A&A, 708, A323 (2026)
https://doi.org/10.1051/0004-6361/202659274
© The Authors 2026

Astronomy & Astrophysics

Large-scale environments of star-forming active galactic nuclei: How black-hole mass, accretion rate, and luminosity connect to dark-matter halos

G. Mountrichas[1,⋆], F. J. Carrera[1], F. Shankar[2], and A. Georgakakis[3]

[1] Instituto de Fisica de Cantabria (CSIC-Universidad de Cantabria), Avenida de los Castros, 39005 Santander, Spain
[2] Department of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[3] Institute for Astronomy & Astrophysics, National Observatory of Athens, V. Paulou & I. Metaxa, Athens 11532, Greece

Received 2 February 2026 / Accepted 10 March 2026

ABSTRACT

Understanding the relative roles of large-scale environment and internal host-galaxy processes in shaping active galactic nuclei (AGN) activity is key to constraining models of black-hole growth and galaxy evolution. In this work, we investigated how the environment of X-ray-selected active galactic nuclei (AGNs) relates to black-hole growth and accretion properties, and whether these introduce a dependence on a large-scale environment beyond that set by the host galaxy itself. By combining the XXL and Stripe 82X surveys, we assembled samples of 427 broad-line AGNs at $0.5 < z < 1.2$ and over 20 000 galaxies, with host-galaxy properties consistently derived using the same spectral energy distribution fitting methodology and assumptions. Dark-matter halo (DMH) masses were inferred from AGN–galaxy cross-correlation functions, while a multivariate nearest-neighbour matching algorithm was applied to isolate trends with black-hole mass ($M_{\rm BH}$), Eddington ratio ($\lambda_{\rm Edd}$), and X-ray luminosity ($L_{\rm X}$) under controlled host-galaxy conditions. Within the statistical uncertainties of the present dataset, we find that X-ray AGNs typically reside in halos of $\log(M_{\rm DMH}/h^{-1}M_\odot) \simeq 13$, with no significant variation as a function of $M_{\rm BH}$, $\lambda_{\rm Edd}$, or $L_{\rm X}$. Neither $M_{\rm BH}$ nor $\lambda_{\rm Edd}$ shows a measurable correlation with large-scale environment, which is consistent with a scenario in which long-term black-hole growth and short-term accretion variability are primarily regulated by internal host-galaxy processes rather than halo-scale mass alone. The absence of a statistically significant $M_{\rm DMH}$–$L_{\rm X}$ trend further indicates that AGN radiative output reflects stochastic or feedback-regulated variability, rather than direct modulation by the large-scale environment. Overall, these results support a self-regulated co-evolution framework in which large-scale structure sets the boundary conditions for gas availability and AGN duty cycle, while the subsequent growth and luminosity evolution of AGNs are determined predominantly by local processes within their host galaxies.

Key words. galaxies: active – galaxies: evolution – cosmology: observations – large-scale structure of Universe

1. Introduction

Active galactic nuclei (AGNs) represent one of the most energetic phenomena in the Universe, powered by accretion of matter onto supermassive black holes (SMBHs) residing at the centres of galaxies. Their radiative output spans many orders of magnitude, influencing both the host galaxy and the surrounding intergalactic medium through feedback processes that can regulate star formation and baryon cycling on multiple scales. Understanding the environmental context of AGN is therefore crucial for constraining how black-hole growth connects to galaxy evolution across cosmic time.

The large-scale environment, traced through the dark-matter halo (DMH) mass, provides a statistical measure of how AGNs are distributed within the cosmic web and how AGN host galaxies populate large-scale structure. The clustering strength of AGNs reflects the typical DMH mass associated with their host galaxies, and its dependence on host and SMBH properties offers insight into whether observed environmental trends arise from a large-scale structure or from internal, secular processes within galaxies. Theoretical and semi-empirical models have also emphasised the role of halo assembly in shaping the long-term growth of SMBHs and the global regulation of feedback (e.g. Shankar et al. 2010, 2013, 2017, 2020; Hopkins et al. 2006, 2008; Croton et al. 2006; Bower et al. 2006; Fanidakis et al. 2012; Hirschmann et al. 2014; Weinberger et al. 2018). In these frameworks, the mass and assembly history of the DMH primarily influence the availability of gas, the probability of triggering AGN episodes, and the duty cycle of accretion, while the instantaneous accretion rate and radiative output are largely governed by baryonic processes internal to the host galaxy, such as gas inflows, disc instabilities, and feedback self-regulation – even if instantaneous accretion rates are determined by local conditions. Quantifying these dependencies is therefore a key step towards disentangling environmental and host-galaxy effects in the co-evolution of galaxies and SMBHs.

It is important to stress that large-scale (>1 Mpc) clustering measurements do not directly trace AGN triggering mechanisms. Instead, they primarily reflect the properties of the host-galaxy population, most notably stellar mass ($M_\star$), and the way AGNs occupy the galaxy population (e.g. Leauthaud et al. 2015; Georgakakis et al. 2019; Allevato et al. 2021; Shankar et al. 2020). In this context, AGNs are known to occupy a broad distribution of DMH masses that peaks at a few $\times 10^{12}\,{\rm M}_\odot$, while the 'typical' halo masses inferred from correlation function and bias measurements are biased towards higher values because

---

[⋆] Corresponding author: `gmountrichas@gmail.com`

Open Access article, published by EDP Sciences, under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
This article is published in open access under the Subscribe to Open model. Subscribe to A&A to support open access publication.

they are sensitive to the high-mass tail of this distribution (e.g. Mountrichas et al. 2013). As a result, differences in the inferred halo masses among AGN samples can naturally arise from selection effects – namely how different observational criteria preferentially select hosts with different $M_\star$ – rather than reflecting distinct AGN triggering mechanisms.

Previous studies have explored how AGN clustering depends on a variety of host and black-hole properties, though results often differ due to sample selection, redshift coverage, and the lack of consistent control of covariates. Regarding host-galaxy properties, a mild dependence of clustering amplitude on $M_\star$ has been reported (Mountrichas et al. 2019), while other analyses found no such trend (Allevato et al. 2019; Viitanen et al. 2019). These latter studies, which probed broad redshift intervals up to $z \approx 3$ and $z \approx 2.5$, suggest that the environment–$M_\star$ connection may weaken at higher redshift or depend on AGN type. In contrast, a dependence on star-formation rate (SFR) has been observed (Mountrichas et al. 2019; Allevato et al. 2019), with the strongest effect seen for the specific star-formation rate (sSFR), where AGNs in less star-forming galaxies occupy more massive halos (Mountrichas et al. 2019).

The situation for X-ray luminosity (L_X) is more complex. A positive correlation between clustering amplitude and L_X has been identified at low redshifts ($z < 0.5$; Krumpe et al. 2010a, 2012), whereas no dependence is observed in the local Universe ($0.007 < z < 0.037$), although differences appear in how low- and high-luminosity AGNs populate halos on small scales (Krumpe et al. 2018). At intermediate redshifts, a mild dependence has been suggested (Powell et al. 2020; Comparat et al. 2023), but other analyses found an inverse relation, whereby more-luminous AGNs reside in less-massive halos (Mountrichas et al. 2016). Such discrepancies are often attributed to differences in luminosity intervals, redshift coverage, and selection effects (Fanidakis et al. 2013b,a; Powell et al. 2020).

Dependence on black-hole mass ($M_{\rm BH}$) appears weak, with only mild trends found up to $z < 0.6$ (Krumpe et al. 2015, 2023). Studies of optical quasars showed that when the sample is divided by $M_{\rm BH}$, clustering differences emerge only for the most massive subset (Shen et al. 2009). In contrast, the Eddington ratio ($\lambda_{\rm Edd}$), which traces instantaneous accretion efficiency, exhibits no clear correlation with environment up to comparable redshifts (Krumpe et al. 2015, 2023). Using its observational proxy, the specific black-hole accretion rate ($L_X/M_\star$), no significant dependence is detected up to $z \sim 2.5$ (Viitanen et al. 2019) or $z \sim 3$ for type 2 AGNs (Allevato et al. 2019).

Obscuration introduces additional ambiguity: some studies report no clustering difference between obscured and unobscured AGNs (Viitanen et al. 2023), while others find unobscured systems to be more clustered (Allevato et al. 2014) or obscured AGN to reside in more massive halos (Powell et al. 2018); in the local Universe, type 1 and type 2 AGNs may even occupy halos differently (Krumpe et al. 2018). A key limitation of many previous studies is that, when investigating the dependence of clustering on one parameter, other potentially relevant host or SMBH properties are not simultaneously controlled. Some analyses account for one or two covariates (e.g. $M_\star$ or $\lambda_{\rm Edd}$), but few match across multiple physical dimensions, which can introduce biases and mask genuine correlations. Moreover, most existing works are confined to low redshifts ($z < 0.6$) or span wide redshift ranges that complicate interpretation due to possible evolutionary mixing.

Within this framework, our goal is not to identify AGN triggering mechanisms, but to quantify how different AGN selections probe the underlying halo-mass distribution once both SMBH and host-galaxy properties are carefully controlled. To this end, we used X-ray-selected AGNs from the XMM–XXL and Stripe 82X surveys to examine the relation between the typical DMH mass and $M_{\rm BH}$, $\lambda_{\rm Edd}$, and L_X.

A multivariate nearest-neighbour matching algorithm was employed to explicitly control for host-galaxy $M_\star$, as well as SFR, sSFR, redshift, and the remaining SMBH parameters not under investigation in each case. This ensures that, when comparing AGN subsets split by a given parameter, their $M_\star$ and other covariate distributions are statistically consistent. As a result, any differences in the inferred halo masses cannot be attributed to systematic variations in host-galaxy properties.

The paper is structured as follows. Section 2 describes the AGN and galaxy samples. Section 3 presents the derivation of physical parameters from spectral energy distribution (SED) fitting, the clustering analysis, and matching methodology. Results of the dependence of halo mass on $M_{\rm BH}$, $\lambda_{\rm Edd}$, and L_X are presented in Section 4. Section 5 discusses the physical interpretation of these results in the context of AGN and galaxy co-evolution, and Section 6 summarises the main conclusions. Throughout this work, we adopted a flat ΛCDM cosmology with $\Omega_M = 0.315$, $\Omega_\Lambda = 0.685$, $h = 0.674$ (i.e. $H_0 = 67.4\,{\rm km\,s^{-1}\,Mpc^{-1}}$), and $\sigma_8(z=0) = 0.811$, which are consistent with the Planck Collaboration VI (2020) cosmological parameters.

2. Data

In this work, we used data from two well-studied extragalactic survey fields, namely XMM–XXL and Stripe 82X. Below, we describe the main characteristics of the X-ray AGNs and galaxies observed in these fields.

2.1. X-ray AGNs

Our analysis is based on X-ray-selected AGNs detected in the XMM–XXL north (hereafter XXL) and Stripe 82X fields. The XXL AGN sample has been extensively used and described in several of our previous works (e.g. Masoura et al. 2018, 2021; Pouliasis et al. 2020; Mountrichas & Shankar 2023; Mountrichas 2023; Mountrichas & Buat 2023; Mountrichas et al. 2024b).

In brief, XMM–XXL (Pierre et al. 2016) is a medium-depth X-ray survey covering a total area of $\sim$50 deg^2, divided into two nearly equal regions: the XMM–XXL north and south fields. The XXL-N dataset includes 8445 X-ray sources, 5294 of which have optical counterparts in the SDSS and 2512 have reliable spectroscopic redshifts (Menzel et al. 2016; Liu et al. 2016). Mid-infrared (MIR) and near-infrared (NIR) counterparts were identified using the likelihood ratio method (Sutherland & Saunders 1992), as implemented in Georgakakis et al. (2011). Details of the XMM-Newton data reduction and infrared counterpart identification are provided in Georgakakis et al. (2017).

Among the 2512 XXL-N X-ray sources with reliable spectroscopy, 1786 were classified as broad-line AGNs (BLAGN1) by Menzel et al. (2016). A source was identified as BLAGN1 when at least one broad emission line had a full width at half maximum (FWHM) greater than 1000 km s^{-1}. Liu et al. (2016) performed spectral fitting of the BOSS spectroscopy for these 1786 BLAGN1 to estimate single-epoch virial black-hole masses ($M_{\rm BH}$) from the continuum luminosities and broad-line widths, following the approach of Shen et al. (2013).

In short, they measured the continuum luminosities and the FWHMs of the broad emission lines, and then applied the appropriate single-epoch virial mass estimators depending on redshift: Hβ for $z < 0.9$, Mg II for $0.9 < z < 2.2$, and C IV for $z > 2.2$. Previous studies have demonstrated that $M_{\rm BH}$ derived from different broad lines using these fiducial relations are generally consistent with each other, with negligible systematic offsets and small scatter (e.g. Shen et al. 2008, 2011, 2013; Shen & Liu 2012). Liu et al. (2016) confirmed this consistency for the XXL sample. The average uncertainty of their $M_{\rm BH}$ measurements is ~0.5 dex, while sources with higher signal-to-noise ratio spectra have errors below 0.15 dex.

