Summer 2024

The Large-scale Environments of Active Galactic Nuclei

Grayson C. Petter
Dartmouth College, grayson.c.petter.gr@dartmouth.edu

THE LARGE-SCALE ENVIRONMENTS OF ACTIVE GALACTIC NUCLEI

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

in

Physics and Astronomy

by Grayson Charles Petter

Guarini School of Graduate and Advanced Studies
Dartmouth College
Hanover, New Hampshire
May 2024

Examining Committee:

Ryan C. Hickox, Ph.D. (chair)

Brian C. Chaboyer, Ph.D.

John R. Thorstensen, Ph.D.

David M. Alexander, Ph.D.

F. Jon Kull, Ph.D.
Dean of the Guarini School of Graduate and Advanced Studies

Abstract

It is now recognized that the energy released by accreting supermassive black holes observed as Active Galactic Nuclei (AGN) is integral in shaping the dynamics of baryons on up to cosmological scales, and AGN thus play a significant role in regulating the formation and evolution of galaxies. Studying the clustering properties of AGN reveals which environments they release this feedback energy into, testing models of AGN-galaxy coevolution and AGN structure. In this thesis, I leverage wide-area photometric and spectroscopic survey data to measure the clustering properties of various AGN samples containing millions of systems, placing tight constraints on the properties of their host dark matter halos. I first investigate the host halo properties of optically-selected red and blue quasars, finding no significant difference in their environments. Thus, quasars buried under mild dust columns likely do not represent a special phase of AGN-galaxy coevolution. On the other hand, I show that heavily obscured infrared-selected quasars occupy systematically more massive halos than their unobscured counterparts, suggesting that obscured quasars may represent a special phase of AGN-galaxy coevolution in which the black hole and galaxy are fed by common gas streams. Finally, I show that luminous low-frequency radio galaxies are hosted by massive galaxy groups over cosmic time, implying that jet-mode feedback dominates over quasar wind-mode feedback in groups and clusters.

Preface

I firstly want to thank my thesis advisor, **Ryan Hickox**, who has been the best mentor I can imagine. I met Ryan in person at my first academic conference when he rushed over, out of breath from finishing his keynote speech two minutes earlier to offer kind words about the poster I was presenting. After coming to Dartmouth and publishing my first paper concerning radio observations of starburst galaxies with Ryan, I was eager to explore other fields, and was particularly interested in performing large statistical studies. Ryan took a chance and introduced me to the initially daunting world of large-scale structure, which I had little understanding of at the time. Years later, the subject of uncovering the large-scale environments hosting active galactic nuclei has become the most fulfilling intellectual journey I have undertaken, and I am very grateful. Ryan is an astonishing scientific advisor, somehow able without fail to get to the heart of a subject with a question, find the bug in a code, or guide a paper back to the big picture. Ryan is also a exceptionally kind and personable mentor, organizing research group outings, checking in on students, and showing his appreciation for the work we do.

I also want to thank my loving and supportive family, **Mom, Dad**, and **Graham** for raising me to value learning and hard work. I credit a lot of my love of exploring ideas to my mom reading to and with me from the time I was two years old and later having long talks about anything and everything on the porch. Thank you for supporting me in following something I am passionate about.

I want to thank my love, **Sarah Nichols**. We fortuitously met when I arrived in Vermont, and it is not exaggeration to say that I would not have made it through this experience without you. Thank you for keeping me somewhat sane, making me laugh, and being my best friend.

Finally, I am thankful to my professors, to my friends in the Hickox research group, to

all of my scientific collaborators especially **Dave Alexander** and **Adam Myers**, to my previous undergraduate advisors **Gregory Rudnick** and **Amanda Kepley**, and to my thesis committee for their advice throughout my graduate career and their time and feedback on this thesis document.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 AGN Are Cosmologically Relevant

The Λ-cold dark matter (Λ-CDM) cosmological model is remarkably successful in reproducing a wide array of astronomical observations, including the power spectrum of the cosmic microwave background (CMB) (e.g., Planck Collaboration et al. 2020a), the features of the power spectrum of galaxies (e.g., Alam et al. 2017), and the acceleration history of the universe (Riess et al. 1998). Galaxies form in a Λ-CDM universe when baryons gravitationally coupled to the dark matter density field collect in collapsed dark matter structures (halos) and condense into stars, and then grow hierarchically through repeated mergers (e.g., White & Rees 1978; White & Frenk 1991). Modern N-body simulations make exceptionally precise predictions of the distribution of dark matter as a function of time (e.g., Maksimova et al. 2021), as the underlying physics is purely gravitational and thus relatively simple. After identifying virialized halos in these simulations, one might expect to be able to predict the distribution and number density of galaxies with a very simple prescription: given the density of normal matter is $\sim 1/6$th that of dark matter, every halo should eventually form a galaxy with stellar mass order $\sim 1/6$th that of the host

halo. However, baryons interact with each other nongravitationally, giving rise to a highly nonlinear emergent physics which governs how galaxies form and evolve beyond purely cosmological accretion of gas.

A number of discrepancies began to emerge approximately three decades ago between observed galaxy populations and Λ-CDM predictions of galaxy formation and evolution. In particular, the stellar mass function of galaxies is increasingly suppressed relative to the mass function of dark matter halos at stellar masses larger than $\sim 10^{11} M_{\odot}$ (Balogh et al. 2001; Benson et al. 2003), implying gas in halos above this mass scale is increasingly inefficient at cooling onto galaxies to form stars. Contemporaneously, the launch of sensitive X-ray telescopes such as *Chandra* revealed the hot gas filling the most massive groups and clusters (halo masses $\gtrsim 10^{13} M_{\odot}$), discovering an overabundance of hot gas over what purely cosmological accretion-driven shock-heating should be capable of sustaining (Edge et al. 1992; Cavagnolo et al. 2008; Dunn & Fabian 2008). Together, these observations imply that in sufficiently massive halo environments, the cooling of cosmologically accreted gas is stifled by an injection of heating energy from within.

Active galactic nuclei (AGN), the most luminous persistent sources of radiation in the universe had earlier been recognized to be powered by accretion onto supermassive black holes (SMBHs) at the centers of galaxies (Salpeter 1964; Lynden-Bell 1969). However, the discovery that all massive galaxies in the local Universe harbor a SMBH at their core (Magorrian et al. 1998), the existence of tight correlations between this BH mass and the host galaxy properties (Kormendy & Ho 2013), and the discovery of X-ray gas cavities collocated with AGN jets (Boehringer et al. 1993; Carilli et al. 1994) revealed that the energy released by AGN must be capable of coupling strongly to the surrounding galaxy and gaseous environment. By implementing a prescription for energy released by AGN kinetically coupling to the surroundings, semi-analytic and hydrodynamical simulations of galaxy formation began to successfully reproduce the observed stellar mass functions and

X-ray halo properties (Croton et al. 2006; Bower et al. 2006; McCarthy et al. 2010). It is now widely recognized that AGN feedback is a critical process in regulating the formation and evolution of galaxies.

The Schwarzschild radius of a supermassive black hole of $10^9 M_\odot$ is just $\sim$ 10 AU, while accretion onto such a black hole appears to be capable of shaping the dynamics of gas on up to $\sim$ Mpc scales, with the record "giant radio galaxies" spanning $\sim$ 5 Mpc (Oei et al. 2022). Thus, AGN feedback involves a dynamic range in scale of up to $\sim$10 orders of magnitude, analogous to the difference in size between the Earth and a grain of salt. This remarkable fact in part motivates studying the *AGN-halo connection*, as the halo gravitational potential well is the source drawing the gaseous fuel which will ultimately power the central engine. Studying which large-scale environments AGN lie within is a particularly clean test of AGN feedback models, as it uncovers which halos this energy is being deposited into.

Black holes, once thought to be a mathematical artifact of General Relativity are now recognized to be real, ubiquitous at the centers of galaxies, and capable of materially altering the distribution of matter on cosmological scales. This thesis aims to contribute to our understanding of this feedback process by revealing which halo environments AGN release their energy into.

1.2 AGN Host Halos Probe Feedback and Structure

The spectral energy distributions (SEDs) of AGN vary considerably, and thus a literature has spawned sorting them into a wide array of observational classes (see review by Padovani et al. 2017). These include dividing by AGN luminosity, obscuration, presence or lack of a synchrotron jet, and presence broad or narrow emission lines, among other properties.

Many of these classes have been shown to represent different observational manifestations

of the same underlying system, under the *unified model* of AGN (Urry & Padovani 1995). Under this model, many observed differences between spectral types are attributed simply to the chance inclination at which we happen to view the axisymmetric accretion structure.

