



|                                  |                                                                                                                                                                                     |
|----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Publication Year</b>          | 2019                                                                                                                                                                                |
| <b>Acceptance in OA</b>          | 2020-12-18T10:57:58Z                                                                                                                                                                |
| <b>Title</b>                     | Chandra COSMOS Legacy Survey: Clustering dependence of Type 2 active galactic nuclei on host galaxy properties                                                                      |
| <b>Authors</b>                   | Allevato, V., Viitanen, A., Finoguenov, A., Civano, F., Suh, H., Shankar, F., BONGIORNO, ANGELA, Ferrara, A., GILLI, Roberto, Miyaji, T., Marchesi, S., Cappelluti, N., Salvato, M. |
| <b>Publisher's version (DOI)</b> | 10.1051/0004-6361/201936191                                                                                                                                                         |
| <b>Handle</b>                    | <a href="http://hdl.handle.net/20.500.12386/28996">http://hdl.handle.net/20.500.12386/28996</a>                                                                                     |
| <b>Journal</b>                   | ASTRONOMY & ASTROPHYSICS                                                                                                                                                            |
| <b>Volume</b>                    | 632                                                                                                                                                                                 |

# Chandra COSMOS Legacy Survey: Clustering dependence of Type 2 active galactic nuclei on host galaxy properties

V. Allevato<sup>1,2,3</sup>, A. Viitanen<sup>2</sup>, A. Finoguenov<sup>2</sup>, F. Civano<sup>4</sup>, H. Suh<sup>5</sup>, F. Shankar<sup>6</sup>, A. Bongiorno<sup>7</sup>, A. Ferrara<sup>1</sup>,  
R. Gilli<sup>3</sup>, T. Miyaji<sup>8</sup>, S. Marchesi<sup>3,9,1</sup>, N. Cappelluti<sup>10</sup>, and M. Salvato<sup>11</sup>

<sup>1</sup> Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy  
e-mail: [viola.allevato@sns.it](mailto:viola.allevato@sns.it)

<sup>2</sup> Department of Physics and Helsinki Institute of Physics, University of Helsinki, Gustaf Hållströmin katu 2a,  
00014 Helsinki, Finland

<sup>3</sup> INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti, 93/3, 40129 Bologna, Italy

<sup>4</sup> Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA

<sup>5</sup> Subaru Telescope, National Astronomical Observatory of Japan (NAOJ), 650 North A'ohoku Place, Hilo, HI 96720, USA

<sup>6</sup> Department of Physics and Astronomy, University of Southampton, Highfield SO17 1BJ, UK

<sup>7</sup> INAF-Osservatorio Astronomico di Roma, Via Frascati 33, 00040 Monteporzio Catone, Italy

<sup>8</sup> Instituto de Astronomía sede Ensenada, Universidad Nacional Autónoma de México, KM 103, Carret. Tijuana-Ensenada,  
Ensenada 22860, Mexico

<sup>9</sup> Department of Physics and Astronomy, Clemson University, Kinard Lab of Physics, Clemson, SC 29634, USA

<sup>10</sup> Physics Department, University of Miami, Knight Physics Building, Coral Gables, FL 33124, USA

<sup>11</sup> Max Planck-Institute for Extraterrestrial Physics, PO Box 1312, Giessenbachstr. 1., 85741 Garching, Germany

Received 26 June 2019 / Accepted 14 October 2019

## ABSTRACT

**Aims.** We perform clustering measurements of 800 X-ray selected *Chandra* COSMOS Legacy (CCL) Type 2 active galactic nuclei (AGN) with known spectroscopic redshift to probe the halo mass dependence on AGN host galaxy properties, such as galaxy stellar mass  $M_{\text{star}}$ , star formation rate (SFR), and specific black hole accretion rate (BHAR;  $\lambda_{\text{BHAR}}$ ) in the redshift range  $z = [0-3]$ .

**Methods.** We split the sample of AGN with known spectroscopic redshifts according to  $M_{\text{star}}$ , SFR and  $\lambda_{\text{BHAR}}$ , while matching the distributions in terms of the other parameters, including redshift. We measured the projected two-point correlation function  $w_p(r_p)$  and modeled the clustering signal, for the different subsamples, with the two-halo term to derive the large-scale bias  $b$  and corresponding typical mass of the hosting halo.

**Results.** We find no significant dependence of the large-scale bias and typical halo mass on galaxy stellar mass and specific BHAR for CCL Type 2 AGN at mean  $z \sim 1$ , while a negative dependence on SFR is observed, i.e. lower SFR AGN reside in richer environment. Mock catalogs of AGN, matched to have the same X-ray luminosity, stellar mass,  $\lambda_{\text{BHAR}}$ , and SFR of CCL Type 2 AGN, almost reproduce the observed  $M_{\text{star}}-M_h$ ,  $\lambda_{\text{BHAR}}-M_h$  and SFR- $M_h$  relations, when assuming a fraction of satellite AGN  $f_{\text{AGN}}^{\text{sat}} \sim 0.15$ . This corresponds to a ratio of the probabilities of satellite to central AGN of being active  $Q \sim 2$ . Mock matched normal galaxies follow a slightly steeper  $M_{\text{star}}-M_h$  relation, in which low mass mock galaxies reside in less massive halos than mock AGN of similar mass. Moreover, matched mock normal galaxies are less biased than mock AGN with similar specific BHAR and SFR, at least for  $Q > 1$ .

**Key words.** galaxies: active – large-scale structure of Universe – quasars: general – dark matter

## 1. Introduction

Almost every galaxy in the Universe hosts a supermassive black hole (BH) at its center. During active phases, when the BH growth is powered by matter accretion, the galaxy is observed as an active galactic nucleus (AGN). Black hole masses tightly correlate with several properties of their host galaxy including stellar mass, velocity dispersion, and galaxy environment. These correlations suggest the existence of a fundamental link among BH growth, host galaxy structure and evolution, and cosmic large-scale structure, although the relative importance of the underlying physical processes is not yet fully understood.

Galaxies, and hence AGN, are not randomly distributed in space. On small scales, baryonic matter settles in the potential wells of virialized dark matter structures, the so-called halos. On large scales, the Universe displays coherent structures, with groups of galaxies sitting at the intersections of matter filaments,

i.e., “the cosmic web”. Clustering is commonly described as the distribution of AGN pairs as a function of their spatial separation, and it is a quantitative measure of the cosmic web topology. It also provides an indirect measurement of hosting dark matter halo masses, statistically classifies the typical AGN environment, and quantifies how active BHs populate halos.

Clustering measurements of different types of AGN have been carried out by several groups exploiting data from multiple surveys in diverse wavebands. Different typical hosting halo masses have been found in studies at different bands, ranging from  $\sim 10^{12}$  solar mass for optically selected quasars (e.g., Croom et al. 2005; Porciani & Norberg 2006; Shen et al. 2013; Ross et al. 2009) to dense environment typical of galaxy groups for X-ray selected AGN (Hickox 2009; Cappelluti et al. 2010; Allevato et al. 2011, 2012, 2014; Krumpe et al. 2010; Mountrichas et al. 2013; Koutoulidis et al. 2013; Plionis et al. 2018). However, this difference in the typical halo mass of X-ray

compared to optically selected AGN may not be present at low redshift (e.g., [Krumpe et al. 2012](#)) and cannot be explained at present.

Observational biases might be responsible for these different results. Recent studies (e.g., [Georgakakis et al. 2014](#); [Mendez et al. 2016](#); [Powell et al. 2018](#); [Mountrichas et al. 2019](#)) have suggested that AGN clustering can be entirely understood in terms of galaxy clustering and its dependence on galaxy parameters, such as stellar mass and star formation rate (SFR); and AGN selection effects. In detail, [Mendez et al. \(2016\)](#) compared the clustering of X-ray, radio, and infrared PRIMUS and DEEP2 AGN with matched galaxy samples designed to have the same stellar mass, SFR, and redshift and found no difference in the clustering properties. In this scheme, AGN selected using different techniques represent separate galaxy populations; the difference in the hosting dark matter halos is mainly driven by host galaxy properties. Clustering studies of large samples of AGN with known host galaxy properties have become then crucial to understand clustering of AGN.

The clustering dependence on galaxy stellar mass and then the relation between stellar/halo mass has been extensively studied in the last decade for normal galaxies at  $z \sim 1$  ([Zheng et al. 2007](#); [Wake et al. 2008](#); [Meneux et al. 2009](#); [Mostek et al. 2013](#); [Coil et al. 2017](#)) and at higher redshift ([Bielby et al. 2013](#); [Legrand et al. 2018](#)). In the sub-halo abundance matching technique, the number density of galaxies (from observations) and dark matter halos (from simulations) are matched to derive the stellar-to-halo mass relation at a given redshift (see, e.g., [Behroozi et al. 2013](#); [Behroozi & Silk 2018](#); [Mostek et al. 2013, 2018](#); [Shankar et al. 2016](#)). Moreover, other studies have used a halo occupation distribution modeling ([Zheng et al. 2007](#); [Leauthaud et al. 2010](#); [Coupon et al. 2015](#)), in which a prescription for how galaxies populate dark matter halos can be used to simultaneously predict the number density of galaxies and their spatial distribution.

The relation between the stellar mass content of a galaxy and the mass of its dark matter halo is still to investigate for active galaxies at all redshift. [Viitanen et al. \(2019\)](#) found no clustering dependence on host galaxy stellar mass and specific BHAR for a sample of XMM-COSMOS AGN in the range  $z = [0.1-2.5]$ . These authors also argued that the observed constant halo-galaxy stellar mass relation is due to a larger percentage of AGN in satellites (and then in more massive parent halos) in the low  $M_{\text{star}}$  bin compared to AGN in more massive host galaxies. [Mountrichas et al. \(2019\)](#) suggested that X-ray selected AGN and normal galaxies matched to have the same stellar mass, SFR, and redshift distributions, reside in similar halos and have similar dependence on clustering properties. They also found a negative clustering dependence on SFR, as also suggested in [Coil et al. \(2009\)](#), with clustering amplitude increasing with decreasing SFR (see also [Mostek et al. 2013](#) for non-active galaxies).

