HOD

In Section 4.4, we derived the occupation functions for this AGN sample (Fig. 6). However, HOD functions are simply a stepping stone towards the broader goal of improving observational constraints on mechanisms that fuel AGN activity. While a full discussion is beyond the scope of this paper (for a related discussion using hydrodynamic simulations, see Chatterjee et al. 2012), in this section we provide a qualitative assessment of how AGN HODs relate to theoretical models of AGN activity.

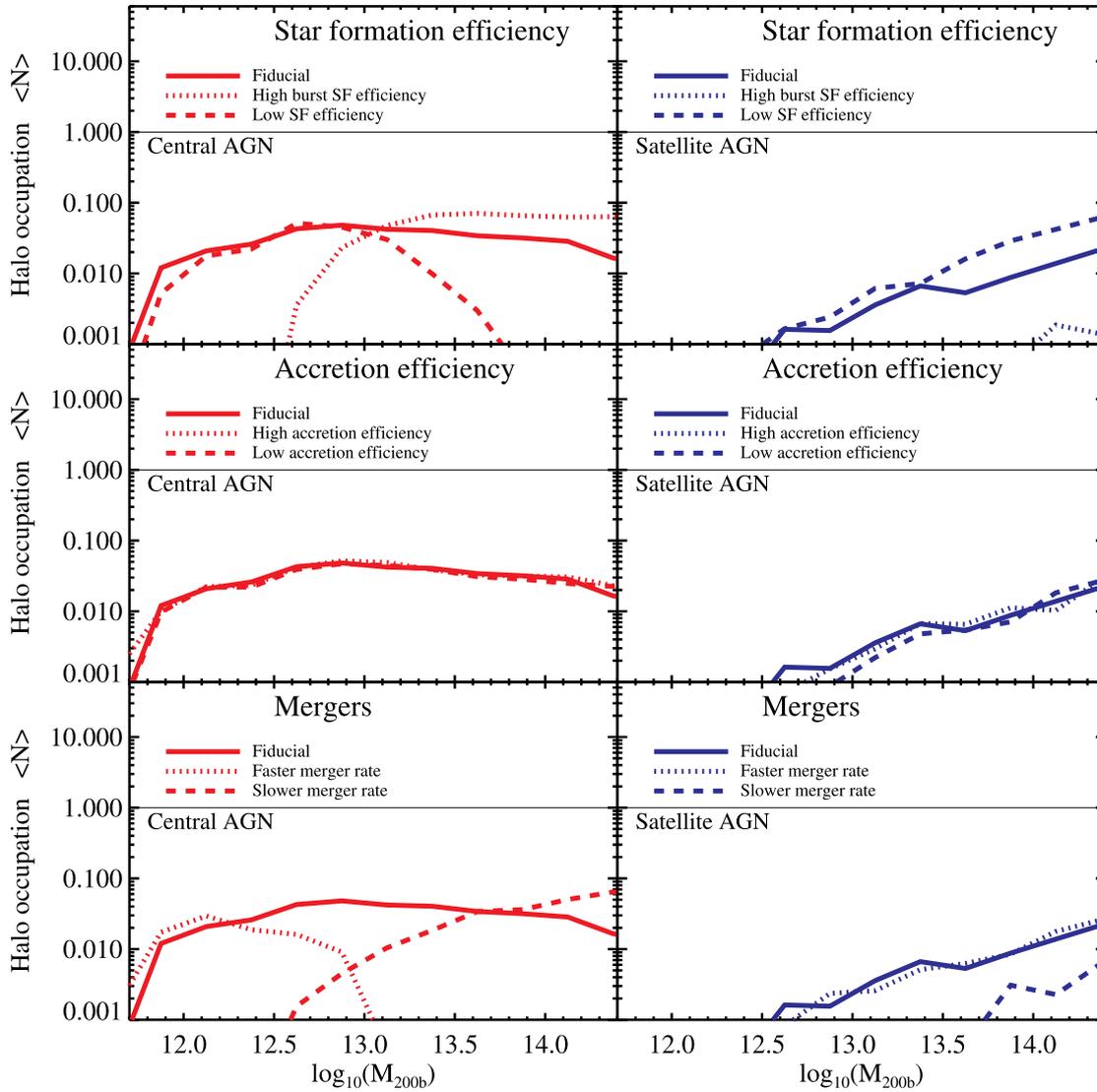
Fig. 9 plots Galacticus SHMRs for X-ray AGN at  $z \sim 0.6$  with the relation derived from the COSMOS data overlaid. Given that the global SHMR in Galacticus does not match the data, we do not attempt a direct comparison but simply show how physically informative parameters of the SAM might be constrained by the AGN HODs. We vary the following key parameters related to AGN activity in Galacticus and show their relative impact.