Alternatively, AGN of different spectral types may exhibit different properties because they represent fundamentally different objects. Or rather, they may represent the same object viewed by an observer at different times in a coevolutionary stage with the host galaxy and environment. Such a hypothesis is characterized in Figure 1.1. Hickox et al. (2009) showed that AGN selected in different wavebands reside in different halo environments. They find that radio galaxies with powerful jets occupy more massive halos than X-ray-selected AGN, which in turn occupy more massive halos than optical quasars. As halos grow over cosmic time due to hierarchical accretion, a simple model would suggest that a given galaxy harboring a central SMBH would evolve from an optical quasar phase, to a X-ray AGN phase, then to a radio AGN phase as its host halo mass increases through hierarchical structure growth.

Such a scheme may be expected given the well-understood physics of both cosmological gas accretion onto halos, as well as accretion modes of gas onto a SMBH. The cooling timescale of gravitationally shocked-heated gas accreted onto a halo scales predominantly with the halo mass, such that gas in halos of mass $M_h \gtrsim 10^{13} M_\odot$ reach cooling timescales longer the Hubble time (Donahue & Voit 2022). These "hot halos" thus are incapable of feeding a central galaxy with cold streams of gas to fuel both star formation and efficient SMBH growth. On the other hand, the gas in halos of $M_h \lesssim 10^{12} M_\odot$ is able to cool rapidly and inhomogeneously. The resulting cold gas streams are expected to be able to propagate down to the central region, fueling radiatively-efficient accretion onto the SMBH (Shakura & Sunyaev 1973), where gas is able to accrete at high Eddington rates, forming a thin disk and radiating $\sim 10\%$ of the input rest-mass energy in a quasar-like mode. Meanwhile, accretion in a massive, hot halo will be inefficient and homogeneous, fueling a central

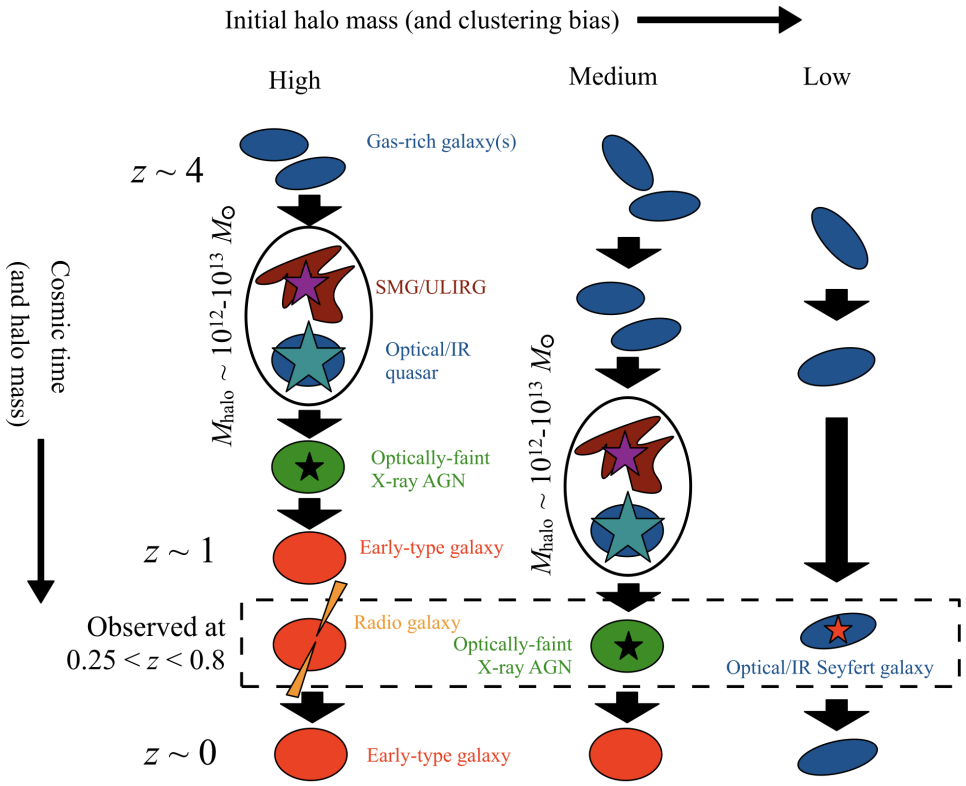

Figure 1.1: A schematic evolutionary sequence of AGN classifications with host halo mass and redshift, adapted from Hickox et al. (2009). Here, the observation that radio galaxies cluster more strongly than X-ray AGN, which in turn cluster more strongly then optical quasars is interpreted as an evolutionary sequence. In this model, as a galaxy's halo grows in mass over cosmic time, the central SMBH presents itself as different AGN classes to an observer.

SMBH in an advection-dominated accretion flow at relatively low Eddington rates and preferentially releasing energy in the form of mechanical jets (Narayan & Yi 1994). In this model, radiative AGN and jet-mode AGN are expected to be evolutionary phases driven by the host halo mass, and I test this picture in Chapter 4.

Estimating the host halo properties of different spectral types of AGN is a powerful test of whether the differences observed in the systems can be solely attributed to an orientation effect in the unified scheme, or if a possible evolutionary link can be drawn between the two classes. In other words, the $\sim$Mpc large-scale environment of an AGN can surprisingly reveal information about the $\sim$pc-scale accretion structure. If properly controlling for confounding factors such as host galaxy properties, two populations of AGN belonging to different classifications should occupy identical large-scale environments statistically **if** their differences are driven only by an orientation effect with the observer, under the Copernican principle. Conversely, stages in galaxy evolution are known to be connected to host halo mass (e.g., quenched elliptical galaxies cluster more strongly than star-forming spiral galaxies of the same mass; Zehavi et al. 2011). Thus, if two AGN types represent distinct phases in galaxy-AGN coevolution, a difference in their host halo properties is expected. Therefore, detecting a difference in the host halo properties of two properly-controlled samples of AGN effectively rules out an orientation-based explanation of their differing properties.

The role that orientation and evolution play in AGN *obscuration* is heavily debated. Obscured AGN lack soft X-ray emission and the distinguishing features of AGN in the optical band relative to unobscured systems, like a blue power law disk continuum and broad emission lines, while still exhibiting the warm dust SED in the mid-infrared characteristic of AGN tori (Hickox & Alexander 2018). This is interpreted in the unified model as an AGN viewed at a line-of-sight through the dusty torus, which extincts the emission from the accretion disk, corona, and broad-line region. Alternatively, evolutionary models suggest

that some obscured quasars may represent a relatively brief coevolutionary phase with the host galaxy, where the central SMBH and the host galaxy are both fed by rapidly cooling streams of gas, obscuring the central source by dusty star-forming clouds or circumnuclear gas (Sanders et al. 1988b; Hopkins et al. 2005; Alexander & Hickox 2012). In this scenario, matched samples of obscured and unobscured quasars may not be expected to occupy the large-scale structure in the same manner, and we test this hypothesis in Chapters 2 and 3.

1.3 Inferring Host Halo Properties

Once a sample of AGN has been selected, the properties of the host halos can be probed by measuring its clustering properties either through correlation function or gravitational lensing measurements. A correlation function measures the excess probability of detecting a pair of galaxies at a given separation (Peebles 1980) and can be performed spatially if accurate redshifts are known to each object, otherwise the weaker projected clustering can be recovered in angular space on the sky. In each case, the correlation function is measured by counting pairs of galaxies in the survey in separation scale bins, and comparing these counts to pair counts of a *random catalog*, which is a Poisson realization of the survey selection function. A correlation function is simply the Fourier transform of the power spectrum; thus correlation functions can be compared to the power spectrum of matter predicted by the cosmological model to infer how a sample of galaxies are tracing the underlying matter field.

Likewise, a lensing measurement entails correlating the projected density field of the galaxy/AGN sample on the sky with an independently derived map of integrated mass inferred by the amplitude of gravitational lensing for each sightline. In this thesis, I study the clustering of AGN samples which are visible out to $z \sim 3$ owing to their extreme luminosities and space density evolution. Therefore, a lensing signal induced by their host halos will only be observed in radiation sources in the background, emitted at higher redshift than

the AGN themselves. The cosmic microwave background radiation serves as an excellent background for this purpose. The CMB photons emitted at $z \sim 1100$ act as a backlight for nearly all of the structure in the universe, particularly sensitive to structure at $z \sim 1-3$, fortuitously the same epoch where quasars were most numerous. CMB experiments such as *Planck* can detect the average CMB photon deflection by searching for non-Gaussian modes in the temperature maps, or for *B*-modes in the polarization maps.

Massive galaxies and AGN are "biased" tracers of the underlying dark matter-dominated density field, such that they only tend to form in the more overdense regions, or equivalently, the more massive halos. On large scales governed by linear structure growth, the power spectrum of a galaxy sample P_{gg} tends to be excellently described by a simple linear scaling of the matter power spectrum $P_{mm}(k,z)$:

$$P_{gg}(k,z) = b_g^2(M,z)P_{mm}(k,z), \tag{1.1}$$

where b_g is the galaxy bias and is canonically a function only of halo mass and redshift. As both galaxy correlation functions and lensing cross-spectra are transforms of P_{gg}, one can fit these signals on large scales for a bias factor. Then by comparing to bias factors for halos of various masses recovered in *N*-body simulations, one can infer the effective mass of the halos hosting the galaxies.