The Stripe 82X survey (LaMassa et al. 2013a,b, 2015) was designed to exploit the extensive multiwavelength coverage of Stripe 82, with the main goal of identifying high-luminosity AGNs at high redshift. The survey includes both *Chandra* and XMM–Newton observations. In this work, we used the Data Release 3 (DR3) of Stripe 82X, presented by LaMassa et al. (2024). This latest release combines the X-ray detections from the initial Stripe 82X catalog (DR1; LaMassa et al. 2016) with the multi-wavelength identifications and photometric redshifts from Ananna et al. (DR2; 2017), as well as spectroscopic redshifts from independent surveys, the dedicated SDSS-IV eBOSS program (LaMassa et al. 2019), and additional ground-based observations obtained with the Palomar and Keck telescopes.

The Stripe 82X DR3 catalogue covers a total area of 31.3 deg^2 and contains 6181 X-ray sources. Spectroscopic redshifts are available for 3457 of them, corresponding to an overall spectroscopic completeness of 56%. Within the contiguous regions of homogeneous X-ray coverage and for sources brighter than $r = 22$ mag, the spectroscopic completeness increases to 90%. The catalogue adds 343 new spectroscopic redshifts compared to DR2 and includes $M_{\rm BH}$ estimates derived from SDSS spectra of 1297 type 1 AGNs, which we used in our analysis and describe below.

Compared to earlier releases, DR3 introduces additional columns that facilitate AGN identification based on X-ray luminosity. Sources are classified as X-ray AGNs when their rest-frame 2–10 keV luminosity exceeds 10^{42} erg s^{-1} (Brandt & Hasinger 2005). The rest-frame luminosities were k-corrected using a power-law spectral model with slopes $\Gamma = 2.0$ (soft band) and $\Gamma = 1.7$ (hard and full bands), following the procedure outlined by LaMassa et al. (2013b, 2016). Detection thresholds correspond to significance levels of 4.5σ for *Chandra* and 5σ for *XMM-Newton* sources. When a source is not detected in the hard band, its 2–10 keV luminosity is estimated by scaling the full- or soft-band luminosities, as detailed in LaMassa et al. (2019). Finally, we note that in our analysis, for both X-ray AGN datasets, we used the observed –i.e. not corrected for absorption– value of 2–10 keV $L_{\rm X}$ provided in the two catalogues.

2.2. Galaxy samples

To investigate the large-scale environments of X-ray AGNs, we estimated the mass of their host DMHs through the two-point cross-correlation function. For this purpose, we cross-correlated the AGN samples with galaxy catalogues that spatially overlap with the X-ray survey regions.

For the XXL AGNs, we followed the methodology described in Mountrichas et al. (2016) and Mountrichas et al. (2019) and use the spectroscopic galaxy catalogue from the VIMOS Public Extragalactic Redshift Survey (VIPERS; Garilli et al. 2014; Guzzo et al. 2014). This dataset has also been employed in several of our previous (non clustering) studies (e.g. Pouliasis et al. 2020; Mountrichas & Buat 2023; Mountrichas et al. 2024b).

In summary, the galaxy sample used in our analysis is drawn from the second public data release (PDR-2; Scodeggio et al. 2018) of VIPERS. The survey was carried out with the Visible Multi-Object Spectrograph (VIMOS; Le Fèvre et al. 2003) on the ESO Very Large Telescope (VLT) and covers an effective area of ≈23.5 deg^2, split between two regions (W1 and W4) within the Canada–France–Hawaii Telescope Legacy Survey–Wide (CFHTLS-Wide) fields. Spectroscopic targets were selected down to a magnitude limit of $i_{\rm AB} < 22.5$ using the T0006 release of the CFHTLS photometric catalogues. An optical colour–colour pre-selection, defined as $(r - i) > 0.5(u - g)$ or $(r - i) > 0.7$, was applied to remove low-redshift ($z < 0.5$) galaxies, achieving over 98% completeness for $z > 0.5$ (see Guzzo et al. 2014).

The PDR-2 catalogue contains 86 775 galaxies with spectroscopic redshifts. Each spectrum is assigned a quality flag that quantifies the reliability of the redshift measurement. Following the standard VIPERS convention, only galaxies with flags between 2 and 9 are considered to have reliable redshifts and are included in our analysis (Garilli et al. 2014; Scodeggio et al. 2018). Restricting the sample to the redshift range $0.5 < z < 1.2$ yielded a final set of 45 180 galaxies.

For the Stripe 82X field, we made use of galaxies located within the Stripe 82 region on the celestial equator, which spans approximately 300 deg^2, with right ascension ranging from 20$^{\rm h}$ to 04$^{\rm h}$ and a declination of ±1.25°. Stripe 82 is part of the Sloan Digital Sky Survey (SDSS; Jiang et al. 2014) and benefits from deep, repeated imaging that has been co-added to produce significantly deeper data than single-epoch SDSS observations. The co-added images reach 2–3 mag deeper, depending on the photometric band, allowing the detection of faint galaxies and studies of the low-surface-brightness Universe (Annis et al. 2014).

In addition to the SDSS optical data, the Stripe 82 region has extensive multi-wavelength coverage, including NIR imaging from the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007), MIR observations from the *Spitzer Space Telescope*, and far-infrared (FIR) coverage from *Herschel* (Viero et al. 2014). Galaxy redshifts are primarily spectroscopic, supplemented where necessary by high-quality photometric redshifts derived from the extensive multi-wavelength coverage available in this field. For the purpose of our clustering analysis, we restricted the galaxy sample to the regions of Stripe 82 that overlap with the X-ray footprint of the Stripe 82X survey (see Section 3.2) and to the same redshift range used for the AGN–galaxy cross-correlation measurements, ensuring consistency with the XXL analysis. The resulting redshift distribution closely matches that of the XXL/VIPERS galaxy sample, with similar median redshifts (0.70 for Stripe 82 and 0.65 for XXL/VIPERS), ensuring that systematic differences in redshift sampling do not bias the inferred halo masses.

3. Analysis

In this section, we present the SED fitting procedure and the clustering analysis used to derive the host-galaxy properties and the DMH masses of our samples. We also describe the selection criteria applied to define the final X-ray AGN and galaxy datasets.

3.1. SED fitting

To derive the host galaxy properties and estimate the $\lambda_{\rm Edd}$ of the AGNs, we constructed and fitted their SEDs. To ensure reliable results, we followed the methodology established in our previous studies (e.g. Mountrichas et al. 2021a,b, 2022c,a,b, 2023, 2024a; Mountrichas & Georgantopoulos 2024), requiring all AGNs and galaxy sources to have detections in the optical, near-infrared (NIR), and mid-infrared (MIR) photometric bands.

The SED construction and fitting for most datasets have been presented in our previous works. Specifically, the SEDs of the XXL AGNs were built, fitted, and analysed in Mountrichas & Shankar (2023), Mountrichas (2023), Mountrichas & Buat (2023) and Mountrichas et al. (2024b) while those of the VIPERS galaxies were described in Mountrichas & Buat (2023) and Mountrichas et al. (2024b). The SEDs of galaxies in the Stripe82 region were modelled in Mountrichas et al. (2025). Further details on the photometric coverage, parameter grids, and fitting configuration can be found in these references. In the present work, we only performed SED fitting for the X-ray AGNs detected in Stripe 82X.

For the SED fitting, we adopted the same set of templates and parameter grid as in our previous analyses. In summary, the galaxy component is described by a delayed star-formation history (SFH) of the form $\mathrm{SFR}(t) \propto t \times \exp(-t/\tau)$. A secondary burst component was included as a constant star-formation episode lasting 50 Myr, following the prescription of Małek et al. (2018) and Buat et al. (2019). Stellar emission was modelled using the single stellar population (SSP) templates of Bruzual & Charlot (2003) and attenuated according to the Charlot & Fall (2000) law. Nebular emission was modelled using templates based on Villa-Velez et al. (2021), while the dust emission heated by stars is represented by the Dale et al. (2014) models, excluding any AGN contribution. The AGN emission is included using the `SKIRTOR` models (Stalevski et al. 2012, 2016), which account for both the torus and the central engine components. The complete list of modules and the adopted parameter space used for the SED fitting can be found, for example, in Table 1 of Mountrichas et al. (2022c,b).

3.2. Selection criteria and final datasets

As discussed in the previous section, we require all datasets to have photometric coverage in the optical, NIR, and MIR bands. This criterion reduces the number of VIPERS galaxies to 14 128 sources (see Sect. 2.2 in Mountrichas & Buat 2023), while 320 319 galaxies in Stripe82 fulfil the same requirement (see Sect. 2.2 in Mountrichas et al. 2025). Among the X-ray samples, 687 XXL AGN and 731 Stripe 82X AGN meet this photometric criterion and have available $M_{\rm BH}$ measurements.

Following our previous studies (e.g. Mountrichas et al. 2021b; Buat et al. 2021; Koutoulidis et al. 2022; Pouliasis et al. 2020), we excluded sources with poor SED fits or unreliable host-galaxy estimates. Specifically, we applied a reduced chi-square threshold of $\chi^2_{\rm r} \leq 5$ (e.g. Masoura et al. 2018; Buat et al. 2021) and removed systems for which `CIGALE` could not constrain key physical parameters such as the SFR and $M_\star$. To identify these cases, we adopted the same reliability criteria used in our recent analyses (e.g. Mountrichas et al. 2021a, 2022a,b; Buat et al. 2021; Koutoulidis et al. 2022; Pouliasis et al. 2022). For each parameter estimated by `CIGALE`, two values are provided: one corresponding to the best-fitting model and another representing the Bayesian-likelihood-weighted mean. Large discrepancies between these values indicate a complex likelihood distribution and significant uncertainties. We therefore only retained sources satisfying the conditions $1/5 \leq \mathrm{SFR_{best}}/\mathrm{SFR_{bayes}} \leq 5$ and $1/5 \leq M_{*,\rm best}/M_{*,\rm bayes} \leq 5$, where the subscripts 'best' and 'bayes' denote the best-fit and Bayesian estimates, respectively. After applying these criteria, the remaining sample consists of 181 361 galaxies in Stripe82 and 10 150 in VIPERS. For the X-ray samples, 584 AGNs in XXL and 550 in Stripe 82X satisfy these quality constraints.

Furthermore, to ensure that the galaxy catalogues represent non-active systems, we excluded potential AGN candidates, i.e. galaxies that are not detected in X-rays but exhibit a strong AGN component in their SEDs. We removed all galaxies with an AGN fractional contribution of $\mathrm{frac_{AGN}} > 0.2$ (e.g. Mountrichas et al. 2021b, 2022c,a,b, 2024b), where $\mathrm{frac_{AGN}}$ denotes the fraction of the total infrared emission attributed to AGN heating. This criterion excludes approximately 45% of VIPERS galaxies and 35% of Stripe 82 galaxies, consistent with the fractions reported in our previous analyses of the Boötes, COSMOS, and eFEDS fields, at similar redshifts (Mountrichas et al. 2021b, 2022a,c).

In addition, we restricted all samples to the $0.5 \leq z \leq 1.2$ redshift range, consistently with the redshift coverage of VIPERS (see Sect. 2.2). The final datasets comprise ~4000 galaxies in VIPERS and ~40 000 in Stripe 82, along with 240 X-ray AGN in XXL and 290 in Stripe 82X within the same redshift interval. The much larger galaxy sample in Stripe 82 arises from the substantially deeper optical imaging, which reaches ~2–2.5 mag fainter than the CFHTLS–W1 photometry used in the XXL area. This allows Stripe 82 to detect a far larger number of low-mass and moderately star-forming galaxies at $z \sim 1$, whereas X-ray AGN counts are governed primarily by the depth of the X-ray surveys. Since the XMM coverage in XXL and Stripe 82X is similar in sensitivity and sky area, the number of detected AGNs is comparable across the two fields despite the large difference in galaxy counts.

To prepare for the clustering analysis, we ensured that AGNs and galaxies occupy the same effective survey regions within each field by matching their spatial footprints using a custom `Python` routine. For each AGN, we identified galaxies located within a projected comoving separation of $r_p \leq 20\,h^{-1}$ Mpc and a line-of-sight separation of $\pi \leq 60\,h^{-1}$ Mpc. This three-dimensional window was chosen to be sufficiently large to guarantee that, for all transverse scales used in the $w_p(r_p)$ measurements ($r_p \leq 30\,h^{-1}$ Mpc), galaxy neighbours can be measured around each AGN without being affected by survey edges or masked regions.

Active galactic nuclei that have no galaxies within this broad geometric window were removed from the analysis. Such cases indicate that the AGN lies outside the effective galaxy footprint (e.g. in masked or sparsely sampled areas) rather than reflecting a physically underdense environment. Importantly, no minimum number of galaxies is required: the criterion is purely geometric and does not depend on the local galaxy density in a physical sense. The statistical robustness of the clustering signal is handled at the level of pair counts, fitting ranges, and jackknife error estimation.