(1) *Star formation efficiency in bursts.* Models with efficiencies 10 and 1/10 times the fiducial efficiency are considered. More efficient star formation in bursts tends to reduce the net accretion of gas on to BHs (since gas becomes more quickly locked up in stars or ejected from the galaxy in winds).

(2) *Accretion efficiency on to BHs from the hot atmosphere.* Models with accretion from the hot atmosphere (which drives the



**Figure 9.** Variations of the Galacticus SHMR with the star formation efficiency and galaxy merger time-scales. Errors on the SHMR indicate the  $1\sigma$  scatter in the model relation at fixed stellar mass and are shown only on the fiducial model for visual purposes. In this SAM, modifying the accretion efficiency on to BHs in radio mode only has a minor effect on the SHMR and so these parameter variations are not displayed.



**Figure 10.** Impact of a few key parameters on the HODs of AGN selected in the Galacticus SAM that have been roughly selected to mimic our COSMOS sample. Note that we do not necessarily expect the Galacticus HODs to match the ones derived in this paper because our baseline Galacticus model does not reproduce the COSMOS SHMR. Our goal in this figure is to compare qualitative variations in the occupation functions.

‘radio mode’ AGN) at rates 10 and 1/10 times the fiducial rate are considered. The radio mode accretion rate controls the efficacy of AGN feedback in the Galacticus model. We find however, that modifying the accretion efficiency on to BHs in radio mode does not affect the SHMR because the strength of radio model feedback in this SAM is well above that required to completely shut down cooling in high-mass haloes. As such, reducing the accretion rate (and, therefore, the feedback power) by a large factor still leaves enough feedback power to shut down cooling, and increasing feedback power makes no real difference (since once cooling is shut down, more feedback can have no additional effect).

(3) *Galaxy merger time-scales.* Time-scales for galaxy–galaxy mergers (driven by dynamical friction and with time-scales computed using the fitting formula of Jiang et al. 2008) are varied by a factor of 10 above and below the time-scales in the fiducial model. Rapid merging of galaxies leads to more rapid build up of BH masses (both by direct merging of BHs and by driving gas into the spheroid where it may be accreted by the central BH).

Fig. 10 displays the effects of varying these four parameters on the Galacticus AGN HOD. One immediate point of interest here is that the general *shapes* of the HODs from Galacticus match those found in this paper quite well. Most of the HODs displayed in Fig. 10 would be relatively well described by equations (2) and (3) for  $N_{\text{cen}}$  and  $N_{\text{sat}}$ . However, although the shape of  $N_{\text{sat}}$  is well described by a power law, the overall amplitude of  $N_{\text{sat}}$  is quite low. As discussed already in the previous section, this is a manifestation of the fact that the satellite fraction for AGN is lower than for galaxies in this SAM.

Fig. 10 shows that radio-mode accretion efficiency has almost no effect on the AGN occupation functions. This is because in this SAM, radio-mode accretion mainly dominates for AGN with lower luminosities. On the other hand, star formation efficiency and mergers have a large impact on the characteristic halo mass scales of the central occupation functions as well as on the amplitude of the satellite occupation function. Although we have not explored this aspect in great detail, it is possible that a higher star formation efficiency reduces the HOD in lower mass haloes because gas is

efficiently used up by star formation instead of accreting on to the BHs. A low star formation efficiency in bursts may result in an enhanced satellite contribution because there is now more gas left in satellite spheroids to accrete on to their BHs. Finally, a low merger rate for galaxies may reduce the central galaxy HOD in low-mass haloes due to the lack of major mergers which drive gas on to the BHs.