The goal of this work is to extend the clustering measurements performed in [Viitanen et al. \(2019\)](#) on XMM-COSMOS AGN to the new *Chandra*-COSMOS Legacy catalog, building one of the largest samples of X-ray selected AGN detected in a contiguous field. This sample of  $\sim 800$  CCL Type 2 AGN with available spectroscopic redshift in the range  $z = [0-3]$ , allows us to study the AGN clustering dependence on host galaxy stellar mass, specific BHAR, and SFR while matching the distributions in terms of the other parameters, including redshift.

Throughout this paper we use  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $\sigma_8 = 0.8$  and all distances are measured in comoving coordinates and are given in units of  $\text{Mpc } h^{-1}$ , where  $h = H_0/100 \text{ km s}^{-1}$ . The symbol  $\log$  signifies a base 10 logarithm. In the calculation of the

X-ray luminosities we fix  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  (i.e.,  $h = 0.7$ ). This is to facilitate comparison with previous studies that also follow similar conventions.

## 2. *Chandra* COSMOS Legacy Catalog

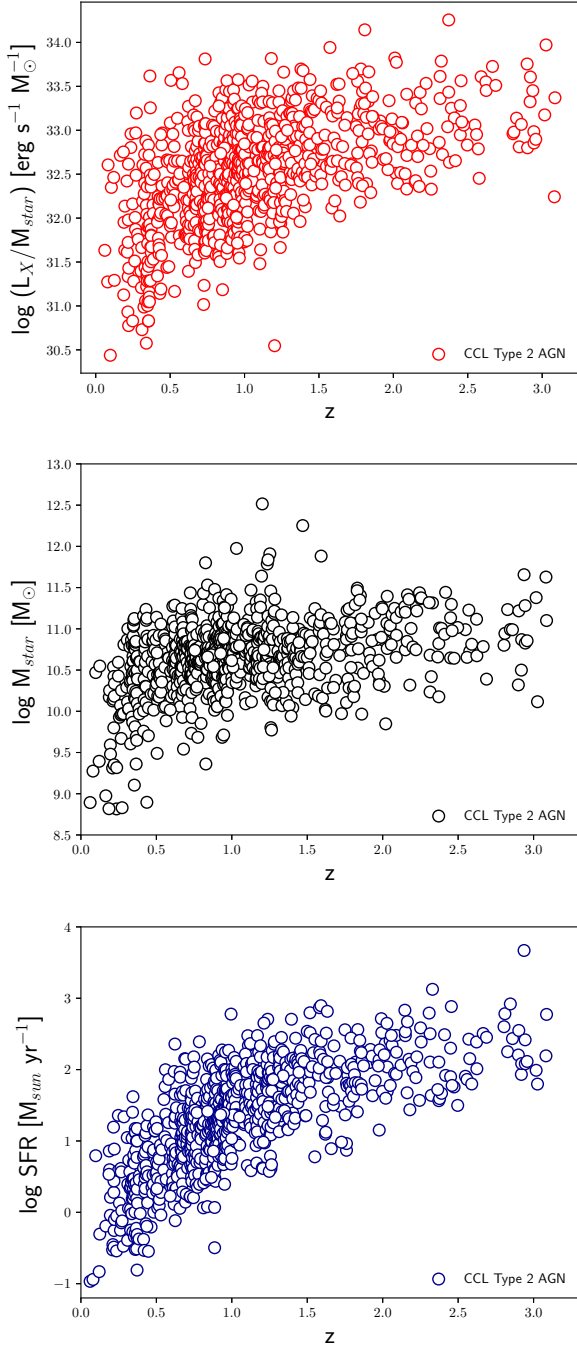
The *Chandra* COSMOS Legacy Survey CCL ([Civano et al. 2016](#)) is a large area, medium-depth X-ray survey covering  $\sim 2 \text{ deg}^2$  of the COSMOS field obtained by combining the 1.8 Ms *Chandra* COSMOS survey (C-COSMOS; [Elvis et al. 2009](#)) with 2.8 Ms of new *Chandra* ACIS-I observations. The CCLS is one of the largest samples of X-ray AGN selected from a single contiguous survey region, containing 4016 X-ray point sources, detected down to limiting fluxes of  $2.2 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$ ,  $1.5 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ , and  $8.9 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$  in the soft (0.5–2 keV), hard (2–10 keV), and full (0.5–10 keV) bands. As described in [Civano et al. \(2016\)](#) and [Marchesi et al. \(2016\)](#), 97% of CCL sources were identified in the optical and infrared bands and therefore photometric redshifts were computed. Thanks to the intense spectroscopic campaigns in the COSMOS field,  $\sim 54\%$  of the X-ray sources have been spectroscopically identified and classified. The full catalog of CCLS was presented by [Civano et al. \(2016\)](#) and [Marchesi et al. \(2016\)](#), including X-ray and optical/infrared photometric and spectroscopic properties.

The host galaxy properties of 2324 Type 2 CCL AGN have been studied in the redshift range  $z = [0-3]$  in [Suh et al. \(2017; 2019\)](#). These sources are classified as non-broad-line and/or obscured AGN (hereafter, “Type 2” AGN), i.e., they show only narrow emission-line and/or absorption-line features in their spectra or their photometric spectral energy distributions (SEDs) are best fitted by an obscured AGN template or a galaxy template. Making use of the existing multiwavelength photometric data available in the COSMOS field, these authors performed multicomponent modeling from far-infrared to near-ultraviolet using a nuclear dust torus model, a stellar population model, and a starburst model of the SEDs. Through detailed analyses of SEDs, they derived stellar masses in the range  $9 < \log M_{\text{star}}/M_\odot < 12.5$  with uncertainties of  $\sim 0.19$  dex. Moreover, SFR are estimated by combining the contributions from UV and IR luminosity. The total sample spans a wide range of SFRs ( $-1 < \log \text{SFR } [M_\odot \text{ yr}^{-1}] < 3.5$  with uncertainties of  $\sim 0.20$  dex).

For this study we selected 1701/2324 CCL Type 2 AGN detected in the soft band and focused on 884/1701 sources with known spectroscopic redshift up to  $z = 3$ . In order to study the AGN clustering dependence on host galaxy properties, we divided the sample according to the galaxy stellar mass, SFR, and specific BHAR  $\lambda_{\text{BHAR}} = L_X/M_{\text{star}}$ . The latter defines the rate of accretion onto the central BH scaled relative to the stellar mass of the host galaxy. To the extent that a proportionality between the BH mass and the host galaxy mass can be assumed, this ratio gives a rough measure of the Eddington ratio. Following [Bongiorno et al. \(2012, 2016\)](#) and [Aird et al. \(2018\)](#),

$$\lambda_{\text{Edd}} = \frac{k_{\text{bol}} \cdot A}{1.3 \times 10^{38}} \times \frac{L_X}{M_{\text{star}}}, \quad (1)$$

where  $L_X$  is the intrinsic X-ray luminosity in  $\text{erg s}^{-1}$ ,  $k_{\text{bol}}$  is a bolometric correction factor, and  $M_{\text{star}}$  is the total stellar mass of the AGN host galaxy in units of  $M_\odot$ . The factor  $A$  is a constant if the BH mass can be related to the host galaxy mass through scaling relations (with  $A \approx 500-1000$ ; [Magorrian et al. 1998](#); [Haring & Rix 2004](#)). Thus, for a mean bolometric correction of  $k_{\text{bol}} = 25$  and a constant host stellar to BH mass ratio of



**Fig. 1.** Specific BHAR (upper panel), host galaxy stellar mass (middle panel), and SFR (lower panel) as a function of spectroscopic redshifts for 884 CCL Type 2 AGN with known spectroscopic redshifts.

$A = 500$ , a ratio of  $L_X/M_{\text{star}} = 10^{34} [\text{erg s}^{-1} M_{\text{star}}]$  would approximately correspond to the Eddington limit. The specific BHAR distribution can then be regarded as a tracer of the distribution of Eddington ratios, i.e.  $\lambda_{\text{Edd}} \propto \lambda_{\text{BHAR}}$  for a fixed  $k_{\text{bol}}$ . Figure 1 shows the distribution of the specific BHAR, host galaxy stellar mass, and SFR for the sample of 884 CCL Type 2 AGN with known spectroscopic redshift in the range  $z = [0-3]$ .

We then made subsamples in bins of galaxy stellar mass, specific BHAR, and SFR. Specifically, to avoid selection effects between the different bins of stellar mass (see Powell et al. 2018) we defined two bins of  $M_{\text{star}}$  and then for each bin, we randomly selected  $N$  AGN of the sample, with  $N$  the larger

number of sources in the bin, to match the distributions in redshift,  $\lambda_{\text{BHAR}}$  and SFR. Similarly, we defined two bins in specific BHAR (SFR) and randomly selected subsamples with similar redshift, stellar mass, and SFR ( $\lambda_{\text{BHAR}}$ ) distributions. The final subsamples consist of (a) 362 low and 374 high galaxy stellar mass AGN using a cut at  $\log(M_{\text{star}}/M_{\odot}) = 10.7$  (see Fig. 2); (b) 339 low and 326 high specific BHAR AGN with a cut at  $\log(\lambda_{\text{BHAR}}/[\text{erg s}^{-1} M_{\text{star}}^{-1}]) = 32.6$  (see Fig. 3); and (c) 262 low and 247 high SFR AGN cutting at  $\log(\text{SFR}/M_{\text{star}} \text{ yr}^{-1}) = 1.4$  (see Fig. 4). The characteristics of these subsamples are summarized in Table 1.

