The Astrophysical Journal, 977:158 (23pp), 2024 December 20 https://doi.org/10.3847/1538-4357/ad8342

Cross-correlation of Luminous Red Galaxies with Machine Learning Selected Active Galactic Nuclei in HSC-SSP: Unobscured AGN Residing in More Massive Halos

Rodrigo Córdova Rosado[1], Andy D. Goulding[1], Jenny E. Greene[1], Grayson C. Petter[2], Ryan C. Hickox[2], Nickolas Kokron[1], Michael A. Strauss[1], Jahmour J. Givans[1,3], Yoshiki Toba[4,5,6], and Cassandra Starr Henderson[7]

[1] Department of Astrophysical Sciences, Peyton Hall, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA; rodrigoc@princeton.edu
[2] Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, NH 03755, USA
[3] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[4] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[5] Academia Sinica Institute of Astronomy and Astrophysics, 11F of Astronomy-Mathematics Building, AS/NTU, No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan
[6] Research Center for Space and Cosmic Evolution, Ehime University, 2-5 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan
[7] Scripps Institution of Oceanography, La Jolla, CA, USA

Received 2024 May 22; revised 2024 September 25; accepted 2024 October 2; published 2024 December 10

Abstract

Active galactic nuclei (AGN) are the signposts of black hole growth, and likely play an important role in galaxy evolution. An outstanding question is whether AGN of different spectral types indicate different evolutionary stages in the coevolution of black holes and galaxies. We present the angular correlation function between an AGN sample selected from Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) optical photometry and Wide-field Infrared Survey Explorer mid-IR photometry and a luminous red galaxy (LRG) sample from HSC-SSP. We investigate AGN clustering strength as a function of luminosity and spectral features across three independent HSC fields totaling ∼600 deg^2, for $z \in 0.6-1.2$ and AGN with $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}. There are ∼28,500 AGN and ∼1.5 million LRGs in our primary analysis. We determine the average halo mass for the full AGN sample ($M_h \approx 10^{12.9}\,h^{-1}M_\odot$), and note that it does not evolve significantly as a function of redshift (over this narrow range) or luminosity. We find that, on average, unobscured AGN ($M_h \approx 10^{13.3}\,h^{-1}M_\odot$) occupy ∼4.5× more massive halos than obscured AGN ($M_h \approx 10^{12.6}\,h^{-1}M_\odot$), at 5σ statistical significance using 1D uncertainties, and at 3σ using the full covariance matrix, suggesting a physical difference between unobscured and obscured AGN, beyond the line-of-sight viewing angle. Furthermore, we find evidence for a halo mass dependence on reddening level within the Type I AGN population, which could support the existence of a dust-obscured phase. However, we also find that quite small systematic shifts in the redshift distributions of the AGN sample could explain current and previously observed differences in M_h.

Unified Astronomy Thesaurus concepts: Active galactic nuclei (16); Galaxy evolution (594); Quasars (1319); Supermassive black holes (1663)

1. Introduction

Supermassive black holes (SMBHs) influence the growth and evolution of the galaxies in which they reside (J. Kormendy & D. Richstone 1995; J. Kormendy & L. C. Ho 2013). Periods of rapid mass accretion onto SMBHs, creating an observable active galactic nucleus (AGN; M. Schmidt 1963), provide a unique opportunity to study their attributes. Tracing the black hole (BH) and galaxy coevolution is critical to describing the role of BH and AGN feedback in galaxy growth (J. Kormendy & L. C. Ho 2013; T. M. Heckman & P. N. Best 2014).

Historically, AGN have been classified into two classes, unobscured (Type I) and obscured (Type II). In strict unification models, all AGN are identical but are seen from different angles. Dusty flattened regions on parsec scales, known as the torus, act as a screen at particular inclinations (R. Antonucci 1993; C. M. Urry & P. Padovani 1995; H. Netzer 2015). However, there have been significant results that challenge a strict unified model of AGN structure, showing that the observed obscuration level is correlated with the evolutionary stage of the host galaxy and obscuration effects (D. B. Sanders et al. 1988; R. C. Hickox et al. 2011; V. Allevato et al. 2014; S. L. Ellison et al. 2019; V. A. Fawcett et al. 2023). There have been hints that observed AGN spectral properties are tied to merger history (see R. C. Hickox & D. M. Alexander 2018 for a recent review). Additional studies are needed to probe whether obscured and unobscured AGN are objects along different points in an evolutionary track (P. F. Hopkins et al. 2008; R. C. Hickox et al. 2009; N. Cappelluti et al. 2012).

Several observational results have shown that obscured AGN are often found in mergers, suggesting that there is a link with AGN activity (J. C. Mihos & L. Hernquist 1994, 1996; A. W. Blain et al. 1999; T. Urrutia et al. 2008; M. Koss et al. 2010; S. L. Ellison et al. 2011, 2013, 2019; E. Glikman et al. 2015; A. D. Goulding et al. 2018; N. J. Secrest et al. 2020; C. Ricci et al. 2021). There is also some evidence that reddening in AGN is significantly impacted by host galaxy properties (A. D. Goulding & D. M. Alexander 2009; A. D. Goulding et al. 2012). Meanwhile, unobscured AGN may be a later stage of the merger evolutionary scenario, once winds and outflows (the potential "blowout" phase) have revealed the AGN broad-line region (P. F. Hopkins et al. 2008; see C. R. Almeida & C. Ricci 2017 for a recent review of AGN obscuration morphology).

Analyzing these trends in single objects is difficult because the timescales for star formation and mergers are much longer than the AGN variability timescales (R. C. Hickox et al. 2014). However, by measuring the clustering of a statistical sample of galaxies, we can infer the properties of the dark matter (DM) halos in which they reside. Clustering measurements from large-scale structure analyses have thus become a preeminent tool to study ensemble properties across samples of AGN and other galaxy populations (P. S. Osmer 1981; P. A. Shaver 1984; T. Shanks et al. 1987; A. Iovino & P. A. Shaver 1988; P. Andreani & S. Cristiani 1992; H. J. Mo & L. Z. Fang 1993; T. Shanks & B. J. Boyle 1994; S. M. Croom & T. Shanks 1996; F. La Franca et al. 1998; S. M. Croom et al. 2005; A. Lidz et al. 2006; Y. Shen et al. 2008; Y. Toba et al. 2017; W. He et al. 2018; J. Arita et al. 2023). They are particularly useful for studying the relationship of galaxies and the halos they are found in, and provide important clues at the population level.

Studies find that luminous AGN often reside in DM halos of typical mass $\log M_h \sim 12.5 - 13 \log h^{-1} M_\odot$ (N. Cappelluti et al. 2012; J. D. Timlin et al. 2018). These analyses also probe the possible halo mass dependence of observed AGN luminosities, and how it relates to BH accretion properties (e.g., S. M. Croom et al. 2005; A. Lidz et al. 2006; L. Koutoulidis et al. 2013; Y. Shen et al. 2013; A. J. Mendez et al. 2016; M. Krumpe et al. 2018; J. D. Timlin et al. 2018).

One way to test the idea that AGN move from a buried (and obscured) growth phase to an unobscured state through blowout could be to find differences in clustering strength between obscured and unobscured AGN. This would also be a test of the pure orientation model. Simple unification cannot explain the observed differences in clustering amplitude between obscured and unobscured AGN that have already been reported (R. C. Hickox et al. 2011; V. Allevato et al. 2014; M. A. DiPompeo et al. 2014, 2016, 2017; N. Jiang et al. 2016; L. Koutoulidis et al. 2018; M. C. Powell et al. 2018; G. C. Petter et al. 2023; Q. Li et al. 2024). Unfortunately, different studies come to different conclusions about whether obscured or unobscured AGN are found in more massive halos. The results depend on the selection method, luminosities, and redshift range of the samples. On the other hand, many of the studies are based on relatively small areas and only hundreds of AGN (see A. L. Coil et al. 2009; R. Gilli et al. 2009; N. Cappelluti et al. 2010; V. Allevato et al. 2011; L. Koutoulidis et al. 2013; M. Krumpe et al. 2018).

In this work, we implement the AGN selection by A. D. Goulding et al. (2024, in preparation), which aims to represent a more complete sampling of the AGN color space. Using an optical and mid-infrared (MIR) color selection based on unsupervised machine learning classification, we construct unobscured and obscured AGN samples to investigate the clustering strength of these populations. Increasing the source number density will also reduce potential biases in AGN subtype samples, and allows for more precise measurements of the clustering on all relevant scales. We carry out our analysis using the cross-correlation function between the Hyper Suprime-Cam (HSC) photometrically selected galaxy sample and our AGN samples, as opposed to the more common AGN autocorrelation. Cross-correlations have the benefit of being less sensitive to systematic uncertainties that are not shared between both samples, and in this case provide a well-understood galaxy sample with which to compare the AGN. They also have higher signal-to-noise ratios (S/N) for sparse samples like AGN. We make use of the individual redshift distribution for each object in our sample, from either photometric or spectroscopic measurements.

This paper is organized as follows. In Section 2, we summarize the data sets used in this analysis, and the subdivisions of the AGN sample. In Section 3, we outline our methodologies for the projected angular correlation function calculation, uncertainty estimation, and parameter fitting and interpretation. We present the results of our autocorrelation of HSC galaxies, the cross-correlation with the full AGN sample, and the cross-correlations with the AGN subtype samples in Section 4. We discuss our results in Section 5, and conclude in Section 6.

Throughout this analysis, we adopt a ΛCDM "Planck 2018"-like cosmology (Planck Collaboration et al. 2020), with $h = H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}} = 0.67$, $\Omega_c = 0.27$, $\Omega_b = 0.045$, $n_s = 0.96$, and $\sigma_8 = 0.83$. Quantities expressed with a log are exclusively $\log_{10}$ values. All magnitudes are in the AB system (J. B. Oke & J. E. Gunn 1983), unless otherwise noted. In the context of galaxy bias and halo mass parameterization, we make use of the J. L. Tinker et al. (2010) formalism with $\Delta = 200$ (the spherical overdensity radius definition). Foreground dust extinction corrections are applied to all magnitudes as supplied in the HSC catalog (H. Aihara et al. 2022) based on D. J. Schlegel et al. (1998).

2. Data

2.1. HSC Survey

HSC is a wide-field prime-focus camera mounted on the 8.2 m Subaru Telescope, located atop Maunakea, Hawaii (S. Miyazaki et al. 2018). The HSC Subaru Strategic Program (SSP; H. Aihara et al. 2018) is designed to make the most of the $1.77\,\mathrm{deg}^2$ field of view by using 330 nights on Subaru to explore the full range of galactic history from the present to $z \sim 7$ across three imaging layers, each with specific scientific goals. The Wide, Deep, and Ultradeep surveys consist of different sky coverage and exposure times in the *grizy* wideband filters, in combination with survey-specific narrowband filters (S. Kawanomoto et al. 2018). We refer the reader to the most recent (PDR3) release of HSC-SSP data as described in H. Aihara et al. (2022). We make use of the $670\,\mathrm{deg}^2$ full-depth full-color Wide imaging within PDR3 to constrain the properties of galaxies positioned across ~2% of the sky. The sensitivity is limited to the $i \lesssim 26$ band, with the point-spread function (PSF) at a 5σ point-source depth. The Wide survey is split between four disjoint fields with varying coverage, two equatorial strips for spring and fall, as well as the higher-decl. field of HECTOMAP and the AEGIS calibration field.

Galaxy photometric redshifts (photo-z) are measured from the spectral energy distribution (SED) of each object, following the methods outlined in M. Tanaka et al. (2018) and the template-fitting algorithm `Mizuki` (M. Tanaka 2015). This procedure recovers $p(z)$, the redshift probability distribution, for each object. We also have spectroscopic redshifts (spec-z) for a small fraction (<2%) of objects, but these are not distributed homogeneously over the HSC fields. In HSC PDR3, a wide array of publicly available spectroscopic catalogs are combined to maximize the number of spectroscopic redshifts for observed sources, as detailed in Section 4.1 of H. Aihara et al. (2022). We model the $p(z)$ for the spec-z as a narrow Gaussian distribution, with $\sigma = 0.01 \times (1 + z)$ (wider than typical spec-z modeling to overcome numerical limitations in our halo model clustering tools).

2.2. Masking

The HSC PDR3 data have a bright-star mask already applied to them, as described in detail by J. Coupon et al. (2018) and H. Aihara et al. (2022). They used the Gaia DR2 bright-star catalog (Gaia Collaboration et al. 2018) to identify bright sources and remove affected sky regions from the total survey area. These subtractions were calculated from constraining the extent of stellar halos (the extended PSF of a bright source), ghosts (optical reflections inside the camera that are displayed in the recorded image), blooming (electron spillover into neighboring pixels), and channel stops (diffraction patterns perpendicular to blooming) in HSC imaging. This source mask is already applied to our HSC galaxy catalog, and we apply it to the AGN catalog used here.

Details of the source mask constructed based on the Wide-field Infrared Survey Explorer (WISE) imaging data (E. L. Wright et al. 2010; R. M. Cutri et al. 2012) can be found in A. D. Goulding et al. (2024, in preparation), but we provide a brief summary here. Using the WISE catalog, we select all objects with CC_FLAGS=“H `or` D `or` X `or` P”. These catalog objects are those identified from the pipeline to be affected by data issues. We use the distribution of these objects to create a 2D sky density map in order to flag regions of the sky with potentially spurious and/or inaccurately measured sources, removing $\sim$56 deg^2. These regions are identified as $\geqslant 7\sigma$ overdensities (relative to the average number density of the field) on a map smoothed with a $0\overset{\circ}{.}5$ boxcar kernel. In addition to this flag map, we identify regions of striping within the remaining unmasked objects caused by significantly deeper data, moonlight contamination, or additional artifacts. These extremely high-density regions are also excluded as part of the final source mask.

Additionally, we remove point sources at the catalog level that are not included in the bright-star mask by identifying and excluding objects whose `CModel` magnitudes differ by $\leqslant$0.06 in g, r, i, or z from the PSF magnitude estimate. This difference reflects how consistent the PSF model for a point source is with the more robust `CModel` magnitude for an extended object (J. Bosch et al. 2018). Not surprisingly, the excluded objects follow the stellar locus in color–color space.

2.3. HSC Galaxy Sample

We use a magnitude-limited galaxy sample ($i < 24$) from the HSC-SSP Wide survey, which will be cross-correlated with our (HSC-derived) AGN sample. We perform the experiment on the three largest HSC PDR3 equatorial fields with full filter coverage, XMM-LSS (hereafter referred to as XMM), VVDS, and GAMA. These fields encompass 2.3×10^7 galaxies in our analysis redshift range ($z \in 0.6$–1.2, see Section 3.3), after applying all relevant survey masks (see Section 2.2). The mean number density of galaxies is $\sim$40,000 deg^{-2}, for a total area of 568.55 deg^2 (post-masking). Using either spectroscopic or photometric information, we recover the $p(z)$ for each source.

Prior HSC analyses have selected samples with smaller photo-z uncertainty in order to perform statistical measurements, and found the most effective means of doing so is to use a luminous red galaxy (LRG) sample (e.g., M. M. Rau et al. 2023). Additionally, the strongly clustered LRGs provide a high-S/N comparison for cross-correlations. We similarly find it necessary to isolate the LRG population in our galaxy sample to reduce the $p(z)$ uncertainty in our sample, given tests of the galaxy autocorrelation analysis for the complete ($i < 24$) HSC galaxy sample. We select the LRG population in the HSC sample with an optical color–color cut ($g - r > 1.2$ and $r - i > 1.0$), and use these objects throughout our clustering analysis cross-correlations. See Appendix A for further details of the LRG selection and preliminary autocorrelation analysis. In total, there are 1.6×10^6 LRGs in our analysis, with an average number density of $\sim$2900 deg^{-2}. Figure 1 illustrates the LRG population and its density across the HSC fields in our analysis and Table 1 gives the number of objects in each field.

2.4. WISE- and HSC-selected AGN Sample

We present an AGN sample that combines optical and MIR photometry from HSC-SSP and WISE (E. L. Wright et al. 2010). The details of this method are explained in A. D. Goulding et al. (2024, in preparation), and we give a high-level overview here. Using a maximum likelihood estimator, HSC *grizy* photometry is matched to sources detected with $\mathrm{S/N} > 5$ in their W1 photometry in the allWISE (E. L. Wright et al. 2019) and unWISE (UnWISE Team 2021) catalogs (A. Mainzer et al. 2011; E. F. Schlafly et al. 2019; R. M. Cutri et al. 2021). After combining the source catalogs, we additionally require sources to have $\mathrm{S/N} > 4$, 3, and 3 in their g, W2, and W3 photometry.

Utilizing an unsupervised dimensionality reduction technique, the uniform manifold approximation and projection (UMAP) algorithm (L. McInnes et al. 2018), we distill the multidimensional color, magnitude, and source size space down to a single interpretable 2D manifold. UMAP incorporates the input multidimensional photometric information from HSC and WISE to construct a neural network that identifies objects with similar properties, grouping them together, while simultaneously placing objects with dissimilar properties far apart. The result is a simple 2D space in which we can identify the region occupied by AGN. Inputting a test sample with known labels after the training process has been completed, but providing no prior knowledge of the intrinsic source properties to UMAP, A. D. Goulding et al. (2024, in preparation) show that UMAP segregates known stars from galaxies, and known obscured or unobscured AGN/quasars from inactive galaxies. The final UMAP manifold contains 16 distinct clusters, two of which are dominated by AGN. The AGN nature of these clusters is further validated via spectroscopic follow-up by R. E. Hviding et al. (2024), who find the UMAP AGN classification matches the spectral categorization with emission-line fitting and continuum ratios indicative of AGN activity. These UMAP AGN subtype classifications are not part of the principal AGN identification detailed in Section 2.4.1. We use this step to solely identify the AGN.

A. D. Goulding et al. (2024, in preparation) further provide photometric redshift distributions and spectroscopic redshifts (where available) for all of the UMAP-classified AGN. These photometric redshifts are determined by utilizing the full g through W3 photometry to train an augmented random forest algorithm, which they show significantly outperforms the `Mizuki` photo-z that are primarily designed for inactive galaxies. These random forest based photometric redshifts perform equally well on both Type I and Type II AGN out to $z \sim 3$ with an average precision of $\delta z/(1 + z) \sim 0.02$ and 0.03, respectively.