After applying this footprint-matching requirement, the samples used in the clustering analysis consist of 203 and 245 X-ray AGNs, and 2 428 and 19 141 galaxies in XXL and Stripe 82X, respectively. These matched samples probe identical sky regions and redshift intervals, ensuring a consistent and unbiased basis for the AGN–galaxy cross-correlation measurements.

Finally, we restricted our analysis to AGNs hosted by star-forming galaxies. Quiescent galaxies are known to reside in

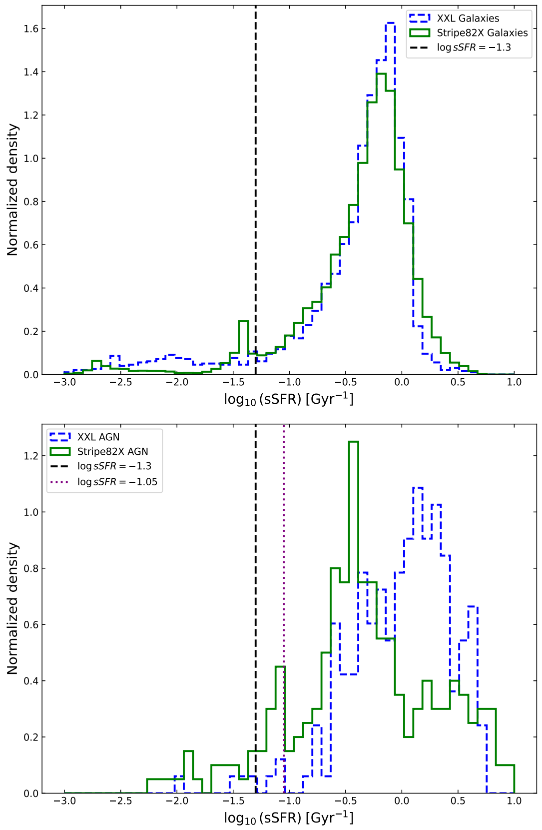


Fig. 1. Distributions of sSFR (sSFR = SFR/$M_\star$) for galaxies (top panel) and X-ray AGNs (bottom panel) for the two fields, as indicated in the legend.

more massive halos and have higher clustering amplitudes than star-forming systems at the same $M_\star$ (e.g. Coil et al. 2017), which would introduce an environmental bias unrelated to black-hole mass. By only selecting star-forming hosts, all AGN subsamples are drawn from a homogeneous population, allowing us to isolate environmental trends linked specifically to BH properties.

To identify quiescent systems, we examined the distributions of sSFR, defined as sSFR = SFR/$M_\star$ (Fig. 1). The sSFR distributions of the galaxy samples exhibit a prominent peak at high sSFR values and a secondary peak at lower values, corresponding to quiescent systems (e.g. Mountrichas et al. 2021b, 2022b). Using the galaxy samples, owing to their larger statistics, we identify the position of this secondary peak at $\log(\mathrm{sSFR}/\mathrm{Gyr}^{-1}) \approx -1.3$ (top panel of Fig. 1). We note that the AGN samples display a similar behaviour, with the secondary peak appearing at comparable sSFR values (bottom panel of Fig. 1). However, the sSFR distributions of the two AGN samples differ, with XXL AGNs appearing, on average, more star-forming than their Stripe 82X counterparts. This difference can arise even though the two surveys have comparable X-ray flux limits, because the effective selection of the final AGN samples also depends on the identification of optical/IR counterparts and on the availability and quality of redshifts and multi-wavelength photometry used for SED fitting, which differ between fields. In particular, sSFR estimates are sensitive to the photometric coverage and depth (e.g. UV/optical constraints on recent star formation and IR constraints on obscured star formation), which are not identical across XXL and Stripe 82. In addition, modest differences in the underlying redshift and stellar-mass distributions, as well as field-to-field variance, can shift the observed sSFR distributions.

Table 1. Final numbers of sources used in the clustering analysis after applying all photometric, quality, and redshift selection criteria, restricted to overlapping AGN–galaxy regions in each field and excluding quiescent systems from the X-ray datasets.

Field	X-ray AGN	Galaxies
XMM–XXL	199	2 428
Stripe 82X	228	19 141

Importantly, these differences do not bias our main conclusions. When testing for a dependence on a given AGN property, we applied multivariate matching between the AGN sub-samples themselves (Sect. 3.4), ensuring that their redshift and host-galaxy properties (including $M_\star$, SFR, and sSFR) are statistically consistent before comparing their clustering measurements.

Based on the sSFR distributions, we classified as quiescent all systems with $\log(\mathrm{sSFR}/\mathrm{Gyr}^{-1}) < -1.3$. Approximately 10% of the galaxies and 5% of the AGNs in both fields fall below this limit. However, since galaxies serve solely as the tracer population in our cross-correlation analysis, we only excluded quiescent systems from the AGN subsets. Removing them from the galaxy samples would reduce their number density and unnecessarily increase the statistical uncertainties of the clustering measurements. The final number of sources included in the clustering analysis is listed in Table 1. The distributions of black-hole and host-galaxy properties are shown in Fig. 2.

3.3. Clustering analysis

To quantify the large-scale environments of AGNs, we measured the AGN–galaxy cross-correlation function, $w_p(r_p)$ (e.g. Mountrichas et al. 2009a,b; Hickox et al. 2009; Donoso et al. 2010; Krumpe et al. 2010b; Miyaji et al. 2011; Mountrichas & Georgakakis 2012; Mountrichas et al. 2013; Shen et al. 2013; Georgakakis et al. 2014; Krumpe et al. 2012; Mendez et al. 2016; Shirasaki et al. 2016; Krumpe et al. 2018; Georgakakis et al. 2019). The use of the cross-correlation, instead of the AGN auto-correlation, offers two main advantages. First, it significantly reduces statistical uncertainties, as the galaxy sample provides a denser tracer of the underlying matter distribution. Second, it avoids the need to construct random catalogues for the AGN, whose selection function is typically complex due to spatially varying X-ray sensitivity and survey geometry. Random catalogues are only required for the galaxy populations, whose angular and redshift selection functions are well defined. This approach follows the methodology of Mountrichas et al. (2016) and Mountrichas et al. (2019) who applied the same technique to the XXL AGN and VIPERS galaxies.

The AGN–galaxy cross-correlation approach assumes that any residual angular incompleteness of the AGN sample is uncorrelated with the large-scale density field traced by galaxies. On the scales considered in this work ($r_p \gtrsim 2$–$3\,h^{-1}$ Mpc), this assumption is well justified, as AGN-specific selection effects (e.g. variability, compactness, or optical magnitude limits) are not expected to correlate with large-scale structure. This methodology has been extensively adopted in previous X-ray AGN

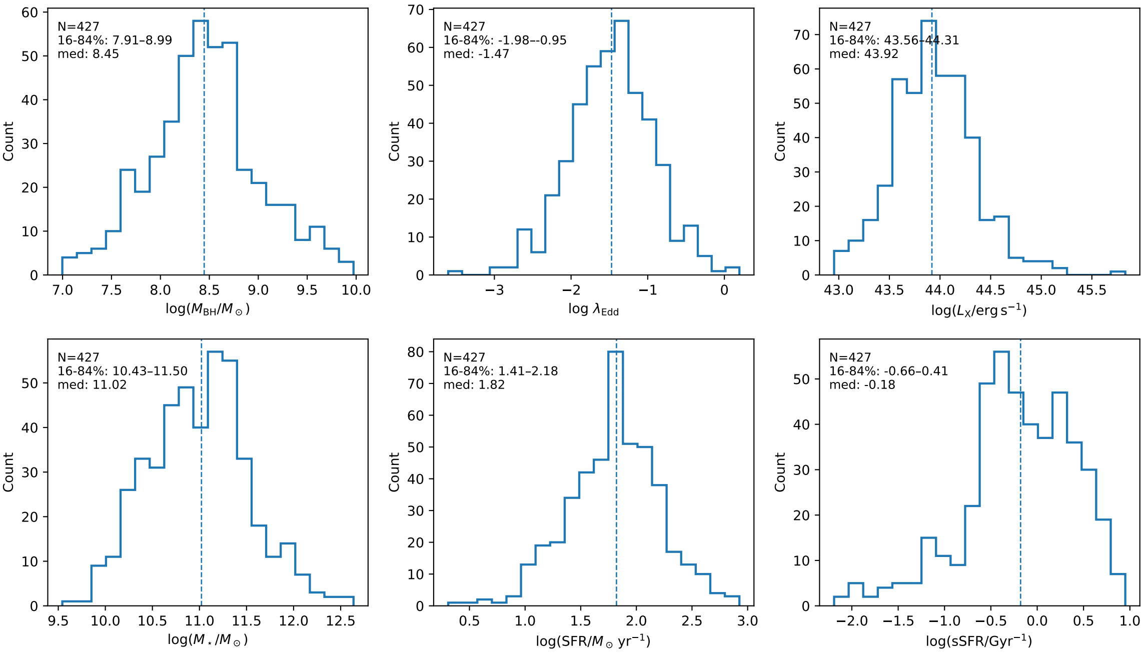


Fig. 2. Distribution of key physical properties for the combined X-ray AGN sample from the XXL and Stripe 82X surveys. From left to right and top to bottom, the panels show the distributions of black-hole mass ($\log M_{\rm BH}/M_\odot$), Eddington ratio ($\log \lambda_{\rm Edd}$), X-ray luminosity ($\log L_{\rm X}/{\rm erg\,s^{-1}}$), stellar mass ($\log M_\star/M_\odot$), SFR ($\log {\rm SFR}/M_\odot\,{\rm yr^{-1}}$), and sSFR ($\log {\rm sSFR/Gyr^{-1}}$). All quantities are shown in logarithmic units.

clustering studies and provides robust estimates of the linear bias and typical DMH mass.

As an additional validation of the angular selection underlying the AGN–galaxy cross-correlation approach, we performed a test in which the AGN angular correlation function is compared to that of random sub-samples of the galaxy catalogue using the same survey mask. The results, presented in Appendix A, show that the AGN angular clustering is fully consistent with expectations from random galaxy sub-samples on the large scales relevant for our bias measurements.

The comoving separation between two objects can be decomposed into components perpendicular (r_p) and parallel (π) to the line of sight. The two-dimensional redshift-space correlation function, $\xi(r_p, \pi)$, was estimated using the Davis & Peebles (1983) estimator:

$$\xi(r_p, \pi) = \frac{DD(r_p, \pi)}{DR(r_p, \pi)} - 1, \tag{1}$$

where DD and DR represent the data–data and data–random pair counts, respectively. For the cross-correlation, DR refers to AGN–random pairs, while for the galaxy auto-correlation it refers to galaxy–random pairs. This estimator only requires a random catalogue for the galaxy sample, which accounts for its angular- and redshift-selection function.

We note that alternative estimators, most commonly the Landy–Szalay (Landy & Szalay 1993) form, can in principle provide lower variance measurements and improved edge correction when accurate random catalogues are available for all samples entering the pair counts. In our case, the dominant practical limitation is the construction of an AGN random catalogue that faithfully captures the spatially varying X-ray sensitivity and complex survey geometry. We therefore adopt the Davis–Peebles estimator. Moreover, our bias measurements were derived by fitting the two-halo regime on scales of $r_p \gtrsim 2$–$3\,h^{-1}$ Mpc, where the correlation function is well described by linear biasing and any residual estimator-dependent effects are sub-dominant for the determination of the large-scale bias and inferred halo mass, which are the quantities of interest here.

The projected correlation function, $w_p(r_p)$, was obtained by integrating $\xi(r_p, \pi)$ along the line of sight:

$$w_p(r_p) = 2 \int_0^{\pi_{\rm max}} \xi(r_p, \pi)\, d\pi. \tag{2}$$

The integration limit $\pi_{\rm max}$ was determined empirically by varying its value and examining the convergence of $w_p(r_p)$. We adopted $\pi_{\rm max} = 40\,h^{-1}$ Mpc for the XXL field and $\pi_{\rm max} = 20\,h^{-1}$ Mpc for the Stripe 82X field, consistent with the saturation scales observed in our tests (see Mountrichas et al. 2019). For the galaxy auto-correlation functions, the adopted $\pi_{\rm max}$ values are 50 and 40 h^{-1} Mpc for XXL and Stripe 82, respectively.

Uncertainties in the projected correlation functions were estimated using the jackknife resampling technique (e.g. Ross et al. 2008). The survey area was divided into $N_{\rm JK} = 32$ and 48 spatial sub-regions for XXL and Stripe 82, respectively. The choice of the number of jackknife regions represents a compromise between two requirements. First, each sub-region must be sufficiently large compared to the maximum projected separation used in the clustering fits so that the jackknife samples capture the variance on the scales of interest. Second, the number of sub-regions must be large enough to allow a stable estimation of the covariance matrix. We verified that varying the number of jackknife regions within reasonable limits does not significantly affect the derived bias values within the statistical uncertainties. For each resampling, one sub-region was excluded and the correlation function recalculated, allowing the construction of the covariance matrix

$$C_{ij} = \frac{N_{\rm JK} - 1}{N_{\rm JK}} \sum_{k=1}^{N_{\rm JK}} \left[w_p^k(r_{p,i}) - \overline{w_p}(r_{p,i}) \right] \left[w_p^k(r_{p,j}) - \overline{w_p}(r_{p,j}) \right]. \tag{3}$$

The resulting covariance matrix captures both the uncertainties and the bin-to-bin correlations in $w_p(r_p)$ and is used when fitting the large-scale (two-halo) component of the correlation function and converting the best-fitting clustering amplitude into a halo mass.