### 3. Projected 2pcf

The most commonly used quantitative measure of large-scale structure is the 2pcf,  $\xi(r)$ , which traces the amplitude of AGN clustering as a function of scale. The quantity  $\xi(r)$  is defined as a measure of the excess probability  $dP$ , above what is expected for an unclustered random Poisson distribution, of finding an AGN in a volume element  $dV$  at a separation  $r$  from another AGN,

$$dP = n[1 + \xi(r)]dV, \quad (2)$$

where  $n$  is the mean number density of the AGN sample (Peebles 1980). Measurements of  $\xi(r)$  are generally performed in comoving space, where  $r$  is measured in units of  $h^{-1} \text{ Mpc}$ .

We measured the projected 2pcf in bins of  $r_p$  and  $\pi$  (distances perpendicular and parallel to the line of sight, respectively) using CosmoBolognaLib, a large set of Open Source C++ numerical libraries for cosmological calculations (Marulli et al. 2016), which counts the number of pairs of galaxies in a catalog separated by  $r_p$  and  $\pi$ . We then integrated along the  $\pi$ -direction to eliminate any redshift-space distortion, and we estimated the so-called projected correlation function  $w_p(r_p)$  (Davis & Peebles 1983) as

$$w_p(r_p) = 2 \int_0^{\pi_{\text{max}}} \xi(r_p, \pi) d\pi, \quad (3)$$

where  $\xi(r_p, \pi)$  is the two-point correlation function in terms of  $r_p$  and  $\pi$ , measured using the Landy & Szalay (1993, LS) estimator

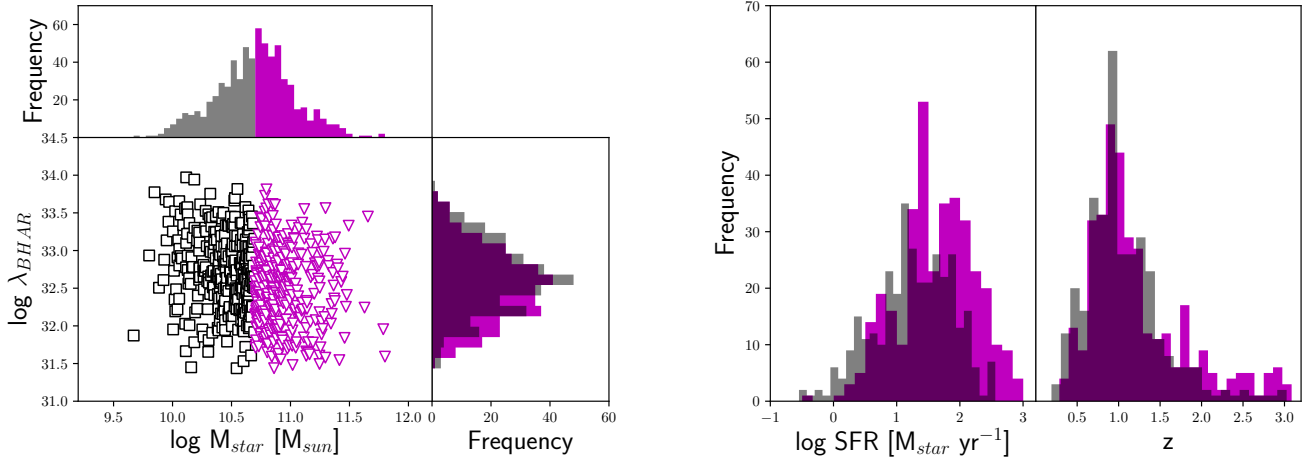
$$\xi = \frac{1}{RR'} [DD' - 2DR' + RR'], \quad (4)$$

where  $DD'$ ,  $DR'$ , and  $RR'$  are the normalized data-data, data-random, and random-random pairs.

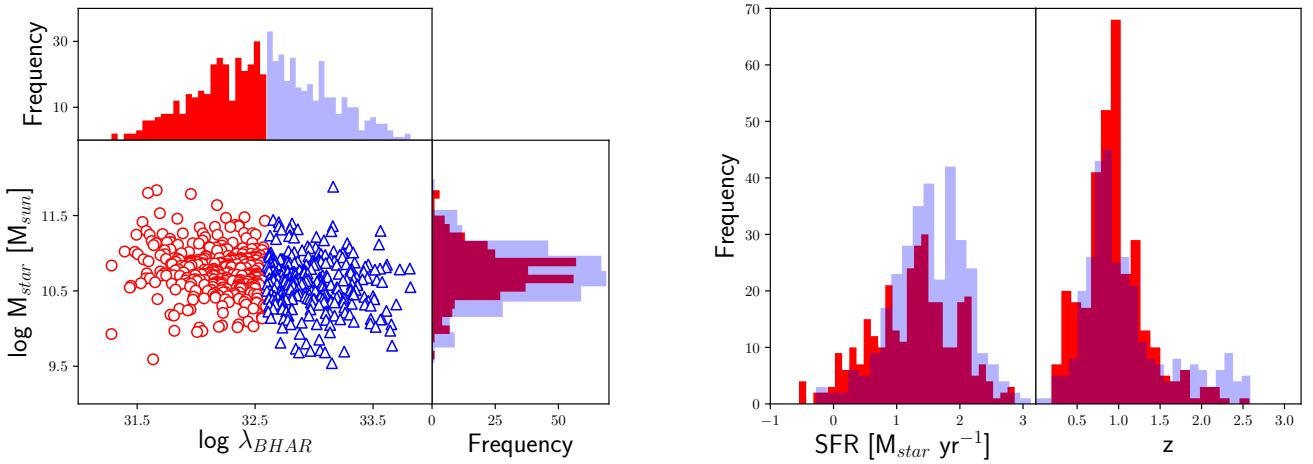
The measurements of the 2pcf requires the construction of a random catalog with the same selection criteria and observational effects as the data. To this end, we constructed a random catalog in which each simulated source is placed at a random position in the sky, with its flux randomly extracted from the catalog of real source fluxes. The simulated source is kept in the random sample if its flux is above the sensitivity map value at that position (Miyaji et al. 2007; Cappelluti et al. 2009). The corresponding redshift for each random object is then assigned based on the smoothed redshift distribution of the AGN sample.

The value of  $\pi_{\text{max}}$  is chosen such that the amplitude of the projected 2pcf converges and gets noisier for any higher values. We calculated the covariance matrix via the jackknife resampling method as follows:

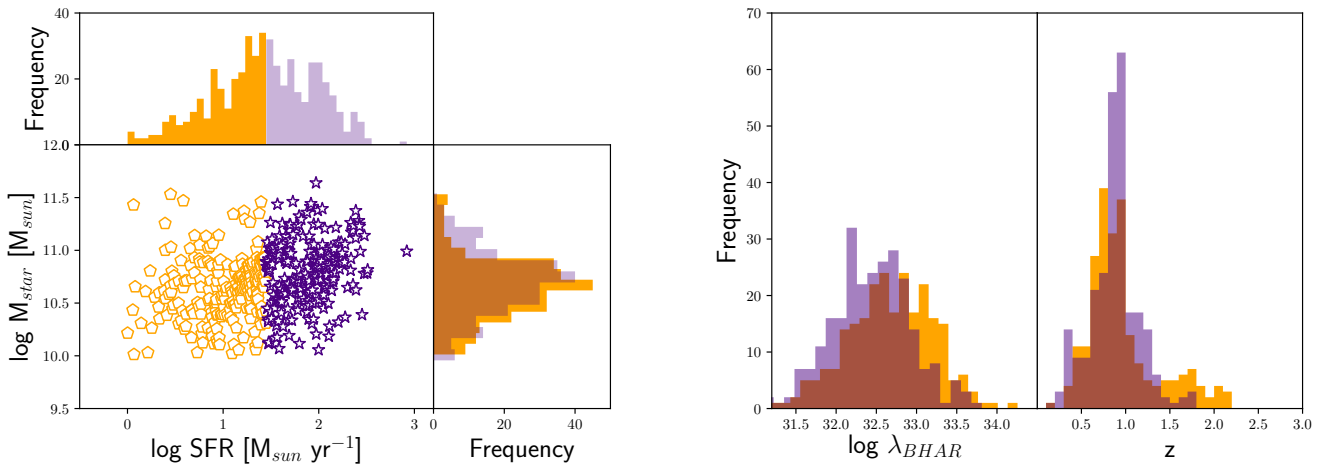
$$C_{i,j} = \frac{M}{M-1} \sum_k^M [w_k(r_{p,i}) - \langle w(r_{p,i}) \rangle] \times [w_k(r_{p,j}) - \langle w(r_{p,j}) \rangle], \quad (5)$$



**Fig. 2.** Host galaxy stellar mass as a function of specific BHAR (*left panel*) for low ( $\log M_{\text{star}}/[M_{\odot}] \lesssim 10.75$ ) and high ( $>10.75$ ) stellar mass subsamples. The corresponding distribution in terms of  $M_{\text{star}}$ ,  $\lambda_{\text{BHAR}}$ , SFR, and spectroscopic redshift (*right panel*) are shown for the two AGN subsets.



**Fig. 3.** Specific BHAR as a function of host galaxy stellar mass (*left panel*) for low ( $\log \lambda_{\text{BHAR}} \lesssim 32.6$ ) and high ( $>32.6$ ) specific BHAR subsamples. The corresponding distribution in terms of  $M_{\text{star}}$ ,  $\lambda_{\text{BHAR}}$ , SFR, and spectroscopic redshift (*right panel*) are shown for the two AGN subsets.