However, as can be seen from Fig. 9, as we vary the star formation efficiency in bursts and the Galaxy merger time-scales, the SHMR also varies. As a result, it is difficult to know how much of the change in the model AGN HODs is due to changing the nature of BH growth and activity and how much is due to simply changing the SHMR. In practice, we would need to only explore models with viable SHMRs to ascertain how these physical processes directly affect the growth and fuelling of BHs. This points to an interesting direction for future research. The solution to this problem will be to first calibrate the Galacticus model to accurately match the measured SHMR. Using MCMC techniques as described in Benson (2014) would allow us to survey the entire model parameter space and locate those regions which adequately match the measured SHMR. Sampling model parameters from these regions of parameter space would then allow us to explore how the AGN HOD depends on model parameters once the SHMR is fixed.

## 6 SUMMARY AND CONCLUSION

In this paper, we have developed a new framework for studying how BH fuelling may be tied to host dark matter haloes by tying measurements of AGN host stellar masses to prior knowledge about the SHMR. In contrast with previous work, which only considered a single effective halo mass scale, the technique presented here can be used to infer the full halo mass distribution for AGN samples.

HOD modelling of AGN populations is fundamentally limited by model degeneracies driven by the fact that AGN may live a wide range of halo masses with an occupation function whose general shape and normalization are poorly known (e.g. Shen et al. 2013). Faced with this difficulty, we propose that instead of trying to constrain a full HOD model from AGN samples, we can ask a more simple, but no less fundamental question: How do AGN samples differ from inactive galaxies of equivalent stellar mass? This can be achieved through a rigorous comparison of the clustering, lensing, and cross-correlation signals of AGN hosts to the fiducial SHMR derived for all galaxies, irrespective of nuclear activity.

The key advantage of this approach is that by using large samples of galaxies that are complete in terms of stellar mass, the SHMR can be built with much higher accuracy than by using any statistic measured from AGN samples alone. Statistics measured from AGN are only used to constrain *deviations* from the fiducial model.

We have applied this technique to a sample of moderate luminosity ( $\langle \log(L_X) \rangle = 42.7$ ) obscured X-ray AGN at  $z < 1$  from the COSMOS field. Despite the small sample size (several hundred AGN), we demonstrate that our method can be used to constrain AGN halo occupation statistics. For the first time, we measure the galaxy–galaxy lensing signal of X-ray-selected obscured AGN. We find excellent agreement between the AGN lensing signal and the prediction based on our fiducial SHMR. There is no evidence from our analysis to suggest that AGN populate dark matter haloes in a different manner compared to galaxies with the same  $M_*$ , regardless of nuclear activity. We discuss how similar tests in future work could equally be well performed for the AGN autocorrelation function, or for cross-correlations between AGN and mass-limited galaxy samples.

In contrast with previous work which typically only provides a single effective halo mass scale, the technique presented here can be used to infer the full halo mass distribution for AGN samples. Contrary to conventional wisdom, our method suggests that most X-ray AGN do not live in medium-sized groups with  $M_h \geq 10^{13} M_\odot$ . Instead, 50 per cent of the AGN in our sample lives in haloes less massive than  $\log_{10}(M_{200b}) \sim 12.5$  and hence in relatively low-mass dark matter haloes. Only  $\sim 60$  per cent of AGN satellites are contained in haloes with  $\log_{10}(M_{200b}) > 13$ . We stress that these values are specific to our particular AGN sample selection and that the lower halo mass limit described here is primarily set by our choice to select an AGN sample with host masses above  $\log_{10}(M_*) > 10.5$ . Our work is consistent with moderate luminosity X-ray AGN occupying a wide range of halo masses. A similar picture is supported for luminous QSO samples by Conroy & White (2013). Taken together, these two results suggest that both QSOs and moderate luminosity X-ray AGN may occupy haloes in a relatively ‘normal’ way, calling into question previous claims for an environmental signature of distinct fuelling modes for QSOs compared to moderate luminosity X-ray AGN.