To derive the AGN bias and the corresponding DMH mass, we fitted the measured projected correlation functions to the theoretical dark-matter correlation function, $w_p^{\rm DM}(r_p)$, computed from the linear matter power spectrum. Following the two-halo approximation, the observed correlation function at large scales can be written as

$$w_p(r_p) = b^2\, w_p^{\rm DM}(r_p), \tag{4}$$

where b is the linear bias parameter. The galaxy auto-correlation function (ACF) and the AGN–galaxy cross-correlation function (CCF) are fitted independently using their corresponding jackknife covariance matrices. The galaxy bias, b_g, was derived from the ACF, while the AGN–galaxy bias, $b_{\rm Ag}$, was derived from the CCF. The AGN bias was then calculated as

$$b_{\rm AGN} = \frac{b_{\rm Ag}}{b_g}. \tag{5}$$

In principle, the uncertainties of the ACF and CCF are correlated because the same galaxy sample and the same underlying large structure are shared in both measurements. However, the galaxy sample is significantly larger than the AGN sample and therefore the uncertainty on b_g is much smaller than that on $b_{\rm Ag}$. Consequently, the contribution of the covariance between the two measurements to the final uncertainty of $b_{\rm AGN}$ is negligible, and the errors can be safely propagated assuming independent estimates. We verified this explicitly: the fractional uncertainty on the galaxy bias is typically $\Delta b_g/b_g \sim 0.08$, compared to $\Delta b_{\rm Ag}/b_{\rm Ag} \sim 0.30$. Therefore, the uncertainty on $b_{\rm AGN}$ is dominated by the error on $b_{\rm Ag}$, and the covariance term between the ACF and CCF has a negligible impact on $\Delta b_{\rm AGN}$.

The fits were performed over the $r_p = 2$–$30\,h^{-1}$ Mpc range, where the signal is dominated by the two-halo term. The linear bias inferred from Eq. (5) was then converted into a characteristic DMH mass using the ellipsoidal collapse model of Sheth et al. (2001), adopting the analytical bias–halo-mass relations of Van den Bosch (2002). We only fitted the two-halo term of the correlation function, as the available AGN samples do not contain sufficient numbers of pairs to robustly constrain the one-halo term. Consequently, our analysis focuses on scales larger than ~2–3 h^{-1} Mpc, where linear biasing provides a valid description of the clustering signal.

Finally, the random catalogues were generated for the galaxy samples, following their exact angular footprint and redshift distribution. Each random catalogue contains 20 times more objects than the corresponding galaxy dataset, ensuring smooth pair statistics. We verified that increasing the size of the random catalogues to 30–50 times the galaxy sample does not lead to a measurable reduction in the statistical uncertainties of the projected correlation functions. This is expected, as the error budget is dominated by cosmic variance and the finite size of the AGN samples, rather than by Poisson noise in the random catalogues. We therefore adopted 20× random catalogues as a conservative and computationally efficient choice.

3.4. Matched sub-samples and parameter control

We investigated the dependence of the AGN large-scale environment on three fundamental physical parameters: $M_{\rm BH}$, $\lambda_{\rm Edd}$, and $L_{\rm X}$. For each parameter, the AGN sample was divided into low and high subsets, and the corresponding DMH masses were estimated.

A key aspect of this analysis is controlling for other host-galaxy and AGN properties known to correlate with clustering amplitude. Previous studies have shown that X-ray AGN clustering depends on quantities such as $L_{\rm X}$, $M_\star$, SFR, and sSFR (e.g. Krumpe et al. 2010b; Georgakakis et al. 2014; Mountrichas et al. 2016, 2019; Allevato et al. 2019; Krumpe et al. 2023). Although the detailed form of these dependencies remains debated (e.g. Krumpe et al. 2018; Viitanen et al. 2019; Powell et al. 2020), we explicitly accounted for them when isolating the impact of SMBH-related parameters on the large-scale environment.

Moreover, $M_{\rm BH}$, $\lambda_{\rm Edd}$, and $L_{\rm X}$ are themselves mutually correlated (e.g. Krumpe et al. 2015; Mountrichas & Buat 2023). To avoid conflating these interdependencies, we explicitly matched the distributions of all parameters not under investigation in each test. For example, when examining clustering as a function of $M_{\rm BH}$, we required the high- and low-$M_{\rm BH}$ sub-samples to have statistically consistent distributions in $\lambda_{\rm Edd}$, $L_{\rm X}$, redshift, and host-galaxy properties.

In addition, we distinguished between two broad luminosity regimes, defining low-to-moderate-luminosity AGNs as systems with $\log(L_{\rm X}/{\rm erg\,s^{-1}}) < 44$ and high-luminosity AGNs as those with $\log(L_{\rm X}/{\rm erg\,s^{-1}}) \geq 44$. This threshold approximately separates moderate-luminosity Seyfert-like AGNs from quasar-level activity and is commonly adopted in X-ray-AGN clustering studies. The luminosity split allows us to test whether any residual environmental trends depend on accretion regime while maintaining sufficient statistics within each sub-sample.

To construct sub-samples with consistent distributions across all control parameters, we applied a multivariate nearest-neighbour matching algorithm implemented in Python. The matching is performed using a custom routine (`multivar_match`) developed for this work. The algorithm performs a greedy nearest–neighbour search in a multi-dimensional parameter space defined by the matching variables (e.g. log(sSFR), log(SFR), $\log(M_\star)$, $\log(L_{\rm X})$, $\log(M_{\rm BH})$, and $\log\lambda_{\rm Edd}$).

The implementation relies on the NumPy numerical library (`numpy`) and operates on Astropy tables. All covariates were standardised using the pooled mean and standard deviation of the two samples, and distances are computed using the Euclidean metric in this standardized parameter space. For each object in the reference sample, the algorithm identifies the closest counterpart in the comparison sample and performs greedy nearest–neighbour matching without replacement, ensuring strict one-to-one matching between objects. The matching routine is available from the authors upon request.

A key parameter of the method is the caliper, which defines the maximum allowable distance between two matched objects in the normalised covariate space. We adopted a caliper value of 1.2, which provides a good balance between sample size and matching precision. Lower values led to a substantial reduction in the number of matched sources, while higher values increase the dispersion of the matched properties and weaken the control over covariates. In practice, this procedure ensures that when comparing low and high subsets of a given parameter (e.g. $M_{\rm BH}$), that parameter is allowed to differ by construction, while all other covariates remain closely matched within the adopted caliper distance in the multivariate parameter space.

For validation, we also experimented with alternative approaches, such as applying hard cuts on the control parameters and matching the distributions of, for instance, $M_\star$, SFR, and sSFR between sub-samples. However, these methods resulted in significantly smaller subsets, limiting their statistical usefulness. The multivariate matching approach therefore offers the most efficient compromise between parameter control and sample size. The resulting matched sub-samples are consistent in redshift, host-galaxy and SMBH properties, allowing a robust and controlled investigation of how $M_{\rm BH}$, $\lambda_{\rm Edd}$, and $L_{\rm X}$ independently influence AGN clustering.

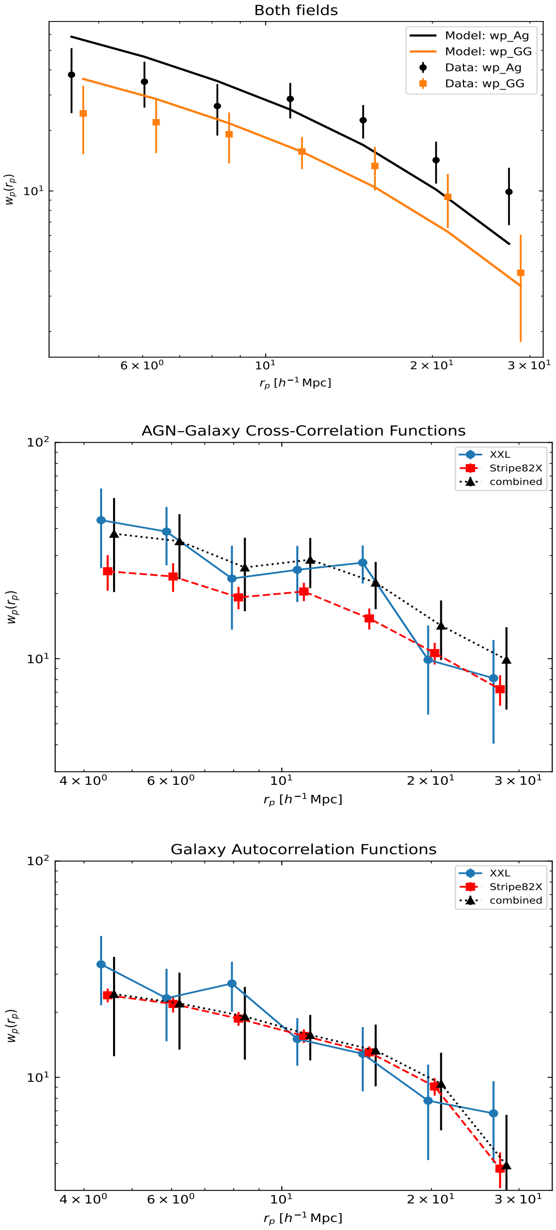
4. Results

This section presents the derived DMH masses of the full X-ray AGN population. We further investigated how the clustering strength, and hence the DMH mass, varies with $M_{\rm BH}$, $\lambda_{\rm Edd,}$ and $L_{\rm X}$. We emphasise that all clustering measurements and bias estimates presented below are derived exclusively from the large-scale ($r_p \gtrsim$ 2–3 h^{-1} Mpc) two-halo regime, where the effects of the one-halo term are negligible and linear biasing provides a valid description of the clustering signal. Given the limited AGN statistics after multivariate matching, our binning strategy is intentionally conservative and optimised for the stability of the inferred large-scale bias rather than for the visual smoothness of $w_p(r_p)$.

4.1. Clustering and DMH mass of the full X-ray AGN population

We first estimated the DMH masses of the total AGN (i.e. including quiescent hosts) and galaxy populations in the two fields independently, and then combined the pair counts (DD and DR) from both surveys to derive the overall AGN–galaxy cross-correlation function and the galaxy auto-correlation function (e.g. Mountrichas et al. 2013; Georgakakis et al. 2014). This combined measurement provides the characteristic DMH mass of the X-ray AGN population and of the galaxy tracer sample across both fields.

Figure 3 (top panel) presents the AGN–galaxy cross-correlation function and the galaxy auto-correlation function for the combined XXL and Stripe 82X samples, together with their best-fitting models. The fitting is performed over scales of 4–30 h^{-1} Mpc; however, we verified that the derived results remain statistically consistent when adopting alternative fitting ranges, such as 2–25 h^{-1} Mpc. The middle and bottom panels show the AGN–galaxy and galaxy auto-correlation functions, respectively, measured separately in each field, as well as the combined results. The properties of the AGN and galaxy datasets, together with the corresponding DMH mass estimates for each field and for the merged sample, are listed in Table 2. As already mentioned, the combined cross-correlation function is obtained by summing the DD and DR pair counts from the individual fields, using the same estimator and the same random-to-data ratio in all cases. In this formulation, $w_p^{\rm comb}(r_p)$ is a DR-weighted average of the individual field measurements, and small deviations where the combined curve lies slightly above or below the individual ones are fully consistent with the statistical uncertainties.



Both fields

AGN–Galaxy Cross-Correlation Functions

Galaxy Autocorrelation Functions

Fig. 3. Correlation function measurements. Top panel: AGN–galaxy cross-correlation function (black points) and galaxy auto-correlation function (orange points) for the combined XXL and Stripe 82X samples, shown together with their best-fitting models (solid lines). Middle and bottom panels: Same measurements performed separately for each field, compared with the combined results. The corresponding DMH mass estimates for each field and for the merged dataset are listed in Table 2. Data points have been offset in the x-axis for clarity.

Our results show that X-ray AGNs reside in DMHs with typical masses of $\sim 10^{13}\,M_\odot$, consistent with previous clustering studies over a wide range of redshifts up to $z \sim 3$ (e.g. Allevato et al. 2011; Krumpe et al. 2012; Mountrichas & Georgakakis 2012; Mountrichas et al. 2013; Koutoulidis et al. 2013; Georgakakis et al. 2014; Krumpe et al. 2018; Viitanen et al. 2019). Powell et al. (2020) investigated the clustering of X-ray-selected AGNs in the same fields used here (Stripe 82X and XMM–XXL north) by measuring the AGN auto-correlation function in two X-ray luminosity intervals: a low-luminosity bin with $10^{43} \leq L_{\rm X}\,(2-10\,{\rm keV}) < 10^{44.5}\,{\rm erg\,s^{-1}}$ at $z \approx 0.8$, and a high-luminosity bin with $L_{\rm X}\,(2–10\,{\rm keV}) > 10^{44.5}\,{\rm erg\,s^{-1}}$ at $z \approx 1.8$. Their results for the low-luminosity subset yielded a halo mass of $\log(M_{\rm DMH}/h^{-1}M_\odot) = 12.83^{+0.25}_{-0.41}$ (see their Table 2), in excellent agreement with our measurements.