**Fig. 4.** Star formation rate as a function of host galaxy stellar mass (*left panel*) for low ( $\log \text{SFR} \lesssim 1.4$ ) and high ( $>1.4$ ) SFR subsamples. The corresponding distribution in terms of  $M_{\text{star}}$ ,  $\lambda_{\text{BHAR}}$ , SFR, and spectroscopic redshift (*right panel*) are shown for the two subsets.

where we split the sample into  $M = 9$  sections of the sky, and computed the cross-correlation function when excluding each section ( $w_k$ ). We quote the errors on our measurement as the square root of the diagonals,  $\sigma_i = \sqrt{C_{i,i}}$ .

In the halo model approach, the large-scale amplitude signal is due to the correlation between objects in distinct halos and the bias parameter defines the relation between the large-scale clustering amplitude of the AGN correlation function and the

**Table 1.** Properties of the AGN samples.

|                              | Sample | $z$<br>median | $\log L_X/M_{\text{star}}$<br>( $\text{erg s}^{-1}/M_{\odot}$ ) | $\log M_{\text{star}}$<br>( $M_{\odot}$ ) | $\log \text{SFR}$<br>( $M_{\odot} \text{ yr}^{-1}$ ) | $b$                    | $\log M_{\text{typ}}$<br>( $\text{h}^{-1} M_{\odot}$ ) |
|------------------------------|--------|---------------|-----------------------------------------------------------------|-------------------------------------------|------------------------------------------------------|------------------------|--------------------------------------------------------|
| Low $M_{\text{star}}$        | 362    | 0.96          | 32.7                                                            | 10.4                                      | 1.4                                                  | $2.14^{+0.08}_{-0.08}$ | $12.93^{+0.06}_{-0.06}$                                |
| High $M_{\text{star}}$       | 374    | 1.01          | 32.5                                                            | 10.9                                      | 1.6                                                  | $2.31^{+0.09}_{-0.09}$ | $13.03^{+0.07}_{-0.07}$                                |
| Low $\lambda_{\text{BHAR}}$  | 300    | 0.91          | 32.2                                                            | 10.7                                      | 1.4                                                  | $2.32^{+0.11}_{-0.11}$ | $13.02^{+0.08}_{-0.08}$                                |
| High $\lambda_{\text{BHAR}}$ | 306    | 1.03          | 32.9                                                            | 10.6                                      | 1.5                                                  | $2.40^{+0.10}_{-0.10}$ | $13.11^{+0.06}_{-0.06}$                                |
| Low SFR                      | 251    | 0.92          | 32.5                                                            | 10.7                                      | 1.0                                                  | $2.43^{+0.06}_{-0.06}$ | $13.14^{+0.06}_{-0.06}$                                |
| High SFR                     | 260    | 1.09          | 32.6                                                            | 10.6                                      | 1.9                                                  | $2.16^{+0.08}_{-0.08}$ | $12.94^{+0.06}_{-0.06}$                                |

DM two-halo term as follows:

$$b^{2-h}(r_p) = (w_{\text{AGN}}(r_p)/w_{\text{DM}}^{2-h}(r_p))^{1/2}. \quad (6)$$

We first estimated the DM two-halo term at the median redshift of the sample, using

$$w_{\text{DM}}^{2-h}(r_p) = r_p \int_{r_p}^{\infty} \frac{\xi_{\text{DM}}^{2-h}(r) r dr}{\sqrt{r^2 - r_p^2}}, \quad (7)$$

where

$$\xi_{\text{DM}}^{2-h}(r) = \frac{1}{2\pi^2} \int P^{2-h}(k) k^2 \left[ \frac{\sin(kr)}{kr} \right] dk. \quad (8)$$

The expression  $P^{2-h}(k)$  is the Fourier transform of the linear power spectrum. In particular, we based our estimation of the linear power spectrum on Eisenstein & Hu (1999), which is also implemented in CosmoBolognaLib.

## 4. Results

### 4.1. Large-scale bias versus $M_{\text{star}}$

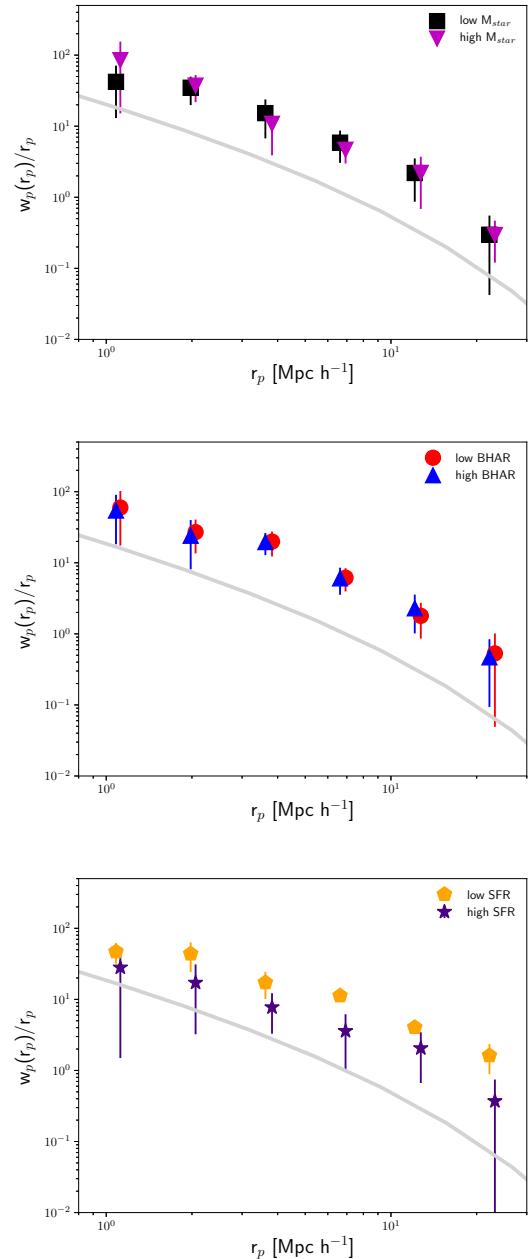
The goal of this paper is to study the AGN clustering dependence on host galaxy stellar mass, specific BHAR and SFR. In this section, we focus on the clustering properties of CCL Type 2 AGN as a function of host stellar mass  $M_{\text{star}}$ .

As shown in Fig. 5 (upper panel), the projected 2pcf  $w_p(r_p)$  of 362 (374) AGN with  $\log(M_{\text{star}}/M_{\odot}) < 10.7$  ( $\geq 10.7$ ) was measured in the  $r_p$  range  $1\text{--}30 \text{ h}^{-1} \text{ Mpc}$ , following Eq. (3). The typical value of  $\pi_{\text{max}}$  used in clustering measurements of both optically selected luminous quasars and X-ray selected AGN is  $\sim 20\text{--}100 \text{ h}^{-1} \text{ Mpc}$  (e.g., Zehavi et al. 2005; Coil et al. 2009; Krumpe et al. 2010; Allevato et al. 2011). We found the optimal  $\pi_{\text{max}}$  value to be  $\sim 60 \text{ h}^{-1} \text{ Mpc}$ , by deriving the value at which the amplitude of the signal appears to level off.

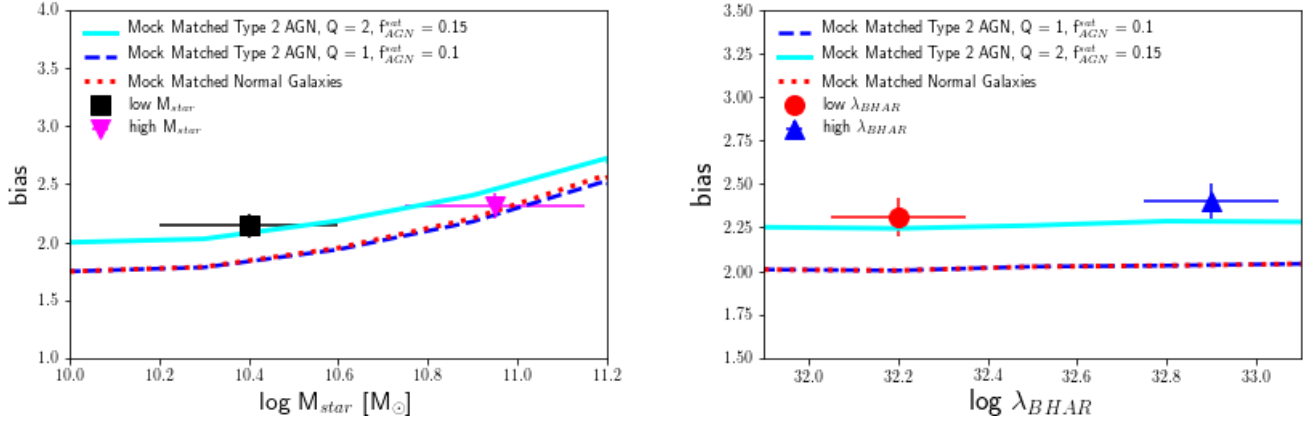
The  $1\sigma$  errors on  $w_p(r_p)$  are the square root of the diagonal components of the covariance matrix (Miyaji et al. 2007; Krumpe et al. 2010; Allevato et al. 2016) estimated using the jackknife method. The latter quantifies the level of correlation between different bins.

Following Eq. (6), we derived the best-fit bias using a  $\chi^2$  minimization technique with one free parameter  $\chi^2 = \Delta^T M_{\text{cov}}^{-1} \Delta$ . In detail,  $\Delta$  is a vector composed of  $w_p(w_p) - w_{\text{mod}}(r_p)$ ,  $\Delta^T$  is its transpose, and  $M_{\text{cov}}^{-1}$  is the inverse of covariance matrix. The latter full covariance matrix is used in the fit to take into account the correlation between errors.