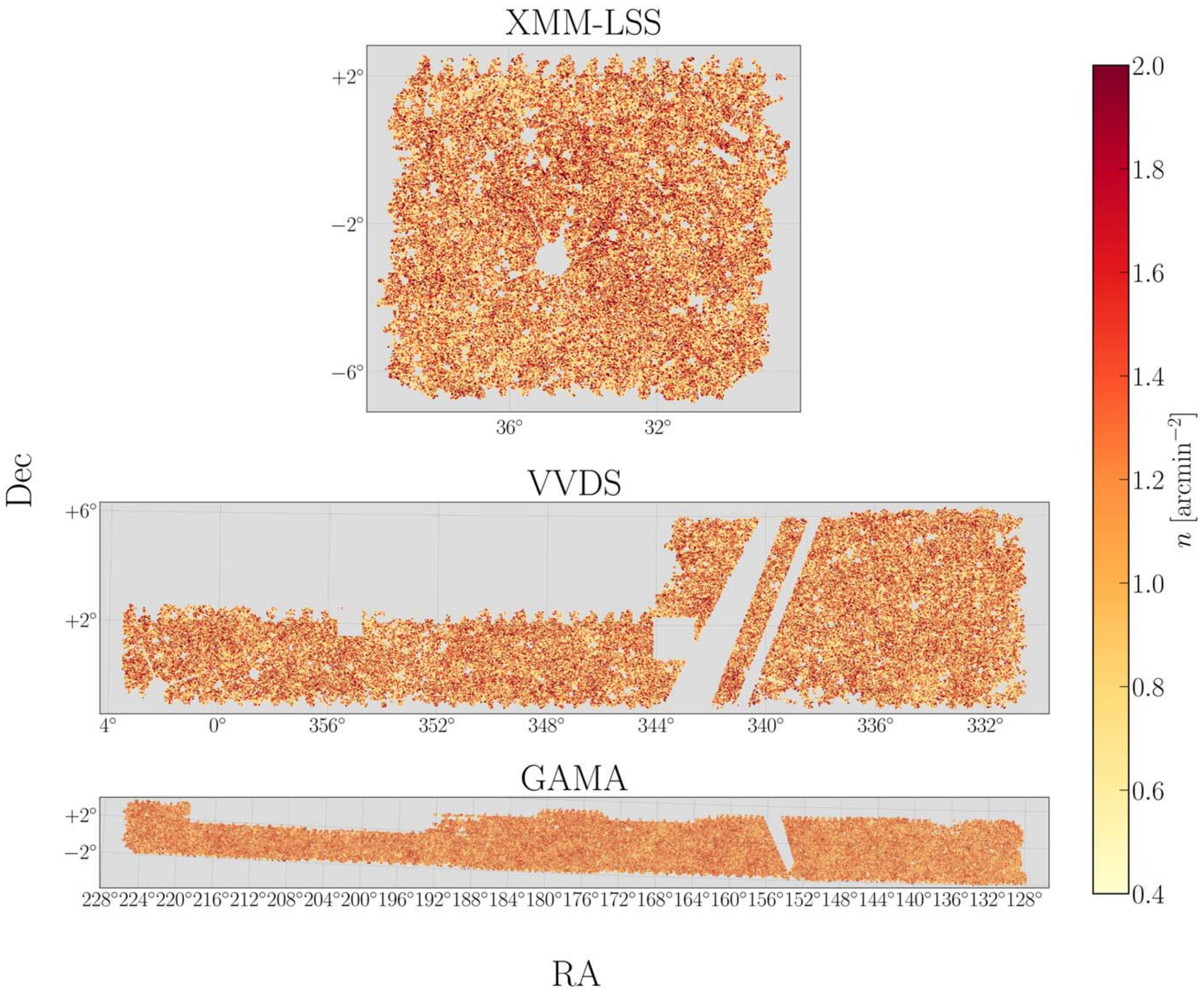

Figure 1. The HSC galaxy sample used in this analysis after isolating the LRG population with an optical color–color cut, split by HSC field. The fields (and approximate central R.A., decl. coordinates) are XMM-LSS (34°, −2°), VVDS (346°, 2°), and GAMA (176°, 0°). The bespoke (see Section 2.2) masking strategy, visualized here by stripes that have been removed from the HSC fields, is reflective of areas where there are correlated photometry errors in the WISE data. Other removed areas are primarily due to bright stars and Galactic foreground objects. This figure is produced with the plotting tool `skymapper` (P. Melchior 2021).

Table 1
Field Properties and Number of Objects in Our Redshift Range ($z \in 0.6$–1.2)

	GAMA	VVDS	XMM
Area [deg^2]	397.18	100.95	70.42
$N_{\rm obj}$ (LRGs)	1,165,141	303,143	204,849
	(15,444)[b]	(11,469)[b]	(16,396)[b]
$N_{\rm obj}$ (AGN)[a]	23,244 (1981)[b]	5447 (1632)[b]	5457 (1667)[b]
$N_{\rm obj}$ (unobscured AGN)[a]	9197 (1506)[b]	1446 (1074)[b]	1466 (1018)[b]
$N_{\rm obj}$ (reddened AGN)[a]	5957 (335)[b]	1150 (392)[b]	1129 (364)[b]
$N_{\rm obj}$ (obscured AGN)[a]	7536 (140)[b]	2774 (166)[b]	2803 (285)[b]

Notes.
[a] Luminous AGN sample ($L_{6\ \mu{\rm m}} > 3 \times 10^{44}$ erg s^{-1}).
[b] Number of objects with a measured spectroscopic redshift indicated in parentheses.

The total redshift distributions for the AGN samples from the photometric and spectroscopic measurements are illustrated in panel (a) of Figure 2. We account for each object's full $p(z)$ in the clustering measurements, as described in Appendix B.

2.4.1. AGN Subsamples

Accurate AGN classification is paramount to appropriately estimating the objects' properties. Previously, AGN subtypes have been split based on an optical–MIR color, such as $r-$ W2 (see G. C. Petter et al. 2023). However, a more robust classification scheme should consider the SED evolution with redshift, and empirically derive the minima in the AGN color–redshift space to sort the subtypes, which we now proceed to do.

From the subset of UMAP-classified AGN with spectroscopy available from the Sloan Digital Sky Survey (SDSS), A. D. Goulding et al. (2024, in preparation) train a k-nearest neighbor (KNN) algorithm to probabilistically label these AGN as unobscured, reddened, or obscured, based on their position in $g-$ W3 versus redshift space (the outputs of these distributions are shown in Figure 2). The unobscured and reddened training samples are spectroscopically characterized by broad emission lines, but the reddened AGN have significant dust obscuration and thus a red continuum in the rest-frame optical. We will typically combine these in our

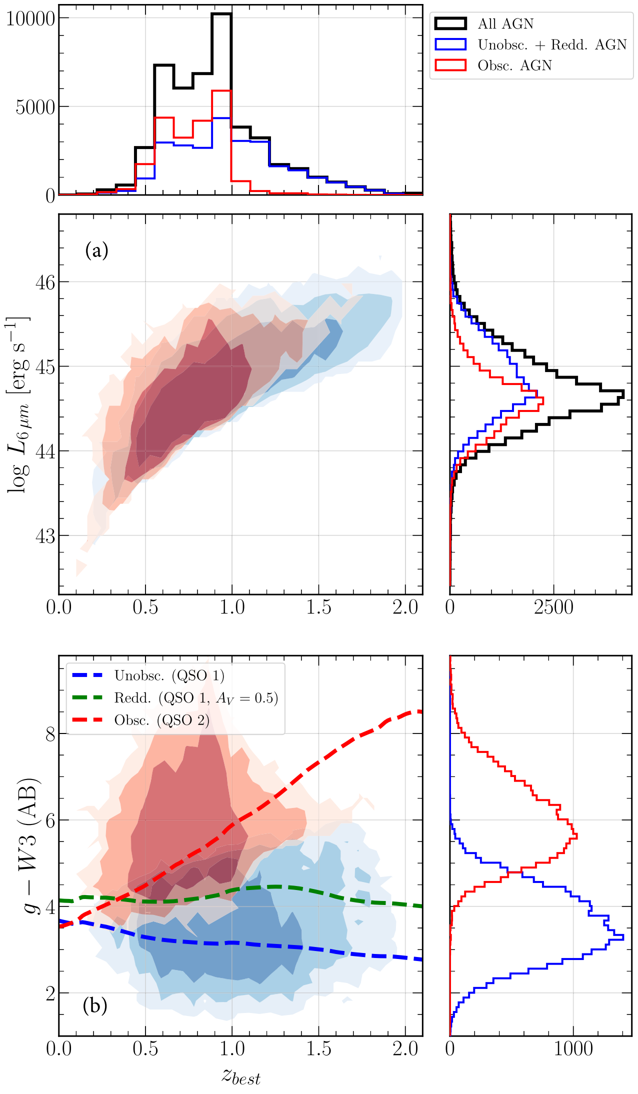

Figure 2. (a) The luminosity distribution of the unobscured + reddened (blue contours, objects we infer to have broad emission lines) and obscured (red contours, objects inferred to have narrow emission lines) AGN samples as a function of the fiducial redshift ($z_{\rm best}$) for each object. The contours are set at the (0.25, 0.5, 0.75) quantiles of each distribution. (b) The $g - {\rm W3}$ (AB magnitude) color vs. the $z_{\rm best}$ distribution for unobscured + reddened and obscured AGN, following the same color scheme and contour quantile levels in the upper panel. We divide the AGN sample into unobscured, reddened, and obscured classes based on their distribution in this color–redshift space (A. D. Goulding et al. 2024, in preparation). We convert the WISE magnitudes in the Vega system to AB with $m_{\rm AB} = m_{\rm Vega} + \Delta m$, where $\Delta m({\rm W3}) = 5.174$ (T. H. Jarrett et al. 2011). The dashed lines are the $g - {\rm W3}$ color–redshift evolution for three AGN templates from M. Polletta et al. (2007). We use their unobscured quasar template (QSO 1, blue line), their unobscured quasar with a B. T. Draine (2003) extinction law applied with $A_V = 0.5$ and $R_V = 3.1$ (to create a reddened quasar template; green line), and their obscured quasar template (QSO 2, red line). Histograms show the individual distributions for a given quantity, where colors indicate the AGN sample specified.

analysis as an unobscured + reddened sample. Meanwhile, the obscured AGN training sample is made up of objects that are both heavily dust-obscured and have only narrow emission lines. This process recovers ∼48,000 AGN in our redshift range (i.e., at least 3% of the $p(z)$ is in $z \in 0.6$–1.2), across the three HSC fields used here. Of these, ∼15,000 objects are unobscured AGN, ∼11,000 are reddened objects, ∼21,000 are obscured AGN, and <1% of objects do not have a spectral type assigned. Probabilistic estimates of AGN/galaxy classifications using machine learning methodologies that encompass a wide range of source properties inherently come with caveats, but on the whole are deemed to be more complete and reliable than standard singular or 2D demarcations. Hence, we choose to exploit the available data to its reasonable limits to classify our AGN.

The $g - {\rm W3}$ distribution of the unobscured + reddened and obscured objects in our analysis as a function of redshift is shown in panel (b) of Figure 2. The redshift evolution for the $g - {\rm W3}$ color of SWIRE library templates (M. Polletta et al. 2007) for the three types of AGN is shown as dashed lines in panel (b) of Figure 2. We will investigate the clustering amplitude of the sets of unobscured + reddened and obscured AGN, as well as the difference between unobscured and reddened AGN.

We note that these color cuts are a function of redshift, given the minima of the KNN probability density distribution, and are not like previously defined selection functions for AGN types using only optical and MIR colors (e.g., D. Stern et al. 2005, 2012; R. C. Hickox et al. 2009; R. J. Assef et al. 2013; R. E. Hviding et al. 2022). These new classifications recover a significant fraction of spectroscopically confirmed ${\rm W1} - {\rm W2} < 0.8$ AGN (below the D. Stern et al. 2012 limit), generating a wider (and more complete) sampling of the unobscured + reddened and obscured AGN color space for our analysis (AGN spectra from SDSS DR16; R. Ahumada et al. 2020). From follow-up spectra of 178 objects in our sample with $i < 22.5$, R. E. Hviding et al. (2024) confirm the classification for 85% of randomly selected unobscured + reddened AGN, and similarly 65% of the obscured AGN sample, finding a contamination rate of 3% and 15%, respectively, and bolstering our confidence in the classification.

We present the subsample redshift and luminosity distributions in Figure 2. We note the excess of unobscured + reddened objects at high redshift, which are driving the excess of high-luminosity AGN in the $L_{6\,\mu{\rm m}}$ plot (see panel (a) of Figure 2). We obtain rest-frame $L_{6\,\mu{\rm m}}$ luminosity measurements for each AGN in the sample via the standard power-law fitting to the MIR photometry from WISE. In order to compare the clustering across AGN subtypes, we divide the analysis into a series of redshift bins (Section 3.3). Their luminosity distributions are illustrated in Figures 3 and 4. We establish a lower limit for the AGN luminosity at $L_{6\,\mu{\rm m}} > 3 \times 10^{44}$ erg s^{-1}, close to the peak of the distribution, where we are confident that we are complete out to $z \sim 1.2$ (see Figure 2). This has the added benefit of allowing an analysis comparing AGN subtypes where their luminosity distribution is more consistent: above our threshold. This results in a total of 34,144 luminous AGN in our redshift range that we will use in our analysis. The number of luminous AGN (and their spectral classes) in each field is shown in Table 1, including the number of objects for which we have spectroscopic redshifts.

Additionally, we define a low- and high-$L_{6\,\mu{\rm m}}$ range to investigate if there is any evidence of an inferred halo mass trend with luminosity. These additional cuts are set at $3 \times 10^{44} < L_{6\,\mu{\rm m}} < 10^{45}$ erg s^{-1} for the low-$L_{6\,\mu{\rm m}}$ sample, and at $L_{6\,\mu{\rm m}} > 10^{45}$ erg s^{-1} for the high-$L_{6\,\mu{\rm m}}$ sample.

3. Methodology

In the following section, we describe our methods to measure the clustering statistic and compare with DM halo models to ascertain galaxy bias and halo mass estimates.

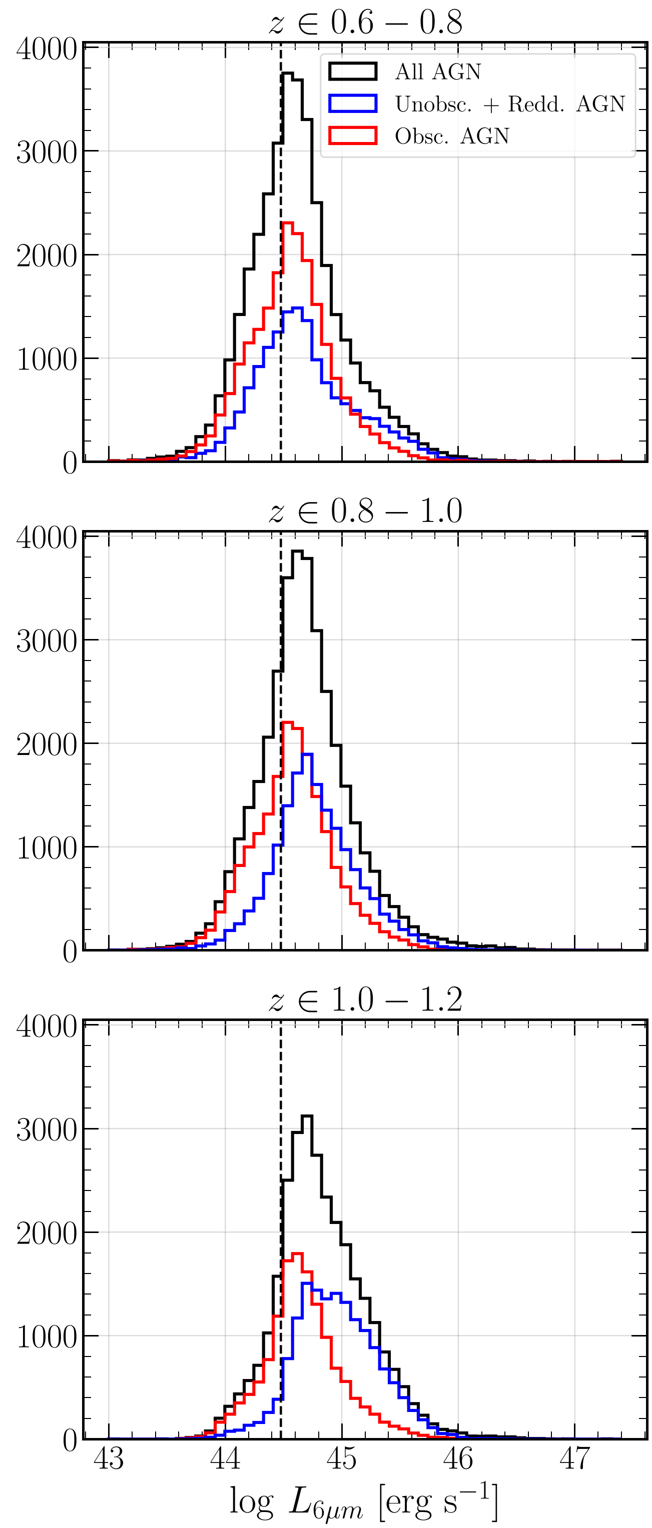

Figure 3. $L_{6\ \mu\mathrm{m}}$ luminosity distributions for the specified redshift bins for this analysis combining all considered HSC fields. The vertical dashed line indicates the $L_{6\ \mu\mathrm{m}}$ lower limit for this analysis at 3×10^{44} erg s^{-1}. The different colors indicate the different AGN subsamples.