We compare our results with previous halo mass estimates inferred from X-ray clustering. We globally find that our predicted effective halo mass (measured in a consistent fashion as with clustering studies) is lower than previous work. However, we also caution that sample selection effects may be non-negligible when performing such comparisons and need to be considered carefully. Studies of X-ray AGN clustering are typically limited to bright hosts simply due to the fact that such measurements require spectroscopic redshifts. As a result, samples from previous work are typically limited to hosts with  $I_{AB} < 21.5\text{--}23$ . In detail, there are important variations in the selection functions applied to AGN samples between different studies. Differences in the dark matter halo distributions between various groups are in fact expected – hosts with brighter luminosities on average live in larger dark matter haloes. A fixed  $I$ -band cut will also probe different host stellar masses at different redshifts. We stress that these (sometimes implicit) cuts on host properties need to be accounted before meaningful comparisons can be made.

We derive the halo occupation functions for our sample and show that they are well described by the same functional form for galaxies but with an overall amplitude normalization of  $f_A \sim 0.028$ . At group scales, the satellite occupation is well described by a power law with a slope of  $\alpha = 1$ .

Finally, we investigate some simple models from the Galacticus SAM and find broadly consistent shapes for AGN HODs. However, in contrast with our lensing results, the SAMs predict an AGN satellite fraction that is an order of magnitude lower compared to the overall galaxy population. This suggests a tension worth investigating in future work using higher S/N weak lensing and clustering measurements for AGN host galaxies.

## ACKNOWLEDGEMENTS

We thank Phil Hopkins, Surhud More, and John Silverman for insightful discussions while preparing this paper. We also thank Ed Turner for valuable discussions related to statistical methods. We are grateful to Ian Harnett for editing this manuscript. We also thank Nikos Fanidakis for clarifications regarding halo mass values in Fanidakis et al. (2013). This work was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. AK is supported by the National Science Foundation Graduate Research Fellowship, grant no. DGE-1148900. FC acknowledges financial support by the NASA grant AR1-12012X.

RM is supported by a Royal Society University Research Fellowship. JR was supported by JPL, which is run by Caltech under a contract for NASA. ALC acknowledges support from NSF CAREER award AST-1055081.