As shown in Table 1, we derived for the low and high  $M_{\text{star}}$  subsamples  $b = 2.14^{+0.08}_{-0.08}$  and  $b = 2.31^{+0.09}_{-0.09}$  at mean  $z \sim 1$ , respectively. Following the bias–mass relation  $b(M_h, z)$  described in



**Fig. 5.** Projected 2pcf of CCL COSMOS Type 2 AGN as a function of *panel a:* host galaxy stellar mass,  $\log M_{\text{star}} < 10.7$  (black squares), and  $\geq 10.7$  (magenta triangles); *panel b:* specific BHAR,  $\log \lambda_{\text{BHAR}} < 32.6$  (red circles), and  $\geq 32.6$  (blue triangles); *panel c:*  $\log \text{SFR} < 1.4$  (orange pentagons) and  $\geq 1.4$  (purple stars). The gray line shows the DM projected 2pcf at mean  $z \sim 1$ .



**Fig. 6.** Large-scale bias evolution as a function of host galaxy stellar mass (*left panel*) and specific BHAR (*right panel*) for CCL Type 2 AGN and mock AGN. These are matched to have the same host galaxy properties of CCL Type 2 AGN, for  $Q = 1$  and  $=2$  (according to the legend), which correspond to a relative fraction of satellite AGN  $f_{\text{AGN}}^{\text{sat}} = 0.1$  and  $=0.15$ , respectively. The error bars on  $x$ -axis represent the typical error on the stellar mass and specific BHAR estimates in COSMOS. For comparison, the red dotted lines show the large-scale bias as a function of  $M_{\text{star}}$  and  $\lambda_{\text{BHAR}}$  for mock matched normal galaxies.

van den Bosch (2002) and Sheth et al. (2001), the large-scale bias values correspond to typical masses of the hosting halos of  $\log(M_{\text{h}}/M_{\odot} h^{-1}) = 12.93^{+0.06}_{-0.06}$  and  $13.03^{+0.07}_{-0.07}$ , for the low and high stellar mass subsamples, respectively. It is worth noting that these two AGN subsets have similar distributions in terms of specific BHAR, SFR, and redshift. Throughout this work, we refer to typical halo mass as the DM halo mass that satisfies  $b = b(M_{\text{halo}})$  (e.g., Hickox 2009; Allevato et al. 2014, 2016; Mountrichas et al. 2019).

#### 4.2. AGN bias versus $\lambda_{\text{BHAR}}$ and SFR

In this section we investigate the clustering dependence of CCL Type 2 COSMOS AGN on specific BHAR and SFR. For this purpose, we estimated the projected 2pcf of 300 low and 306 high specific BHAR AGN using a cut at  $\log(\lambda_{\text{BHAR}} \text{ erg s}^{-1} M_{\text{star}}^{-1}) = 32.6$  (Fig. 5, middle panel) and (b) 251 low and 260 high SFR AGN cutting at  $\log(\text{SFR}/M_{\text{star}} \text{ yr}^{-1}) = 1.4$  (Fig. 5, lower panel).

In detail, we found a large-scale bias  $b = 2.32^{+0.11}_{-0.11}$  and  $b = 2.40^{+0.10}_{-0.10}$  for the low and high  $\lambda_{\text{BHAR}}$ , respectively. These results suggest no bias evolution with specific BHAR, with a corresponding typical mass of the hosting halos of  $\log(M_{\text{h}}/M_{\odot} h^{-1}) = 13.02^{+0.08}_{-0.08}$  and  $13.11^{+0.06}_{-0.06}$ , respectively. On the contrary, AGN with low SFR are more clustered and reside in more massive dark matter halos ( $\log M_{\text{h}}/M_{\odot} h^{-1} = 13.14^{+0.06}_{-0.06}$ ) compared to high SFR objects ( $\log M_{\text{h}}/M_{\odot} h^{-1} = 12.94^{+0.06}_{-0.06}$ ).

## 5. Discussion

### 5.1. Clustering dependence on stellar mass

We have performed clustering measurements as a function of host galaxy stellar mass, specific BHAR, and SFR using CCL Type 2 AGN at mean  $z \sim 1$ . In particular, our results suggest a constant bias evolution as a function of  $M_{\text{star}}$  for Type 2 AGN in the particular stellar mass and redshift range investigated in this work (Fig. 6).

As shown in Fig. 7, our results are in agreement with the  $M_{\text{star}}-M_{\text{h}}$  relation found in Viitanen et al. (2019) for XMM-COSMOS AGN at similar redshift. It is worth noting that although our study includes XMM-COSMOS AGN, we only focused on Type 2 AGN. Moreover, our larger CCL COSMOS

catalog allows us to split the sample according to the galaxy stellar mass while having matched SFR, specific BH accretion rate, and  $z$  distributions.

Our results are also in agreement with Mountrichas et al. (2019), albeit they inferred a slightly steeper  $M_{\text{star}}-M_{\text{h}}$  relation at mean  $z \sim 0.8$ . In detail, they found that high stellar mass XMM-XXL AGN (both Type 1 and Type 2 sources) resides in slightly more massive halos than low stellar mass objects, when applying a cut at  $\log(M_{\text{star}}/M_{\odot}) = 10.8$  (10.7 in our work). They also showed that the same  $M_{\text{star}}-M_{\text{h}}$  relation is observed for a matched sample of normal non-active galaxies.

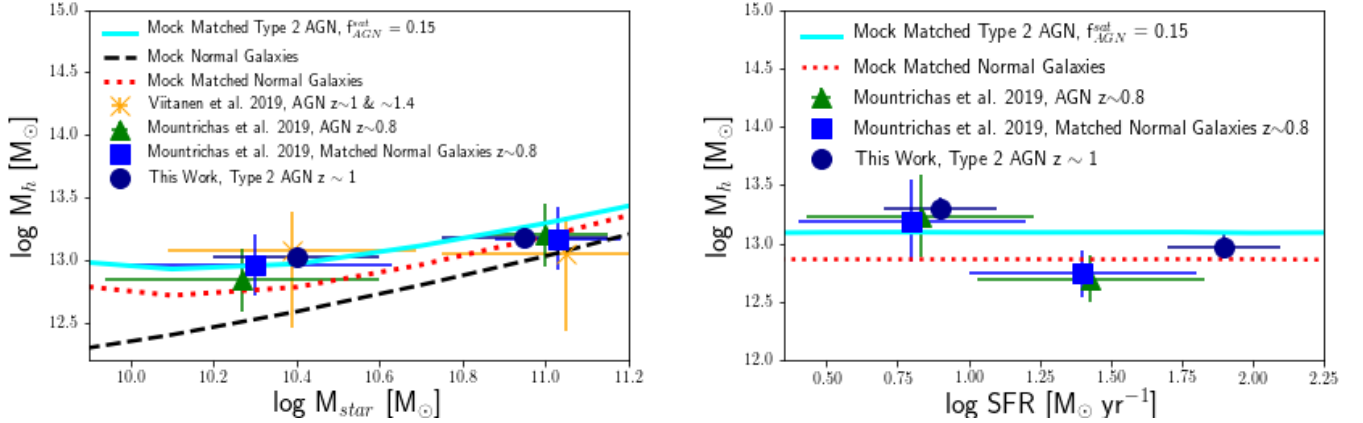
The relation between galaxies and their hosting halos has been derived for normal galaxies in different clustering studies (Zheng et al. 2007; Wake et al. 2008; Mostek et al. 2013; Coil et al. 2007) at similar ( $z \sim 1$ ) and higher redshift (Bielby et al. 2013; Legrand et al. 2018). These results suggest a more steeper  $M_{\text{star}}-M_{\text{h}}$  relation for non-active galaxies than found for CCL Type 2 AGN and from previous clustering measurements of AGN and for matched normal galaxies.

### 5.2. Clustering dependence on $\lambda_{\text{BHAR}}$ and SFR

We found a constant typical dark matter halo mass as a function of specific BHAR,  $\lambda_{\text{BHAR}}$ . The same trend has been observed at similar redshift for XMM-COSMOS AGN by Viitanen et al. (2019) and at lower redshift ( $0.2 < z < 1.2$ ) by Mendez et al. (2016), using PRIMUS and DEEP2 redshift surveys. Assuming a bolometric correction, the specific BHAR can be considered as a proxy of the Eddington ratio (Eq. (1)). Krumpe et al. (2015) found no statistically significant clustering dependence on  $\lambda_{\text{Edd}}$  for RASS Type 1 AGN at  $0.16 < z < 0.36$ . Our results suggest the same constant bias evolution with Eddington ratio for both RASS Type 1 and COSMOS Type 2 AGN at different redshifts.

On the contrary, we found a negative bias dependence on SFR (see Fig. 7), with lower SFR AGN more clustered than higher SFR objects. This clustering result suggests that splitting the sample according to SFR leads to larger hosting halo mass being associated with lower SFR AGN; i.e., given the same galaxy stellar mass distribution, AGN in low SFR galaxies reside in more massive halos than AGN in high SFR hosts.

Clustering analysis of AGN as a function of host galaxy SFR has been performed by Mountrichas et al. (2019) for X-ray selected XMM-XXL AGN at  $z = [0.5-1.2]$ . In particular, they



**Fig. 7.** Galaxy stellar mass as a function of dark matter halo mass as derived for CCL Type 2 AGN (dark blue circles) at mean  $z \sim 1$ , XMM-COSMOS AGN (orange stars) at  $z \sim 1$  (low  $M_{\text{star}}$ ) and  $\sim 1.4$  (high  $M_{\text{star}}$ ) in Viitanen et al. (2019), XMM-XXL AGN (green triangles), and matched normal galaxies (blue squares) in Mountrichas et al. (2019) at mean  $z \sim 0.8$ . The error bars on  $x$ -axis represent typical measurement error on the stellar mass estimates of each subsample. The  $M_{\text{star}}-M_h$  and  $M_{\text{star}}-\text{SFR}$  relations for matched mock AGN (continuous gray line) and mock matched normal galaxies (dotted red line) are shown. For comparison, the halo-stellar mass relation is shown for the full sample of mock galaxies as a dashed black line.

split the AGN sample using a cut at  $\log (\text{SFR}/M_{\text{star}} \text{ yr}^{-1}) = 1.1$  (1.4 in our study) and derived typical halo masses of  $\log (M_h/M_\odot h^{-1}) = 13.08^{+0.38}_{-0.32}$  and  $12.54^{+0.22}_{-0.18}$  at  $z \sim 0.8$ , for the low and high SFR subsets, respectively. These results agree very well with our findings for CCL Type 2 AGN at higher redshift, suggesting a negative clustering dependence as a function of SFR.