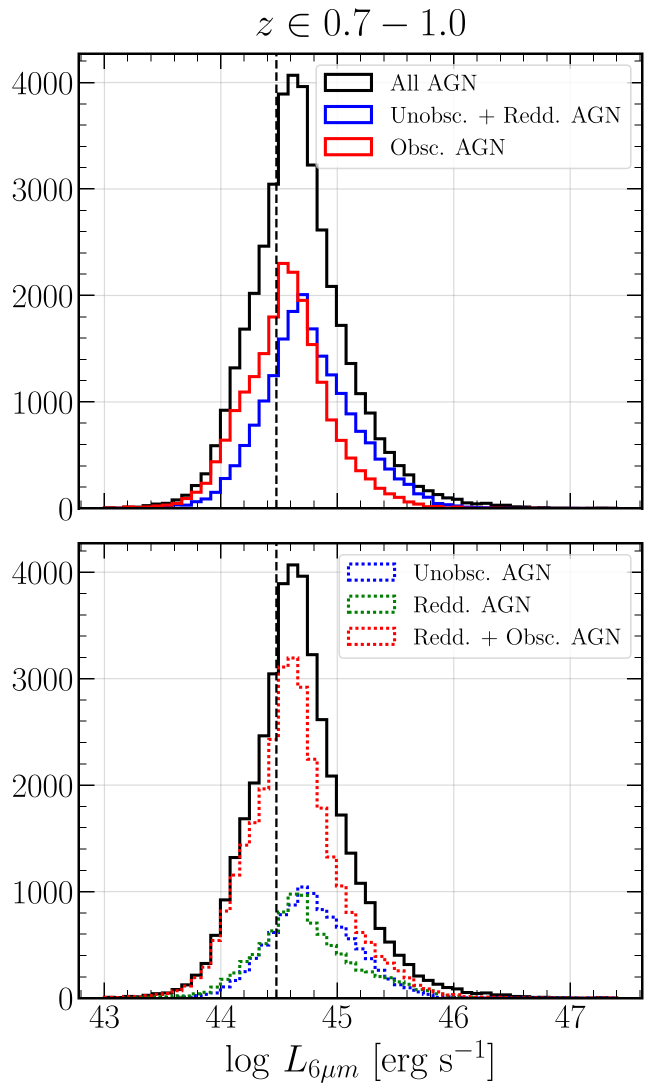

Figure 4. $L_{6\ \mu\mathrm{m}}$ luminosity distributions for the wide redshift bin for this analysis combining all considered HSC fields. The vertical dashed line indicates the $L_{6\ \mu\mathrm{m}}$ lower limit for this analysis at 3×10^{44} erg s^{-1}. The different colors indicate the different AGN subsamples that are considered in the wide redshift bin analysis.

3.1. Two-point Statistics to Assess Galaxy Clustering

3.1.1. Angular Correlation Function

The two-point angular clustering statistic is defined as the excess probability of a pair of objects being separated by an angle θ above a Poisson (random) distribution (P. J. E. Peebles 1973). Spatial correlation functions are a critical tool with which to analyze the clustering properties of galaxies on a wide range of angular scales. With only photometric redshift information in hand for every object in our sample, we limit this analysis to a projected clustering statistic, rather than one in 3D. In order to reduce shot noise bias, we employ the S. D. Landy & A. S. Szalay (1993) estimator of the angular two-point function:

$$\omega(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \quad (1)$$

where we further define the data–data, data–random, and random–random pair counting operations DD, DR, and RR as

$$XY(\theta) = \frac{XY'(\theta)}{N_X N_Y} \quad (2)$$

to ensure proper normalization given the number of objects N considered in the operation. While D is drawn from the observational catalogs, R is synthesized as a set of random points. We generate these randoms such that they match the survey footprint and have all area masks applied. This is critical in order to ensure that any excess clustering from the data is measured in the same survey geometry as the random points that establish the comparison. We will be using a weighted pair counting and binning statistic in this analysis, as detailed in Appendix B.

We follow H. Awan & E. Gawiser (2020) in formalizing the pair counting operation as

$$XY'(\theta) = \sum_{i}^{N_X} \sum_{j\neq i}^{N_Y} \bar{\Theta}_{ij,k} \tag{3}$$

where

$$\bar{\Theta}_{ij,k} = \Theta(\theta_{ij} - \theta_{\min,k})[1 - \Theta(\theta_{ij} - \theta_{\max,k})], \tag{4}$$

and

$$\Theta(x) = \begin{cases} 0\,, & \text{if}\ x < 0 \\ 1\,, & \text{if}\ x \geqslant 0 \end{cases}. \tag{5}$$

The Heaviside step function details the binning operation that counts the number of objects at an angular distance from another object. The parameter θ_{ij} is the separation between galaxy i and galaxy j in the sample of N_X and N_Y galaxies. The per-bin counting operator $\bar{\Theta}_{ij,k}$ counts the number of galaxy pairs at separation $\theta_{\min,k} \leqslant \theta_{ij} < \theta_{\max,k}$ for the kth bin. This formalism is equally applicable to the auto- or cross-correlation of galaxy catalogs. We use 24 spatial bins for the calculation, logarithmically spaced from $s = 0.01\,h^{-1}$ Mpc to $s = 100\,h^{-1}$ Mpc. We convert these projected scales to angular bins with a standard angular diameter distance conversion (from the comoving distance), with the median of the sample's dN/dzas the fiducial z.

3.1.2. Uncertainty Estimation

We use a standard jackknife procedure to estimate the total statistical and systematic error in a given per-field analysis. Splitting each of the three fields into 25 equal-area regions, we calculate the clustering signal from a set of 24 regions, removing one region per iteration until all 25 have been removed once. This removes approximately 4% of the data in each calculation. The per-bin error (which will be used as the 1σ error bars in each of the measured correlation functions) is then calculated from the square root of the diagonal of the covariance matrix as defined in P. Norberg et al. (2009):

$$C_{jk}(x_i, x_j) = \frac{N-1}{N} \sum_{k=1}^{N} (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j), \tag{6}$$

where x_i is the ith measure of the statistic, i.e., $\omega(\theta)$, out of a total of $N = 25$ measurements, and

$$\bar{x}_i = \sum_{k=1}^{N} x_i^k / N. \tag{7}$$

Previous analyses have adopted two different approaches to quantifying the uncertainties: R. C. Hickox et al. (2011), V. Allevato et al. (2011, 2014), P. Laurent et al. (2017), and M. Krumpe et al. (2018) use the full covariance matrix while L. Koutoulidis et al. (2013, 2018) and M. A. DiPompeo et al. (2014, 2016, 2017) use only the square root of the diagonal terms of C_{jk} as the 1D uncertainty. The most recent analysis of G. C. Petter et al. (2023), which we compare with here, uses the diagonal terms. For ease of comparison with prior work, we adopt the diagonal treatment as our default approach. We also evaluate and report the resulting best-fit galaxy bias and inferred halo mass via least-squares minimization of the data–model residual with the full C_{jk} matrix. In doing so, we capture the contribution of the real off-diagonal bin-to-bin correlations from the jackknife, and more accurately capture the precision of our measurements. In presenting the results from the full covariance matrix treatments, we show that our conclusions do not qualitatively change when either approach is implemented (see Appendix D). Additionally, the systematic uncertainty as a function of field-to-field variability is constrained by comparing the different fitted bias (b) and halo mass (M_h) values from each field (defined in full in the following subsection). We combine the bias and halo mass distribution from all fields to estimate the average value and uncertainty from the complete sample via an inverse variance weighted mean.

3.2. Clustering Interpretation

We use a halo model to infer the clustering properties from the measured excess probability $\omega(\theta)$. The relationship between galaxy clustering and DM halo clustering is formalized with the multiplicative ratio term of the galaxy bias, where we are working within the Eulerian framework of peak background split theory (R. K. Sheth & G. Tormen 1999). The bias term b describes the excess clustering signal in galaxies relative to the DM halo clustering model. We use the measured redshift distribution from our sample to model the DM halo clustering signal, which we parameterize with Limber's equation (D. N. Limber 1953; E. J. Groth & P. J. E. Peebles 1977; J. A. Peacock 1991; D. J. Eisenstein & M. Zaldarriaga 2001):

$$\omega(\theta) = \pi \int_{z=0}^{\infty} \int_{k=0}^{\infty} \frac{\Delta^2(k, z)}{k^2} J_0[k\ \theta\ \chi(z)] \times \left(\frac{dN}{dz}\right)_1 \left(\frac{dN}{dz}\right)_2 \left(\frac{dz}{d\chi}\right) dk\ dz, \tag{8}$$

where we use $\Delta^2(k, z) = \frac{k^3}{2\pi^2} P_{\rm HF}(k, z)$, $P_{\rm HF}(k, z)$ being the DM power spectrum from linear to nonlinear scales from `halofit` as implemented in the Core Cosmology Library (CCL; R. Takahashi et al. 2012; N. E. Chisari et al. 2019). J_0 is the zeroth-order Bessel function, $\chi(z)$ is the comoving distance in units of h^{-1} Mpc, and $dz/d\chi$ is defined as $H(z)/c = (H_0/c)[\Omega_m(1+z)^3 + \Omega_\Lambda]^{1/2}$ for a flat cosmology in $h\,{\rm Mpc}^{-1}$.

This parameterization takes a model for the 3D clustering of the halos and the dN/dz from our particular data set, properly accounting for the expected halo clustering over the redshift range being probed. This synthesis of the individual galaxies' redshift probability distributions $p(z)$ is usually constructed via a simple sum of the individual source $p(z)$,

$$\frac{dN}{dz} = \sum_i p_i(z). \tag{9}$$

As detailed in Appendix B, we use Equation (B5) to perform a weighted sum that considers an object's full $p(z)$. The benefit of this weighted method is that it includes the probability of an object being both within and outside a defined redshift bin, rather than standard tomographic approaches that include objects whose fiducial redshift is in the bin. We find that the results from this method are broadly consistent with the standard tomographic method, but opt for our weighted method in order to reduce possible effects on the systematic biases inherent in photometric redshift fitting.

We fit the measured clustering with a forward model via a linear least-squares fit with the simple multiplicative scalar value b^2, using a Monte Carlo (MC) sampling estimator. For our cross-correlation, the fitted multiplicative value is $b_G b_A$, i.e., the product of the bias from each of the two data sets (galaxies and AGN) being correlated. We calculate the galaxy bias from the autocorrelation, as well as the cross-correlation between the galaxy and AGN samples. In this analysis, we are able to divide the cross bias term by the recovered galaxy bias b_G and isolate the AGN bias b_A. We judge the goodness of fit via a χ^2 test, recording the reduced χ^2 statistic, $\chi^2_\nu = \chi^2/n_{\mathrm{d.o.f.}}$, for the total degrees of freedom (dof; X data points being fit – 1 fitting parameter, b or M_h). Using the combined nonlinear and linear DM halo model defined in `halofit`, we establish our fitting range to include smaller scales than what could be constrained with the linear model alone (I. Zehavi et al. 2004; R. Takahashi et al. 2012). We do not model the nonlinear one-halo term in this work, as that would require a halo occupation distribution (HOD) treatment (e.g., A. A. Berlind & D. H. Weinberg 2002), and does not contribute to the linear bias estimate. We thus only fit over scales $\gtrsim 3'$ in the measured correlation function, avoiding points on smaller scales in which the more complex HOD modeling would be necessary. This angular extent is chosen such that we fit for $s > 1\,h^{-1}$ Mpc in all redshift bins. Given these parameters and our 24 logarithmically spaced angular bins, we have a total of 11 data points from each correlation function to fit over, giving us 10 dof. We use the implementation of Limber's equation in the LSST Dark Energy Science Collaboration (DESC) CCL[8] to include the measured dN/dz in the `halofit` model (N. E. Chisari et al. 2019). This method produces the standard bias analysis in the (mostly) linear regime, and allows for an extension to estimate the mass of the average halo leading to the measured bias.

With this model and fitting algorithm, we infer the average halo mass of our samples, M_h, from the measured linear clustering bias. We use the J. L. Tinker et al. (2010) parameterization of the halo mass function to infer the halo masses traced by the measured bias values, as do P. Laurent et al. (2017). Following M. Tegmark & P. J. E. Peebles (1998), we can use the analytical form of $b(z, M)$ such that we may replace $P_{\mathrm{HF}}(k, z)$ in Equation (8) with $b^2(z, M) P_{\mathrm{HF}}(k, z)$ in an autocorrelation, fitting for the parameter M using an MC analysis. From this estimate of the galaxy bias and halo mass, we may then calculate the AGN halo mass by replacing $P_{\mathrm{HF}}(k, z)$ in Equation (8) with $b_G\, b_A(z, M)\, P_{\mathrm{HF}}(k, z)$, where b_G is known from the autocorrelation. While more uncertain than the bias measurement due to the several assumptions made for the halo mass function and HOD, a mass inference will isolate the halo properties from the implicit redshift evolution in the recovered bias. Moreover, halo mass comparisons across analyses are complicated by subtle differences in the precise formalism of the bias–halo mass translation used, and can introduce significant systematic shifts.

Following P. Laurent et al. (2017), we match the measured linear bias with the MC-derived halo mass best fit to estimate $M_{h,\mathrm{min}}$, i.e., a model for a strict lower limit in M_h. For every measured bias value in our chain, we calculate the $M_{h,\mathrm{min}}$ by solving for this value in

$$b_1(z, M_{h,\mathrm{min}}) = \frac{\int_{M_{h,\mathrm{min}}}^{\infty} \frac{dn}{dM} \langle N(M) \rangle\, b(z, M)\, dM}{\int_{M_{h,\mathrm{min}}}^{\infty} \frac{dn}{dM} \langle N(M) \rangle\, dM}, \tag{10}$$

where b_1 is the measured clustering bias, dn/dM is the halo mass function, the effective bias function $b(z, M)$ is as defined in J. L. Tinker et al. (2010), and $\langle N(M) \rangle$ is the average halo occupation, which we assume is 1 for our sample (i.e., every AGN is in the central galaxy of its halo).

This defines the lower bound of the mass range of halos, which we can then use to estimate the average halo mass, $\langle M_h \rangle$, folding in the halo mass and redshift distribution of our sample:

$$\langle M_h \rangle = \frac{\int_{M_{h,\mathrm{min}}}^{\infty} \int_z M \frac{dn}{dM} \frac{dN}{dz} \langle N(M) \rangle\, dz\, dM}{\int_{M_{h,\mathrm{min}}}^{\infty} \int_z \frac{dn}{dM} \frac{dN}{dz} \langle N(M) \rangle\, dz\, dM}. \tag{11}$$

This average halo mass estimate is more representative of the complete halo mass distribution that leads to the measured galaxy bias. Inferring the average host halo mass through this weighted average properly accounts for the cosmological distribution of halo masses in the Universe, as opposed to assuming all AGN are hosted in halos of a single representative mass M_{eff}. We discuss the intercomparison of different results in light of possible choices for the b–M_h connection formalism in Section 5.1.

3.3. Redshift Bins

Following previous work (A. Nicola et al. 2020; M. M. Rau et al. 2023), we subdivide our total sample into broad redshift bins where we evaluate the properties of the clustering as a function of cosmic time. How these bins are constructed becomes particularly relevant in photometric redshift surveys, where the $p(z)$ is much broader than that for a spectroscopic measurement.

Our finalized redshift bins are selected such that we have significant overlap in redshift between the HSC galaxy sample and the AGN, at $z > 0.5$. The median uncertainty of the LRG photo-z estimates, given their $p(z)$, is $\sigma_z \sim 0.1$. This uncertainty leads us to choose a bin width of $\Delta z = 0.2$ in order to capture the full 1σ distribution of an object in the center of our bin. We define three narrow redshift bins over which to conduct the complete AGN sample analysis. The first bin is delimited at $z \in 0.6$–0.8, the second bin is $z \in 0.8$–1.0, and the third is defined for $z \in 1.0$–1.2. We construct an additional wider bin with which to investigate the clustering across AGN subtypes, defined for $z \in 0.7$–1.0.

4. Results

In this section, we will discuss the measurement of the angular correlation functions of HSC galaxies. First we present the galaxy autocorrelations, then we turn to the AGN–galaxy cross-correlations.

4.1. HSC Galaxy Autocorrelation

We compute the LRG autocorrelation using the methods outlined in Section 3. Our estimate of the LRG galaxy bias will be used to isolate the AGN bias from the cross-correlation in

---

[8] https://github.com/LSSTDESC/CCL

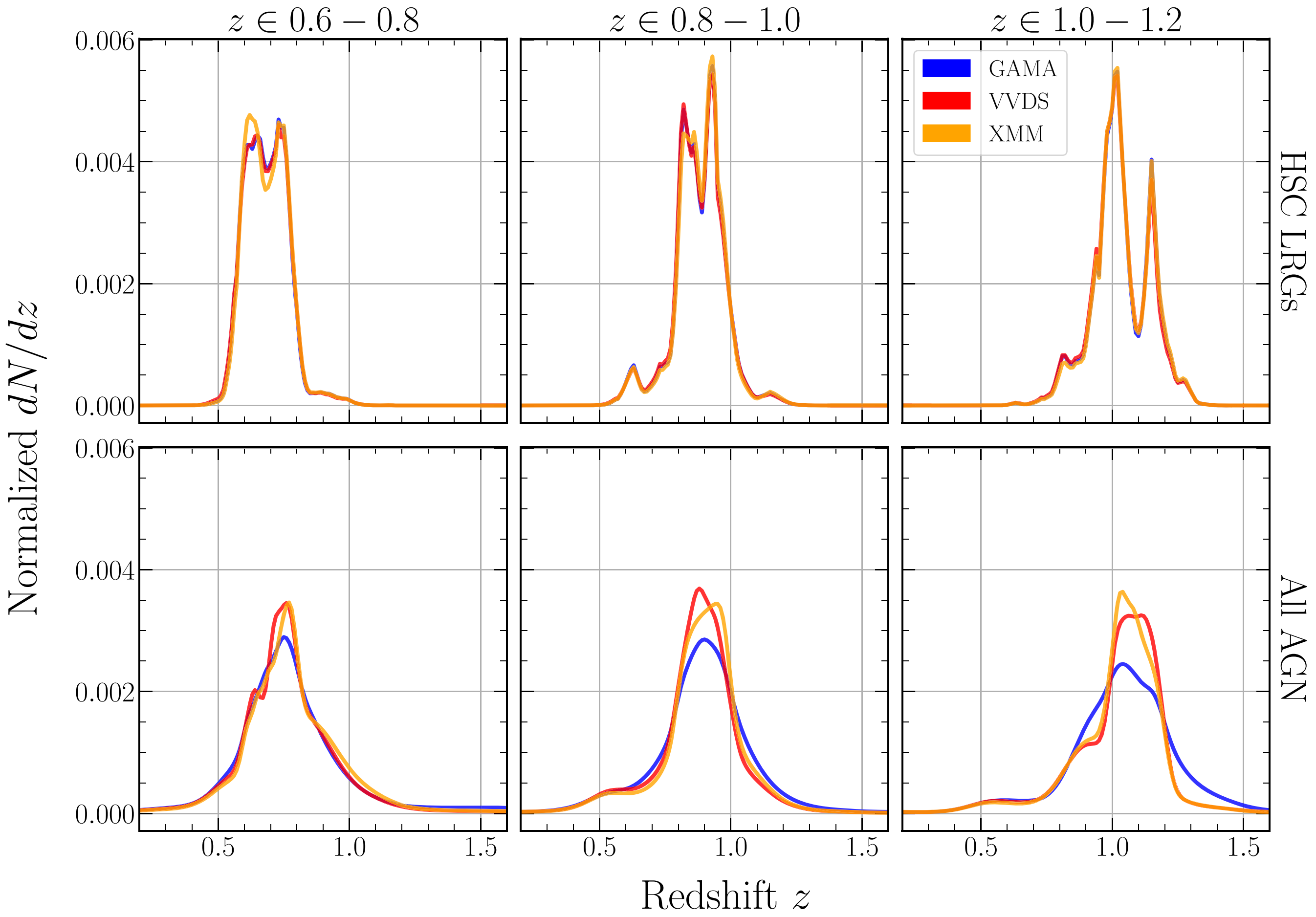

Figure 5. Normalized dN/dz measured from the LRG (top) and full AGN samples (bottom) for auto- and cross-correlations in the three narrow redshift bins, across different HSC fields. These are constructed following the procedure outlined in Appendix B.