Fitting for a bias parameter on linear scales allows inference of the effective (mean) host halo mass of a sample, but galaxies are not in reality hosted by halos of a delta function in mass. More realistic modeling of galaxies' observed clustering and lensing signals has been achieved through the development of the *halo model* (e.g., Seljak 2000; Cooray & Sheth 2002; Asgari et al. 2023), which provides an analytical method to approximate the statistics of the dark matter density field on both linear and nonlinear scales. This model assumes that all dark matter is confined within virialized dark matter halos, then approximates the power spectrum of matter (and its luminous galaxy tracers) as sums of the intra-halo (one-

halo) and inter-halo (two-halo) power. Then, knowing the linear matter power spectrum, the mass function of halos (e.g., Tinker et al. 2008), the linear bias of halos as a function of mass (derived from halos identified in simulations; e.g., Tinker et al. 2010), and the density profile of halos (Navarro et al. 1997), one can predict the power spectrum of halos of a given mass analytically. Finally, by specifying a connection between a sample of galaxies and the distribution of dark matter halo masses that host them, known as a halo occupation distribution (HOD), the clustering of that galaxy population can be predicted on both one and two-halo scales. We leverage this technique to study the full distribution of halos hosting obscured and unobscured quasars in Chapter 3, and also those hosting luminous radio galaxies in Chapter 4.

Many of the AGN samples I study in this thesis are selected photometrically, and thus we do not know the redshifts to most of the objects. However, their host halo properties can still be recovered using clustering and lensing measurements in angular space on the sky. However, interpreting the angular clustering or lensing signal induced by a sample of galaxies or AGN depends sensitively on knowing their statistical redshift distribution. Estimating this for a photometrically-selected sample often entails cross-matching to objects in a "deep field" overlapping the wide-field survey, which contains a significant investment of legacy multiwavelength photometry and spectroscopy in the literature (e.g., the COSMOS field; Scoville et al. 2007). By collecting photometric and/or spectroscopic redshifts in the deep field for matches to the photometric wide-field catalog, one can estimate the redshift distribution, and assume it applies over the wide field under the cosmological principle. However, this can be a poor assumption if the deep field is small enough that cosmic variance is important, or if photometric redshifts have systematics unaccounted for.

One technique I leverage throughout the thesis is the combination of angular clustering and CMB lensing measurements to mitigate systematic uncertainties in the redshift distributions of photometric samples. Clustering and lensing measurements both probe the same

bias parameter of the sample on linear scales, and thus should yield bias measurements within statistical uncertainty of each other when performed correctly. These measurements are completely independent and have different dependencies on the redshift distribution. Clustering is sensitive to the width of the redshift distribution while lensing is less-so, and CMB lensing is increasingly insensitive to structure at $z < 1$. Therefore, an agreement between the halo parameters from an angular clustering and CMB lensing measurement is a strong verification that our inferred redshift distribution is correct.

1.4 Outline

In this thesis, I present three works probing the properties of dark matter halos hosting AGN. In Chapter 2, I study the clustering properties of optically-selected spectroscopic quasars as a function of color. Reddened quasars were previously shown to exhibit radio properties in conflict with their colors being explained by an orientation effect. I thus test if this difference could be an evolutionary effect tied to the large-scale environment but find no such evidence. In Chapter 3, I study the clustering properties of more than one million infrared-selected quasars, separated into obscured and unobscured samples. I show that obscured quasars at $z \sim 1.5$ occupy systematically more massive halo environments than unobscured quasars, ruling out a model where quasar obscuration is purely an orientation effect, and suggesting obscured quasars represent a distinct phase of AGN-galaxy coevolution. In Chapter 4, I study the host halo properties of luminous low-frequency radio galaxies since $z < 2$, showing that jet-mode feedback dominates over quasar wind-mode feedback in galaxy groups and clusters for the past 10 Gyr. I also include an appendix describing the Python software packages I developed over the course of my thesis work to measure and model the observations presented in Chapters 2, 3, & 4. Finally, I conclude in Chapter 5 and outline future works which should advance our understanding of the AGN-halo connection.

Chapter 2

Host Dark Matter Halos of SDSS Red and Blue Quasars: No Significant Difference in Large-scale Environment

by Grayson C. Petter (Dartmouth College), Ryan C. Hickox (Dartmouth College), David M. Alexander (Durham University), James E. Geach (University of Hertfordshire), Adam D. Myers (University of Wyoming), David J. Rosario (Durham University, Newcastle University), Victoria A. Fawcett (Durham University), Lizelke Klindt (Durham University), Kelly E. Whalen (Dartmouth College)

This work first appeared in The Astrophysical Journal 2020, volume 927, page 16.

2.1 Introduction

Quasars are the most luminous class of active galactic nuclei (AGN), objects powered by accretion of interstellar material onto supermassive black holes near the centers of galaxies (Salpeter 1964; Lynden-Bell 1969). Since their discovery, AGN and quasars have been

observationally classified into a taxonomy including many species (e.g., Padovani et al. 2017). However, a comprehensive physical picture has yet to emerge for many of these variations. In particular, the origin of "red" quasars which exhibit redder continuum spectra in the optical through mid-infrared regime than their more numerous blue counterparts remains unexplained. Unified models of AGN (Antonucci 1993; Urry & Padovani 1995; Netzer 2015) attempt to explain the gamut of AGN varieties in terms of a few parameters intrinsic to the system, stipulating that the array of observed AGN classes are inherently similar yet exhibit differing spectral energy distributions (SEDs) due to chance alignment of the accretion disk along the line-of-sight. This model would attribute red quasars' color to extinction arising from a moderate viewing angle of the dusty "torus" (e.g., Wilkes et al. 2002; Rose et al. 2013). Alternatively, evolutionary models suggest that AGN may vary in their observed properties over the course of their activity through interaction with their broader environments. One such model of interest in explaining quasars' colors links quasar activity and star formation in a feedback-driven co-evolutionary scheme, which may produce quasars reddened by their host galaxies during merger-triggered starburst events (e.g., Sanders et al. 1988a; Canalizo & Stockton 2001; Hopkins et al. 2005, 2006, 2008; Urrutia et al. 2008; Alexander & Hickox 2012; Glikman et al. 2012, 2015; Banerji et al. 2015; Hickox & Alexander 2018; Perrotta et al. 2019). A powerful method to distinguish between these schema is to test whether certain classes of AGN exhibit differences across a parameter which cannot feasibly be connected to orientation. Thus, this work aims to probe the nature of optically-selected red and blue quasars by estimating the characteristic dark matter halo mass that each class resides within.

Despite many investigations (e.g., Webster et al. 1995; Wilkes et al. 2002; Richards et al. 2003; Hopkins et al. 2004; Krawczyk et al. 2015; Kim & Im 2018; Klindt et al. 2019; Calistro Rivera et al. 2021), the primary driver of optical quasar color has yet to be conclusively determined. The recent work of Klindt et al. (2019), Rosario et al. (2020), and Fawcett et al. (2020) has demonstrated that optically red quasars exhibit an excess of

radio-emission compared to their blue counterparts in a manner that contradicts a purely orientation-based explanation of their colors. Specifically, this work has found that red quasars display higher radio detection fractions, driven primarily by compact radio sources near the radio-quiet/radio-loud threshold. Crucially, this trend is the opposite of that expected if quasars' colors are dominantly determined by their orientation with the observer's line-of-sight, as jets emanating from an edge-on reddened accretion disk should appear more extended and less bright owing to a lack of relativistic beaming toward the observer. Motivated by these results, we set out to test whether optically red and blue quasars display fundamental differences across other properties, particularly their surrounding large-scale structure (LSS).

As quasars form in overdense regions in the Universe, they are "biased" tracers of the underlying matter distribution. In this study, we estimate how red and blue quasars selected with the Sloan Digital Sky Survey (SDSS) sample the matter distribution by measuring both their two-point correlation functions as well as the gravitational lensing of the cosmic microwave background (CMB) induced by their host dark matter halos. This bias can be interpreted with the halo model (e.g., Seljak 2000; Cooray & Sheth 2002) in a Λ-Cold Dark Matter (Λ-CDM) framework to estimate the characteristic mass of the dark matter halos in which red and blue quasars reside (e.g., Sheth & Tormen 1999; Tinker et al. 2010).