A similar trend has been observed for normal non-active galaxies in different studies. Mostek et al. (2013) studied stellar mass-limited samples of DEEP2 galaxies and found that the clustering amplitude increases with decreasing SFR. Similarly, Coil et al. (2007) suggested a strong evolution of the large-scale bias with SFR (and sSFR) in PRIMUS and DEEP2 surveys.

Our measured  $\text{SFR}-M_h$  relation agrees with the environment quenching picture, where most galaxies in clusters are passive, regardless of their mass (e.g., Peng et al. 2010a,b, 2012). In particular, the passive fraction in both central and satellite galaxies strongly correlates with the halo mass at fixed stellar mass. Above a characteristic halo mass ( $\sim 10^{12} M_\odot$ ) cooling times are long, and the gas that accretes onto the galaxy is hot, so star formation is inefficient (e.g., Gabor & Dave 2015; Birnboim & Dekel 2003, see also Peng et al. 2015, for observational evidence for quenching via gas exhaustion, or “strangulation”). However, CCL Type 2 AGN also cover the redshift range in which a difference in galaxy SFR properties with environments is ceasing (e.g., George et al. 2013; Erfanianfar et al. 2016 at  $z \lesssim 1.2$ ); the main difference is the fraction of bulge-dominated galaxies as a function of halo mass. Then the SFR cut might also lead to higher weight of bulge-dominated galaxies, which reside in more massive halos.

### 5.3. Comparison with AGN mock catalogs

#### 5.3.1. Methodology

Several semi-analytical models and hydrodynamical simulations (e.g., Springel et al. 2005; Hopkins et al. 2006; Menci et al. 2008) have been developed in recent years to describe the main mechanisms that fuel the central BHs. With suitable adjustment of parameters, these models can explain many aspects of AGN phenomenology (e.g., Hopkins et al. 2006, 2008). However, our scant knowledge of the key processes imposes a heavy parameterization of the physics regulating the cooling, star formation,

feedback, and merging of baryons. Thus, current models present serious degeneracies, i.e., they reproduce similar observables by invoking very different scenarios (Lapi & Cavaliere 2011). Instead, in semiempirical models (SEMs) variables such as galaxy stellar mass, BH mass, and AGN luminosity are assigned through a combination of observational and theoretical scaling relations. These SEMs represent an original and competitive methodology, which is fast, flexible and relies on just a few input assumptions. Mock catalogs of galaxies and their BHs can be created via semiempirical relations starting from large samples of dark matter halos extracted from N-body simulations. Therefore, AGN mock catalogs effectively provide a complementary approach to more complex models of AGN and galaxy evolution (e.g., Conroy & White 2013), which can be directly compared with current and future clustering measurements. The SEM for the large-scale distribution of AGN has been used to make realistic predictions for the clustering signal in future experiments and test observational selection effects and biases (e.g., Georgakakis et al. 2018; Comparat et al. 2019).

In this study, we compared our observational results in COSMOS with mock catalogs of active and non-active galaxies created via SEMs based on large N-body simulations. The full description of numerical routines to create mock catalogs of galaxies and their BHs using SEMs is given in Allevato et al. (in prep.). In this work we only describe the important steps in the generation of the AGN mock catalogs and we refer to Allevato et al. for more details. First, we extracted a large catalog of dark matter halos and subhalos from MultiDark<sup>1</sup>-Planck 2 (MDPL2; Riebe et al. 2013) at the redshift of interest, which currently provides the largest publicly available set of high-resolution and large volume N-body simulations (box size of  $1000 h^{-1} \text{ Mpc}$ , mass resolution of  $1.51 \times 10^9 h^{-1} M_\odot$ ). The ROCKSTAR halo finder (Behroozi et al. 2013) has been applied to the MDPL2 simulations to identify halos and flag those (subhalos) that lie within the virial radius of a more massive host halo. The mass of the dark matter halo is defined as the virial mass in the case of host halos and the infall progenitor virial mass for sub-halos. In the analysis that follows, we use the simulation snapshot at  $z = 1$ , which corresponds to the mean redshift of our AGN clustering measurements.

<sup>1</sup> [www.cosmosim.org](http://www.cosmosim.org)

We assign to each halo: (a) a galaxy stellar mass deduced from the [Grylls et al. \(2019\)](#) semiempirical relation, inclusive of intrinsic (0.15 dex) and measurement scatter (0.2 dex for stellar mass estimates in our COSMOS sample). Satellites are assigned a stellar mass at the redshift of infall; (b) a BH mass assuming empirical BH-galaxy mass relation derived in [Shankar et al. \(2016\)](#), inclusive of a stellar mass dependent scatter; (c) an X-ray luminosity following the observationally deduced specific BHAR distribution described by a Schechter function as suggested in [Bongiorno et al. \(2012, 2016\)](#), [Aird et al. \(2012\)](#), and [Georgakakis et al. \(2017\)](#); (d) a SFR following the SFR – stellar mass relation described in [Tomczak et al. \(2016\)](#) for main sequence galaxies, including intrinsic (0.2 dex) and measurement (0.2 dex for our COSMOS sample) scatter; (e) a hydrogen column density  $N_{\text{H}}$  assigned following the [Ueda et al. \(2014\)](#) empirical distribution such that AGN can be classified into Type 2 obscured, Type 1 unobscured, and Compton thick AGN; and (f) a duty cycle, i.e., a probability for each BH of being active, following [Schulze et al. \(2015\)](#).

The main difference between our approach and recent studies based on SEMs (e.g., [Georgakakis et al. 2019](#); [Comparat et al. 2019](#)) is that the latter assign an Eddington ratio to each mock BH, and consider *active* the objects with a specific BHAR above a given value. On the contrary, we assigned to each BH an Eddington ratio combined with a probability of being active, which depends on the BH mass. A comparison between the different SEMs is beyond the scope of this paper and is discussed in [Allevato et al. \(in prep.\)](#).

We then selected a sample of mock AGN, matched to have the same galaxy stellar mass, specific BHAR, SFR, and X-ray luminosity distributions of CCL Type 2 AGN, including only Type 2 objects selected according to  $N_{\text{H}}$  values. This corresponds to mock AGN living in dark matter halos (parent halos for satellite galaxies) with  $\log M_{\text{h}}[M_{\odot}] > 12$ . To each parent halo mass  $M_{\text{halo}}$ , a bias is assigned that satisfies  $b = b(M_{\text{halo}})$ , following the same bias-mass relation used for CCL Type 2 AGN ([van den Bosch 2002](#); [Sheth et al. 2001](#)). Similarly, we selected a sample of normal non-active galaxies with the same galaxy stellar mass, SFR, and X-ray luminosity distributions of CCL Type 2 AGN.

### 5.3.2. Bias for mock AGN

We follow the formalism of [Shankar et al. \(2019\)](#) to derive the bias of mock AGN and normal galaxies as a function of galaxy stellar mass, specific BHAR, and SFR using the relative probabilities of central and satellite galaxies of being active  $Q = U_{\text{s}}/U_{\text{c}}$ .

The bias of mock objects with stellar mass in the range  $M_{\text{star}}$  and  $M_{\text{star}} + dM_{\text{star}}$  is estimated as

$$b(M_{\text{star}}) = \frac{1}{N_{\text{bin}}} \sum_{i=1}^{N_{\text{bin}}} b[M_{\text{h},i}(M_{\text{star}})], \quad (9)$$

where for each bin of stellar mass the sum runs over all host parent halos. If the probabilities for galaxies to be active, i.e., the AGN duty cycle  $U(M_{\text{star}}) = [U_{\text{c}}(M_{\text{star}})N_{\text{c}}(M_{\text{star}}) + U_{\text{s}}(M_{\text{star}})N_{\text{s}}(M_{\text{star}})]/N(M_{\text{star}})$  is included, the generalized formula is written as

$$b = \frac{\left[ \sum_{i=1}^{N_{\text{c}}} U_{\text{c},i}(M_{\text{star}})b_{\text{c},i}(M_{\text{star}}) + \sum_{i=1}^{N_{\text{s}}} U_{\text{s},i}(M_{\text{star}})b_{\text{s},i}(M_{\text{star}}) \right]}{\left[ \sum_{i=1}^{N_{\text{cen}}} U_{\text{c},i}(M_{\text{star}}) + \sum_{i=1}^{N_{\text{s}}} U_{\text{s},i}(M_{\text{star}}) \right]}, \quad (10)$$

where  $U_{\text{c}}(M_{\text{star}}) = U(M_{\text{star}})N(M_{\text{star}})/[N_{\text{c}}(M_{\text{star}}) + QN_{\text{s}}(M_{\text{star}})]$  is the duty cycle of central AGN,  $U_{\text{s}}(M_{\text{star}}) = QU_{\text{c}}(M_{\text{star}})$  is the duty cycle of satellite AGN, and  $N(M_{\text{star}}) = N_{\text{c}}(M_{\text{star}}) + N_{\text{s}}(M_{\text{star}})$  is the number of central and satellite galaxies in the stellar mass bin  $M_{\text{star}}$  and  $M_{\text{star}} + dM_{\text{star}}$ .