Table 2. Median properties of AGN and galaxy samples in the XXL, Stripe 82X, and combined fields.

Field	Sample	$N_{\rm src}$	$\langle z \rangle$	$\langle \log(M_\star/M_\odot) \rangle$	$\langle \log({\rm SFR}/M_\odot\,{\rm yr}^{-1}) \rangle$	$\langle \log({\rm sSFR}/{\rm Gyr}^{-1}) \rangle$	$\langle \log(L_{\rm X}/{\rm erg\,s}^{-1}) \rangle$	$\langle \log(M_{\rm BH}/M_\odot) \rangle$	$\langle \log \lambda_{\rm Edd} \rangle$	$\log(M_{\rm DMH}/h^{-1}M_\odot)$
XXL	AGN	203	0.82	10.81	1.81	0.07	43.79	8.28	−1.36	13.06 ± 0.27
	Galaxies	2428	0.65	11.02	1.67	−0.27	–	–	–	12.98 ± 0.07
Stripe 82X	AGN	245	0.82	11.17	1.83	−0.38	44.04	8.60	−1.58	12.83 ± 0.12
	Galaxies	19 141	0.68	10.89	1.51	−0.29	–	–	–	12.68 ± 0.03
Combined	AGN	448	0.82	11.02	1.82	−0.18	43.92	8.45	−1.47	13.17 ± 0.22
	Galaxies	21 569	0.68	10.91	1.52	−0.29	–	–	–	12.74 ± 0.16

Notes. Columns: field, sample type, number of sources ($N_{\rm src}$), median values of redshift ($\tilde{z}$), stellar mass ($M_\star$), SFR, sSFR, X-ray luminosity ($L_{\rm X}$), black-hole mass ($M_{\rm BH}$), Eddington ratio ($\lambda_{\rm Edd}$), and measurements of the effective DMH mass ($M_{\rm DMH}$). Dashes indicate quantities not applicable to galaxies.

Table 3. Median properties of the matched AGN subsets used to investigate the dependence of environment on black-hole mass.

Subset	N	$\langle z \rangle$	$\langle \log {\rm sSFR} \rangle$	$\langle \log {\rm SFR} \rangle$	$\langle \log M_\star \rangle$	$\langle \log L_{\rm X} \rangle$	$\log(M_{\rm BH}/M_\odot)$	$\langle \log \lambda_{\rm Edd} \rangle$	$b_{\rm AGN}$	$\log M_{\rm DMH}$
low–$L_{\rm X}$, low–$M_{\rm BH}$	62	0.66	−0.15	1.64	10.89	43.64	8.01	−1.39	1.75 ± 0.40	13.10 ± 0.42
low–$L_{\rm X}$, high–$M_{\rm BH}$	62	0.83	−0.20	1.75	10.96	43.80	8.48	−1.52	1.62 ± 0.50	12.70 ± 0.65
high–$L_{\rm X}$, low–$M_{\rm BH}$	51	0.89	+0.02	1.98	10.93	44.19	8.43	−1.32	2.33 ± 0.35	13.29 ± 0.24
high–$L_{\rm X}$, high–$M_{\rm BH}$	51	0.97	+0.03	1.98	11.09	44.24	8.84	−1.53	2.66 ± 0.20	13.39 ± 0.11
all–$L_{\rm X}$, low–$M_{\rm BH}$	123	0.71	−0.14	1.75	10.89	43.85	8.17	−1.43	1.92 ± 0.22	13.18 ± 0.20
all–$L_{\rm X}$, high–$M_{\rm BH}$	123	0.93	−0.13	1.94	11.13	44.11	8.62	−1.69	2.02 ± 0.30	13.03 ± 0.29

Notes. Median values of sSFR, SFR, stellar mass ($M_\star$), X-ray luminosity ($L_{\rm X}$), redshift (z), black-hole mass ($M_{\rm BH}$), Eddington ratio ($\lambda_{\rm Edd}$), the AGN linear bias parameter ($b_{\rm AGN}$), and the inferred DMH mass ($M_{\rm DMH}$) for the matched AGN subsets. The bias parameter was derived from the AGN–galaxy cross-correlation analysis, and halo masses were obtained by converting $b_{\rm AGN}$ to $M_{\rm DMH}$ using the Sheth–Tormen formalism assuming $\sigma_8 = 0.811$. For the four subsamples split simultaneously into luminosity and black-hole mass, 'low–high' divisions correspond to $\log L_{\rm X} \leq 44$ and $\log M_{\rm BH} \leq 8.3$ (low) versus $\log L_{\rm X} > 44$ and $\log M_{\rm BH} > 8.6$ (high). For the 'all–$L_{\rm X}$' rows, the samples are divided using a single black-hole mass threshold at $\log(M_{\rm BH}/M_\odot) = 8.5$.

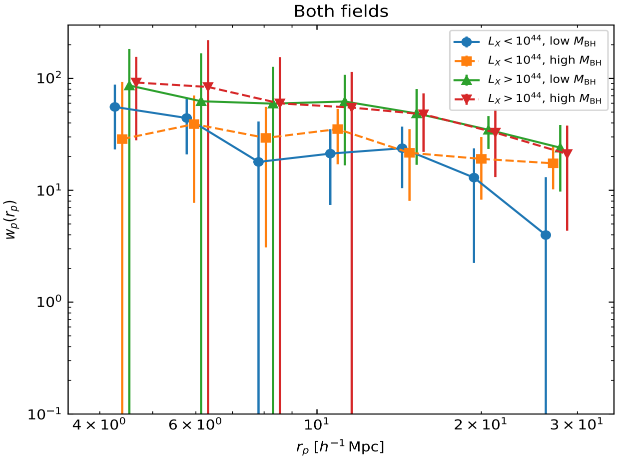


Fig. 4. AGN–galaxy cross-correlation functions as a function of $M_{\rm BH}$, in two $L_{\rm X}$ bins, as indicated in the legend. The corresponding DMH mass estimates are listed in Table 3.

4.2. Relation between black-hole mass and large-scale environment

In this section, we examined how the typical DMH mass of X-ray-selected AGNs depends on their $M_{\rm BH}$. For this purpose, we used the datasets presented in Table 1, i.e. excluding quiescent systems from the AGN samples (see Sect. 3.2) and applied the multivariate nearest-neighbour matching algorithm described in Sect. 3.4.

Ideally, one would like to divide the AGN into low- and high-$M_{\rm BH}$ subsets while leaving a buffer region between the two to account for the typical uncertainties in $M_{\rm BH}$, which are of the order of 0.3–0.5 dex. However, the available sample size does not allow for such fine sub-division. Instead, we used the median $M_{\rm BH}$ values of the combined X-ray datasets in the low- moderate-, and high-luminosity bins, as defined by the cut at $\log(L_{\rm X}/{\rm erg\,s}^{-1}) = 44$ (see Sect. 3.4). Specifically, we adopt thresholds of $\log(M_{\rm BH}/M_\odot) = 8.3$ and 8.6 for the low-, moderate-, and high-$L_{\rm X}$ AGN, respectively.

When running the matching algorithm, we use sSFR, SFR, and $\lambda_{\rm Edd}$ as covariates. Including additional parameters such as $M_\star$ or $L_{\rm X}$ would substantially reduce the number of matched sources and compromise the statistical robustness of the clustering measurements. Nevertheless, matching on sSFR, SFR, and $\lambda_{\rm Edd}$ results in AGN subsets with highly similar distributions in $M_\star$, redshift, and $L_{\rm X}$, as verified a posteriori in Table 3.

The AGN–galaxy cross-correlation functions for the four subsets are shown in Fig. 4, while the corresponding bias measurements and DMH mass estimates are listed in Table 3. In addition to the subsets split simultaneously in $L_{\rm X}$ and $M_{\rm BH}$, the table also includes results for samples divided only by black-hole mass across the full luminosity range ("all–$L_{\rm X}$" rows), providing a complementary comparison that maximises the available statistics.

Due to the small number of sources available in each subset, the fitting is performed over scales of 4–30 h^{-1} Mpc. Based on our measurements, the large-scale environment of X-ray AGNs shows no significant dependence on $M_{\rm BH}$, for either the low-, moderate-, or the high- $L_{\rm X}$ sub-samples. We note that individual $w_p(r_p)$ measurements exhibit modest bin-to-bin fluctuations, which are expected given the limited AGN statistics and the correlated jackknife uncertainties. The inferred bias and halo masses were obtained from fits over a broad range of scales and are robust to reasonable variations in the fitting range.

Table 4. Median properties of the matched AGN subsets used to investigate the dependence of environment on $\lambda_{\rm Edd}$.

Subset	N	$\langle z \rangle$	$\langle \log \rm sSFR \rangle$	$\langle \log \rm SFR \rangle$	$\langle \log M_\star \rangle$	$\langle \log L_{\rm X} \rangle$	$\langle \log M_{\rm BH} \rangle$	$\langle \log \lambda_{\rm Edd} \rangle$	$b_{\rm AGN}$	$\log M_{\rm DMH}$
low–$L_{\rm X}$, low–$\lambda_{\rm Edd}$	66	0.70	−0.19	1.70	10.93	43.80	8.39	−1.71	1.63 ± 0.30	12.94 ± 0.38
low–$L_{\rm X}$, high–$\lambda_{\rm Edd}$	66	0.83	−0.17	1.80	10.93	43.81	8.27	−1.29	2.09 ± 0.27	13.16 ± 0.23
high–$L_{\rm X}$, low–$\lambda_{\rm Edd}$	53	0.81	+0.09	1.94	11.01	44.27	8.71	−1.67	2.21 ± 0.36	13.32 ± 0.28
high–$L_{\rm X}$, high–$\lambda_{\rm Edd}$	53	0.91	+0.03	2.00	10.90	44.32	8.51	−1.30	3.14 ± 0.84	13.72 ± 0.36
all–$L_{\rm X}$, low–$\lambda_{\rm Edd}$	146	0.74	−0.16	1.73	10.93	43.96	8.19	−1.54	1.91 ± 0.26	13.14 ± 0.23
all–$L_{\rm X}$, high–$\lambda_{\rm Edd}$	146	0.91	−0.11	1.99	11.14	44.16	7.92	−0.71	2.52 ± 0.35	13.41 ± 0.20

Notes. Median values of sSFR, SFR, stellar mass ($M_\star$), X-ray luminosity ($L_{\rm X}$), redshift (z), black-hole mass ($M_{\rm BH}$), Eddington ratio ($\lambda_{\rm Edd}$), the AGN linear bias parameter ($b_{\rm AGN}$), and the inferred DMH mass ($M_{\rm DMH}$) for the matched AGN subsets. 'Low–high' divisions correspond to $\log L_{\rm X} \leq 44$ and $\log \lambda_{\rm Edd} \leq -1.50$ (low) versus $\log L_{\rm X} > 44$ and $\log \lambda_{\rm Edd} > -1.50$ (high). The same Eddington-ratio threshold was used for the 'all–$L_{\rm X}$' subsamples. The last column lists the DMH masses derived from the AGN–galaxy cross-correlation analysis.

Krumpe et al. (2015) used RASS/SDSS-detected AGNs at redshifts $z = 0.2$–0.3 and reported a weak dependence of clustering strength on $M_{\rm BH}$. It is important to note, however, several key differences between their study and ours. Their X-ray sample includes sources with systematically lower $M_{\rm BH}$ (see their Table 3). More crucially, Krumpe et al. (2015) did not account for differences in host-galaxy properties, which can influence the inferred DMH masses (e.g. Mountrichas et al. 2019; Allevato et al. 2019).

Shen et al. (2009) analysed optical quasars from the SDSS Data Release 5 (DR5) and found no significant difference in the clustering strength when splitting their sample into high- and low-$M_{\rm BH}$ sub-samples. However, when they compared the most massive 10% of black holes to the rest of the population, the quasars hosting the largest $M_{\rm BH}$ exhibited stronger clustering. It is worth noting, however, that their dataset consists exclusively of optically selected quasars, in contrast to the X-ray–detected AGNs studied here. In addition, their analysis covers a broad redshift range ($0.4 < z < 2.5$) and therefore does not account for potential redshift evolution. Finally, their approach does not control for differences in other black-hole (e.g. $\lambda_{\rm Edd}$, luminosity) or host-galaxy properties, which are likely to vary between the low- and high-$M_{\rm BH}$ sub-samples.