We do not detect any trends of host halo mass with optical quasar color through either method, finding that optically-selected quasars occupy similar dark matter halo environments of $M_h \sim 3 \times 10^{12}\ h^{-1} M_\odot$ across the optical color spectrum. We also confirm fundamental differences in the radio properties between red and blue quasars by performing a stacking analysis of FIRST data. Taken together, these analyses suggest that optically red quasars' colors and enhanced incidence of radio emission are not primarily linked with their surrounding LSS nor their torus' orientation with the line-of-sight. Instead, red quasars' colors likely stem from nuclear-galactic scale processes, perhaps arising from obscuration

by a nuclear dusty wind launched by the quasar system itself (e.g., Elvis 2000; Calistro Rivera et al. 2021; Rosario et al. 2021).

Throughout this work, we adopt a "Planck 2018" CMB+BAO Λ-CDM concordance cosmology (Planck Collaboration et al. 2020a), with $h = H_0/100$ km s^{-1}Mpc^{-1} = 0.6766, $\Omega_m = 0.3111$, $\Omega_\Lambda = 0.6888$, $\sigma_8 = 0.8102$, and $n_s = 0.9665$.

2.2 Data

2.2.1 Quasar Samples

This work aims to probe the dark matter halo environments of quasars as a function of color through two independent measurements, the two-point correlation function and the gravitational lensing of the CMB. However, the lensing signal sourced from typical individual quasar host halos is orders of magnitude below the noise level of current measurements with Planck, meaning that enormous samples of quasars are required to derive significant results. For the purposes of this work, more sources are required than the largest spectroscopic quasar samples available to date. We therefore elect to use similar but distinct samples for the two analyses. In particular, we use the latest Extended Baryon Oscillation Spectroscopic Survey (eBOSS, Dawson et al. 2016) LSS catalogs for the correlation function measurements and the larger XDQSOz photometric quasar candidate catalog (DiPompeo et al. 2015a) for the lensing analysis. These samples were both optically-selected with SDSS imaging data and have similar color distributions, magnitude limits, and redshift distributions (Figure 2.1). The results from the two independent analyses and samples can thus be approximately compared. We further describe these samples in the following text, and their properties are summarized in Table 2.1.

Table 2.1. Optically-selected quasar sample properties

Sample	Type	$N_{\rm QSO}$	Analysis
XDQSO	Photometric	656,899	CMB Lensing
eBOSS	Spectroscopic	343,708	Correlation Functions

Quasar Sample For Lensing Analysis

As large samples of quasars are required to yield statistically significant lensing measurements, we begin with the photometric "XDQSOz" catalog (DiPompeo et al. 2015a), which contains 5,537,436 optically-selected quasar candidates, or 3,874,639 quasars weighted by probability. This catalog was constructed using extreme deconvolution (XD; Bovy et al. 2011, 2012) to model spectroscopically confirmed quasars in flux-redshift space utilizing GALEX ultraviolet through WISE mid-infrared photometry, while optimally incorporating photometric uncertainties and non-detections. This model was then used to calculate quasar probabilities and photometric redshifts for all point sources in the 8th data release of the SDSS (DR8). The XDQSOz catalog contains all point sources with a quasar probability $P_{\rm QSO} > 0.2$.

We make several enhancements and quality cuts to this photometric catalog to ensure its usefulness in a statistical study of quasars. First, we match the XDQSOz catalog to the "DR16Q" catalog of spectroscopically confirmed quasars (Lyke et al. 2020), which contains $\sim 750,000$ bona fide quasars, greater than seven times more objects than appeared in the spectroscopic sample originally used to train the XDQSOz algorithm. We replace the photometric redshifts in the XDQSOz catalog with accurate spectroscopic redshifts and update probabilities to unity for the $\sim 700,000$ matches to spectroscopically confirmed quasars. We also update quasar probabilities to zero for the objects which were targeted by SDSS spectroscopic campaigns and subsequently confirmed not to be quasars. Next, we restrict the catalog by applying the **good** flag in the XDQSOz catalog which mimics cuts used by the BOSS to remove sources with unreliable photometry. This removes approximately

half of the sources. We further require the photometric redshift probability density function to contain only one peak, indicating that the photometric redshift should be reliable. This cut removes an additional quarter of the sample. We thus use the peak redshift estimate in all subsequent analysis, where necessary. Finally, we restrict the catalog to those sources which the XDQSOz algorithm deems high-likelihood quasars, with $P_{\rm QSO} > 0.9$. This sample therefore consists of uniformly-selected photometric quasar candidates which should be $> 90\%$ pure, and is uniformly enhanced by spectroscopic data when available.

Before splitting this sample by color for cross-correlation with CMB lensing measurements, we make two final cuts. The first selects only quasars with a best redshift estimate between $0.8 < z < 2.2$. This is chosen to match the redshift distribution of the eBOSS sample used for correlation function measurements, such that results from the two analyses may be approximately compared. This redshift range also fortuitously overlaps with the CMB lensing "kernel" peak at $z \sim 1-2$ (e.g., Cooray & Hu 2000; Song et al. 2003), and avoids the low quasar selection efficiency at $z > 2.5$ due to degeneracy with stellar colors (Richards et al. 2002, 2003). Finally, this redshift cut avoids low-luminosity quasars at low redshift, which may have colors dominated by host-galaxy contamination rather than dust extinction (Klindt et al. 2019; Calistro Rivera et al. 2021). We have thus constructed a catalog of 656,899 uniformly-selected high-probability quasar candidates (58% of which have been spectroscopically confirmed) suitable for statistical study. We refer to this sample as the "XDQSO sample" throughout the remainder of this work.

Samples For Two-Point Correlation Function Analyses

We also probe the bias of quasars as a function of color with two-point cross-correlation functions. We perform these measurements with data from the completed eBOSS survey (Dawson et al. 2016). eBOSS was a spectroscopic survey carried out as part of the SDSS to measure the baryonic acoustic oscillations in the correlation functions of biased tracers of the underlying matter distribution, including quasars, luminous red galaxies (LRGs)

and emission line galaxies (ELGs). Following the survey's completion, LSS catalogs for each tracer sample were publicly released along with carefully constructed random catalogs which are designed to match each survey's selection function, a necessary component of performing a clustering measurement. In this work, we estimate the halo bias for quasars of a given color by measuring their cross-correlation with the entire eBOSS quasar sample, as well as the cross-correlation with LRGs and ELGs. We thus utilize the publicly available quasar and LRG LSS catalogs (Ross et al. 2020) along with the ELG LSS catalog (Raichoor et al. 2021). The eBOSS quasar catalog contains 343,708 spectroscopically confirmed quasars which were selected uniformly using XD probabilities (Bovy et al. 2012) along with an optical-mid-infrared color cut (Myers et al. 2015). We refer the reader to these publications for details of the catalogs' construction.

2.2.2 Planck Lensing Convergence Map

To estimate the average gravitational deflection of CMB photons induced by a given sample of quasar host halos, we utilize the 2018 release of the Planck lensing convergence map[1] (Planck Collaboration et al. 2020b). We adopt the minimum-variance (MV) estimate combining both temperature and polarization data. We do not use the joint reconstruction with CMB and cosmic infrared background (CIB) data, as our sample of high-redshift quasars may contribute to the CIB and contaminate the lensing estimate. We also make use of the simulated lensing noise maps to estimate uncertainties on our measurements.

To recover signal at the scales of interest, we smooth the map with a $15'$ FWHM Gaussian beam which removes the highest l-modes containing virtually no signal. We also apply an inverse top-hat filter to remove noisy l-modes of $l < 100$ (Geach et al. 2019). This filtered map can then be "stacked" at the positions of quasars in a given sample to probe the average CMB lensing amplitude generated by their host halos.

[1]The lensing products used throughout this work can be accessed at https://pla.esac.esa.int/