Section 4.2. The dN/dz for the LRGs in each redshift bin, which inform the DM model fit to the data, are shown in the top row of Figure 5. The autocorrelations are shown in Figure 6. To emphasize differences between the per-field data and the models, we scale the $\omega(\theta)$ functions by a power of θ. In general, recovered $\omega_{GG}(\theta)$ values agree between fields. The per-angular bin average across all fields is shown as white symbols with a colored outline.

As described in Section 3.2, we recover the bias from these galaxy populations by fitting our `halofit` DM model to the measured points and uncertainties for each field. On smaller scales ($s < 0.3\, h^{-1}$ Mpc) there is a characteristic rise in clustering that deviates in shape from our DM model due to multiple subhalos within the central halo (known as the "one-halo" term). We find the model describes the data well ($\chi^2_\nu \approx 1$) down to arcminute scales, and we limit the fits to linear scales $s > 1\, h^{-1}$ Mpc. Fitting with $s > 2\, h^{-1}$ Mpc yields similar results. We also test whether our results change when excluding larger angular scales, and find that the measured bias values shift by $<1\sigma$ when fitting on scales $1 < s < 20\, h^{-1}$ Mpc. On larger scales, the error bars are significantly larger, but the measurements agree within the uncertainties. We use the field-to-field variability in measured bias as an estimate of the total systematic uncertainty. This is exhibited in Figure 7, where the fitted values of the bias from each field are plotted for a single redshift bin, and the total distribution is used to estimate the median value and its uncertainty. The recovered bias and χ^2_ν values are summarized in Table 2. For the three narrow redshift bins, we find the recovered biases are (with increasing z) 1.80 ± 0.03, 2.24 ± 0.03, and 2.90 ± 0.04. For our wide redshift bin ($z \in 0.7$–1.0), the measured bias is 1.92 ± 0.03. We find the reduced χ^2 values for the linear bias fit to be $\lesssim 1$ in the two lower-redshift bins. In the $z \in 1.0-1.2$ bin, however, we observe a significant amount of excess clustering on large scales relative to the linear DM model. Based on our preliminary galaxy clustering tests, detailed in Appendix A, we interpret this excess as the result of projection effects from lower-redshift objects still being present in our higher-redshift bin. We also perform these fits using the full covariance matrix, finding the measured bias in the two lower-redshift bins shifts by $<1\sigma$, while the highest-redshift bin has a significantly lower bias ($b_G \sim 2.4$), reflective of the uncertainty in the $z \in 1.0$–1.2 bin. The constrained values and uncertainties are reported in Table 3.

We observe a characteristic rise in linear bias as a function of redshift across our three narrow redshift bins, consistent with the findings of recent LRG studies like R. Zhou et al. (2021). This is expected for a population of relatively constant M_h (and at a fixed magnitude limit) at greater lookback time, as has been shown by prior LRG studies (e.g., S. Ishikawa et al. 2021). With bias values in hand for our galaxy population, we now turn to the galaxy–quasar cross-correlations.

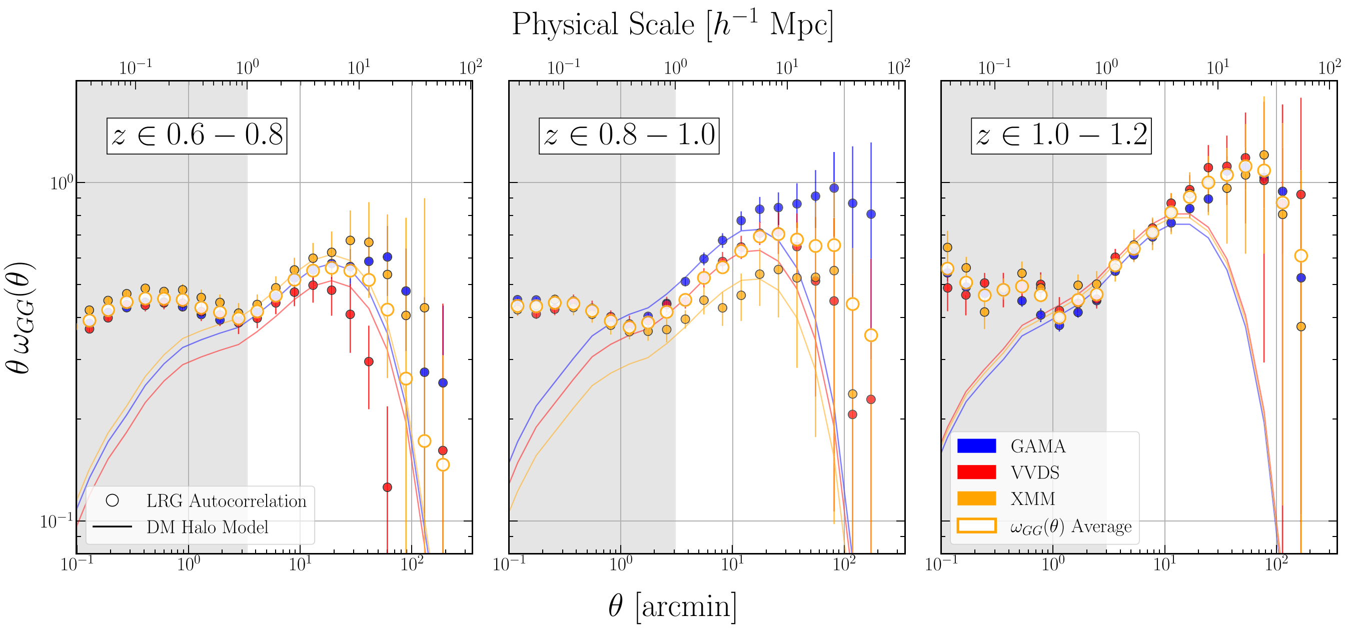

Figure 6. The measured HSC LRG projected angular autocorrelation in our three HSC fields across three redshift bins in $\theta\omega(\theta)$. We use this scaling by a power of θ to reduce the dynamic range of the plot and more clearly see the differences in the measured points. The open symbols are the per-bin average correlation function across fields. The 1σ uncertainties are drawn from the square root of the diagonal of the jackknife covariance matrix for the sample. The solid lines represent the fitted `halofit` DM model to each field. We fit points for physical scales $s > 1\ h^{-1}$ Mpc, while the gray region contains the interior points excluded from the fit. Note that the model does not include the one-halo term, which is why the data rise significantly above the model on small scales.

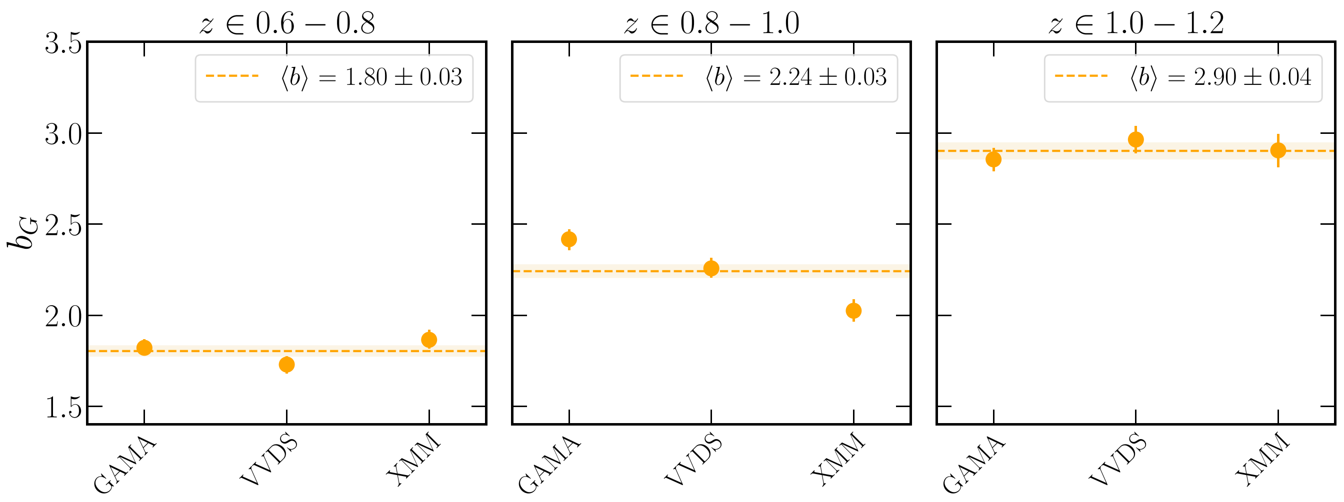

Figure 7. Recovered galaxy bias from the autocorrelation of HSC-selected LRGs, in three redshift bins, for each analysis field. The inverse variance weighted mean and 1σ uncertainty are plotted with a dashed line and shaded region.

4.2. Full AGN Sample Cross-correlations

Recovering the clustering amplitude from the AGN population permits us to ask questions about any trends with luminosity or redshift, and infer the halo mass that the AGN occupy following Section 3.2. The cross-correlations of the LRG sample and the complete AGN catalog, $\omega_{GA}(\theta)$, are shown in Figure 8. We perform the cross-correlation analysis for AGN that lie above the luminosity threshold of $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}, as established in Section 2.4.1 (see bottom panels of Figure 5 for the samples' normalized redshift distributions). These correlation functions are calculated in each HSC field, for each of the three narrow redshift bins. We see broad agreement of the recovered $\omega_{GA}(\theta)$ across fields, but note they are noisier than the correlation for $\omega_{GG}(\theta)$. While consistent within the per-bin errors, the autocorrelations vary at ~10% relative to the average over $3' < \theta < 40'$, and here we see ~20% over the same range for the lowest two redshift bins, and ~50% variability for the highest-redshift bin.

From the measured cross-correlation functions, we fit for the linear bias. Using the per-field autocorrelation bias found in Section 4.1, we divide out the galaxy bias contribution in the measured $b_G b_A$ to isolate the quasar bias, b_A. We perform a `halofit` model fit for scales $s > 1\ h^{-1}$ Mpc using CCL, and infer the halo mass as described in Section 3.2 above. The average bias and halo mass are recorded in Table 2, where they have been calculated with an inverse variance weighting of the results from individual fields. The recovered biases across redshift bins are consistent within 1σ (see Table 2), showing no significant evidence for bias evolution with redshift across this

Table 2
Angular Correlation Function Fit Results

Subset	N_{obj}	Weighted N_{obj}	$\langle L_{6\ \mu\text{m}} \rangle$ (log erg s^{-1})	$\langle z \rangle$	$\langle \chi^2 \rangle$ (10 dof)	b	$\langle M_h \rangle$ (log $h^{-1}\ M_\odot$)
$z \in 0.6$–0.8							
LRGs	1,288,589	879,258.8	⋯	0.7 ± 0.1	9.8	1.80 ± 0.03	⋯
All AGN	22,988	5804.3	$44.7^{+0.4}_{-0.2}$[a]	$0.8^{+0.1}_{-0.2}$	5.4	1.5 ± 0.1	12.8 ± 0.1
$z \in 0.8$–1.0							
LRGs	851,117	440,970.6	⋯	0.9 ± 0.1	6.6	2.24 ± 0.03	⋯
All AGN	26,264	10,381.3	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	5.8	1.4 ± 0.1	12.7 ± 0.1
$z \in 1.0 - 1.2$							
LRGs	324,790	98,498.4	⋯	1.0 ± 0.1	18.7	2.90 ± 0.04	⋯
All AGN	25,235	7204.7	$44.9^{+0.4}_{-0.3}$[a]	1.1 ± 0.2	4.9	1.6 ± 0.1	12.6 ± 0.1
$z \in 0.7$–1.0							
LRGs	1,509,905	843,166.6	⋯	0.8 ± 0.1	15.5	1.92 ± 0.03	⋯
All AGN	28,494	13,898.8	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	6.7	1.6 ± 0.1	12.9 ± 0.1
Unobscured AGN	8266	3942.0	$44.9^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	8.9	2.2 ± 0.1	13.3 ± 0.1
Unobscured + Reddened AGN	15,156	6745.71	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	5.4	1.9 ± 0.1	13.1 ± 0.1
Reddened AGN	6890	2803.6	$44.8^{+0.5}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	3.1	1.6 ± 0.1	$12.8^{+0.1}_{-0.2}$
Reddened + Obscured AGN	19,675	9893.4	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	5.0	1.4 ± 0.1	12.7 ± 0.1
Obscured AGN	12,785	7089.8	$44.7^{+0.3}_{-0.2}$[a]	0.9 ± 0.2	6.9	1.3 ± 0.1	12.6 ± 0.1
High-$L_{6\ \mu\text{m}}$ AGN	7760	2492.2	$45.1^{+0.4}_{-0.2}$[b]	$1.0.^{+0.1}_{-0.2}$	3.6	$1.4^{+0.1}_{-0.2}$	$12.5^{+0.2}_{-0.3}$
Low-$L_{6\ \mu\text{m}}$ AGN	20,734	11,406.5	$44.8^{+0.3}_{-0.2}$[c]	$0.9^{+0.1}_{-0.2}$	6.0	1.7 ± 0.1	13.0 ± 0.1

Notes.
[a] Primary luminous AGN selection ($L_{6\ \mu\text{m}} > 3 \times 10^{44}$ erg s^{-1}).
[b] Higher-luminosity AGN selection ($L_{6\ \mu\text{m}} > 10^{45}$ erg s^{-1}).
[c] Lower-luminosity AGN selection ($3 \times 10^{44} < L_{6\ \mu\text{m}} < 10^{45}$ erg s^{-1}).

narrow redshift range. The average inferred $\log M_h$ for our $L_{6\ \mu\text{m}} > 3 \times 10^{44}$ erg s^{-1} range in the three narrow redshift bins is 12.8 ± 0.1, 12.7 ± 0.1, and 12.6 ± 0.1 log($h^{-1}\ M_\odot$). We also calculate the correlation function for the full AGN sample in our wider redshift bin ($z \in 0.7$–1.0, Figure 9)—see the upper leftmost panel of Figure 10. The inferred halo mass is 12.9 ± 0.1 log($h^{-1}\ M_\odot$).

We additionally test for any evidence of halo mass evolution with luminosity by making successive subsamples with a higher threshold of $L_{6\ \mu\text{m}}$. The inferred mass for the $L_{6\ \mu\text{m}} > 10^{45}$ erg s^{-1} sample is $12.5^{+0.2}_{-0.3}$ log($h^{-1}\ M_\odot$), while the halo mass for the $3 \times 10^{44} < L_{6\ \mu\text{m}} < 10^{45}$ erg s^{-1} bin is 13.0 ± 0.1 log($h^{-1}\ M_\odot$). As visualized in the upper panels of Figure 11, we find no significant evidence for an evolution in halo mass as a function of $L_{6\ \mu\text{m}}$ (also see Table 2). These bias and halo mass values are consistent with the results when using the full covariance matrix treatment, reported in Table 3.

4.3. AGN Subtype Cross-correlations

Following the same procedure as that for the full AGN sample, we calculate the cross-correlation for the AGN subtypes and estimate the physical properties of the halos in which these AGN reside. We investigate the obscured and unobscured AGN categories in a single, broader redshift bin ($z \in 0.7$–1.0) to provide a large enough sample—see the total and weighted number of objects in Table 2. Utilizing the same $L_{6\ \mu\text{m}}$ threshold, the luminosity distributions of the full unobscured + reddened and obscured AGN resemble each other, as seen in Figure 4. From the total number density of 50.1 deg^{-2} ($L_{6\ \mu\text{m}} > 3 \times 10^{44}$ erg s^{-1}) of AGN in the $z \in 0.7$–1.0 bin, there are 26.7 deg^{-2} unobscured + reddened objects, and 22.5 deg^{-2} obscured objects. The unobscured + reddened AGN include 14.6 deg^{-2} unobscured and 12.1 deg^{-2} reddened AGN.