4.3. Relation between Eddington ratio and large-scale environment

In this section, we investigate how the DMH mass of X-ray-selected AGNs depends on their $\lambda_{\rm Edd}$, using the datasets summarised in Table 1 and applying the multivariate nearest-neighbour matching procedure described in Sect. 3.4. The $\lambda_{\rm Edd}$ for each AGN is computed following the standard definition,

$$\lambda_{\rm Edd} = \frac{L_{\rm bol}}{L_{\rm Edd}}, \qquad L_{\rm Edd} = 1.26 \times 10^{38} \left(\frac{M_{\rm BH}}{M_\odot} \right) \rm erg\,s^{-1}. \tag{6}$$

In the implementation used here, $L_{\rm bol}$ was derived from the SED fitting.

Similarly to the previous section, we adopted a luminosity threshold at $\log(L_{\rm X}/\rm erg\,s^{-1}) = 44$ to define low-, moderate-, and high-luminosity AGNs. Within each luminosity bin, the median $\log \lambda_{\rm Edd}$ of the combined XXL and Stripe 82X datasets is −1.5, and this value was used to separate AGN into low- and high-$\lambda_{\rm Edd}$ subsets. The multivariate matching algorithm is applied to the star-forming-AGN populations within each luminosity bin, using $\log \rm sSFR$, $\log \rm SFR$, and $\log M_{\rm BH}$ as matching covariates. The resulting matched sample properties are listed in Table 4.

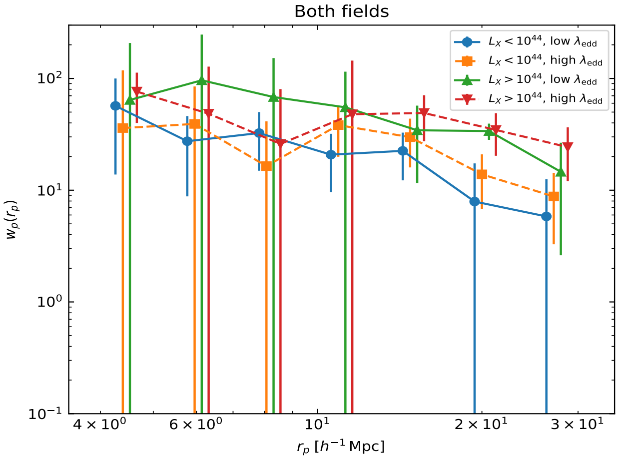


Fig. 5. AGN–galaxy cross-correlation functions as a function of $\lambda_{\rm Edd}$, in two $L_{\rm X}$ bins, as indicated in the legend. The corresponding DMH mass estimates are listed in Table 4.

The AGN–galaxy cross-correlation functions for the four matched subsets are shown in Fig. 5. The calculated $b_{\rm AGN}$ and the corresponding DMH mass estimates are listed in Table 4. The fits are performed over scales of 4–30 h^{-1} Mpc. We find no statistically significant dependence of the typical halo mass on $\lambda_{\rm Edd}$, with any apparent trend being consistent with statistical fluctuations at the $\sim 1\sigma$ level. Previous clustering studies reported little or no correlation between halo mass and $\lambda_{\rm Edd}$ (e.g. Krumpe et al. 2015), or with its empirical proxy, the specific black-hole accretion rate ($L_{\rm X}/M_\star$; Allevato et al. 2019; Viitanen et al. 2019).

4.4. Relation between X-ray luminosity and large-scale environment

In this section, we examine how the DMH mass of X-ray-selected AGNs depends on their $L_{\rm X}$. The (e.g. Shankar et al. 2013; Aird et al. 2012; Allevato et al. 2011; Georgakakis et al. 2019) multivariate nearest-neighbour matching algorithm (see Sect. 3.4) was applied to the AGN samples using $\log \rm sSFR$, $\log \rm SFR$, $\log \lambda_{\rm Edd}$, and $\log M_{\rm BH}$ as covariates. The median properties of the resulting matched subsets are listed in Table 5.

The AGN–galaxy cross-correlation functions for the matched samples are shown in Fig. 6. The calculated $b_{\rm AGN}$ and the corresponding DMH mass estimates are presented in Table 5. Our measurements show a hint that more luminous AGNs tend to reside in more massive DMHs compared to their less luminous counterparts; however, the statistical significance of this trend is statistically insignificant.

Table 5. Median properties of the matched AGN subsets used to investigate the dependence of environment on L_X.

Subset	N	$\langle z \rangle$	$\langle \log \mathrm{sSFR} \rangle$	$\langle \log \mathrm{SFR} \rangle$	$\langle \log M_\star \rangle$	$\langle \log L_X \rangle$	$\langle \log M_{\mathrm{BH}} \rangle$	$\langle \log \lambda_{\mathrm{Edd}} \rangle$	b_{AGN}	$\log M_{\mathrm{DMH}}$
low–L_X	115	0.79	−0.14	1.84	10.96	43.82	8.52	−0.91	2.04 ± 0.39	13.21 ± 0.31
high–L_X	115	0.83	−0.03	1.87	10.97	44.19	8.54	−0.86	2.41 ± 0.24	13.42 ± 0.14

Notes. Median values of sSFR, SFR, stellar mass ($M_\star$), X-ray luminosity (L_X), redshift (z), black hole mass (M_{BH}), Eddington ratio (λ_{Edd}); and measurements of the AGN linear bias parameter (b_{AGN}) and the effective DMH mass (M_{DMH}) for the two matched AGN subsets. 'Low–high' divisions correspond to $\log L_X \leq 44$ (low) versus $\log L_X > 44$ (high).

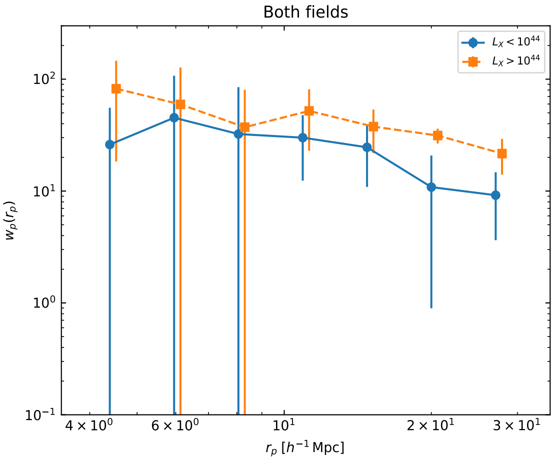


Fig. 6. AGN–galaxy cross-correlation functions as a function of L_X. The corresponding DMH mass estimates are listed in Table 5.

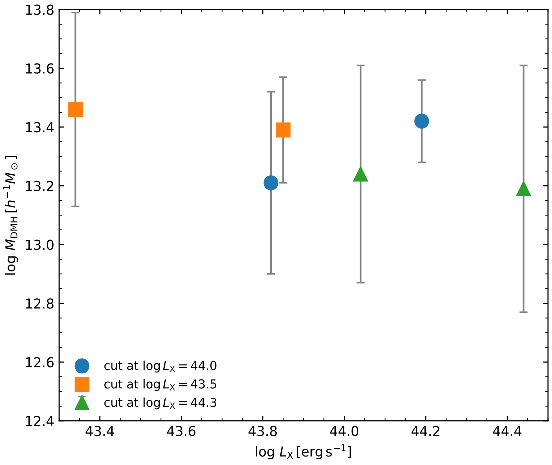


Fig. 7. Dependence of the characteristic DMH mass on the L_X threshold used to divide the AGN sample. Each symbol type corresponds to a different luminosity cut applied to split them. Because the sSFR bins are defined by quantiles, the statistical uncertainties are of a similar order across bins; the highest sSFR bin is open-ended and therefore represents a broader range of star-formation activity. The sample is split into 'low' and 'high' L_X sub-samples: $\log L_X = 43.5$ (squares), 44.0 (circles), and 44.3 (triangles). For each luminosity threshold, two data points are shown, representing the low-L_X and high-L_X sub-samples defined by that cut. The horizontal position of each point corresponds to the median X-ray luminosity of the respective sub-sample. Error bars indicate the 1σ uncertainties on the derived DMH mass.

Previous studies have reported mixed results regarding the dependence of AGN clustering on luminosity. Several works have found a positive correlation, with more luminous AGNs inhabiting more massive halos (e.g. Krumpe et al. 2010b, 2012; Koutoulidis et al. 2013), while others reported no significant dependence on large scales, but differences emerging at small scales (Krumpe et al. 2018). As discussed earlier, the limited size of our current datasets does not allow us to probe the small-scale regime of the correlation function. Interestingly, an opposite (negative) luminosity dependence has also been observed, with higher luminosity AGNs found in less massive halos (Mountrichas et al. 2016). That discrepancy was attributed to differences in the luminosity ranges explored across various studies, an interpretation supported by semi-analytic model predictions (Fanidakis et al. 2013b,a); see also Fig. 7 of Mountrichas et al. (2016). Finally, more recent studies have claimed that the clustering differences between different AGN-selection techniques are dominated by selection biases, and not due to a dependence on AGN luminosity (Powell et al. 2020).

To further assess whether the inferred halo masses depend on AGN luminosity, we recomputed the clustering measurements after applying additional luminosity thresholds at $\log L_X = 43.5$ and $\log L_X = 44.3$. A more stringent cut at $\log L_X = 44.5$ was not pursued, as the multivariate matching leaves too few AGN for a reliable measurement. The results for all luminosity-defined subsets are presented in Fig. 7.

Measurement uncertainties in $\log L_X$, together with the multivariate matching procedure, can introduce partial cross-talk between the luminosity sub-samples at the population level. As a result, formal correlation coefficients or regression analyses between halo mass and L_X would not provide meaningful constraints, since the effective number of independent degrees of freedom is reduced. In this context, the appropriate comparison is the qualitative and quantitative consistency of the halo-mass estimates across different luminosity cuts, rather than attempting to fit a continuous trend.

Across all luminosity thresholds explored, the inferred halo masses remain consistent within their uncertainties, with no systematic increase or decrease towards higher L_X. This stability across overlapping luminosity selections indicates that, within our sample, the large-scale environments of X-ray AGNs show no statistically significant correlation with L_X.

5. Discussion

The analysis presented in this paper was performed within a well-defined and physically motivated region of parameter space. By focusing on X-ray-selected AGNs at fixed redshift ($0.5 < z < 1.2$) and controlling for host-galaxy properties, we probed the typical regime of massive, star-forming AGN hosts rather than the full galaxy population. This approach minimises evolutionary mixing and isolates secondary dependencies beyond those set by the host galaxy.

Based on our results, we find no statistically significant correlation between DMH mass and $M_{\rm BH}$ across the $7.5 \lesssim \log(M_{\rm BH}/M_\odot) \lesssim 10$ range. Given the limited dynamic range in $M_{\rm BH}$ and the modest AGN statistics after multivariate matching, this result should be interpreted as the absence of a detectable trend within the uncertainties of the present dataset, rather than as evidence for the intrinsic absence of a coupling between $M_{\rm BH}$ and halo mass.

The relatively limited dynamic range in $M_{\rm BH}$ (e.g. Shankar et al. 2013; Aird et al. 2012; Allevato et al. 2011; Georgakakis et al. 2019 primarily reflects the deliberate choice to focus on a narrow redshift interval. In this regime, the analysis was optimised to test for the presence of strong or systematic differences in halo mass at fixed host-galaxy properties, rather than to constrain weak or continuous dependencies that would require substantially larger samples and broader parameter coverage. In this context, the absence of statistically significant variations is itself informative, indicating that any residual dependence of halo mass on black-hole or accretion properties must be sub-dominant relative to the effects of the host galaxy.

Several recent studies have suggested that halo mass may play a more fundamental role in regulating black-hole growth than host-galaxy properties alone. For example, clustering and residual analyses indicate that $M_{\rm BH}$ may correlate with halo mass at fixed $M_\star$ or velocity dispersion (e.g. Powell et al. 2020; Shankar et al. 2025), while semi-analytic models predict a close connection between black-hole growth and halo assembly history (e.g. Shankar et al. 2010, 2013, 2016, 2017; Menci et al. 2023). Our analysis does not contradict these findings; rather, it indicates that any such dependence is likely a weak or second-order dependence on the large scales probed here and difficult to detect with current X-ray-AGN samples.

Hydrodynamical simulations such as ILLUSTRISTNG and EAGLE suggest that the mass of the central black hole correlates most strongly with the depth of the central gravitational potential of the host galaxy (e.g. its binding energy or velocity dispersion), while the halo mass influences black-hole growth more indirectly by regulating gas supply, AGN duty cycle, and long-term evolutionary pathways (e.g. Booth & Schaye 2010; Dubois et al. 2012; Bower et al. 2017; Weinberger et al. 2018; Terrazas et al. 2020; Habouzit et al. 2022; Menci et al. 2023).

In this framework, halo mass primarily sets the boundary conditions for black-hole growth through its impact on galaxy assembly and gas accretion histories, whereas the instantaneous accretion rate and observed AGN properties are governed predominantly by baryonic processes operating on galactic and sub-galactic scales. Our results are consistent with this picture, in which halo mass modulates black-hole growth indirectly via host-galaxy evolution rather than through a tight, one-to-one correspondence between halo mass and black-hole mass. (e.g. Shankar et al. 2013; Aird et al. 2012; Allevato et al. 2011; Georgakakis et al. 2019).