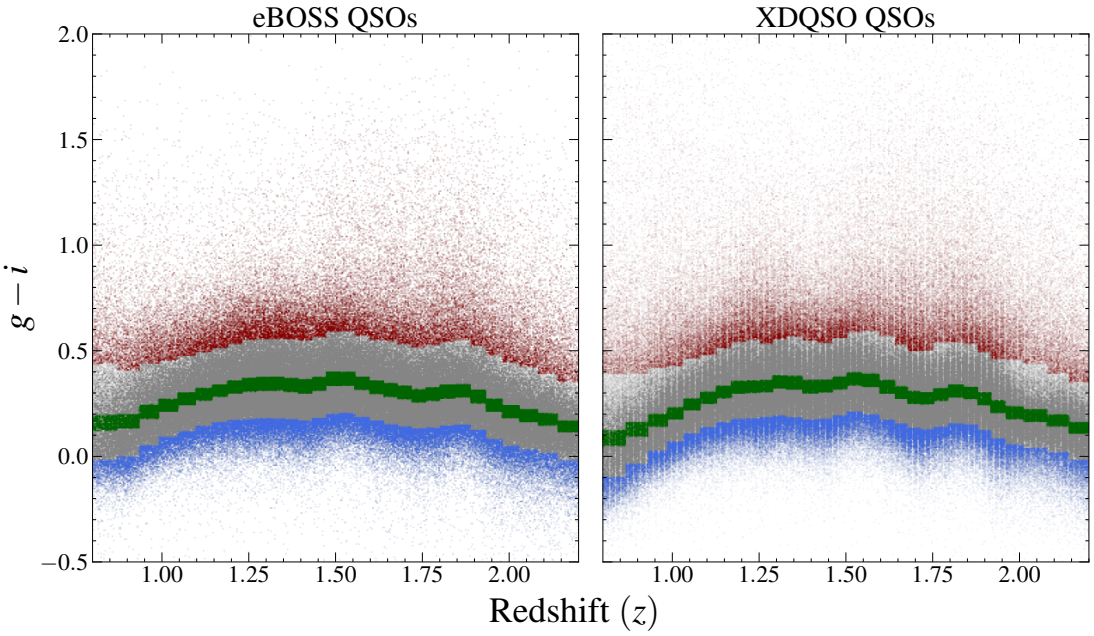

Figure 2.1: The $g-i$ distribution (corrected for Galactic extinction) for the spectroscopic eBOSS quasars (left panel) and photometric XDQSO quasars (right panel) as a function of redshift. The redshift-evolving color binning scheme described in Section 2.3.1 is demonstrated by highlighting the quasars with $g-i$ color indices belonging to the highest, middle, and lowest 14.3% quantiles with red, green, and blue, respectively. This demonstrates that the spectroscopic and photometric samples have similar color and redshift distributions and can thus both be used to study the dependence of halo bias on optical quasar color at $0.8 < z < 2.2$.

2.2.3 FIRST Data

Finally, we utilize data from the Faint Images of the Radio Sky at Twenty-Centimeters (FIRST) survey (Becker et al. 1995) undertaken with the Jansky Very Large Array (JVLA) to confirm key differences in the radio properties of red and blue quasars. FIRST was a radio survey at 1.4 GHz over $\sim 10^4$ deg^2 of the northern sky to ~ 0.15 mJy RMS depth at $\sim 5'$ resolution, which overlaps the SDSS imaging regions and covers $\sim 84\%$ of the eBOSS quasar footprint. We extract FIRST imaging cutouts at the positions of quasars using `astroquery` (Ginsburg et al. 2019) to stack and derive median radio-loudness parameters.

2.3 Measurements

2.3.1 Binning Quasars by Color

We adopt a similar technique to that used in Klindt et al. (2019) in order to separate quasars by color in a redshift-dependent manner. We thus first produce a $g-i$ color distribution (corrected for Galactic reddening; Schlegel et al. 1998; Schlafly & Finkbeiner 2011) for each sample in 30 bins of redshift containing equal numbers of sources. Within each redshift bin, we then divide the $g-i$ distribution into 7 bins containing an approximately equal number of objects. We use 7 bins rather than the 10 used by Klindt et al. (2019) in order to ensure significant detections of the CMB lensing in each bin. This redshift-evolving color selection effectively accounts for the shifting of spectral lines in and out of observed bandpasses at various redshifts and is thus analogous to performing a k-correction to the observed colors before binning. The color as a function of redshift for the reddest, bluest, and median bin of each sample is displayed in Figure 2.1. It is clear that the two samples have similar redshift distributions and color bin definitions, implying that we can approximately compare the results from the two samples and analyses.

As the definition of a "red" quasar varies substantially in the literature, it is important to quantify our delineation. We therefore compute a "color offset" for each quasar for the purpose of comparison with other results, given as the difference between the observed color (corrected for Galactic extinction) and a template color for a "typical" SDSS quasar using the Vanden Berk et al. (2001) (VB) template:

$$\Delta(g-i) = (g-i)_{\rm obs} - (g-i)_{\rm VB}. \tag{2.1}$$

We use the VB template as the reference color rather than the median of our sample (e.g., Klindt et al. 2019), such that the reader may compare our results with other samples without having to reconstruct the samples used here. We will display the median color offset of each color bin in the following results for this purpose.

2.3.2 Weighting Scheme

We develop weights which can be used in both the lensing analysis and the correlation functions to control for any luminosity differences between the different color samples, as halo bias may be correlated with quasar luminosity (Geach et al. 2019), albeit weakly at most (Shen et al. 2009; Eftekharzadeh et al. 2015). We elect to use the mid-infrared to trace the bolometric luminosity of our optically-selected quasars, as this emission should be minimally attenuated by dust. We thus compute rest-frame 1.5 μm luminosities for each quasar in our sample by interpolating/extrapolating observed 3.4 and 4.6 μm fluxes from Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) data, assuming a power law spectrum. We use 1.5 μm emission as the bolometric luminosity tracer rather than the 6 μm used by Klindt et al. (2019) so that we do not lose statistical power by requiring that each source is detected in the far less sensitive WISE W3 band. The normalized 1.5 μm luminosity distributions for each of the color samples are displayed in Figure 2.2, demonstrating that the luminosity distributions vary slightly between quasars of differing

color. In order to control for these luminosity differences, we first generate normalized 2D distributions of rest-frame 1.5 μm luminosity and redshift for each color sample. We then compute the minimum value in each luminosity-redshift bin across all of the color samples. Finally, weights are assigned to quasars in each color sample as the ratio of the minimum luminosity-redshift distribution to the distribution of that color sample. This weighting scheme ensures that all of the color samples are matched in redshift and bolometric luminosity.

For the CMB lensing analysis, we apply a second weighting scheme to account for the fact that some of the sources in the XDQSO sample will be contaminants of stars or galaxies. We thus weight candidates by the quasar probability output by the XDQSO algorithm, $P_{\rm QSO}$.

2.3.3 Stacking Lensing Map

Cross-correlations between LSS tracers and CMB lensing convergence maps are typically calculated in Fourier space, but can equivalently be performed in real space, through a "stacking" procedure (e.g., Madhavacheril et al. 2015, 2020; Baxter et al. 2015; Geach & Peacock 2017; Geach et al. 2019), which is simpler to implement for samples with complex selection functions. Stacking is the process of averaging over many maps to statistically reveal signals buried beneath the noise of individual maps. Here, we perform a weighted stack of the Planck lensing convergence map at the positions of quasars, where the weights aim to control for differences in properties between samples aside from optical quasar color. For each layer in the stack, we reproject the Planck map using the Lambert azimuthal equal-area projection (with `healpy`, Zonca et al. 2019) spanning 4 degrees across centered at the position of each quasar. We then stack by computing a pixel-wise weighted average across all of the projections, ignoring masked pixels.

The result of stacking the Planck lensing convergence map at the positions of the bluest,

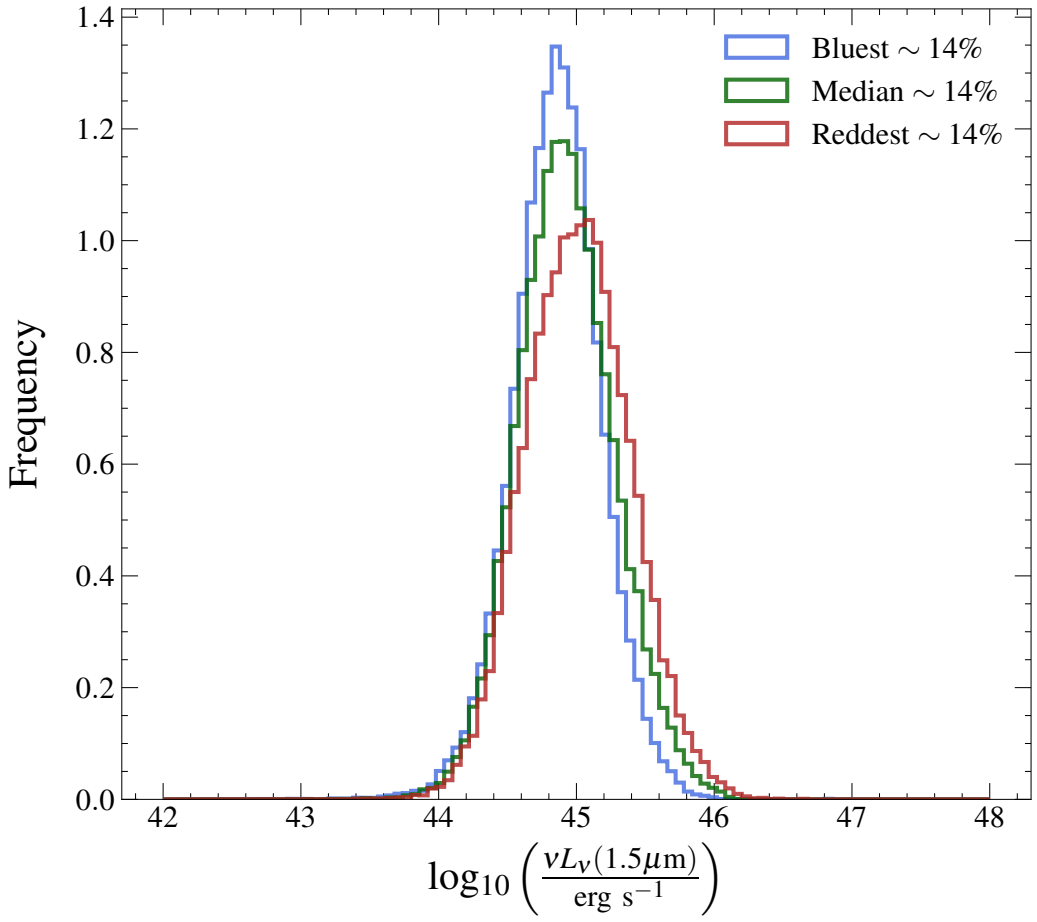

Figure 2.2: The rest-frame 1.5 μm luminosity distributions of the reddest, median, and bluest $\sim$ 14% bins of quasars in the eBOSS sample. We adopt the mid-infrared as a bolometric luminosity indicator as red quasars are likely more affected by dust extinction in the optical. This demonstrates that SDSS quasars with different optical colors exhibit similar but not identical infrared/bolometric luminosity distributions. We assign weights to quasars for all subsequent analyses to control for any effects these small luminosity differences may impart on halo bias measurements.