The measured cross-correlation functions for the complete, unobscured + reddened, and obscured samples are illustrated in the top row of Figure 10, and display a fair amount of field-to-field variability, while being consistent within the errors. We account for these variations and illustrate a mean value as we do for the full AGN sample. From these correlation functions, we note how the average recovered clustering shifts above and below the reference of the complete AGN sample (black dashed line). These shifts in bias are (nonlinearly) correlated with the underlying halo mass distribution of where these AGN reside. We exploit this via the methods outlined in Section 3.2 to recover the average mass of the halos in which these AGN are found. We again observe the rise at small angular scales ($s < 0.3\ h^{-1}$ Mpc) relative to the `halofit` model from the one-halo term. As before, we fit the correlation function with the model beyond $1\ h^{-1}$ Mpc, avoiding the nonlinear-dominated regime. The measured correlation functions are well described by the model in our fitting range (see χ^2 values in Table 2). As an additional test of the (in)consistency between the unobscured and obscured measured $\omega(\theta)$ values, we fit the same CCL-derived $\omega(\theta)$ DM model to both. They are fit with the (per-field) full AGN sample DM halo model (see solid lines in the top left panel of Figure 10). We find that the fitted b^2 value for the LRG × unobscured AGN $\omega(\theta)$ (5.0 ± 0.5) is statistically inconsistent at $>5\sigma$ with the value inferred for the LRG × obscured AGN (2.2 ± 0.4) measurement. This is the

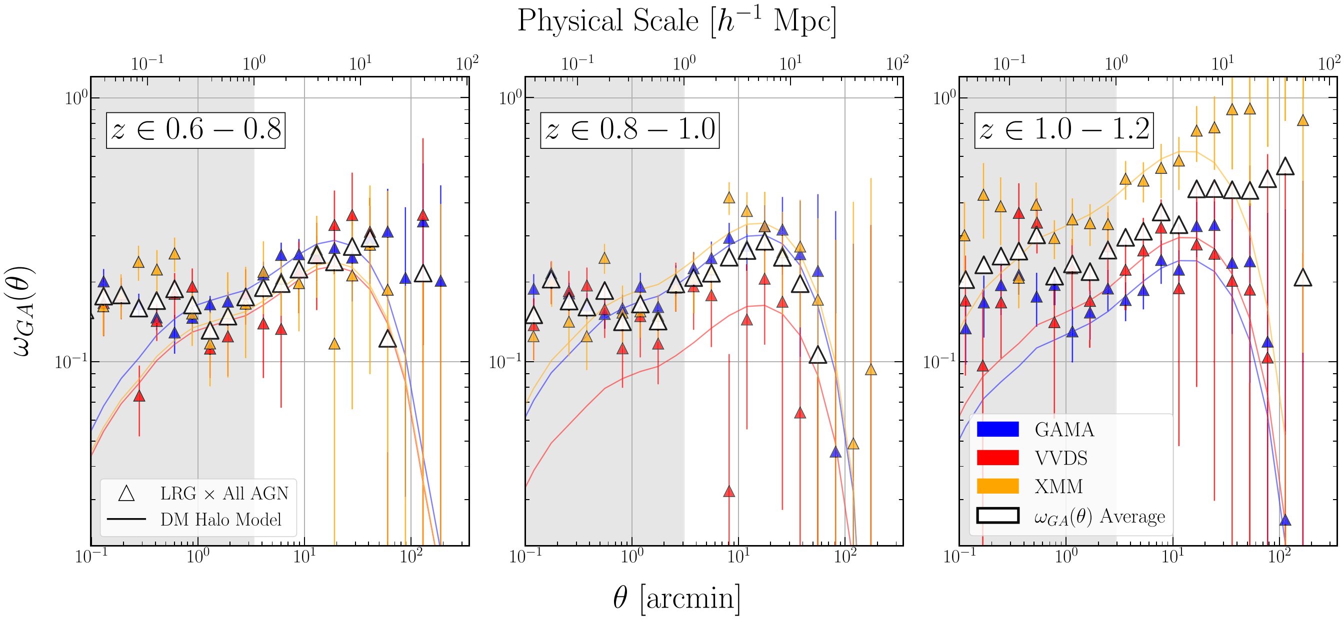

Figure 8. Cross-correlation between HSC LRGs and all AGN measured, fitted with CCL in different fields across redshift bins, shown in $\theta\omega(\theta)$ to reduce the plotted dynamic range. The AGN sample is limited to $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}. The open symbols are the per-bin average $\omega_{GA}(\theta)$ across fields. The 1σ uncertainties are drawn from the square root of the diagonal of the jackknife covariance matrix for the AGN sample. The solid lines represent the fitted `halofit` DM model to each field. The points considered in this analysis are those past the gray shaded region, at $s > 1\ h^{-1}$Mpc ($\theta \gtrsim 3'$).

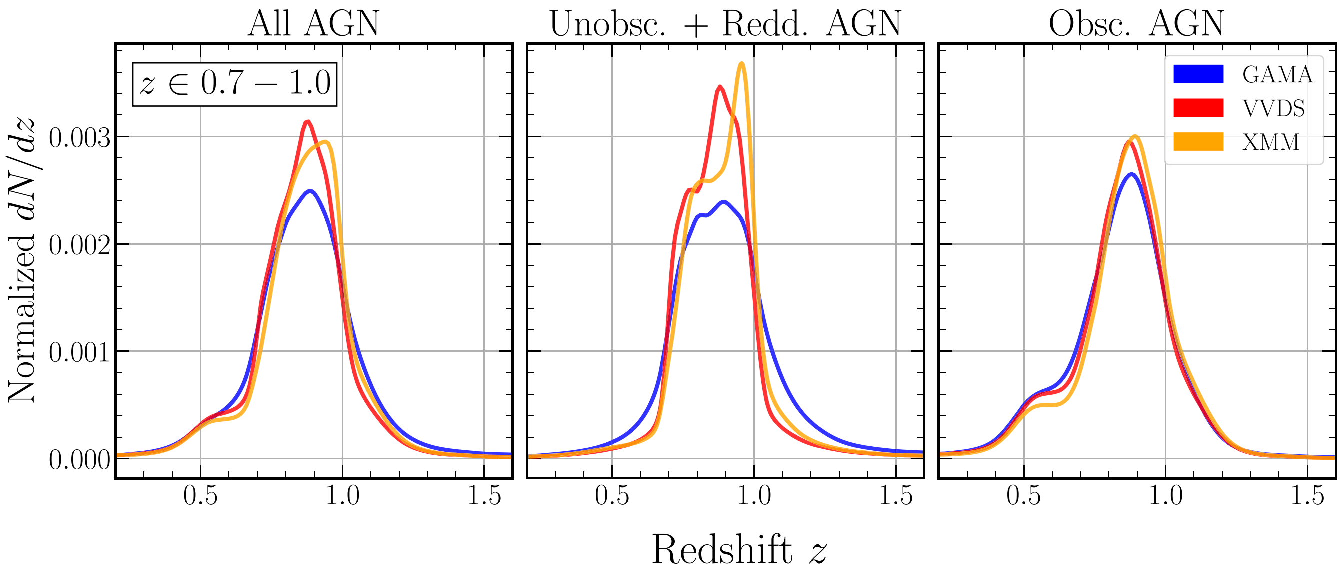

Figure 9. Normalized dN/dz measured from the AGN samples for cross-correlations in the wide redshift bin ($z \in 0.7$–1.0) for the $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1} threshold, for the three fields of interest.

case when fitting with the 1D uncertainties, as well as with the full covariance matrix. This shows that the measured amplitude of the clustering signal from both of these samples is substantially different.

With these correlation functions, we seek to measure the clustering strength of our different AGN subsamples. We estimate the galaxy bias and inferred halo mass for each subtype to investigate whether the different samples have the same or different characteristic M_h. We present the AGN subtype halo masses in Figure 11 and in Table 2, and we present the full covariance matrix treatment's inferred M_h and significance in parentheses throughout the text and in Appendix D (in Table 3). The M_h values inferred with the full covariance matrix treatment shift by $<1\sigma$ for all AGN subtypes. The average unobscured + reddened AGN halo mass is 13.1 ± 0.1 ($13.2^{+0.1}_{-0.2}$) $\log(h^{-1}\,M_\odot)$. We find the M_h for obscured AGN is $12.6 \pm 0.1(12.6^{+0.2}_{-0.3})$ $\log(h^{-1}\,M_\odot)$. The unobscured + reddened average M_h is higher than the obscured halo mass by a factor of ∼3. Given the 1D statistical uncertainties, this is a 3σ difference (2.8σ with the full covariance matrix).

We also investigate what happens to the clustering and the inferred halo mass when we split the unobscured + reddened AGN sample into its reddened and unobscured AGN components. The

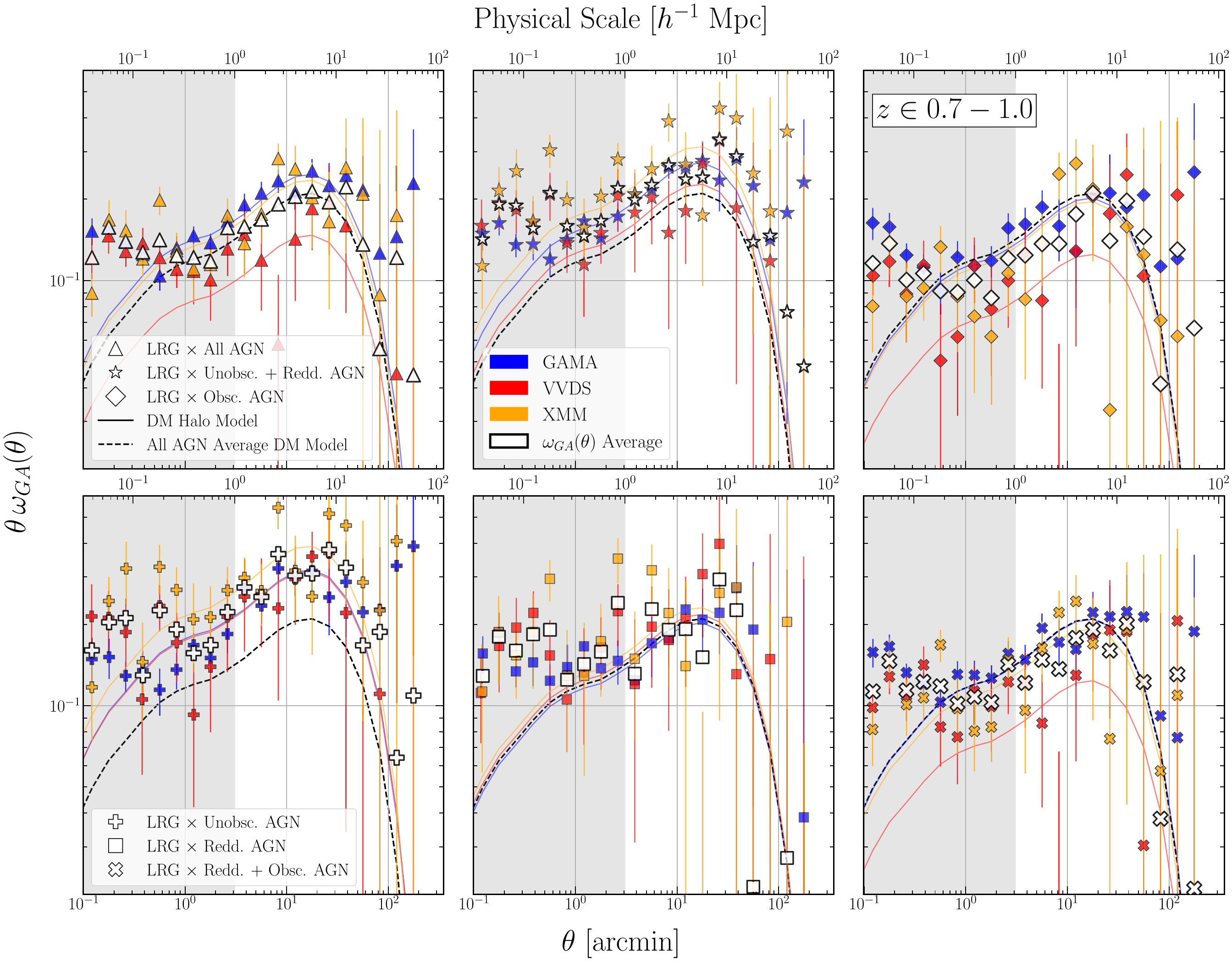

Figure 10. Cross-correlation between HSC galaxies and the full AGN sample and subtype samples in the wide redshift bin ($z \in 0.7$–1.0) and with $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}, shown in $\theta\omega(\theta)$. The open symbols are the per-bin average correlation function across fields. The 1σ uncertainties are drawn from the square root of the diagonal of the jackknife covariance matrix for each AGN subtype sample. The solid lines represent the fitted `halofit` DM model to each field. The points considered in this analysis are those past the gray shaded region, at $s > 1\ h^{-1}$ Mpc, where we are confident the model is able to describe the data. We note the model still agrees with the measured points well within the shaded region. We observe a departure from the DM model on small scales, as expected for the one-halo term. The dashed line represents the full AGN sample average best-fit `halofit` model, and is repeated in all panels as a reference line for ease of comparison. Top: the cross-correlation of the full, unobscured + reddened, and obscured samples of AGN across three fields. Bottom: the cross-correlation of the unobscured, reddened, and reddened + obscured samples with the LRG sample. The symbol shapes used for each sample here match those in Figure 11.

correlation functions for these samples can be seen in the bottom panels of Figure 10. First we compare the average halo mass for the unobscured and reddened AGN, finding $\log M_h = 13.3 \pm 0.1 (13.4 \pm 0.1)$ and $12.8^{+0.1}_{-0.2}(13.0^{+0.1}_{-0.2}) \log(h^{-1}\,M_\odot)$, respectively. Here, our sample of unobscured objects are more massive than their reddened counterparts by a factor of ∼3, at 3σ (2.9σ with the full C_{jk}). Next we contrast blue objects (unobscured AGN) and reddened objects (reddened AGN + obscured AGN), finding that there is a significant difference, with the former being ∼4× more massive, at 5σ (3.9σ). We directly compare the unobscured and obscured AGN, and find a 5σ (3.5σ) statistical difference in the inferred average halo mass, where the unobscured AGN reside in ∼4.5× more massive halos.

These measurements may suggest that AGN populations with different obscuration levels occupy halos with different average inferred masses. However, we must consider possible systematic uncertainties in our photometric redshift values and possible misclassifications before making any physical interpretations.

4.4. *Potential Effects of Unconstrained* dN/dz *Uncertainties*

We investigate how systematic shifts in our dN/dz would impact the interpretation of our measured clustering functions. As we detailed in Section 3.2, the halo model is directly informed by the overlap of the input dN/dz in a cross-correlation. The smaller the overlap of the two samples' dN/dz, the lower the expected clustering amplitude of the halo model. Consequently, the measured bias relative to the halo model will be higher if these dN/dz are shifted away from each other.

Here we use a toy scenario to estimate the magnitude of the required shift in dN/dz to resolve the currently observed difference in average halo mass estimated between our unobscured + reddened and obscured AGN. We shift the dN/dz for our AGN samples and subsamples for each field by

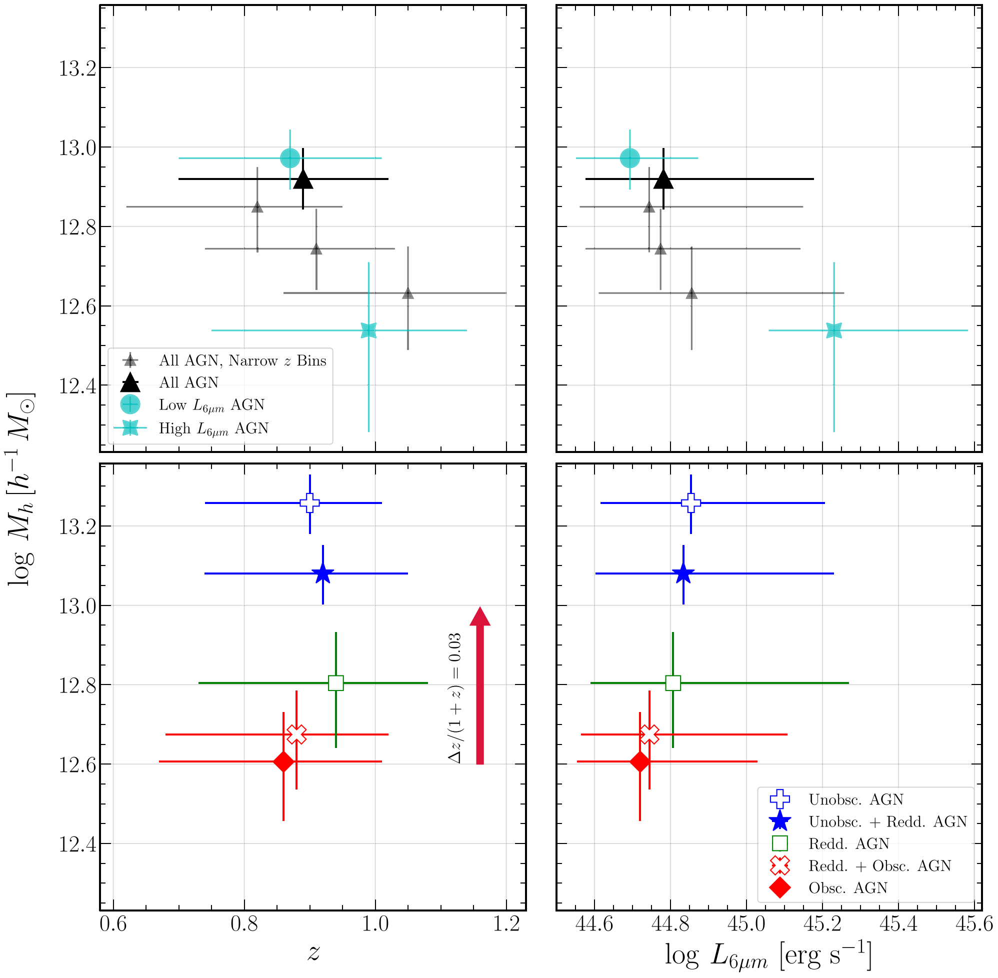

Figure 11. Summary of recovered halo masses (M_h) from the AGN cross-correlation of the full AGN sample (top row) and the AGN subtype samples (bottom row). The left panels illustrate this as a function of redshift. The right panels show the halo masses as a function of estimated $L_{6\ \mu\mathrm{m}}$. The position and uncertainties in z and $L_{6\ \mu\mathrm{m}}$ are drawn from the 16%, 50%, and 84% quantiles of each bin's underlying distribution. We find that unobscured + reddened AGN are found in significantly more massive halos than obscured AGN. We additionally estimate the effect of a systematic redshift error in the dN/dz on our inferred halo masses. As indicated by the red arrow, an upward z shift in the dN/dz by $\Delta z/(1+z) = 0.03$ would lead to a boosting of the inferred M_h by a magnitude of 0.4 dex, a linear factor of ~2.5. The values represented are also found in Table 2.

$\Delta z/(1+z) = \pm 0.03$. We find that shifting the dN/dz to lower redshifts has a minimal impact on the measured bias or inferred M_h. This is expected given the shift to lower redshift does not significantly change the overlap in dN/dz between the LRG and the AGN. Shifting the distribution to higher z (further separating it from the bulk of the LRG dN/dz) shifts the average inferred halo mass up by a factor of ~2.5 (0.4 dex) for any sample/subsample of AGN. This shift is represented by the red arrow in the bottom left panel of Figure 11. The b_A is shifted higher by 0.6 and 0.3 for the unobscured + reddened and obscured samples, respectively.

If the true dN/dz for the obscured objects were to be systematically shifted in this way relative to our measurement (while the unobscured + reddened objects' dN/dz is considered accurate), then the significant difference we measure here would be resolved. As R. C. Hickox et al. (2011) show, systematic shifts in redshifts for photometrically determined obscured AGN have been a topic of considerable uncertainty given the abundance of spectroscopic information on unobscured AGN (see Y. Shen et al. 2009; B. W. Lyke et al. 2020) and the relative dearth of obscured AGN spectra. This is sensible given obscured objects are fainter and will suffer photo-z fitting degeneracies similar to those of galaxies at similar redshifts. Prior studies from HSC have also highlighted the possibility of these systematic shifts in the photometric redshift distributions (see R. Dalal et al. 2023).