## REFERENCES

- Aird J. et al., 2012, *ApJ*, 746, 90  
 Allevato V. et al., 2011, *ApJ*, 736, 99  
 Allevato V. et al., 2012, *ApJ*, 758, 47  
 Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C. G., 1999, *MNRAS*, 305, L21  
 Behroozi P. S., Conroy C., Wechsler R. H., 2010, *ApJ*, 717, 379  
 Benson A. J., 2012, *New Astron.*, 17, 175  
 Benson A. J., 2014, *MNRAS*, 444, 2599  
 Benson A. J., Babul A., 2009, *MNRAS*, 397, 1302  
 Bongiorno A. et al., 2012, *MNRAS*, 427, 3103  
 Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53  
 Brusa M. et al., 2010, *ApJ*, 716, 348  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Bundy K. et al., 2006, *ApJ*, 651, 120  
 Bundy K. et al., 2008, *ApJ*, 681, 931  
 Bundy K. et al., 2010, *ApJ*, 719, 1969  
 Capak P. et al., 2007, *ApJS*, 172, 99  
 Cappelluti N. et al., 2009, *A&A*, 497, 635  
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718  
 Chatterjee S., Degraf C., Richardson J., Zheng Z., Nagai D., Di Matteo T., 2012, *MNRAS*, 419, 2657  
 Civano F. et al., 2012, *ApJS*, 201, 30  
 Civano F., Fabbiano G., Pellegrini S., Kim D., Paggi A., Feder R., Elvis M., 2014, *ApJ*, 790, 16  
 Coil A. L. et al., 2009, *ApJ*, 701, 1484  
 Conroy C., White M., 2013, *ApJ*, 762, 70  
 Cooray A., Sheth R., 2002, *Phys. Rep.*, 372, 1  
 Croom S. M. et al., 2005, *MNRAS*, 356, 415  
 da Ângela J. et al., 2008, *MNRAS*, 383, 565  
 DeGraf C., Di Matteo T., Khandai N., Croft R., 2012, *ApJ*, 755, L8  
 Edgar R., 2004, *New Astron. Rev.*, 48, 843  
 Elvis M. et al., 2009, *ApJS*, 184, 158  
 Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, *MNRAS*, 410, 53  
 Fanidakis N. et al., 2013, *MNRAS*, 435, 679  
 Ferrarese L., Merritt D., 2000, *ApJ*, 539, L9  
 Gebhardt K. et al., 2000, *ApJ*, 539, L13  
 Gilli R. et al., 2009, *A&A*, 494, 33  
 Häring N., Rix H.-W., 2004, *ApJ*, 604, L89  
 Hearin A. P., Watson D. F., Becker M. R., Reyes R., Berlind A. A., Zentner A. R., 2014, *MNRAS*, 444, 729  
 Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, *ApJ*, 613, 109  
 Hopkins P. F., Richards G. T., Hernquist L., 2007, *ApJ*, 654, 731  
 Hopkins P. F., Hickox R., Quataert E., Hernquist L., 2009, *MNRAS*, 398, 333  
 Hütsi G., Gilfanov M., Sunyaev R., 2014, *A&A*, 561, A58  
 Ilbert O. et al., 2009, *ApJ*, 690, 1236  
 Jiang C. Y., Jing Y. P., Faltenbacher A., Lin W. P., Li C., 2008, *ApJ*, 675, 1095  
 Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, *A&A*, 440, 775  
 Kauffmann G., Heckman T. M., 2009, *MNRAS*, 397, 135  
 Kayo I., Oguri M., 2012, *MNRAS*, 424, 1363  
 Koekemoer A. M. et al., 2007, *ApJS*, 172, 196  
 Koutoulidis L., Plionis M., Georgantopoulos I., Fanidakis N., 2013, *MNRAS*, 428, 1382  
 Krumpe M., Miyaji T., Coil A. L., Aceves H., 2012, *ApJ*, 746, 1  
 Leauthaud A. et al., 2007, *ApJS*, 172, 219  
 Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, *ApJ*, 738, 45  
 Leauthaud A. et al., 2012a, *ApJ*, 744, 159 (L12)  
 Leauthaud A. et al., 2012b, *ApJ*, 746, 95  
 Li C., Kauffmann G., Wang L., White S. D. M., Heckman T. M., Jing Y. P., 2006, *MNRAS*, 373, 457  
 Lusso E. et al., 2011, *A&A*, 534, A110  
 McCracken H. J. et al., 2010, *ApJ*, 708, 202  
 Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, *MNRAS*, 368, 715  
 Mandelbaum R., Li C., Kauffmann G., White S. D. M., 2009, *MNRAS*, 393, 377  
 Martini P., Weinberg D. H., 2001, *ApJ*, 547, 12  
 Massey R. et al., 2007, *ApJS*, 172, 239  
 Miralda-Escude J., 1991, *ApJ*, 370, 1  
 Miyaji T., Krumpe M., Coil A. L., Aceves H., 2011, *ApJ*, 726, 83  
 More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, *MNRAS*, 392, 801  
 Moster B. P., Somerville R. S., Maubetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, 710, 903  
 Mountrichas G. et al., 2013, *MNRAS*, 430, 661  
 Nandra K. et al., 2007, *ApJ*, 660, L11  
 Narayan R., Yi I., 1994, *ApJ*, 428, L13  
 Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493  
 Rezzolla L., Barausse E., Dorband E. N., Pollney D., Reisswig C., Seiler J., Husa S., 2008, *Phys. Rev. D*, 78, 044002  
 Rhodes J. D. et al., 2007, *ApJS*, 172, 203  
 Richardson J., Chatterjee S., Zheng Z., Myers A. D., Hickox R., 2013, *ApJ*, 774, 143  
 Ross N. P. et al., 2009, *ApJ*, 697, 1634  
 Salvato M. et al., 2009, *ApJ*, 690, 1250  
 Salvato M. et al., 2011, *ApJ*, 742, 61  
 Schulze A., Wisotzki L., 2010, *A&A*, 516, A87  
 Scoville N. et al., 2007, *ApJS*, 172, 1  
 Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337  
 Shanks T., Croom S. M., Fine S., Ross N. P., Sawangwit U., 2011, *MNRAS*, 416, 650  
 Shen Y. et al., 2007, *AJ*, 133, 2222  
 Shen Y. et al., 2013, *ApJ*, 778, 98  
 Starikova S. et al., 2011, *ApJ*, 741, 15  
 Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878  
 Tinker J. L., Leauthaud A., Bundy K., George M. R., Behroozi P., Massey R., Rhodes J., Wechsler R. H., 2013, *ApJ*, 778, 93  
 Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031  
 Wilms J., Allen A., McCray R., 2000, *ApJ*, 542, 914  
 Wilson G., Kaiser N., Luppino G. A., Cowie L. L., 2001, *ApJ*, 555, 572  
 Yang X., Mo H. J., van den Bosch F. C., 2008, *ApJ*, 676, 248  
 Yang X., Mo H. J., van den Bosch F. C., 2009, *ApJ*, 695, 900