median, and reddest bins of quasars is shown in Figure 2.3. As the color samples are constructed to have identical redshift distributions, the peak lensing convergence amplitude (κ, a unitless surface mass density) directly traces a sample's average halo bias (Eq. 2.13). No clear trend of peak κ with color is apparent, but we quantify this by fitting these maps in Section 2.4.

2.3.4 Stacking FIRST Data

Prompted by the previous finding that red and blue quasars exhibit fundamental differences in their radio properties (Klindt et al. 2019; Rosario et al. 2020; Fawcett et al. 2020), we also perform stacks of 1.4 GHz imaging data from the FIRST survey for our quasar samples, where the data is available ($\sim$ 84% of the eBOSS footprint). We adopt a median stacking procedure rather than an average stack as per the recommendation of White et al. (2007). We thus retain the $\sim$ 2% of FIRST-detected sources in our samples, as the median is insensitive to outliers. We also apply the correction for "CLEAN bias", which is a systematic underestimation of flux by a factor of 1.4 in stacked FIRST images (White et al. 2007). We elect to use the eBOSS spectroscopic quasar sample for this analysis to avoid any contamination by non-quasars and ensure accurate redshift information. It should be noted that we also test stacking radio images at the positions of the XDQSO sample, which gives similar results except for a $\sim 1.5\sigma$ lower flux in the reddest bin. This may indicate that the reddest candidates in the XDQSO sample suffer from higher contamination fractions or poorer photometric redshifts.

The results of stacking the bluest, median and reddest $\sim$ 14% of quasars in the eBOSS sample is shown in Figure 2.4, which demonstrates that optically-selected red quasars exhibit an excess of radio-emission, consistent with previous findings (Klindt et al. 2019; Rosario et al. 2020; Fawcett et al. 2020). We will quantify this result's effect on the median radio-loudness as a function of quasar color in Section 2.5. We define the radio-loudness as the logarithmic ratio between the 1.4 GHz rest-frame luminosity, and the 1.5 μm luminosity

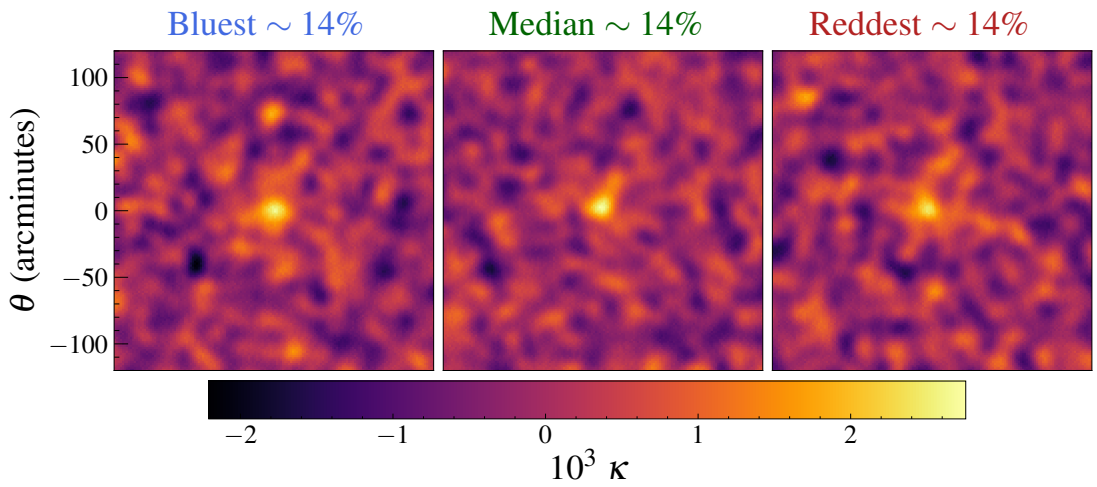

Figure 2.3: The stacked Planck CMB lensing convergence (κ) map at the positions of the XDQSO quasars belonging to the bluest, median and reddest bins. Any differences in bolometric luminosity between the samples have been controlled for by stacking with weights described in Section 2.3.2. No trend of lensing convergence with color is readily apparent, and we quantify the implications for the host halo properties of red and blue quasars in the subsequent analysis.

which is our bolometric luminosity tracer:

$$R = \log_{10}\left(\frac{\nu L_\nu(1.4\ \mathrm{GHz})}{\nu L_\nu(1.5\ \mu\mathrm{m})}\right). \tag{2.2}$$

We compute median rest-frame luminosities at 1.4 GHz from the stacked flux images using the median redshift of our sample, and by assuming a power law spectrum ($S_\nu \propto \nu^{-\alpha}$) with a power law spectral index typical of quasars, $\alpha = 0.5$.

2.3.5 Two-Point Correlation Functions

We measure the spatial clustering of eBOSS quasars as a function of color by cross-correlating them with three samples of LSS tracers. First, we measure the cross-correlation of the quasars in each color bin with the entirety of the eBOSS quasar sample. We choose this technique rather than performing autocorrelations of the quasars within each bin in order to mitigate systematics. This is because autocorrelations would require random cat-

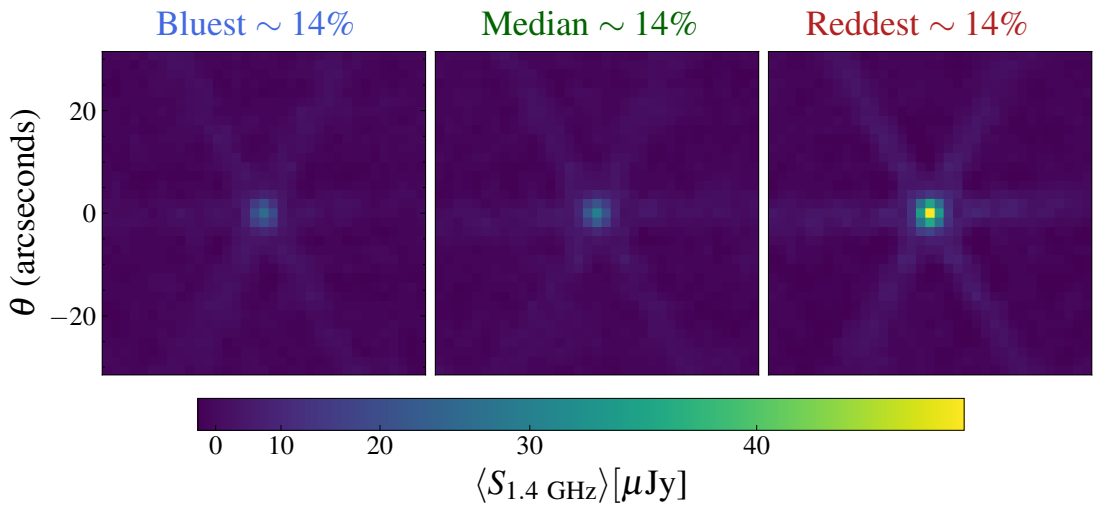

Figure 2.4: The median stacks of FIRST images at the positions of the bluest, median, and reddest quasars in the eBOSS sample. We have applied a power-law stretch to the color scale to emphasize the "source" rather than interferometric artifacts. The flux values have been corrected to account for the "clean bias" described by White et al. (2007). It is apparent that the reddest quasars exhibit an excess of 1.4 GHz emission, consistent with previous findings. These stacks have not been controlled for 1.5 μm luminosity, as we subsequently use them to quantify the median radio-loudness as a function of quasar color.

alogs representative of the selection function of each color bin, which would be difficult to reconstruct given the nontrivial color-based selection of eBOSS quasars (Myers et al. 2015). Instead, a cross-correlation is dominantly sensitive to the selection function of the "tracer" population of the entire eBOSS quasar sample, which has been robustly characterized and may be replicated through adoption of the provided random catalogs and weighting schemes. Next, we measure the cross-correlations of each bin of quasars with both LRGs and ELGs, providing two extra independent probes of the quasars' clustering strength at different effective redshifts. Finally, we must measure autocorrelation functions of the tracer populations of quasars, LRGs and ELGs in order to constrain the absolute bias from the cross-correlations.