We also investigate the effect of having uncorrelated objects (any object that is at a different redshift but is incorrectly in our redshift bin) in the analysis. We estimate how including increasing numbers of randomly distributed objects (i.e., those with incorrect redshifts) in the unobscured and obscured AGN samples affects the measured correlation functions. We find that if 10% of the unobscured AGN sample were in fact uncorrelated random objects (that are then unaccounted for in the dN/dz), they would depress the inferred bias value such that it would resolve the difference in bias with the obscured AGN. The magnitude of this effect is reproduced when including random objects in the obscured AGN sample. We infer from this test that if we had a similar fraction of catastrophic photo-z misattributions only in our obscured AGN sample, it could produce the bias difference we measure in this analysis. Clearly, photometric redshifts provide a systematics floor to our results that can only be resolved with large spectroscopic samples. We discuss the possibility of these redshift failures and other potential contaminating sources in our sample in Section 5.2.

5. Discussion

The use of correlation functions has become an efficient means of investigating the properties of DM halos in which AGN reside. From illustrating the redshift evolution of bias for AGN samples (see N. P. Ross et al. 2009; Y. Shen et al. 2009; V. Allevato et al. 2011; L. Koutoulidis et al. 2013; M. C. Powell et al. 2020) to informing the bias and inferred halo mass of different AGN subtypes (e.g., R. C. Hickox et al. 2009; V. Allevato et al. 2014; M. A. DiPompeo et al. 2016, 2017; P. Laurent et al. 2017; G. C. Petter et al. 2023; A. Viitanen et al. 2023), spatial statistics have served to describe underlying properties of the distribution and abundance of AGN.

We study the halo properties of our HSC + WISE AGN sample using cross-correlation with LRGs. We find that the AGN population shows no evidence for halo mass evolution with redshift, over the narrow redshift range studied here. Focusing on AGN with $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}, where we are relatively complete up to $z \sim 1.2$, we likewise find no evidence for variation in M_h as a function of luminosity. This is equivalent to $L_{\mathrm{bol}} > 3.4 \times 10^{45}$ erg s^{-1}, following the IR correction from P. F. Hopkins et al. (2007). However, we do see a dramatic difference in the inferred halo mass between the unobscured + reddened and obscured AGN samples. The masses differ by a factor of ∼3, with the unobscured + reddened AGN being more massive at >3σ significance relative to the obscured objects.

5.1. Luminosity Dependence of Inferred Halo Properties

Many prior papers have found a lack of luminosity dependence in clustering as found here. Direct comparisons between samples can be complicated by differences in selection method, redshift distribution, and so on. We thus limit our attention to samples with $\langle z \rangle \approx 1$, with photometric or spectroscopic redshifts, and that utilize angular correlation functions. We compare with X-ray (A. L. Coil et al. 2009; V. Allevato et al. 2011; L. Koutoulidis et al. 2013; M. C. Powell et al. 2020), optical/UV (Y. Shen et al. 2009; P. Laurent et al. 2017), and MIR (R. C. Hickox et al. 2011) selected samples. The left panel of Figure 12 includes bias tracks as a function of redshift for a given halo mass, following J. L. Tinker et al. (2010). The right panel shows the reported M_h from these analyses as a function of bolometric luminosity. To convert each sample to a bolometric luminosity in a uniform manner, we use the corrections from P. F. Hopkins et al. (2007). We find the average halo mass for our AGN sample ($\approx$13 $\log h^{-1}M_{\odot}$) is consistent with the AGN halo mass inferred by the studies we compare with here.

We reiterate that comparing halo mass estimates across analyses is imperfect due to different treatments of the galaxy bias to halo mass connection, which may introduce significant shifts. Attention must be paid to whether the inferred M_h is calculated with either the effective redshift or the complete dN/dz of the sample, as well as the precise b–M_h connection used (e.g., R. K. Sheth & G. Tormen 1999; R. K. Sheth et al. 2001; J. L. Tinker et al. 2010). Moreover, distinct cosmological codes make different assumptions in standard galaxy bias–halo mass treatments, such as the mass definition adopted, which in the case of spherical overdensity based masses picks a value of Δ as defined in J. L. Tinker et al. (2010). We have assumed, like P. Laurent et al. (2017) and the CCL default (N. E. Chisari et al. 2019), that $\Delta = 200$. Additionally, there is not an agreed-upon convention for the definition of halo mass (and therefore, a universal value of Δ). This plurality of different approaches can introduce significant shifts on the order of 0.1–1.0 dex, and makes any comparison across analyses difficult. To ensure robustness in our investigation, we check our (N-body derived) halo mass inference from J. L. Tinker et al. (2010) against the analytic formulation from R. K. Sheth & G. Tormen (1999). While the recovered halo masses for all our subsamples shift to ∼0.2 dex larger values, we find that the relative halo mass differences between AGN subtypes are consistent across these b–M_h parameterizations. We also verify that these halo mass differences are preserved when choosing a different halo mass function formalism. We choose a parameterization that integrates over the halo mass function so as to recover the $\langle M_h \rangle$ (see P. Laurent et al. 2017 for another implementation of this method). Other analyses such as G. C. Petter et al. (2023) have opted for an M_{eff} approach, where one reports the halo mass at which the $b(z, M)$ matches the measured galaxy bias, for a given b–M_h connection. While there are physical implications to each of these approaches, we find that the estimated M_h differences and significances are present irrespective of the chosen formalism. We are not so focused on the precise value of the inferred M_h, but rather on the significance of the differences. Given all these possible analysis choices, we conclude that while relative halo masses within any investigation are informative, contrasting absolute values between analyses can be misleading.

As we noted previously, we find no evidence for bias or halo mass evolution as a function of luminosity. Our result is consistent with Y. Shen et al. (2009), who find that over a wide range of redshifts, the recovered halo mass of optically selected quasars does not show any trends with luminosity (see also S. M. Croom et al. 2005; A. Lidz et al. 2006; A. L. Coil et al. 2009; R. C. Hickox et al. 2011; A. J. Mendez et al. 2016). Similarly, Y. Shen et al. (2013) find there is a poor correlation between L and M_h at $\bar{z} \sim 0.5$.

Several analyses have constrained the AGN halo mass range within $M_h \in 12.5 - 13 \log h^{-1} M_{\odot}$ (P. F. Hopkins et al. 2008; N. Cappelluti et al. 2012; M. Krumpe et al. 2014; J. D. Timlin et al. 2018). It has been suggested that the halo mass scale at which the studies presented here converge is unique for AGN triggering due to the low relative galaxy velocities within a group (e.g., P. F. Hopkins et al. 2008; R. C. Hickox et al. 2011). However, studies of more detailed semiempirical models have suggested that the underlying distribution of halo masses is broad, but the combination of higher AGN fractions in higher-mass star-forming galaxies, low-L AGN sample incompleteness, and relatively high satellite fractions among AGN leads to an apparent constant M_h (A. Georgakakis et al. 2019; J. Aird & A. L. Coil 2021).

We note that some analyses have found evidence for a correlation between quasar luminosity and inferred halo mass (L. Koutoulidis et al. 2013). M. Krumpe et al. (2018) study X-ray AGN at $\langle z \rangle < 0.3$. They find no significant difference between their high- and low-L_{X} bins on linear scales (two-halo contribution), but find a significant difference (3σ) when comparing the HOD-derived clustering strength for the one-halo term. This difference may be driven by the inclusion of the intrahalo small-scale clustering (the one-halo term), given that prior studies have noted this higher clustering amplitude due to

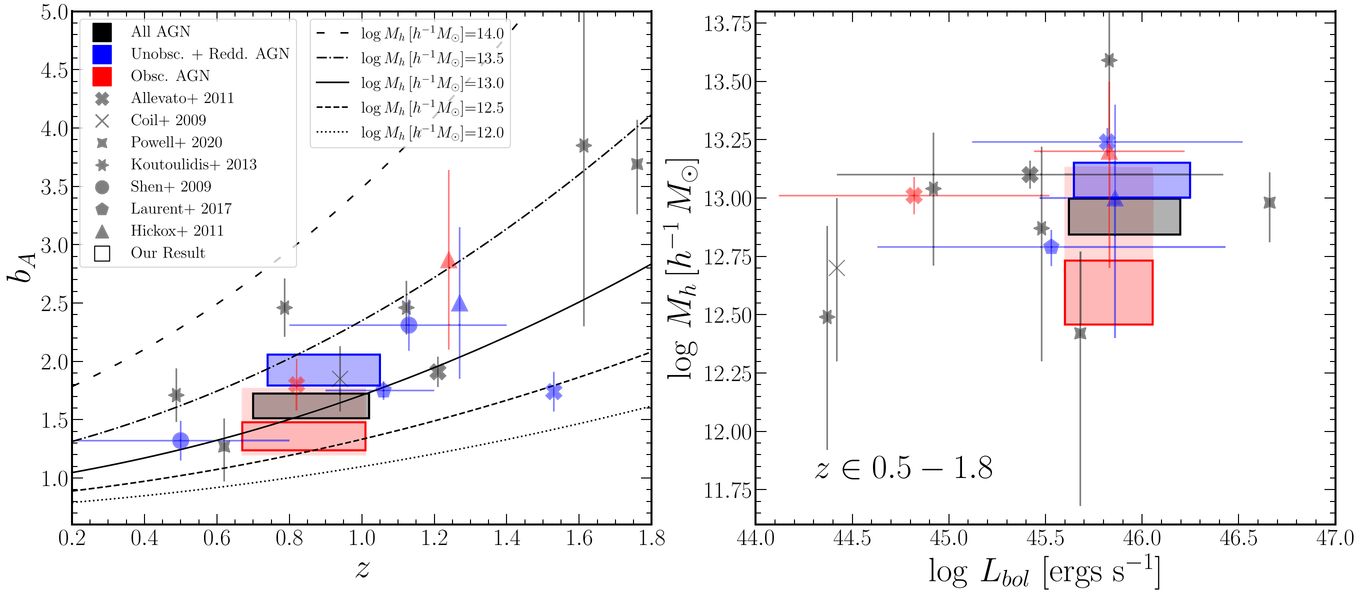

Figure 12. Our measured AGN bias and inferred halo masses for the full AGN sample and unobscured + reddened and obscured samples in comparison with other analyses. Left: the recovered AGN bias as a function of redshift in comparison with X-ray, optical, and IR selected AGN clustering studies. Marker styles represent the analysis used in the comparison. Marker color indicates whether it is an analysis of all the AGN in the sample, or of a subsample for unobscured + reddened or obscured AGN. The colored boxes illustrate our measurements and statistical uncertainties (solid outlines), and we include our estimate of the possible magnitude of inferred bias and halo mass shifts were there to be a systematic bias in the dN/dz for $\Delta z/(1+z) = 0.03$ (lower-opacity red rectangle). Plotted lines illustrate the nominal $b(z, M_h)$ tracks for different halo masses (J. L. Tinker et al. 2010). Right: inferred halo masses as a function of bolometric luminosity ($L_{\rm bol}$); markers are the same as those in the left panel. We convert studies' recorded L_ν to $L_{\rm bol}$ using the correction from P. F. Hopkins et al. (2007). We compare, where available, with values from projected angular correlation function analyses, rather than projected real-space derived values (see R. C. Hickox et al. 2011). This halo mass comparison serves as a visualization of other results in the field and the halo mass differences between subtypes they find, but absolute value comparisons are difficult since the studies presented here use different b–M_h formulations.

the presence of merging systems (J. F. Hennawi et al. 2006; W. Serber et al. 2006; P. F. Hopkins et al. 2008). Since our analysis only considers large scales, we are not surprised at the lack of a luminosity difference in our results. This absence of an L dependence allows us to analyze differences as a function of AGN type (Section 5.2), without requiring identical L distributions.

5.2. Inferred Average Halo Mass Differences

Previous studies have disagreed on whether obscured/Type II and unobscured/Type I AGN reside in DM halos of different masses. For instance, V. Allevato et al. (2011) find a trend where (X-ray selected, optically classified) broad-line AGN are in more massive halos than narrow-line AGN at a range of redshifts ($z_{\rm spec} \sim 0.8$–2.0), and repeat this finding at $z_{\rm spec} \sim 3$ in V. Allevato et al. (2014). However, it is difficult to compare with X-ray analyses given smaller sample sizes and different redshift ranges. M. Krumpe et al. (2018) find excess clustering of (X-ray column depth defined) Type II AGN relative to Type I AGN using spectroscopic redshifts for $\langle z_{\rm spec}\rangle < 0.04$. N. Cappelluti et al. (2010) find the opposite, showing X-ray Type I AGN to be an order of magnitude more massive than their Type II counterparts for $\langle z_{\rm spec}\rangle < 0.06$, again split by their column depth.

Previous optical- and IR-selected AGN studies have also seen a divergence of results. Using a projected real-space correlation analysis of optical−MIR color selected quasars, R. C. Hickox et al. (2011) show that obscured AGN (with primarily photometric redshifts) have statistically consistent clustering with unobscured AGN (with primarily spectroscopic redshifts), but note that the difference may be underestimated due to photo-z uncertainties. A. J. Mendez et al. (2016) also find no significant difference between AGN subtypes across IR selections with spectroscopic redshifts for each object. Other IR analyses have focused on using representative dN/dz distributions from a subset of their IR-selected AGN sample, which complicates direct comparison. G. C. Petter et al. (2023) find, as have previous photometric AGN surveys from IR selection (E. Donoso et al. 2014; M. A. DiPompeo et al. 2017), that obscured AGN are in significantly higher mass halos than their unobscured counterparts. But comparisons with these studies are not straightforward since many of these analyses (such as R. Gilli et al. 2009; E. Donoso et al. 2014; M. A. DiPompeo et al. 2016, 2017; G. C. Petter et al. 2023) span redshift ranges ($z \in 0$–3) that we do not probe in this paper. In addition to differences in redshift range and approach, we also utilize a different color selection, which also precludes direct comparisons.

We find that the $g - {\rm W3}$ color versus redshift selected unobscured + reddened AGN in our sample are in more massive halos than the obscured AGN, by a factor of ∼3, with a statistical significance $>3\sigma$ (2.8σ with the full covariance matrix). We also find that the host halos of unobscured AGN are more massive than the halos of reddened AGN by a factor of ∼3, at a statistical significance of 3σ (2.9σ with the C_{jk}), and the unobscured AGN are $\sim 4.5\times$ more massive than the obscured AGN at 5σ (3.5σ), where the central values of the inferred AGN subtype M_h shift by $<1\sigma$ in all cases. We note that the systematic uncertainty in our inferred average M_h associated with our reliance on photometric redshifts can be as large as the significant difference we see between our subsamples with a $\Delta z \sim +0.06$ systematic shift in the dN/dz (as visualized by the low-opacity red rectangle in Figure 12).

We cannot rule out that such small systematic shifts in the dN/dz could drive our and previous (photo-z based) results regarding M_h for obscured versus unobscured AGN (see Section 4.4).

We leave open the possibility that some objects can be misclassified and included in our AGN sample. Objects such as lower-mass star-forming galaxies (K. N. Hainline et al. 2016) could enter our selection, but only if the redshift solution is incorrect such that the estimated luminosity is above our threshold, and if they are sufficiently bright in the MIR to appear to be similar to an AGN. These tight observational constraints significantly limit the parameter space from which non-AGN may enter our sample. Though we are confident our classification methods are robust, further observations are key to reducing the possibility of this potential systematic effect.

We conclude that additional spectroscopic measurements are necessary to determine whether the differences in halo mass by obscuration are real. If so, we may have found evidence linking obscuration level with different phases in AGN–host galaxy coevolution, as others have proposed. Such differences may arise naturally if (for instance) earlier phases of galaxy and BH growth are associated with lower-mass halos and higher levels of obscuration (e.g., P. F. Hopkins et al. 2007; R. C. Hickox et al. 2009; V. A. Fawcett et al. 2023). While photometric surveys provide large sample statistics from which to perform clustering measurements, spectroscopic information is key to robust AGN identification, classification, and redshift determination. Large-format spectroscopic surveys, like the upcoming survey of the Prime Focus Spectrograph (M. Takada et al. 2014; J. Greene et al. 2022) on the Subaru Telescope, will be essential in further investigating the ensemble properties of AGN and reducing the effect of systematics.

6. Conclusions

We present a correlation analysis between LRGs observed by HSC and AGN selected from HSC and WISE photometry at angular scales of $0\overset{\prime}{.}1 < \theta < 200'$. These AGN are selected with a combination of HSC optical and WISE MIR photometric colors, and their classification has been shown to be robust with spectroscopic confirmation (R. E. Hviding et al. 2024). Using three HSC fields totaling $\sim$600 deg^2, we have a total of 1.7×10^6 LRGs and $\sim$34,000 AGN in the full redshift and luminosity range we analyze ($z \in 0.6$–1.2, $L_{6\,\mu\mathrm{m}} > 3 \times 10^{44}$ erg s^{-1}). For the AGN subtype cross-correlation clustering analysis, we use a single redshift bin ($z \in 0.7$–1.0) containing 1.5×10^6 LRGs and $\sim$28,500 $L_{6\,\mu\mathrm{m}}$-limited AGN. We fit these correlation functions with a linear + nonlinear DM halo model at physical scales $s > 1\,h^{-1}$ Mpc ($\theta \gtrsim 3'$), and interpret the clustering strength with physical parameters. Our principal conclusions are as follows.

1. We find no significant evidence for luminosity dependence on the inferred halo mass where AGN reside.
2. The host halos of unobscured + reddened AGN are $\sim 3\times$ more massive than those of obscured AGN, at a 3σ statistical difference (2.8σ with the full covariance matrix). We also directly compare our samples of unobscured AGN (i.e., Type I quasars) and obscured AGN (i.e., Type II) and find the former are, on average, in $\sim 4.5\times$ more massive halos with $\log M_h = 13.3 \pm 0.1 \log h^{-1} M_\odot$, a 5σ statistical difference. If we use the full covariance matrix instead, we find that the M_h values are unchanged, with a 3.5σ statistical difference between unobscured and obscured AGN.
3. We find that reddened AGN (which we expect to have broad lines) are in halos of intermediate mass between unobscured and obscured AGN, at $\log M_h = 12.8^{+0.1}_{-0.2} \log h^{-1} M_\odot$. As such, the halos of unobscured objects are $\sim 3\times$ more massive than those of reddened AGN, at 3σ (given statistical uncertainties from the diagonal of the covariance matrix). This result requires additional spectroscopic follow-up to better characterize the AGN samples, but could point to an evolutionary sequence between these AGN that is traced by the average halo mass.