We also note that the low AGN duty cycle expected at $z \lesssim 1$ implies that only a small fraction of massive halos host an active nucleus at any given epoch (e.g. Allevato et al. 2011; Aird et al. 2012; Shankar et al. 2013; Georgakakis et al. 2019). As a result, even if halo mass plays a fundamental role in black-hole growth over cosmic time, large-scale clustering measurements of active systems are expected to show only weak dependences once host-galaxy properties are controlled.

Recent weak-lensing studies have also explored the connection between $M_{\rm BH}$ and DMH mass. For example, Li et al. (2024) measured the halo masses of optically selected SDSS AGN using weak gravitational lensing from the UNIONS survey and reported a positive correlation between halo mass and $M_{\rm BH}$. However, their analysis is based primarily on lower-redshift AGN and employs a different halo-mass estimator, making a direct comparison with our X-ray-selected sample and clustering-based measurements non-trivial. In addition, the inferred dependence of halo mass on AGN properties may be sensitive to correlations with host-galaxy parameters, which we explicitly checked for via multivariate matching.

Tracing the instantaneous efficiency of accretion, $\lambda_{\rm Edd}$ shows only a weak and statistically insignificant dependence on halo mass. This suggests that the short-term fuelling efficiency of AGNs is largely independent of the surrounding DMH potential. Large-scale environment may regulate whether a galaxy has gas available for accretion in a statistical sense, but it does not control how efficiently that gas is converted into radiation once it reaches the central regions. Our results are consistent with previous studies showing similar halo masses across wide $\lambda_{\rm Edd}$ ranges (e.g. Krumpe et al. 2015; Viitanen et al. 2019), and together they indicate that AGN variability is dominated by stochastic or secular processes within galaxies rather than by their position in the cosmic web. The weak, marginal trend towards higher halo masses at larger $\lambda_{\rm Edd}$, if real, could reflect the tendency of gas-rich group environments to promote more efficient inflows or longer duty cycles, but this effect is at best secondary.

The comparison between halo mass and $L_{\rm X}$ reveals no statistically significant correlation. This agrees with previous clustering analyses (e.g. Krumpe et al. 2018; Powell et al. 2020), which similarly found that AGNs with different X-ray powers occupy halos of comparable masses. This behaviour suggests that once an AGN phase is triggered, its radiative output is primarily governed by internal processes, such as stochastic fuelling, accretion-rate variability, or self-regulating feedback, rather than by the large-scale environment. In the luminosity regime probed here ($43 \lesssim \log(L_{\rm X}/{\rm erg\,s^{-1}}) \lesssim 45$), most systems likely operate in a moderate-accretion or 'maintenance-mode' regime, where black holes are sustained by intermittent inflows of cold gas but limited by feedback that stabilises the host's gas reservoir. As a result, fluctuations in $L_{\rm X}$ are expected to trace short-term variability or duty-cycle modulation rather than systematic differences in halo potential.

Recent clustering measurements of X-ray AGNs in the COSMOS field by Ikeda et al. (2025) probed a similar redshift range ($0.6 < z < 1.4$), but at significantly lower X-ray luminosities ($\langle \log L_{\rm X} \rangle \sim 42.7$). They reported a characteristic DMH mass of $\log M_{\rm DMH} \sim 11.8\,h^{-1}M_\odot$ (AGN bias of ~1.2), which is substantially lower than the values derived in our analysis ($\log M_{\rm DMH} \sim 13.2\,h^{-1}M_\odot$ and AGN bias of ~2.1). Although the comparison is not fully controlled because their sample includes both type 1 and type 2 AGNs and does not match host-galaxy properties, the difference is consistent with the possibility of increasing halo mass with AGN luminosity when considering a broader $L_{\rm X}$ range. We note, however, that halo masses inferred from clustering depend on the adopted bias–halo mass relation. Ikeda et al. (2025) used the prescription of Tinker et al. (2005), whereas in this work we converted bias to halo mass using the Sheth–Tormen formalism. Consequently, part of the difference in the reported halo masses may reflect the different bias–halo mass conversions, in addition to differences in the underlying AGN selection and luminosity ranges.

An additional parameter that may influence measured halo masses is AGN obscuration. Earlier works suggested that obscured AGNs may reside in denser environments than unobscured ones (e.g. Allevato et al. 2011), although this has not been universally confirmed (e.g. Viitanen et al. 2023). Our sample consists exclusively of type 1 AGNs, and although optical classification does not perfectly correspond to X-ray absorption, both the XXL and Stripe 82X surveys are relatively shallow wide-area fields that contain few heavily obscured sources (Liu et al. 2016; Peca et al. 2023). Specifically, only 17 (~8%) and 12 (~6%) of the sources in the XXL and Stripe 82X samples, respectively, have $N_{\rm H} > 10^{22}\,{\rm cm}^{-2}$. Consequently, the limited presence of obscured AGNs likely minimises any environmental bias that could otherwise arise from obscuration-dependent selection effects.

It is important to emphasise that our conclusions are restricted to the large-scale environment probed by the clustering analysis ($\gtrsim$ 1Mpc). While we find no evidence that large-scale environment plays a dominant role in regulating AGN activity once host-galaxy properties are controlled, this does not exclude the importance of environmental processes operating on smaller scales. Physical mechanisms such as galaxy interactions, mergers, or cluster-specific processes occur on $\lesssim$Megaparcec scales, which are not accessible with the present dataset. Numerous studies have demonstrated that such small-scale environmental effects can be relevant in dense environments, particularly in galaxy groups and clusters. Our results therefore indicate that, on large scales, AGN activity is primarily linked to host-galaxy properties, while environmental effects on smaller scales may still contribute to AGN triggering in specific regimes.

Our results naturally integrate the environmental trends identified in Mountrichas et al. (2019), where clustering amplitude was found to (mildly) increase with $M_\star$ but decrease with both SFR and sSFR. That earlier study demonstrated that dense environments preferentially host massive yet less-star-forming galaxies, implying that large-scale structure regulates the availability of cold gas and the efficiency of star formation. The present findings extend this picture by showing that once galaxies enter an active accretion phase, the large-scale environment no longer determines the level of AGN activity. Instead, the efficiency and luminosity of accretion appear to be internally controlled, while the halo environment mainly influences how often and for how long such phases can occur. Taken together, the two studies suggest an evolutionary continuity: during the star-forming phase, the halo influences gas supply and the likelihood of triggering AGN activity, whereas once accretion begins, the subsequent growth of the black hole and its radiative output are predominantly driven by internal processes within the host.

It is worth emphasising that, once host-galaxy properties such as $M_\star$ are controlled, the range of halo masses probed by AGN samples is expected to be relatively narrow in a statistical sense. This reflects the underlying monotonic relation between $M_\star$ and halo mass established for the general galaxy population, albeit with substantial intrinsic scatter. In this framework, AGN clustering measurements are not designed to identify triggering mechanisms, but rather to test whether different AGN selections preferentially populate distinct regions of the underlying halo-mass distribution beyond what is already set by their hosts. The absence of significant clustering differences at fixed host properties therefore indicates that $M_{\rm BH}$, $\lambda_{\rm Edd,}$ and radiative output do not introduce an additional dependence on halo mass beyond that already encoded by the host galaxy.

Overall, our results point towards a coherent physical picture in which the typical DMH mass of X-ray AGNs is $\sim 10^{13}\,h^{-1}M_\odot$, with no statistically significant variation as a function of $M_{\rm BH}$, $\lambda_{\rm Edd}$, or $L_{\rm X}$ within the uncertainties and parameter ranges probed here. While we do not detect a clear $M_{\rm halo}$–$M_{\rm BH}$ trend, this should not be interpreted as evidence that such a relation is intrinsically absent. Given the relatively narrow dynamic range in $M_{\rm BH}$ explored by our sample, combined with observational uncertainties and intrinsic scatter, such a correlation could plausibly remain hidden in large-scale clustering measurements.

Similarly, the absence of a statistically significant dependence of DMH mass on $\lambda_{\rm Edd}$ suggests that the instantaneous accretion state is not tightly coupled to halo mass on the scales probed here. Instead, large-scale environment appears to regulate AGN activity primarily through its impact on gas supply, star-formation history, and duty cycle, rather than directly controlling short-term accretion efficiency. When considered alongside previous evidence for a strong inverse relation between star-formation activity and clustering strength, our findings support a self-regulated co-evolution framework in which halo mass sets broad boundary conditions for galaxy evolution, while black-hole growth and accretion variability are governed predominantly by baryonic processes operating on galactic and sub-galactic scales.

6. Summary

We investigated the relation between the large-scale environment of X-ray-selected AGNs and fundamental black-hole and accretion properties, combining the wide-area XXL and Stripe 82X surveys. The joint datasets includes more than 400 AGNs with reliable spectroscopic redshifts ($0.5 \lesssim z \lesssim 1.2$) and $M_{\rm BH}$ measurements and about 20 000 non-AGN galaxies. Physical parameters for all sources were derived through consistent SED fitting using the CIGALE code.

To quantify the large-scale environment of AGNs, we measured the AGN–galaxy cross-correlation and galaxy auto-correlation functions and inferred the corresponding DMH masses via standard bias conversion. The analysis was performed separately in bins of X-ray luminosity, black-hole mass, and Eddington ratio. To ensure that any detected trends reflect genuine physical dependencies rather than host or selection biases, we implemented a multivariate nearest-neighbour matching technique. This method pairs AGN subsets while controlling for key covariates such as $M_\star$, SFR, and sSFR (and, when relevant, $M_{\rm BH}$, $\lambda_{\rm Edd,}$ or $L_{\rm X}$), thereby isolating the role of the targeted parameter under fixed host conditions.

The main results can be summarised as follows.

- The typical DMH mass of an X-ray AGN is $\sim 10^{13}, h^{-1}M_\odot$, which is consistent across both *XXL* and *Stripe82X*, indicating that moderate-luminosity AGNs inhabit group-sized DMHs.
- No statistically significant correlation is found between DMH mass and $M_{\rm BH}$ over the $7.5 \lesssim \log(M_{\rm BH}/M_\odot) \lesssim 10$ range, suggesting that black-hole growth is largely decoupled from the large-scale halo environment once galaxies reach the massive, star-forming regime.
- The $\lambda_{\rm Edd}$ shows no measurable correlation with halo mass, indicating that accretion efficiency is largely independent of large-scale environment.
- The relation between halo mass and X-ray luminosity is statistically flat, indicating that instantaneous AGN power is not strongly correlated with the mass of the surrounding DMH.

Taken together, these results support a coherent evolutionary picture in which the environment primarily sets the initial gas supply and the likelihood of triggering nuclear activity, while the subsequent growth and variability of the black hole are governed by processes internal to the host. Once an AGN phase is triggered, feedback and short-term accretion variability dominate over large-scale environmental effects, yielding similar halo masses across the AGN population. This self-regulated framework links galaxy quenching and black-hole fuelling as consecutive manifestations of the same baryon-cycle processes operating across different scales.

Future wide-field surveys will enable decisive progress in this direction. In particular, the eROSITA all-sky survey (Merloni et al. 2024) is providing the largest homogeneous samples of X-ray-selected AGNs to date, extending clustering studies over wide luminosity and redshift ranges. On the spectroscopic side, large-scale programs such as the SDSS-V Black Hole Mapper (Anderson et al. 2023; Aydar et al. 2025), the Dark Energy Spectroscopic Instrument (DESI) survey (Aghamousa 2016), the Subaru Prime Focus Spectrograph (PFS) (e.g. Takada et al. 2014), and the forthcoming 4MOST surveys (e.g. de Jong et al. 2012) will deliver extensive redshift measurements and broad-line AGN classifications over thousands of square degrees.

The combination of these X-ray and spectroscopic datasets will provide the statistical power and uniformity needed to probe the dependence of AGN clustering on black-hole mass, accretion rate, and luminosity with unprecedented precision. Such samples will allow environmental trends to be examined as a continuous function of accretion state and redshift, offering a definitive test of the role of large-scale structure in black-hole–galaxy co-evolution.

Acknowledgements. GM acknowledges funding from grant PID2021-122955OB-C41 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF/EU". This work was partially supported by the European Union's Horizon 2020 Research and Innovation program under the Maria Sklodowska-Curie grant agreement (No. 754510). This publication is part of the R&D&I project PID2024-155779OB-C31, funded by MICIU/AEI/10.13039/501100011033 and co-funded by FEDER, EU. We also acknowledge partial support from the European Union's Horizon 2020 research and innovation pro- gramme under the Marie Skłodowska-Curie grant agreement No 860744 (Bid4BESt; grant coordinator F. Shankar).