The two-point autocorrelation function $\xi(r)$ is defined as the excess probability above that expected from an unclustered Poisson distribution of finding an object within a volume element dV at a separation r from a randomly chosen initial object in a field with average

number density n (Peebles 1980):

$$dP = n[1 + \xi(r)]dV. \tag{2.3}$$

A cross-correlation function is analogously the excess probability of finding an object in one sample at a separation r from a randomly chosen object in another sample. To estimate the cross-correlations between quasars in a color bin and a given LSS tracer, we adopt the Davis & Peebles (1983) estimator:

$$\xi = \frac{Q_c T}{Q_c R} - 1, \tag{2.4}$$

where $Q_c T$ represents weighted pair-counts between quasars in a given color bin and the tracer population of either eBOSS quasars, LRGs or ELGs, and $Q_c R$ represents weighted counts between the quasars and the randoms representative of the selection function of the tracer population. This estimator is preferred for this analysis as it is independent of the selection function of the quasar sample of interest (e.g., Coil et al. 2007), which varies with quasar color in a nontrivial fashion.

We adopt the weights provided within the eBOSS LSS catalogs to ensure the randoms are representative of the selection function of each LSS tracer. These weights incorporate corrections for imaging systematics ($w_{\rm sys}$), redshift failures ($w_{\rm noz}$), fiber collisions ($w_{\rm cp}$) and optimizations for the signal-to-noise at the BAO scale ($w_{\rm FKP}$, Feldman et al. 1994). The weighted pair-counts in Eq. 2.4 are calculated by counting each pair by the product of their weights, where we have weighted the LSS tracers and randoms by $w_{\rm sys}$, $w_{\rm noz}$, and $w_{\rm cp}$, while the quasars in each color bin have been weighted according to our scheme to control for bolometric luminosity differences (Section 2.3.2).

As we measure redshifts rather than distances to quasars, peculiar motions with respect to the Hubble flow cause the spatial distribution of sources to appear extended or com-

pressed along radial sightlines (e.g., Kaiser 1987) – dubbed "redshift-space distortions" (RSDs). Thus, rather than measuring the one-dimensional CF, $\xi(r)$, we estimate the two-dimensional analog: $\xi(r_p, \pi)$, where r_p is the projected separation along a transverse axis, and π is the redshift-space separation. As RSDs only affect the line-of-sight component of ξ, one can integrate over this axis to derive the projected correlation function $w_p(r_p)$ which is RSD independent (Davis & Peebles 1983):

$$w_p(r_p) = 2 \int_0^{\pi_{\rm max}} d\pi \, \xi(r_p, \pi). \tag{2.5}$$

In practice, choosing the integration limit $\pi_{\rm max}$ is a balance between incorporating a majority of the redshift-space clustering and avoiding excess noise induced by including physically uncorrelated pairs in the statistic. To determine the optimal integration limits, we visually inspect the two-dimensional correlation functions $\xi(r_p, \pi)$ and determine the π separations within which most of the signal is contained. We thus use limits of $\pi_{\rm max}$ of 20, 25, and 15 h^{-1} Mpc for the autocorrelations of eBOSS quasars, LRGs, and ELGs, and use limits of 20, 15, and 15 h^{-1} Mpc for the cross-correlations between quasars and eBOSS quasars, LRGs, and ELGs, respectively.

These cross-correlation measurements simultaneously probe the bias of the quasars in each bin as well as the LSS tracer population. In order to estimate the absolute bias of each set of quasars, we must first estimate the bias of the reference samples of quasars, LRGs, and ELGs. For this, we perform autocorrelation function measurements using the Landy & Szalay (1993) estimator:

$$\xi = \frac{DD - 2DR + RR}{DD}, \tag{2.6}$$

where *DD*, *DR*, and *RR* are weighted pair counts normalized by number density as a function of separation for data-data pairs, data-random pairs, and random-random pairs, re-

spectively. We weight both the data and randoms by $w_{\rm sys}$, $w_{\rm noz}$, and $w_{\rm cp}$. When measuring the autocorrelations of ELGs and LRGs, we wish to probe the bias at the same effective redshift as the cross-correlation with quasars. We thus apply additional weights in the autocorrelation given as the overlap between the redshift distribution of the quasars and the galaxy sample:

$$w_{\rm overlap} = \sqrt{\frac{dN}{dz}_{\rm QSO} \frac{dN}{dz}_{\rm GAL}}. \tag{2.7}$$

We estimate the uncertainty on the measured auto and cross-correlation functions by adopting a variation of the procedure recommended by Norberg et al. (2009), which found that dividing samples into $N_{\rm sub}$ subvolumes and resampling with the bootstrap method can approximately reproduce the "true" covariance matrix inferred from mock catalogs when the number of resamplings $N_r = 3N_{\rm sub}$. We thus divide each quasar and galaxy sample into subvolumes on the sky using HEALPix, and resample as described above. We then perform a second bootstrap resampling of the individual objects selected from the subvolumes (e.g., Hickox et al. 2011). This technique should recover the uncertainty deriving from both the variance of clustering across different regions of the survey footprint and also from the finite counting statistics of the samples. We recalculate the correlation function with each realization, which gives an estimate of the uncertainty to be used for model fitting in the following section.

2.4 Modeling

2.4.1 Modeling the Lensing Profile

In order to estimate the characteristic halo bias for a sample of quasars given a stacked lensing map, we must be able to model the average lensing convergence signal expected from a sample of idealized dark matter halos. We adopt a similar procedure as described

by Geach & Peacock (2017); Geach et al. (2019), which is summarized here.

The lensing convergence profile expected from a single spherically symmetric mass distribution such as a dark matter halo (the "one-halo term") is given by:

$$\kappa_1(\theta, z) = \frac{\Sigma(\theta)}{\Sigma_{crit}(z)}, \tag{2.8}$$

where the numerator is the mass density profile $\rho(r)$ of the halo projected onto the plane of the sky:

$$\Sigma(R = \theta D_{OL}) = 2\int_R^\infty dr\, \frac{r\rho(r)}{\sqrt{r^2 - R^2}}, \tag{2.9}$$

and the denominator is the critical surface mass density for lensing:

$$\Sigma_{crit}(z) = \frac{c^2}{4\pi G}\frac{D_{OS}(z)}{D_{LS}(z)D_{OL}(z)}, \tag{2.10}$$

where D_{OS}, D_{LS}, and D_{OL} are the angular diameter distances between the observer and source, the lensing system and the source, and the observer and the lens, respectively. The source in this case is the CMB emission at $z = 1100$.

We assume the widely used 'NFW' (Navarro et al. 1997) model for the density profile of a single dark matter halo applicable to Eq. 2.9, which is given by:

$$\rho(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)\left(1 + \frac{r}{r_s}\right)^2}. \tag{2.11}$$

We adopt a standard scale radius $r_s = r_{200}/c$, where r_{200} is the radius at which the enclosed halo is overdense with respect to the universe's critical density by a factor of 200, and c is the "concentration" parameter. We assume the mass-concentration-redshift relation of Ludlow et al. (2016), and convert this to a bias-concentration relation with the Tinker et al.

(2010) model. We thus characterize the profile of Eq. 2.11 in terms of halo bias and redshift rather than ρ_s and r_s.

The observed lensing convergence profile around a halo will be the superposition of the one-halo term and the "two-halo term", the lensing due to correlated LSS, derived using the Limber (1953) approximation (e.g., Oguri & Hamana 2011):

$$\kappa_2(\theta,z) = \frac{\overline{\rho_m}(z)}{(1+z)^3\Sigma_{crit}(z)D_A^2(z)} b_h \int \frac{ldl}{2\pi} J_0(l\theta) P_m(k,z), \tag{2.12}$$

where $\overline{\rho_m}(z)$ is the average matter density of the universe at redshift z, J_0 is the zeroth-order Bessel function, b_h is the linear halo bias, P_m is the matter power spectrum, and $k = l/[(1+z)D_A]$. We specify the linear halo bias in terms of the spherical overdensity halo mass and redshift by adopting the model of Tinker et al. (2010).