## APPENDIX A: THE GALACTICUS SAM

In Galacticus, BHs are assumed to accrete from both the interstellar medium in the spheroid of their host galaxy and the hot atmosphere of gas surrounding the host galaxy at rates governed by Bondi–Hoyle accretion (Edgar 2004) with a multiplicative pre-factor designed to take into account the fact that the model does not resolve the relevant length scales for accretion (Booth & Schaye 2009). The nature of the accretion disc surrounding each BH is determined by the accretion rate on to the BH. At accretion rates below 1 per cent or above 30 per cent of the Eddington accretion rate, the accretion disc is modelled as a radiatively inefficient, geometrically thick ADAF (Narayan & Yi 1994), otherwise a radiatively efficient, geometrically thin Shakura & Sunyaev (1973) solution is used. The evolution of BH spin is also tracked, using the method of Benson &

Babul (2009) to account for spin-up by accretion and spin-down by jet production. During galaxy mergers, BHs are assumed to merge instantaneously. The resulting merged BH has a mass equal to the sum of the masses of its progenitors and a spin computed using the method described by Rezzolla et al. (2008) assuming that the progenitor BHs have randomly aligned spin vectors. In comparison with the BH evolution model of Fanidakis et al. (2011), our model ignores the details of misaligned accretion discs–BH spins, but employs a more detailed model of accretion.<sup>9</sup> In other respects, our model and that of Fanidakis et al. (2011) are comparable in terms of the physics included and level of detail in the modelling.

Each galaxy therefore contains a supermassive BH with known mass, spin, and accretion rate. With the default parameters of our model the correlation between BH mass and spheroid stellar mass (Häring & Rix 2004) is approximately reproduced. The bolometric luminosity is computed from the BH rest mass accretion rate and radiative efficiency (assumed to be  $\epsilon_{\text{rad}} = 1 - E_{\text{ISCO}}(j)$  for a BH of spin  $j$  accreting via a thin accretion disc, where  $E_{\text{ISCO}}$  is the specific energy of material at the innermost stable circular orbit of the BH, while for a BH accreting from a radiatively inefficient

thick accretion flow the radiative efficiency is  $0.01\lambda/\lambda_{\text{thin}}$  where  $\lambda$  is accretion rate in units of the Eddington rate, and  $\lambda_{\text{thin}}$  is the minimum such accretion rate at which a thin disc occurs). An SED for an AGN of this bolometric luminosity is then computed using the model of Hopkins, Richards & Hernquist (2007). From this SED, a broad-band luminosity is computed in each X-ray band assuming a spectrum of the form  $f_{\nu} \propto \nu^{\alpha}$  with  $\alpha = -0.4$  as in the observational analysis.

Our COSMOS AGN sample is expected to be roughly obscured by a mean column density of  $N_{\text{H}} \sim 10^{22} \text{ cm}^{-2}$  with values extending out to  $N_{\text{H}} \sim 10^{23} \text{ cm}^{-2}$  (see fig. 2 in Lusso et al. 2011). Lusso et al. (2011) find a mean value of  $N_{\text{H}} \sim 10^{22} \text{ cm}^{-2}$  for a similarly selected sample of type-2 AGN from the *XMM*–COSMOS sample. In this Galacticus SAM, the X-ray luminosity,  $L_{\text{X}}$ , is attenuated from a fixed overall column density (mimicking a torus + ISM) of  $N_{\text{H}} = 10^{22} \text{ cm}^{-2}$  assuming solar metallicity. The photoelectric absorption cross-section per hydrogen is computed as a function of photon energy used (Wilms, Allen & McCray 2000). Multiplying by the hydrogen column density gives the net absorption as a function of energy to the AGN.

<sup>9</sup> In Fanidakis et al. (2011), a fixed fraction of the available gas mass is funnelled into BHs during each galaxy merger or disc instability event.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.