We investigate the coevolution of SMBHs and their galactic hosts, and find that inferred average halo masses continue to be an effective means of tracing average AGN properties. By holding the luminosity and redshift distributions relatively constant between AGN subsamples, we are able to infer the differences between AGN subtypes as a function of our photometric color and redshift classification. However, possible unconstrained dN/dz systematic uncertainties prevent us from concluding that these inferred differences represent the underlying distribution with certainty, if they were to only affect the lower halo mass results. HOD analyses are essential to exploit the small-scale clustering measurements we have shown here, and will be a promising extension of this study to learn about AGN satellite fractions. Additional work is necessary to continue constraining the properties of the obscured AGN, but the results here show that they may include a significant population of lower halo mass objects, relative to the unobscured + reddened AGN. Evolutionary models and semiempirical approaches like those outlined by J. Aird & A. L. Coil (2021) would be useful in constructing simulations where we may test for possible halo mass differences. Inferring these AGN's characteristics has been an area of active research (G. C. Petter et al. 2022; V. A. Fawcett et al. 2023), and further observational data is key to understanding their morphology and potential place in the AGN evolutionary picture.

The data underlying this analysis are set to improve significantly in the coming years. Upcoming wide-field photometric surveys, like the Rubin Observatory's Legacy Survey of Space and Time (Z. Ivezić et al. 2019) and eROSITA (A. Merloni et al. 2012), and spectroscopic surveys such as the Prime Focus Spectrograph survey (M. Takada et al. 2014; J. Greene et al. 2022) and the Dark Energy Spectroscopic Instrument survey (DESI Collaboration et al. 2016, 2024) will be particularly capable of revealing AGN (and AGN subtype) populations. From disentangling AGN properties as a function of time and evolutionary stage, to better characterizing the spectral properties of AGN and the galaxies in which they reside, new data sets will readily enable extensions of the work presented here.

Acknowledgments

We thank the anonymous referee for the many helpful comments. We thank R. Lupton, P. Melchior, N. E. Chisari, and N. Bahcall for helpful conversations throughout the course of this work.

Computing was performed using the Princeton Research Computing resources at Princeton University. R.C.R. acknowledges support from the Ford Foundation Predoctoral Fellowship from the National Academy of Sciences, Engineering, and Medicine. A.D.G. and R.C.R. gratefully acknowledge support from the NASA Astrophysics Data Analysis Program #80NSSC23K0485. A.D.G. and J.E.G. acknowledge support from the National Science Foundation under grant No. AST-1613744, and J.E.G. acknowledges support from the National Science Foundation under grant No. AST-2306950.

The HSC Collaboration includes the astronomical communities of Japan and Taiwan, and Princeton University. The HSC instrumentation and software were developed by the National Astronomical Observatory of Japan (NAOJ), the Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), the University of Tokyo, the High Energy Accelerator Research Organization (KEK), the Academia Sinica Institute for Astronomy and Astrophysics in Taiwan (ASIAA), and Princeton University. Funding was contributed by the FIRST program from the Japanese Cabinet Office, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Japan Society for the Promotion of Science (JSPS), the Japan Science and Technology Agency, the Toray Science Foundation, NAOJ, Kavli IPMU, KEK, ASIAA, and Princeton University.

This paper makes use of software developed for the Vera C. Rubin Observatory. We thank the Rubin Observatory for making their code available as free software at http://pipelines.lsst.io/. This paper is based on data collected at the Subaru Telescope and retrieved from the HSC data archive system, which is operated by the Subaru Telescope and Astronomy Data Center at NAOJ. Data analysis was in part carried out with the cooperation of the Center for Computational Astrophysics (CfCA), NAOJ.

This publication makes use of data products from WISE, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration.

This publication also makes use of data products from NEOWISE, which is a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the Planetary Science Division of the National Aeronautics and Space Administration.

We are honored and grateful for the opportunity of observing the Universe from Maunakea, which has cultural, historical, and natural significance in Hawaii.

Facilities: Subaru (HSC), WISE, NEOWISE, Sloan.

Software: `Astropy` (Astropy Collaboration et al. 2018, 2022), `Matplotlib` (J. D. Hunter 2007), `NumPy` (S. van der Walt et al. 2011; C. R. Harris et al. 2020), `SciPy` (P. Virtanen et al. 2020), `Corrfunc` (M. Sinha & L. H. Garrison 2020), Core Cosmology Library (N. E. Chisari et al. 2019).

Appendix A
LRG Selection

We have defined our parent HSC galaxy samples as objects brighter than $i = 24$ AB mag, using three narrow photometric redshift bins, and a wide redshift bin. We determine the bin widths based on the median uncertainty from the AGN sample's photometric redshifts (see Section 3.3). As a check on the photometric redshift accuracy, we calculate the cross-correlation between the galaxy sample in the first and third narrow redshift bins (i.e., before color cuts). If our photometric redshifts and their errors are reliable, we should not detect much signal above the random in the cross-correlation. Nonetheless, by estimating the amplitude of the cross-bin correlation relative to the autocorrelations, we find that there is ∼30% correlation between the full magnitude-limited galaxy samples in the 0.6–0.8 and 1.0–1.2 redshift bins. In addition, the autocorrelation of this sample displays artificial signal on large scales, where the true signal is about 1% of the small-scale amplitude and we become systematic-limited. These findings require that we prune the galaxy catalog and identify the source of these systematic limitations to our measurement.

To fix this, we adopt a color space sample selection to isolate LRGs. LRGs have been found to have more reliable redshifts, including by prior HSC analyses (D. J. Eisenstein et al. 2001; M. M. Rau et al. 2023). This is primarily due to the presence of

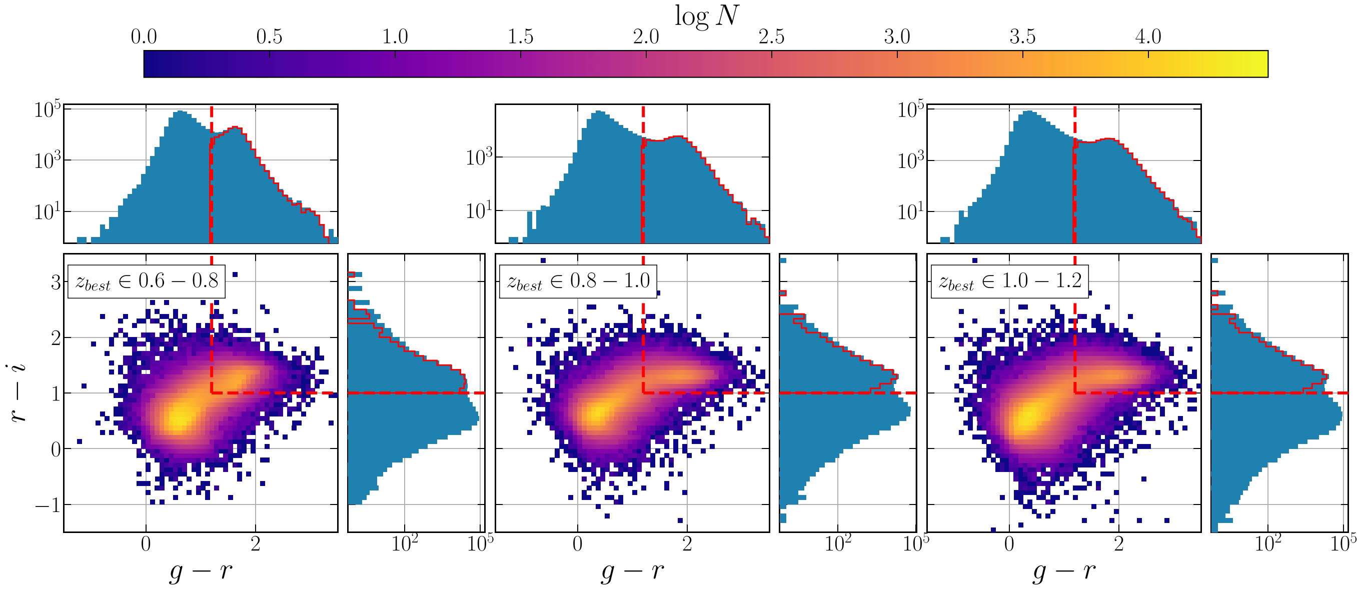

Figure 13. HSC color–color diagram for the HSC galaxy sample in the XMM field, across three narrow redshift bins. As described in Appendix A, we isolate an LRG sample from the whole galaxy catalog via a color–color cut, $g - r > 1.2$ and $r - i > 1.0$ (as indicated by the red dashed lines). The objects are isolated in a specific redshift bin given the fiducial redshift ($z_{\rm best}$) for every object. The histograms illustrate the distribution of galaxies in color space, and illustrate the scale at which we separate the LRG population from the rest of the galaxies. The red solid lines in the histograms highlight the color distributions of the final selection of objects.

the 4000 Å Balmer break in the galaxy SED, characteristic of LRGs, which allows photometric redshift codes to find a more stable solution. Using the measured `CModel` AB magnitude for HSC galaxies, we select objects with $g - r > 1.2$ and $r - i > 1.0$ colors to isolate the LRGs readily seen in color–color diagrams, shown in the color space representation of galaxies in the XMM field in Figure 13. Given the HSC photometric bands, this color selection isolates LRGs whose redshifts are close to the $0.7 \lesssim z \lesssim 1.0$ range. We find the number of objects is reduced significantly, removing between 80% and 95% of the total galaxy sample (for the lowest- and highest-redshift bins, respectively). The inferred stellar mass available in the `Mizuki` catalog for the complete $(i > 24)$ galaxy population peaks at $\sim 9 \times 10^9 M_\odot$, while for the red galaxy sample the peak is at $\sim 6 \times 10^{10} M_\odot$. After this cut, the cross-correlation between the red galaxy samples in the lowest ($z \in 0.6$–0.8) and highest ($z \in 1.0$–1.2) redshift bins is ∼10% the amplitude of the autocorrelation in a single bin. We find this reduction in systematic uncertainty to be sufficient for our analysis and propagate the color cut throughout the rest of the experiment.

The per redshift bin LRG sample dN/dz is illustrated in the top row of Figure 5. The multiple peaks in each redshift bin are reflective of unphysical fitting degeneracies in the HSC photometric redshift code, wherein certain values are overprescribed. There are also secondary peaks outside the redshift bin after the weighted addition, indicative of a significant number of objects whose $p(z)$ is clustered outside (but significantly overlaps with) the bin. We test how these dN/dz shapes affect our forward model, and find a Gaussian-smoothed version of these distributions produces forward models that diverge from the fiducial model by <0.5%. These distributions are representative of the estimated redshifts and their uncertainties in each bin.

Appendix B
Proper Accounting of *p*(*z*) Uncertainties

The standard tomographic method for redshift binning takes into account only those galaxies whose fiducial redshift is in the bin. This procedure would, in a given tomographic redshift bin, exclude objects whose nominal redshift ($z_{\rm best}$) was not inside the bounds of the bin, even if a significant fraction of their $p(z)$ did fall in the bin. We seek to account for the complete redshift uncertainty in our sample, for objects that both scatter into and out of our redshift bins. As such, we implement the following formalism to create a more representative clustering statistic taking this into account.

B.1. Redshift-weighted Correlation Function

We define a weight $\mathcal{W}_i$ based on the fraction of an object's $p(z)$ found within a bin

$$\mathcal{W}_i = \frac{\int_a^b p_i(z)\, dz}{\int_0^\infty p_i(z) dz} \tag{B1}$$

for a redshift bin with edges a and b. We define our binning selection such that $\mathcal{W}_i \geqslant \mathcal{W}_0$, where $\mathcal{W}_0$ is the threshold we establish for inclusion in our redshift bin. In this analysis, any object with greater than 3% of the probability $p(z)$ falling inside the predefined bin is included. This choice reflects a balance of the inclusion of all available redshift space while not making the calculation intractable.

We implement this weight for a pair counting operation such that

$$\mathcal{W}_{XY}(\theta) = \frac{\sum_i^{N_X}\sum_{j\neq i}^{N_Y}\mathcal{W}_{i,X}\mathcal{W}_{j,Y}}{\sum_i^{N_X}\sum_{j\neq i}^{N_Y}\bar{\Theta}_{ij,k}} \tag{B2}$$

and substitute the total number of objects for the summed weights:

$$\mathcal{W}_X = \sum_i^{N_X} \mathcal{W}_{i,X}. \tag{B3}$$

The correlation function element is thus defined as

$$XY(\theta) = \frac{\mathcal{W}_{XY}(\theta) XY'(\theta)}{\mathcal{W}_X \mathcal{W}_Y}. \tag{B4}$$

We substitute this expression into the S. D. Landy & A. S. Szalay (1993) estimator in Equation (1) to account for appropriate object weighting.

B.2. Redshift-weighted Forward Modeling

We also update the dN/dz calculation to include the appropriate per-object weighting, where

$$\frac{dN}{dz} = \sum_i \mathcal{W}_i \cdot p_i(z \mid \mathcal{W}_i \geqslant \mathcal{W}_0). \tag{B5}$$

These are used in the forward model as described in Section 3.2.

Following significant testing and comparison with standard tomographic binning, we find that the two methods return values that are entirely consistent ($<1\sigma$) with each other. Nevertheless we implement our weighted analysis throughout this work, preferring to fold in all the available $p(z)$ information. As a result, this method is less subject to systematic biases inherent in photometric redshift fitting.

Appendix C
All AGN Subtypes∍ *dN*/*dz*

Figure 14 highlights the dN/dz for each of the AGN subtype samples in the principal analysis, for the three HSC fields considered. We find that the AGN dN/dz show multiple peaks and departures from a normal distribution, possibly indicative of fitting degeneracies in the AGN photo-z pipeline. As discussed in Section 3.2, each of these dN/dz is used to calculate the $\omega(\theta)$ DM model for the particular subset, then used to fit the measured correlation.

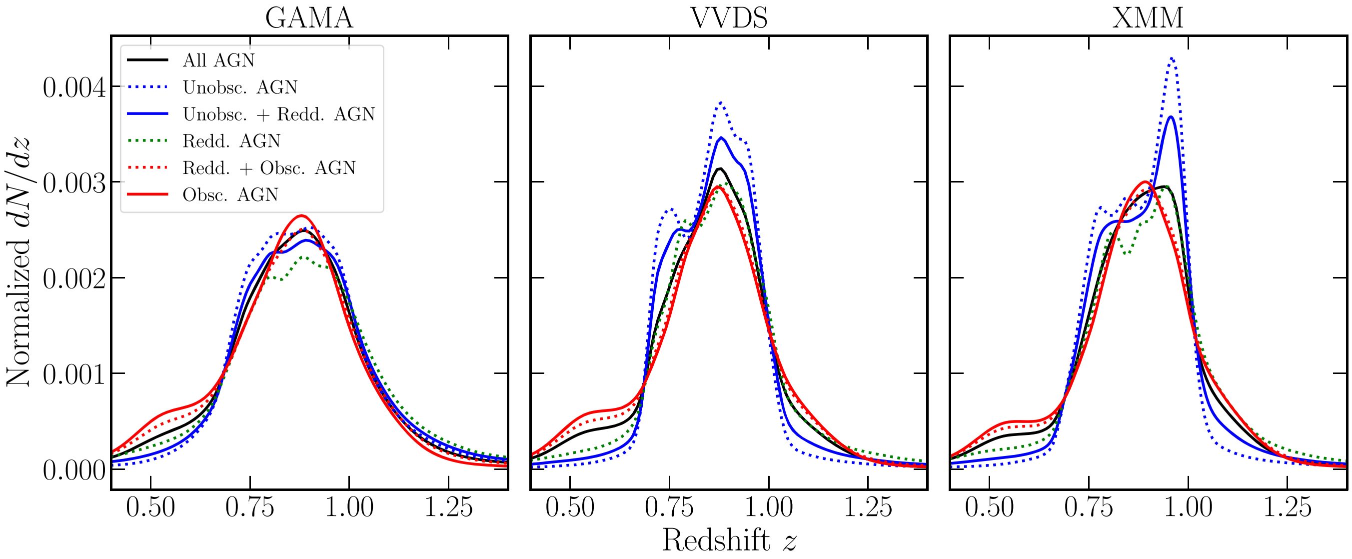

Figure 14. Normalized dN/dz from all AGN subtype samples for cross-correlation in the wide redshift bin ($z \in 0.7$–1.0) for the $L_{6\ \mu m} > 3 \times 10^{44}$ erg s^{-1} threshold, split by HSC field.

Appendix D
Full Covariance Matrix χ^2 Minimization Fit Results

Table 3 shows the fit results for the complete experiment when using the full covariance matrix, rather than 1D uncertainties from the square root of the diagonal of the covariance matrix (as shown in Table 2). We describe in Section 3 how this result incorporates the off-diagonal elements showcasing the bin-to-bin correlations, and thus more accurately represents the precision of our measurements. The off-diagonal elements of the cross-correlations' covariance matrices have amplitudes of order 0.2–0.9 of the value of the diagonal elements. However, for ease of comparison with other analyses of AGN clustering using $\omega(\theta)$ (see L. Koutoulidis et al. 2013, 2018; M. A. DiPompeo et al. 2014, 2016, 2017; G. C. Petter et al. 2023), we represent the main results using the diagonal elements alongside the effect on the statistical significance when using the full covariance.

Table 3
Angular Correlation Function Fit Results (Full Covariance Matrix)

Subset	$N_{\rm obj}$	Weighted $N_{\rm obj}$	$\langle L_{6\ \mu m}\rangle$ (log erg s^{-1})	$\langle z\rangle$	$\langle\chi^2\rangle$ (10 dof)	b	$\langle M_h\rangle$ (log h^{-1} $M_\odot$)
$z \in 0.6$–0.8							
LRGs	1,288,589	879,258.8	⋯	0.7 ± 0.1	32.8	$1.83^{+0.04}_{-0.05}$	⋯
All AGN	22,988	5804.3	$44.7^{+0.4}_{-0.2}$[a]	$0.8^{+0.1}_{-0.2}$	15.8	1.7 ± 0.1	13.2 ± 0.1
$z \in 0.8-1.0$							
LRGs	851,117	440,970.6	⋯	0.9 ± 0.1	28.1	$2.1^{+0.1}_{-0.2}$	⋯
All AGN	26,264	10,381.3	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	29.4	1.5 ± 0.1	12.8 ± 0.1
$z \in 1.0-1.2$							
LRGs	324,790	98,498.4	⋯	1.0 ± 0.1	21.2	$2.43^{+0.01}_{-0.03}$	⋯
All AGN	25,235	7204.7	$44.9^{+0.4}_{-0.3}$[a]	1.1 ± 0.2	10.3	1.6 ± 0.1	12.9 ± 0.1
$z \in 0.7-1.0$							
LRGs	1,509,905	843,166.6	⋯	0.8 ± 0.1	19.2	1.9 ± 0.1	⋯
All AGN	28,494	13,898.8	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	30.7	1.4 ± 0.2	$13.2^{+0.1}_{-0.2}$
Unobscured AGN	8266	3942.0	$44.9^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	22.3	2.3 ± 0.2	13.4 ± 0.1
Unobscured + Reddened AGN	15,156	6745.71	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	53.3	2.1 ± 0.2	$13.2^{+0.1}_{-0.2}$
Reddened AGN	6890	2803.6	$44.8^{+0.5}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	16.7	1.6 ± 0.2	$13.0^{+0.1}_{-0.2}$
Reddened + Obscured AGN	19,675	9893.4	$44.8^{+0.4}_{-0.2}$[a]	$0.9^{+0.1}_{-0.2}$	29.2	1.4 ± 0.2	12.7 ± 0.2
Obscured AGN	12,785	7089.8	$44.7^{+0.3}_{-0.2}$[a]	0.9 ± 0.2	36.7	1.2 ± 0.2	$12.6^{+0.2}_{-0.3}$
High-$L_{6\ \mu m}$ AGN	7760	2492.2	$45.1^{+0.4}_{-0.2}$[b]	$1.0.^{+0.1}_{-0.2}$	10.4	1.3 ± 0.2	$12.6^{+0.2}_{-0.3}$
Low-$L_{6\ \mu m}$ AGN	20,734	11,406.5	$44.8^{+0.3}_{-0.2}$[c]	$0.9^{+0.1}_{-0.2}$	23.4	1.5 ± 0.2	$13.0^{+0.1}_{-0.2}$

Notes.
[a] Primary luminous AGN selection ($L_{6\ \mu m} > 3 \times 10^{44}$ erg s^{-1}).
[b] Higher-luminosity AGN selection ($L_{6\ \mu m} > 10^{45}$ erg s^{-1}).
[c] Lower-luminosity AGN selection ($3 \times 10^{44} < L_{6\ \mu m} < 10^{45}$ erg s^{-1}).