References

Aird, J., Coil, A. L., Moustakas, J., et al. 2012, ApJ, 746, 90
Allevato, V., Finoguenov, A., Cappelluti, N., et al. 2011, ApJ, 736, 99
Allevato, V., Finoguenov, A., Civano, F., et al. 2014, ApJ, 796, 4
Allevato, V., Viitanen, A., Finoguenov, A., et al. 2019, A&A, 632, A88
Allevato, V., Miyaji, T., Georgakakis, A., et al. 2021, AJ, 914, 7
Ananna, T. T., Salvato, M., LaMassa, S., et al. 2017, ApJ, 850, 66
Anderson, S., Eracelous, M., Green, P., et al. 2023, Am. Astron. Soc. Meeting Abstr., 241, 301.03
Annis, J., Soares-Santos, M., Strauss, M. A., et al. 2014, ApJ, 794, 120
Aydar, C., Merloni, A., Dwelly, T., et al. 2025, A&A, 698, A132
Booth, C. M., & Schaye, J. 2010, MNRAS, 405, L1
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Bower, R. G., Schaye, J., Furlong, M., et al. 2017, MNRAS, 465, 32
Brandt, W. N., & Hasinger, G. 2005, ARA&A, 43, 827
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buat, V., Ciesla, L., Boquien, M., Małek, K., & Burgarella, D. 2019, A&A, 632, A79
Buat, V., Mountrichas, G., Yang, G., et al. 2021, A&A, 654, A93
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Coil, A. L., Mendez, A. J., Eisenstein, D. J., & Moustakas, J. 2017, ApJ, 838, 87
Comparat, J., Luo, W., Merloni, A., et al. 2023, A&A, 673, A122
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Dale, D. A., Helou, G., Magdis, G. E., et al. 2014, ApJ, 784, 83
Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465
de Jong, R. S., Bellido-Tirado, O., Chiappini, C., et al. 2012, in Ground-based and Airborne Instrumentation for Astronomy IV, eds. I. S. McLean, S. K. Ramsay, & H. Takami (SPIE), 8446, 84460T
DESI Collaboration (Aghamousa, A., et al.) 2016, arXiv e-prints [arXiv:`1611.00036`]
Donoso, E., Li, C., Kauffmann, G., Best, P. N., & Heckman, T. M. 2010, MNRAS, 407, 1078
Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2012, MNRAS, 420, 2662
Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, MNRAS, 419, 2797
Fanidakis, N., Macciò, A. V., Baugh, C. M., Lacey, C. G., & Frenk, C. S. 2013a, MNRAS, 436, 315
Fanidakis, N., Georgakakis, A., Mountrichas, G., et al. 2013b, MNRAS, 435, 679
Garilli, B., Guzzo, L., Scodeggio, M., et al. 2014, A&A, 562, A23
Georgakakis, A., Coil, A. L., Willmer, C. N. A., et al. 2011, MNRAS, 418, 2590
Georgakakis, A., Mountrichas, G., Salvato, M., et al. 2014, MNRAS, 443, 3327
Georgakakis, A., Salvato, M., Liu, Z., et al. 2017, MNRAS, 469, 3232
Georgakakis, A., Comparat, J., Merloni, A., et al. 2019, MNRAS, 487, 275
Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108
Habouzit, M., Onoue, M., Bañados, E., et al. 2022, MNRAS, 511, 3751
Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891
Hirschmann, M., Dolag, K., Saro, A., et al. 2014, MNRAS, 442, 2304
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1
Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, AJSS, 175, 356
Ikeda, H., Miyaji, T., Ueda, Y., et al. 2025, ApJ, 982, 192
Jiang, L., Fan, X., Bian, F., et al. 2014, ApJS, 213, 12
Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, MNRAS, 428, 1382
Koutoulidis, L., Mountrichas, G., Georgantopoulos, I., Pouliasis, E., & Plionis, M. 2022, A&A, 658, A35
Krumpe, M., Lamer, G., Markowitz, A., & Corral, A. 2010a, ApJ, 725, 2444
Krumpe, M., Miyaji, T., & Coil, A. L. 2010b, 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
Krumpe, M., Miyaji, T., Coil, A. L., & Hector, A. 2018, MNRAS, 474, 1773
Krumpe, M., Miyaji, T., Georgakakis, A., et al. 2023, ApJ, 952, 109
LaMassa, S. M., Cales, S., Moran, E. C., et al. 2013a, MNRAS, 432, 135
LaMassa, S. M., Megan Urry, C., Cappelluti, N., et al. 2013b, MNRAS, 436, 3581
LaMassa, S. M., Cales, S., Moran, E. C., et al. 2015, ApJ, 800, 144
LaMassa, S. M., Megan Urry, C., Cappelluti, N., et al. 2016, ApJ, 817, 21
LaMassa, S. M., Georgakakis, A., Vivek, M., et al. 2019, ApJ, 876, 50
LaMassa, S., Peca, A., Urry, C. M., et al. 2024, ApJ, 974, 235
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, eds. M. Iye, & A. F. M. Moorwood, SPIE Conf. Ser., 4841, 1670
Leauthaud, A., Benson, A. J., Civano, F., et al. 2015, MNRAS, 446, 1874
Li, Q., Kilbinger, M., Luo, W., et al. 2024, ApJ, 969, L25
Liu, Z., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 459, 1602
Małek, K., Buat, V., Roehlly, Y., et al. 2018, A&A, 620, A50
Masoura, V. A., Mountrichas, G., Georgantopoulos, I., et al. 2018, A&A, 618, A31
Masoura, V. A., Mountrichas, G., Georgantopoulos, I., & Plionis, M. 2021, A&A, 646, A167
Menci, N., Fiore, F., Shankar, F., Zanisi, L., & Feruglio, C. 2023, A&A, 674, A181
Mendez, A. J., Coil, A. L., Aird, J., et al. 2016, ApJ, 821, 55
Menzel, M.-L., Merloni, A., Georgakakis, A., et al. 2016, MNRAS, 457, 110
Merloni, A., Liu, T., Predehl, P., et al. 2024, A&A, 682, A34
Miyaji, T., Krumpe, M., Coil, A. L., & Aceves, H. 2011, ApJ, 726, 83
Mountrichas, G. 2023, A&A, 672, A98
Mountrichas, G., & Buat, V. 2023, A&A, 679, A151
Mountrichas, G., & Georgakakis, A. 2012, MNRAS, 420, 514
Mountrichas, G., & Georgantopoulos, I. 2024, A&A, 683, A160
Mountrichas, G., & Shankar, F. 2023, MNRAS, 518, 2088
Mountrichas, G., Sawangwit, U., Shanks, T., et al. 2009a, MNRAS, 394, 2050
Mountrichas, G., Sawangwit, U., & Shanks, T. 2009b, MNRAS, 398, 971
Mountrichas, G., Georgakakis, A., Finoguenov, A., et al. 2013, MNRAS, 430, 661
Mountrichas, G., Georgakakis, A., Menzel, M.-L., et al. 2016, MNRAS, 457, 4195
Mountrichas, G., Georgakakis, A., & Georgantopoulos, I. 2019, MNRAS, 483, 1374
Mountrichas, G., Buat, V., Georgantopoulos, I., et al. 2021a, A&A, 653, A70
Mountrichas, G., Buat, V., Yang, G., et al. 2021b, A&A, 653, A74
Mountrichas, G., Buat, V., Yang, G., et al. 2022a, A&A, 663, A130
Mountrichas, G., Buat, V., Yang, G., et al. 2022b, A&A, 667, A145
Mountrichas, G., Masoura, V. A., Xilouris, E. M., et al. 2022c, A&A, 661, A108
Mountrichas, G., Yang, G., Buat, V., et al. 2023, A&A, 675, A137
Mountrichas, G., Masoura, V. A., Corral, A., & Carrera, F. J. 2024a, A&A, 683, A143
Mountrichas, G., Siudek, M., & Cucciati, O. 2024b, A&A, 686, A229
Mountrichas, G., Carrera, F. J., Georgantopoulos, I., et al. 2025, A&A, 700, A234
Peca, A., Cappelluti, N., Urry, C. M., et al. 2023, ApJ, 943, 162
Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1

Planck Collaboration VI. 2020, A&A, 641, A6
Pouliasis, E., Mountrichas, G., Georgantopoulos, I., et al. 2020, MNRAS, 495, 1853
Pouliasis, E., Mountrichas, G., Georgantopoulos, I., et al. 2022, A&A, 667, A56
Powell, M. C., Cappelluti, N., Urry, C. M., et al. 2018, ApJ, 858, 110
Powell, M. C., Urry, C. M., Cappelluti, N., et al. 2020, ApJ, 891, 41
Ross, N. P., Shanks, T., Cannon, R. D., et al. 2008, MNRAS, 387, 1323
Scodeggio, M., Guzzo, L., Garilli, B., et al. 2018, A&A, 609, A84
Shankar, F., Weinberg, D. H., & Shen, Y. 2010, MNRAS, 1959
Shankar, F., Weinberg, D. H., & Miralda-Escudé, J. 2013, MNRAS, 428, 421
Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, MNRAS, 460, 3119
Shankar, F., Sonnenfeld, A., Mamon, G. A., et al. 2017, ApJ, 840, 34
Shankar, F., Allevato, V., Bernardi, M., et al. 2020, Nat. Astron., 4, 282
Shankar, F., Bernardi, M., Roberts, D., et al. 2025, MNRAS, 530, 2174
Shen, Y., & Liu, X. 2012, ApJ, 753, 125
Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, ApJ, 680, 169
Shen, Y., Michael, A. S., Nicholas, P., et al. 2009, ApJ, 697, 1656
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
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
Shirasaki, M., Yoshida, N., Ikeda, H., Nishizawa, A. J., & Suto, Y. 2016, MNRAS, 459, 432
Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popović, L. Č. 2012, MNRAS, 420, 2756
Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288
Sutherland, W., & Saunders, W. 1992, MNRAS, 259, 413
Takada, M., Ellis, R., Chiba, M., et al. 2014, PASJ, 66, R1
Terrazas, B. A., Bell, E. F., Pillepich, A., et al. 2020, MNRAS, 493, 1888
Tinker, J. L., Weinberg, D. H., & Zheng, Z. 2005, ApJ, 631, 41
Van den Bosch, F. C. 2002, MNRAS, 331, 98
Viero, M. P., Asboth, V., Roseboom, I. G., et al. 2014, ApJS, 210, 9
Viitanen, A., Allevato, V., Finoguenov, A., et al. 2019, A&A, 629, A14
Viitanen, A., Allevato, V., Finoguenov, A., et al. 2023, A&A, 674, A214
Villa-Velez, J. A., Buat, V., Theule, P., Boquien, M., & Burgarella, D. 2021, A&A, 654, A153
Weinberger, R., Springel, V., Pakmor, R., et al. 2018, MNRAS, 479, 4056

Appendix A: Validation of the angular selection

As an empirical validation of the assumptions underlying the AGN–galaxy cross-correlation approach, we test whether the angular clustering of the AGN sample is consistent with that expected for random subsamples of the galaxy catalogue once the survey mask is taken into account. We compute the AGN angular two-point correlation function, $w_{\rm AGN}(\theta)$, using the galaxy random catalogue to describe the angular selection function. We then draw N random subsamples from the galaxy catalogue, each containing the same number of objects as the AGN sample and matched to its redshift distribution in coarse bins, and measure $w(\theta)$ for each subsample using the same random catalogue.

Figure A.1 compares $w_{\rm AGN}(\theta)$ to the distribution of $w(\theta)$ obtained from the galaxy subsamples. The AGN measurements lie within the 16–84 per cent range of the galaxy-subsample distribution over the angular scales corresponding to the two-halo regime probed in this work. This indicates that, once the survey mask is accounted for, the angular selection of the AGN sample does not introduce a detectable large-scale systematic in the inferred clustering amplitude. We stress that this test does not demonstrate that the AGN and galaxy selection functions are identical; rather, it provides empirical support for the validity of using the galaxy random catalogue in the AGN–galaxy cross-correlation analysis on the large scales relevant for this study.

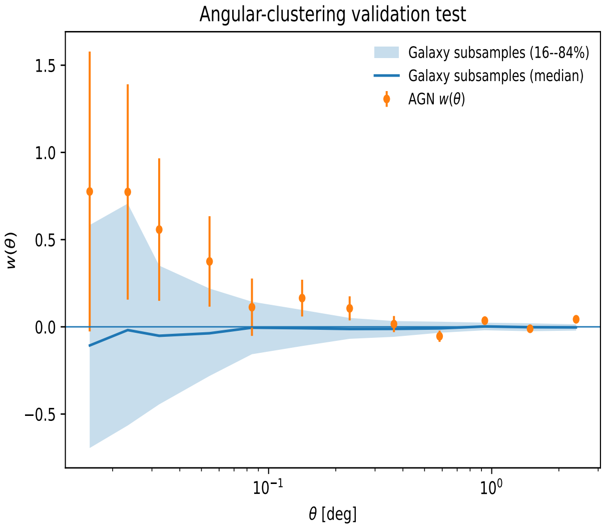


Fig. A.1. Angular correlation functions. Comparison between the AGN angular correlation function, $w_{\rm AGN}(\theta)$, and the distribution of $w(\theta)$ measured from random galaxy subsamples of equal size and matched redshift distribution. The shaded region shows the 16–84 per cent range and the solid line the median of the galaxy subsamples. The AGN measurements are consistent with this distribution on the angular scales relevant for the clustering analysis.