When  $U_{\text{s}} = U_{\text{c}}$ , i.e., if all central and satellite galaxies are active or share equal probabilities of being active, Eq. (10) reduces to Eq. (9). It is important to note that the bias is thus mainly affected by  $Q$  and then by the fraction of AGN in satellite halos  $f_{\text{sat}}^{\text{AGN}}$ , so that  $Q = f_{\text{sat}}^{\text{AGN}}(1 - f_{\text{sat}}^{\text{BH}})/[(1 - f_{\text{sat}}^{\text{AGN}})f_{\text{sat}}^{\text{BH}}]$ , where  $f_{\text{sat}}^{\text{BH}} = N_{\text{s}}/(N_{\text{s}} + N_{\text{c}})$  is the total fraction of (active and non-active) BHs in satellites with host galaxy stellar mass within  $M_{\text{star}}$  and  $M_{\text{star}} + dM_{\text{star}}$ . In our mock of AGN matched to CCL Type 2 AGN, we have a total fraction of satellite galaxies  $f_{\text{sat}}^{\text{BH}} \sim 0.1$ , which corresponds to  $Q = 1$  for a fraction of satellite AGN  $f_{\text{sat}}^{\text{AGN}} \sim 0.1$  and  $Q = 2$  for  $f_{\text{sat}}^{\text{AGN}} \sim 0.15$ , respectively. Similarly, Eq. (10) can be written in bins of specific BHAR and SFR, for  $Q = 1$  and 2, respectively.

### 5.3.3. Predictions from SEMs

Figure 6 shows the mean bias as a function of the host galaxy stellar mass and specific BHAR for mock AGN (Eq. (10)) and for normal galaxies (Eq. (9)), where  $Q = 1$  and  $=2$ , which corresponds to a fraction of satellite AGN  $f_{\text{sat}}^{\text{AGN}} \sim 0.1$  and  $\sim 0.15$ , respectively. The host galaxy stellar masses in the mock catalog are assigned by following [Grylls et al. \(2019\)](#), i.e., are defined as Sersic + exponential model ([Bernardi et al. 2013](#)) with a mass-to-light ratio from [Bell et al. \(2013\)](#). To correct for the different definition in COSMOS ([Bruzual & Charlot 2003](#); [Chabrier 2003](#)) we need to decrease the stellar masses of mock objects by a factor of  $\sim 0.15$  dex ([Grylls et al. 2020](#)).

The bias- $M_{\text{star}}$  relation for mock AGN and matched mock normal galaxies is slightly steeper compared to clustering measurements of COSMOS AGN, when we assume the same probability of satellite and central galaxies of being active ( $Q = 1$ ). In particular, we found that CCL Type 2 AGN host galaxies with low  $M_{\text{star}}$  reside in slightly more massive halos than mock AGN and normal galaxies of similar stellar mass. It is important to note that when  $Q = 1$ , Eq. (10) reduces to the simple case of Eq. (9), and AGN and normal galaxies have the same bias- $M_{\text{star}}$  relation.

Our results for CCL Type 2 AGN can be reproduced if we assume a larger satellite AGN fraction  $f_{\text{sat}}^{\text{AGN}} \sim 0.15$  and  $Q \sim 2$ . For instance, [Leauthaud et al. \(2015\)](#) found  $f_{\text{sat}}^{\text{AGN}} \sim 0.18$  for COSMOS AGN at  $z \lesssim 1$ . The importance of the relative fraction of satellite AGN has also been underlined in [Viitanen et al. \(2019\)](#). They have suggested that when excluding AGN that are associated with galaxy groups, the bias of low  $M_{\text{star}}$  objects decreases, while not affecting the high stellar mass systems. Similar results are found in [Mountrichas et al. \(2013\)](#) for moderate luminosity X-ray AGN at  $z \sim 1$ . This is because at lower stellar mass, AGN are more likely in satellite galaxies hosted by more biased and massive parent dark matter halos. Similarly, the bias as a function of specific BHAR and SFR of CCL Type 2 AGN is better reproduced by mock AGN with  $Q = 2$ . It is worth noting that the bias- $M_{\text{star}}$ , bias- $\lambda_{\text{Edd}}$  and bias-SFR relations of normal non-active mock galaxies are independent of the relative duty cycle of AGN in satellite and central halos,  $Q$ .

Our results thus suggest that for CCL Type 2 AGN at  $z \sim 1$ , the relative probabilities of AGN in satellites is roughly two times larger than in central halos; the fraction of AGN in satellite halos is consistent with  $f_{\text{sat}}^{\text{AGN}} \sim 0.15$ . [Starikova et al. \(2015\)](#) studied the HOD of AGN detected by the *Chandra* X-Ray

Observatory in the Bootes field over a redshift interval  $z = [0.17-3]$ , showing a satellite fraction of  $\sim 10\%$ . Allevato et al. (2012) performed direct measurement of the HOD for COSMOS AGN based on the mass function of galaxy groups hosting AGN and found that the duty cycle of satellite AGN is comparable or even larger than that of central AGN, i.e.,  $Q \gtrsim 1$ . The central locations of the quasar host galaxies are expected in major merger models because mergers of equally sized galaxies preferentially occur at the centers of DM halos (Hopkins et al. 2008). Our predictions from mock matched AGN suggests a high percentage of satellite AGN, in agreement with studies that have found a small fraction of AGN associated with morphologically disturbed galaxies (Cisterans et al. 2011; Schawinski et al. 2011; Rosario et al. 2011) and that have suggested secular processes and bar instabilities are efficient in producing luminous AGN (e.g., Georgakakis et al. 2009; Allevato et al. 2011).

Figure 7 shows the typical dark matter halo mass as a function of galaxy stellar mass, as found for CCL Type 2 AGN at mean  $z \sim 1$ , for XMM-XXL AGN at  $z \sim 0.8$  (Mountrichas et al. 2019, green triangles) and as predicted for matched mock AGN and matched normal galaxies for  $Q = 2$ . The predictions from mock AGN well reproduce the observations in COSMOS, as well as the results for XMM-XXL AGN at similar redshift.

A slightly steeper trend is observed for matched normal mock galaxies; non-active BH reside in less massive parent halos than mock AGN with low host galaxy stellar mass. This is in contrast with the results of Mountrichas et al. (2019), at least for low  $M_{\text{star}}$  galaxies. They suggest that AGN have the same stellar-to-halo mass ratio of matched normal galaxies at all stellar masses. Previous clustering studies of normal galaxies (e.g., Zheng et al. 2007; Coil et al. 2007) have suggested a steeper halo – stellar mass relation at similar redshift. A flatter  $M_{\text{star}}-M_{\text{h}}$  relation found for mock matched galaxies is mainly due to the large scatter (measurement 0.2 dex and intrinsic 0.15 dex) in the input relation (see Sect. 4.3.1) used to create the mock catalog to reproduce the stellar mass measurement error in COSMOS, and to the selections in terms of  $L_X$ ,  $M_{\text{star}}$ , specific BHAR, and SFR applied to match CCL Type AGN hosts. Figure 7 shows that when selecting a subsample of galaxies matched to have the same properties of CCL Type 2 AGN hosts, the galaxy bias is driven up at low stellar mass with respect to the full galaxy population.

Figure 7 also shows the halo mass–SFR relation for mock AGN and matched normal galaxies. The predictions are almost consistent with our results in COSMOS and with previous studies in XMM-XXL at similar redshifts (Mountrichas et al. 2019), but suggest a slightly flatter  $\text{SFR}-M_{\text{h}}$  relation than observed. A constant SFR as a function of the halo mass is a consequence of the almost flat  $M_{\text{star}}-M_{\text{h}}$  relation obtained for mock objects, combined with the input assumption that each mock AGN and galaxy follow a simple main sequence  $\text{SFR}-M_{\text{star}}$  relation.

We stress that given the limited sample of CCL Type 2 AGN in bins of  $M_{\text{star}}$ , SFR and specific BHAR, we can only estimate typical halo masses as a function of AGN host galaxy properties, from the modeling of the two-halo term. The full halo mass distribution and halo occupation (possibly separating the contribution of AGN in central and satellite galaxies) require higher statistics to constrain the clustering signal at small scale. Currently, one possibility to overcome the low statistics is to combine available samples of AGN in X-ray surveys, such as CCL (Civano et al. 2016), AEGIS and 4 Ms CDFS (Georgakakis et al. 2014), and XMM-XXL (Mendez et al. 2016), with robust host galaxy property estimates. Following this approach the number of AGN with known spectroscopic redshift can be almost

doubled with respect to the sample of Type 2 AGN used in this work. In subsequent years, the synergy of eRosita, 4MOST, WISE, as well as Euclid and JWST in the near future, will allow us to derive host galaxy stellar mass and SFR estimates of millions of moderate-high luminosity AGN up to  $z \sim 2$ .

Moreover, our clustering measurements refer to Type 2 AGN only. New CCL data are now available for Type 1 AGN (Suh et al. 2019) and then AGN clustering dependence on host galaxy properties will be probed in the near future as a function of obscuration.

## 6. Conclusions

In this paper, we have performed clustering measurements of CCL Type 2 AGN at mean  $z \sim 1$ , to probe the AGN large-scale bias dependence on host galaxy properties, such as galaxy stellar mass, specific BHAR, and SFR. We found no dependence of the AGN large-scale bias on galaxy stellar mass and specific BHAR, suggesting almost flat  $M_{\text{star}}-M_{\text{h}}$  and  $\lambda_{\text{BHAR}}-M_{\text{h}}$  relations. A negative clustering dependence is instead observed as a function of SFR, with the typical hosting halo mass increasing with decreasing SFR. Mock catalogs of AGN matched to have the same host galaxy properties of CCL Type 2 AGN predict the observed  $M_{\text{star}}-M_{\text{h}}$ ,  $\text{SFR}-M_{\text{h}}$  and  $\lambda_{\text{BHAR}}-M_{\text{h}}$  relations, when assuming a fraction of satellite AGN  $f_{\text{sat}}^{\text{AGN}} \sim 15\%$  and then  $Q = 2$ . Mock matched normal galaxies follow a slightly steeper  $M_{\text{star}}-M_{\text{h}}$  relation, where low mass mock galaxies reside in slightly less massive halos than mock AGN of similar mass. Similarly, mock galaxies reside in less massive hosting halos than mock AGN with similar specific BHAR and SFR, at least for  $Q > 1$ .