Appendix E
Per-field Correlation Results (1D Uncertainties)

Table 4 shows the per-field measurements of the galaxy and AGN bias given the 1D uncertainties, as well as the inferred average halo mass. We take these measurements and combine them with an inverse variance weighted mean to produce our nominal results, as shown in Table 2.

Table 4
Angular Correlation Function Fit Results for Each HSC Field

Subset	Field	$N_{\rm obj}$	Weighted $N_{\rm obj}$	$\langle z \rangle$	χ^2 (10 dof)	b	$\langle M_h \rangle$ ($\log h^{-1} M_\odot$)
$z \in 0.6-0.8$							
All AGN	GAMA	15,494	3877.6	$0.8^{+0.1}_{-0.3}$	5.4	1.6 ± 0.1	12.9 ± 0.1
All AGN	VVDS	3815	1050.9	$0.8^{+0.1}_{-0.3}$	10.4	1.3 ± 0.2	$12.5^{+0.3}_{-0.4}$
All AGN	XMM	3679	875.7	$0.8^{+0.1}_{-0.3}$	2.7	1.3 ± 0.2	$12.5^{+0.3}_{-0.4}$
$z \in 0.8-1.0$							
All AGN	GAMA	17,785	6444.19	$0.9^{+0.1}_{-0.2}$	1.0	1.5 ± 0.1	12.7 ± 0.1
All AGN	VVDS	4221	1936.76	0.9 ± 0.1	7.3	0.8 ± 0.2	$11.2^{+0.5}_{-0.6}$
All AGN	XMM	4258	2000.31	0.9 ± 0.1	5.8	1.7 ± 0.2	12.9 ± 0.2
$z \in 1.0-1.2$							
All AGN	GAMA	17,110	4618.77	1.1 ± 0.2	4.9	1.4 ± 0.1	12.4 ± 0.2
All AGN	VVDS	3907	1200.99	1.0 ± 0.2	1.0	1.5 ± 0.2	$12.5^{+0.3}_{-0.4}$
All AGN	XMM	4218	1384.91	1.0 ± 0.2	13.2	3.1 ± 0.3	13.5 ± 0.3
$z \in 0.7-1.0$							
All AGN	GAMA	19,442	8748.86	0.9 ± 0.2	7.6	1.8 ± 0.1	13.0 ± 0.1
All AGN	VVDS	4540	2595.26	$0.9^{+0.1}_{-0.2}$	6.7	1.1 ± 0.2	$12.1^{+0.3}_{-0.5}$
All AGN	XMM	4512	2554.68	$0.9^{+0.1}_{-0.2}$	5.9	1.7 ± 0.2	12.9 ± 0.2
Unobscured AGN	GAMA	6510	2800.22	$0.9^{+0.1}_{-0.2}$	8.9	2.2 ± 0.2	13.2 ± 0.1
Unobscured AGN	VVDS	863	590.09	0.9 ± 0.1	3.0	2.0 ± 0.3	$13.1^{+0.2}_{-0.3}$
Unobscured AGN	XMM	893	551.78	0.9 ± 0.1	9.4	2.4 ± 0.3	$13.4^{+0.1}_{-0.2}$
Unobscured + Reddened AGN	GAMA	11,610	4761.98	$0.9^{+0.1}_{-0.2}$	5.0	2.0 ± 0.1	13.1 ± 0.1
Unobscured + Reddened AGN	VVDS	1783	1029.33	$0.9^{+0.1}_{-0.2}$	5.4	1.5 ± 0.2	$12.7^{+0.2}_{-0.3}$
Unobscured + Reddened AGN	XMM	1763	954.39	$0.9^{+0.1}_{-0.2}$	9.9	2.1 ± 0.2	13.2 ± 0.2
Reddened AGN	GAMA	5100	1961.77	$1.0^{+0.1}_{-0.2}$	2.7	1.6 ± 0.2	12.9 ± 0.2
Reddened AGN	VVDS	920	439.24	0.9 ± 0.1	3.1	1.6 ± 0.3	$12.8^{+0.3}_{-0.4}$
Reddened AGN	XMM	870	402.62	0.9 ± 0.1	8.2	$1.7^{+0.3}_{-0.4}$	$12.9^{+0.3}_{-0.5}$
Reddened + Obscured AGN	GAMA	12,491	5897.68	0.9 ± 0.2	5.0	1.6 ± 0.1	$12.8^{+0.1}_{-0.2}$
Reddened + Obscured AGN	VVDS	3617	1998.54	$0.9^{+0.1}_{-0.2}$	8.2	1.0 ± 0.2	$11.8^{+0.4}_{-0.6}$
Reddened + Obscured AGN	XMM	3567	1997.2	$0.9^{+0.1}_{-0.2}$	4.2	1.4 ± 0.2	$12.6^{+0.2}_{-0.3}$
Obscured AGN	GAMA	7391	3935.91	0.9 ± 0.2	6.4	1.5 ± 0.1	12.7 ± 0.2
Obscured AGN	VVDS	2697	1559.3	0.9 ± 0.2	9.8	1.0 ± 0.2	$11.8^{+0.4}_{-0.6}$
Obscured AGN	XMM	2697	1594.58	$0.9^{+0.1}_{-0.2}$	6.9	1.4 ± 0.2	$12.6^{+0.2}_{-0.3}$
High-$L_{6\,\mu{\rm m}}$ AGN	GAMA	5796	1631.12	$1.0^{+0.2}_{-0.1}$	4.5	1.3 ± 0.2	$12.4^{+0.3}_{-0.4}$
High-$L_{6\,\mu{\rm m}}$ AGN	VVDS	1153	514.76	0.9 ± 0.1	3.6	$1.3^{+0.3}_{-0.4}$	$12.5^{+0.4}_{-0.7}$
High-$L_{6\,\mu{\rm m}}$ AGN	XMM	811	346.38	1.0 ± 0.1	2.9	$1.7^{+0.3}_{-0.4}$	$12.8^{+0.3}_{-0.5}$
Low-$L_{6\,\mu{\rm m}}$ AGN	GAMA	13,646	7117.74	$0.9^{+0.1}_{-0.2}$	7.4	1.8 ± 0.1	13.0 ± 0.1
Low-$L_{6\,\mu{\rm m}}$ AGN	VVDS	3387	2080.5	$0.9^{+0.1}_{-0.2}$	5.9	1.3 ± 0.2	$12.4^{+0.3}_{-0.4}$
Low-$L_{6\,\mu{\rm m}}$ AGN	XMM	3701	2208.31	$0.9^{+0.1}_{-0.2}$	6.0	1.7 ± 0.2	12.9 ± 0.2

ORCID iDs

Rodrigo Córdova Rosado https://orcid.org/0000-0002-7967-7676
Andy D. Goulding https://orcid.org/0000-0003-4700-663X
Jenny E. Greene https://orcid.org/0000-0002-5612-3427
Grayson C. Petter https://orcid.org/0000-0001-6941-8411
Ryan C. Hickox https://orcid.org/0000-0003-1468-9526
Nickolas Kokron https://orcid.org/0000-0002-5808-4708
Michael A. Strauss https://orcid.org/0000-0002-0106-7755
Jahmour J. Givans https://orcid.org/0000-0002-5870-6108
Yoshiki Toba https://orcid.org/0000-0002-3531-7863
Cassandra Starr Henderson https://orcid.org/0000-0003-4684-608X

References

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3
Aihara, H., AlSayyad, Y., Ando, M., et al. 2022, PASJ, 74, 247
Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
Aird, J., & Coil, A. L. 2021, MNRAS, 502, 5962
Allevato, V., Finoguenov, A., Cappelluti, N., et al. 2011, ApJ, 736, 99
Allevato, V., Finoguenov, A., Civano, F., et al. 2014, ApJ, 796, 4
Almeida, C. R., & Ricci, C. 2017, NatAs, 1, 679
Andreani, P., & Cristiani, S. 1992, ApJ, 398, L13
Antonucci, R. 1993, ARA&A, 31, 473
Arita, J., Kashikawa, N., Matsuoka, Y., et al. 2023, ApJ, 954, 210
Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Awan, H., & Gawiser, E. 2020, ApJ, 890, 78
Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587
Blain, A. W., Jameson, A., Smail, I., et al. 1999, MNRAS, 309, 715
Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S5
Cappelluti, N., Ajello, M., Burlon, D., et al. 2010, ApJL, 716, L209
Cappelluti, N., Allevato, V., & Finoguenov, A. 2012, AdAst, 2012, e853701
Chisari, N. E., Alonso, D., Krause, E., et al. 2019, ApJS, 242, 2
Coil, A. L., Georgakakis, A., Newman, J. A., et al. 2009, ApJ, 701, 1484
Coupon, J., Czakon, N., Bosch, J., et al. 2018, PASJ, 70, S7
Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, MNRAS, 356, 415
Croom, S. M., & Shanks, T. 1996, MNRAS, 281, 893
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2012, yCat, II/281
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, yCat, II/328
Dalal, R., Li, X., Nicola, A., et al. 2023, PhRvD, 108, 123519
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024, arXiv:2404.03000
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
DiPompeo, M. A., Hickox, R. C., Eftekharzadeh, S., & Myers, A. D. 2017, MNRAS, 469, 4630
DiPompeo, M. A., Hickox, R. C., & Myers, A. D. 2016, MNRAS, 456, 924
DiPompeo, M. A., Myers, A. D., Hickox, R. C., Geach, J. E., & Hainline, K. N. 2014, MNRAS, 442, 3443
Donoso, E., Yan, L., Stern, D., & Assef, R. J. 2014, ApJ, 789, 44
Draine, B. T. 2003, ARA&A, 41, 241
Eisenstein, D. J., Annis, J., Gunn, J. E., et al. 2001, AJ, 122, 2267
Eisenstein, D. J., & Zaldarriaga, M. 2001, ApJ, 546, 2
Ellison, S. L., Mendel, J. T., Patton, D. R., & Scudder, J. M. 2013, MNRAS, 435, 3627
Ellison, S. L., Patton, D. R., Mendel, J. T., & Scudder, J. M. 2011, MNRAS, 418, 2043
Ellison, S. L., Viswanathan, A., Patton, D. R., et al. 2019, MNRAS, 487, 2491
Fawcett, V. A., Alexander, D. M., Brodzeller, A., et al. 2023, MNRAS, 525, 5575
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1
Georgakakis, A., Comparat, J., Merloni, A., et al. 2019, MNRAS, 487, 275
Gilli, R., Zamorani, G., Miyaji, T., et al. 2009, A&A, 494, 33
Glikman, E., Simmons, B., Mailly, M., et al. 2015, ApJ, 806, 218
Goulding, A. D., & Alexander, D. M. 2009, MNRAS, 398, 1165
Goulding, A. D., Alexander, D. M., Bauer, F. E., et al. 2012, ApJ, 755, 5
Goulding, A. D., Greene, J. E., Bezanson, R., et al. 2018, PASJ, 70, S37
Greene, J., Bezanson, R., Ouchi, M., Silverman, J. & the PFS Galaxy Evolution Working Group 2022, arXiv:2206.14908
Groth, E. J., & Peebles, P. J. E. 1977, ApJ, 217, 385
Hainline, K. N., Reines, A. E., Greene, J. E., & Stern, D. 2016, ApJ, 832, 119
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
He, W., Akiyama, M., Bosch, J., et al. 2018, PASJ, 70, S33
Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589
Hennawi, J. F., Strauss, M. A., Oguri, M., et al. 2006, AJ, 131, 1
Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625
Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, ApJ, 696, 891
Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9
Hickox, R. C., Myers, A. D., Brodwin, M., et al. 2011, ApJ, 731, 117
Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, ApJS, 175, 356
Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731
Hunter, J. D. 2007, CSE, 9, 90
Hviding, R. E., Hainline, K. N., Goulding, A. D., & Greene, J. E. 2024, AJ, 167, 169
Hviding, R. E., Hainline, K. N., Rieke, M., et al. 2022, AJ, 163, 224
Iovino, A., & Shaver, P. A. 1988, ApJ, 330, L13
Ishikawa, S., Okumura, T., Oguri, M., & Lin, S.-C. 2021, ApJ, 922, 23
Ivezić, Z., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Jarrett, T. H., Cohen, M., Masci, F., et al. 2011, ApJ, 735, 112
Jiang, N., Wang, H., Mo, H., et al. 2016, ApJ, 832, 111
Kawanomoto, S., Uraguchi, F., Komiyama, Y., et al. 2018, PASJ, 70, 66
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581
Koss, M., Mushotzky, R., Veilleux, S., & Winter, L. 2010, ApJL, 716, L125
Koutoulidis, L., Georgantopoulos, I., Mountrichas, G., et al. 2018, MNRAS, 481, 3063
Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, MNRAS, 428, 1382
Krumpe, M., Miyaji, T., Coil, A., et al. 2014, in The X-ray Universe 2014, ed. J.-U. Ness (Madrid: ESA)
Krumpe, M., Miyaji, T., Coil, A. L., & Aceves, H. 2018, MNRAS, 474, 1773
La Franca, F., Andreani, P., & Cristiani, S. 1998, ApJ, 497, 529
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Laurent, P., Eftekharzadeh, S., Goff, J.-M. L., et al. 2017, JCAP, 2017, 017
Li, Q., Kilbinger, M., Luo, W., et al. 2024, ApJ, 969, L25
Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006, ApJ, 641, 41
Limber, D. N. 1953, ApJ, 117, 134
Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8
Mainzer, A., Bauer, J., Grav, T., et al. 2011, ApJ, 731, 53
McInnes, L., Healy, J., & Melville, J. 2018, arXiv:1802.03426
Melchior, P. 2021, Skymapper: Mapping astronomical survey data on the sky, Astrophysics Source Code Library, ascl:2107.007
Mendez, A. J., Coil, A. L., Aird, J., et al. 2016, ApJ, 821, 55
Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv:1209.3114
Mihos, J. C., & Hernquist, L. 1994, ApJ, 425, L13
Mihos, J. C., & Hernquist, L. 1996, ApJ, 464, 641
Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, PASJ, 70, S1
Mo, H. J., & Fang, L. Z. 1993, ApJ, 410, 493
Netzer, H. 2015, ARA&A, 53, 365
Nicola, A., Alonso, D., Sánchez, J., et al. 2020, JCAP, 2020, 044
Norberg, P., Baugh, C. M., Gaztañaga, E., & Croton, D. J. 2009, MNRAS, 396, 19
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Osmer, P. S. 1981, ApJ, 247, 762
Peacock, J. A. 1991, MNRAS, 253, 1P
Peebles, P. J. E. 1973, ApJ, 185, 413
Petter, G. C., Hickox, R. C., Alexander, D. M., et al. 2022, ApJ, 927, 16
Petter, G. C., Hickox, R. C., Alexander, D. M., et al. 2023, ApJ, 946, 27
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
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
Rau, M. M., Dalal, R., Zhang, T., et al. 2023, MNRAS, 524, 5109
Ricci, C., Privon, G. C., Pfeifle, R. W., et al. 2021, MNRAS, 506, 5935
Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, ApJ, 697, 1634
Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988, ApJ, 325, 74
Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schmidt, M. 1963, Natur, 197, 1040
Secrest, N. J., Ellison, S. L., Satyapal, S., & Blecha, L. 2020, MNRAS, 499, 2380
Serber, W., Bahcall, N., Ménard, B., & Richards, G. 2006, ApJ, 643, 68

Shanks, T., & Boyle, B. J. 1994, MNRAS, 271, 753
Shanks, T., Fong, R., Boyle, B. J., & Peterson, B. A. 1987, MNRAS, 227, 739
Shaver, P. A. 1984, A&A, 136, L9
Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, ApJ, 680, 169
Shen, Y., McBride, C. K., White, M., et al. 2013, ApJ, 778, 98
Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, ApJ, 697, 1656
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119
Sinha, M., & Garrison, L. H. 2020, MNRAS, 491, 3022
Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30
Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, ApJ, 761, 152
Tanaka, M. 2015, ApJ, 801, 20
Tanaka, M., Coupon, J., Hsieh, B.-C., et al. 2018, PASJ, 70, S9
Tegmark, M., & Peebles, P. J. E. 1998, ApJ, 500, L79
Timlin, J. D., Ross, N. P., Richards, G. T., et al. 2018, ApJ, 859, 20
Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878
Toba, Y., Nagao, T., Kajisawa, M., et al. 2017, ApJ, 835, 36
UnWISE Team 2021, unWISE Catalog, IPAC, doi:10.26131/IRSA525
Urrutia, T., Lacy, M., & Becker, R. H. 2008, ApJ, 674, 80
Urry, C. M., & Padovani, P. 1995, PASP, 107, 803
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Viitanen, A., Allevato, V., Finoguenov, A., et al. 2023, A&A, 674, A214
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2019, AllWISE Source Catalog, IPAC, doi:10.26131/IRSA1
Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004, ApJ, 608, 16
Zhou, R., Newman, J. A., Mao, Y.-Y., et al. 2021, MNRAS, 501, 3309