*Acknowledgements.* VA acknowledges funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 749348. TM is supported by CONACyT 252531 and UNAM-DGAPA PAPIIT IN111319.

## References

- Aird, J., Coil, A. L., Moustakas, J., et al. 2012, *ApJ*, 746, 90
- Aird, J., Coil, A. L., Georgakakis, A., et al. 2018, *MNRAS*, 474, 1225
- Allevato, V., Finoguenov, A., Cappelluti, N., et al. 2011, *ApJ*, 736, 99
- Allevato, V., Finoguenov, A., Hasinger, G., et al. 2012, *ApJ*, 758, 47
- Allevato, V., Finoguenov, A., Civano, F., et al. 2014, *ApJ*, 796, 4
- Allevato, V., Civano, F., Finoguenov, A., et al. 2016, *ApJ*, 832, 70
- Behroozi, P., & Silk, J. 2018, *MNRAS*, 477, 5382
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, 770, 57
- Bell, E. F., McIntosh, D. H., Katz, N., et al. 2013, *ApJS*, 149, 289
- Bernardi, M., Meert, A., Sheth, R. K., et al. 2013, *MNRAS*, 436, 697
- Bielby, R., Hill, M. D., Shanks, T., et al. 2013, *MNRAS*, 430, 425
- Birnboim, Y., & Dekel, A. 2003, *MNRAS*, 345, 349
- Bongiorno, A., Merloni, A., Brusa, M., et al. 2012, *MNRAS*, 427, 3103
- Bongiorno, A., Schultze, A., Merloni, A., et al. 2016, *A&A*, 588, A78
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, *A&A*, 497, 635
- Cappelluti, N., Aiello, M., Burlon, D., et al. 2010, *ApJ*, 716, 209
- Civano, F., Marchesi, S., Comastri, A., et al. 2016, *ApJ*, 819, 62
- Cisterans, M., Jahnke, K., Inskip, K. J., et al. 2011, *ApJ*, 726, 57
- Chabrier, G. 2003, *ApJ*, 586, L133
- Coil, A. L., Hennawi, J. F., Newman, J. A., et al. 2007, *ApJ*, 654, 115
- Coil, A. L., Georgakakis, A., Newman, J. A., et al. 2009, *ApJ*, 701, 1484
- Coil, A. L., Mendez, A. J., Eisenstein, D. J., & Moustakas, J. 2017, *ApJ*, 838, 87
- Comparat, J., Merloni, A., Salvato, M., et al. 2019, *MNRAS*, 487, 2005
- Conroy, C., & White, M. 2013, *ApJ*, 762, 70
- Coupon, J., Arnouts, S., van Waerbeke, L., et al. 2015, *MNRAS*, 449, 1352
- Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, *MNRAS*, 356, 415
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, 267, 465
- Eisenstein, D. J., & Hu, W. 1999, *ApJ*, 511, 5
- Elvis, M., Civano, F., Vignani, C., et al. 2009, *ApJS*, 184, 158
- Erfanianfar, G., Popesso, P., Finoguenov, A., et al. 2016, *MNRAS*, 455, 2839

- Gabor, J. M., & Dave, R. 2015, *MNRAS*, **447**, 374
- Georgakakis, A., Coil, A. L., Laird, E. S., et al. 2009, *MNRAS*, **397**, 623
- Georgakakis, A., Mountrichas, G., Salvato, M., et al. 2014, *MNRAS*, **443**, 3327
- Georgakakis, A., Aird, J., Schulze, A., et al. 2017, *MNRAS*, **471**, 1976
- Georgakakis, A., Comparat, J., Merloni, A., et al. 2018, *MNRAS*, **487**, 275
- Georgakakis, A., Comparat, J., Merloni, A. et al. 2019, *MNRAS*, **487**, 275G
- George, M. R., Ma, C. P., Bundy, K., et al. 2013, *ApJ*, **770**, 113
- Grylls, P. J., Shankar, F., Zanisi, L., & Bernardi, M. 2019, *MNRAS*, **483**, 2506
- Grylls, P. J., Shankar, F., Leja, J., et al. 2020, *MNRAS*, **491**, 634
- Haring, N., & Rix, H. 2004, *ApJ*, **604**, L89
- Hopkins, P. F., Hernquist, L. C., Thomas, J., et al. 2006, *ApJS*, **163**, 1
- Hopkins, P. F., Hernquist, L. C., Thomas, J., et al. 2008, *ApJ*, **679**, 156
- Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, *ApJ*, **696**, 891
- Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, *MNRAS*, **428**, 1382
- Krumpe, M., Miyaji, T., & Coil, A. L. 2010, *ApJ*, **713**, 558
- Krumpe, M., Miyaji, T., Coil, A. L., & Aceves, H. 2012, *ApJ*, **746**, 1
- Krumpe, M., Miyaji, T., Husemann, B., et al. 2015, *ApJ*, **815**, 21
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, **412**, 64
- Lapi, A., & Cavaliere, A. 2011, *ApJ*, **743**, 127L
- Leauthaud, A., Finoguenov, A., Kneib, J. P., et al. 2010, *ApJ*, **709**, 97
- Leauthaud, A. J., Benson, A., & Civano, F. 2015, *MNRAS*, **446**, 1874
- Legrand, L., McCracken, H. J., Davidzon, I., et al. 2018, *MNRAS*, **486**, 5468
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, **115**, 2285
- Marchesi, S., Civano, F., Elvis, M., et al. 2016, *ApJ*, **817**, 34
- Marulli, F., Veropalumbo, A., & Moresco, M. 2016, *Astron. Comput.*, **14**, 35
- Menci, N., Fiore, F., Puccetti, S. et al. 2008, *ApJ*, **686**, 219
- Mendez, A. J., Coil, A. L., Aird, J., et al. 2016, *ApJ*, **821**, 55
- Meneux, B., Guzzo, L., de la Torre, S., et al. 2009, *A&A*, **505**, 463
- Miyaji, T., Zamorani, G., Cappelluti, N., et al. 2007, *ApJS*, **172**, 396
- Mostek, N., Coil, A. L., Cooper, M., et al. 2013, *ApJ*, **767**, 89
- Moster, B. P., Naab, T., & White, S. D. M. 2013, *MNRAS*, **428**, 3121
- Moster, B. P., Naab, T., & White, S. D. M. 2018, *MNRAS*, **477**, 1822
- Mountrichas, G., Georgakakis, A., Finoguenov, A., et al. 2013, *MNRAS*, **430**, 661
- Mountrichas, G., Georgakakis, A., & Georgantopoulos, I. 2019, *MNRAS*, **483**, 1374
- Peebles, P. J. E. 1980, *The Large Scale Structure of the Universe* (Princeton: Princeton Univ. Press)
- Peng, Y., Lilly, S. J., Kovač, K., et al. 2010a, *ApJ*, **721**, 193
- Peng, Y., Lilly, S. J., Renzini, A., et al. 2010b, *ApJ*, **757**, 4
- Peng, Y., Lilly, S. J., Renzini, A., et al. 2012, *ApJ*, **757**, 4
- Peng, Y., Maiolino, R., & Cochrane, R. 2015, *Nature*, **521**, 192
- Plionis, M., et al. 2018, *A&A*, **371**, 1824
- Porciani, C., & Norberg, P. 2006, *MNRAS*, **371**, 1824
- Powell, M. C., Cappelluti, N., Urry, C. M., et al. 2018, *ApJ*, **858**, 110
- Riebe, K., Partl, A. M., Enke, H., et al. 2013, *AN*, **334**, 691
- Rosario, D. J., McGurk, R. C., Max, C. E., et al. 2011, *ApJ*, **739**, 44
- Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, *ApJ*, **697**, 1634
- Schawinski, K., Treister, E., Urry, C. M., et al. 2011, *ApJ*, **727**, 31
- Schulze, A., Bongiorno, A., Gavignaud, I., et al. 2015, *MNRAS*, **447**, 2085
- Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, *MNRAS*, **460**, 3119
- Shankar, F., Allevato, V., Bernardi, M., et al. 2019, Arxiv e-prints [arXiv:1910.10175]
- Shen, Y., McBride, C. K., White, M., et al. 2013, *ApJ*, **778**, 98
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, **323**, 1
- Springel, V., Di Matteo, T., Hernquist, L., et al. 2005, *ApJ*, **620**, L79
- Starikova, S., Cool, R., Eisenstein, D., et al. 2015, *ApJ*, **741**, 15
- Suh, H., Civano, F., Hasinger, G., et al. 2017, *ApJ*, **841**, 102
- Suh, H., Civano, F., Hasinger, G., et al. 2019, *ApJ*, **872**, 168
- Tomczak, A. R., Quadri, R. F., Tran, K. V. H., et al. 2016, *ApJ*, **817**, 118
- Wake, D. A., Croom, S. M., Sadler, E. M., & Johnston, H. M. 2008, *MNRAS*, **391**, 1674
- Ueda, Y., Akiyama, M., Hasinger, G., et al. 2014, *ApJ*, **786**, 104
- van den Bosch, F. C. 2002, *MNRAS*, **331**, 98
- Viitanen, A., Allevato, V., Finoguenov, A., et al. 2019, *A&A*, **629**, A14
- Zehavi, I., Eisenstein, D. J., Nichol, R. C., et al. 2005, *ApJ*, **621**, 22
- Zheng, Z., Coil, A. L., & Zehavi, I. 2007, *ApJ*, **667**, 760