The total predicted lensing convergence amplitude from a population of dark matter halos and their surrounding LSS is given by the sum of the one and two-halo terms, averaged over the normalized redshift distribution of the lenses, $\frac{dn}{dz}$:

$$\langle\kappa\rangle = \int dz\, (\kappa_1 + \kappa_2) \frac{dn}{dz}. \tag{2.13}$$

The model of Eqs. 2.8-2.13 produces the expected κ profile for a given redshift distribution, and takes only the halo bias as a free parameter. To compare this model to our stacked data, we first filter the model in the same manner as the data, smoothing with a $15'$ Gaussian and removing Fourier modes $l < 100$. We then project the model onto the same grid as the data and measure both the data and the model profile by azimuthally averaging over radial bins. We then vary the halo bias in the model to optimize the fit as follows.

The smoothing applied to the data creates covariance between angular bins. We thus construct a covariance matrix by computing:

$$C_{ij} = \frac{1}{N-1}\sum_{k=0}^{N}(\kappa_i^k - \bar{\kappa_i})(\kappa_j^k - \bar{\kappa_j}), \tag{2.14}$$

where i, j are labels for the angular bins and each k represents a different realization of the stacked noise map. We construct these realizations by filtering and stacking 60 simulated noise maps provided by Planck in the same manner as the data maps.

We can now estimate a best fitting model by optimizing the likelihood function:

$$\chi^2 = (\kappa_{\rm obs} - \kappa_{\rm model})^T C_{ij}^{-1}(\kappa_{\rm obs} - \kappa_{\rm model}). \tag{2.15}$$

To evaluate the uncertainty on the best fit halo bias, we re-perform the fit on 60 maps generated by stacking bootstrap resamplings of the relevant quasar sample. In Figure 2.5, we illustrate an example of the result of this procedure by displaying the measured stacked lensing convergence profile for the middle color bin of the XDQSO quasar sample, along with the best fit model.

2.4.2 Modeling Correlation Functions

We also model the dark matter projected correlation function in our cosmology to fit the observed correlation functions and derive bias values. We thus first generate matter power spectra on a grid across the redshift span of each sample using `CAMB` (Lewis et al. 2000), then Fourier transform the power spectra to derive correlation functions. Next, we project the real space correlation functions $\xi(r)_{dm}$ to projected correlation functions $w_p(r_p)_{dm}$ with the Abel transform:

$$w_p(r_p)_{dm} = 2\int_{r_p}^{\infty} dr\, \frac{r\xi(r)_{dm}}{\sqrt{r^2 - r_p^2}}. \tag{2.16}$$

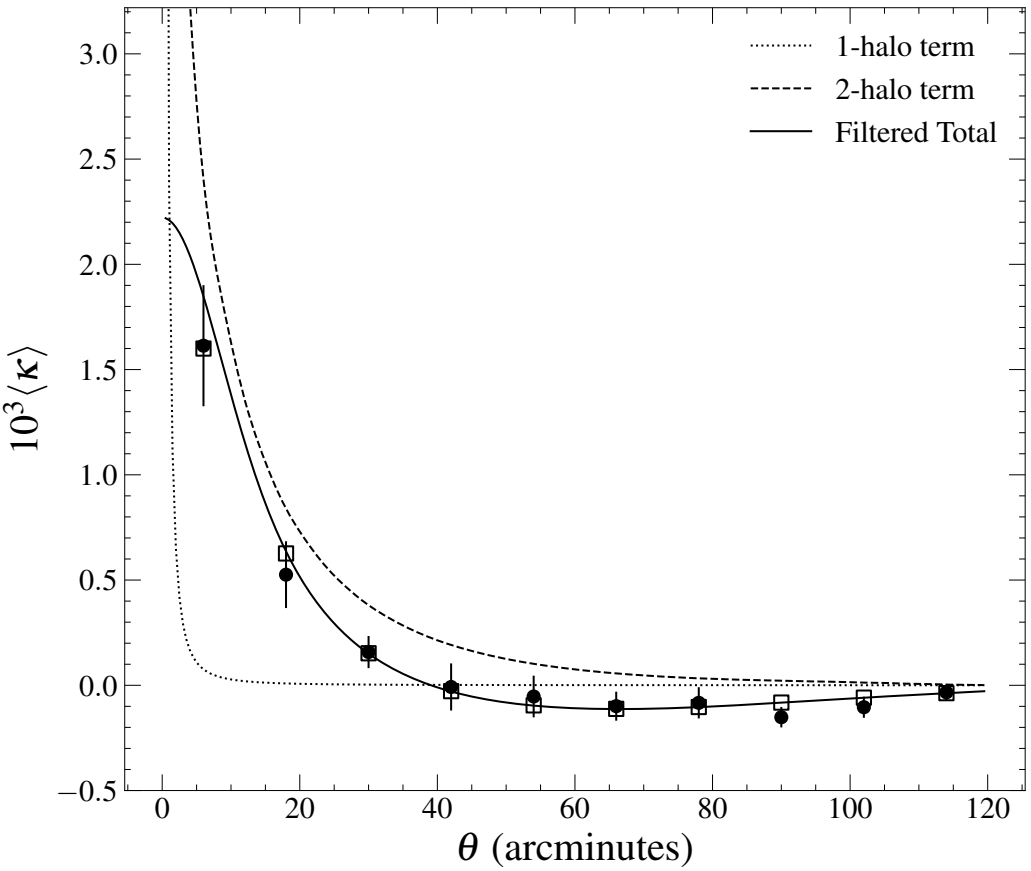

Figure 2.5: The measured lensing convergence profile for the stack of the median $\sim 14\%$ of XDQSO quasars (Figure 2.3, middle panel) and the best fit model. The circular points represent average κ values measured in annular bins of the stacked lensing map. The solid line represents the best fit model comprising the sum of the one and two-halo terms, after filtering. The open squares represent the value of the best fit model "measured" in the same bins as the data.

Finally, we perform a weighted average of the dark matter correlation functions at various redshifts, weighted by the redshift distribution of the sample (or the redshift overlap in the case of a cross-correlation, Eq. 2.7). The resulting correlation function is averaged in the same radial bins as the data and used to fit for a bias value. As we model only the two-halo term of the correlation function for the purposes of this study, we perform all fits on scales of $5 < r_p < 25\ h^{-1}$ Mpc, a regime which is dominated by the two-halo term and is governed by linear structure growth.

In order to estimate the absolute bias of a given set of quasars from a cross-correlation, one must have an estimate of the bias of the tracer population which is being correlated against. We thus first fit the autocorrelation functions of eBOSS quasars, LRGs, and ELGs to estimate their absolute bias. First, we compute the average of correlation functions at different redshifts, weighted by the amplitude of the normalized redshift distribution dn/dz of the corresponding tracer sample. The linear bias b of the sample with respect to the underlying matter distribution can then be fit according to:

$$w_p(r_p) = b^2 w_p(r_p)_{dm}. \tag{2.17}$$

The cross-correlation function between two LSS tracers can be fit by the product of the bias of each tracer with the dark matter correlation function:

$$w_p(r_p)_{QT} = b_1 b_2 w_p(r_p)_{dm}. \tag{2.18}$$

The uncertainties on the clustering measurements derive from bootstrap resampling sub-volumes of the data, described in Section 2.3.5. With these realizations, we construct a covariance matrix in an analogous manner to Eqs. 2.14 & 2.15. However, we find that fitting our data with these matrices give poor results due to noisy cross-terms, a problem noted by similar studies (e.g., Shen et al. 2009). Therefore, we use only the diagonal of the

covariance matrix in our fitting. We fit each realization for the bias parameter and take the variance of the results as our uncertainty for the bias of the full sample.

2.5 Results

In Figure 2.6, we display the autocorrelation function of eBOSS quasars and the best fitting model, with a bias value of $b_q = 2.35 \pm 0.02$. This is in excellent agreement with the finding from Laurent et al. (2017), which used a preliminary subset of the eBOSS quasar sample to estimate the evolution of quasar bias as a function of redshift, which gives $b_q = 2.36 \pm 0.04$ for our sample's redshift distribution. It is also fully consistent with the parameterization as probed by the Two Degree Field Quasar Redshift (2QZ) survey given in Croom et al. (2005), which predicts a bias of $b_q = 2.39 \pm 0.31$. The above measured bias value for the entire eBOSS quasar sample will be combined with the cross-correlation measurements to derive the absolute bias of each color sample. We estimate the bias of LRGs and ELGs in the same manner, which can then be used to probe the absolute bias of quasars through their cross-correlations.

We display the results of all three cross-correlation analyses, including the best fit bias parameter and corresponding characteristic halo mass for each color bin of quasars in Figure 2.7, and record the fit parameters in Table 2.2. These analyses demonstrate that optically-selected quasars of differing color occupy dark matter halos of similar mass. This result appears to hold across all three independent analyses which probe different effective redshifts. The results of each analysis also scatter around the Laurent et al. (2017) prediction for eBOSS quasars of all colors, demonstrating that our results are both self-consistent and consistent with previous findings in the literature.

By utilizing the independent technique of measuring the CMB lensing around XDQSO quasars, we find results entirely consistent with the cross-correlation analyses. Figure 2.8 displays the bias and halo mass derived from measuring the gravitational lensing signal

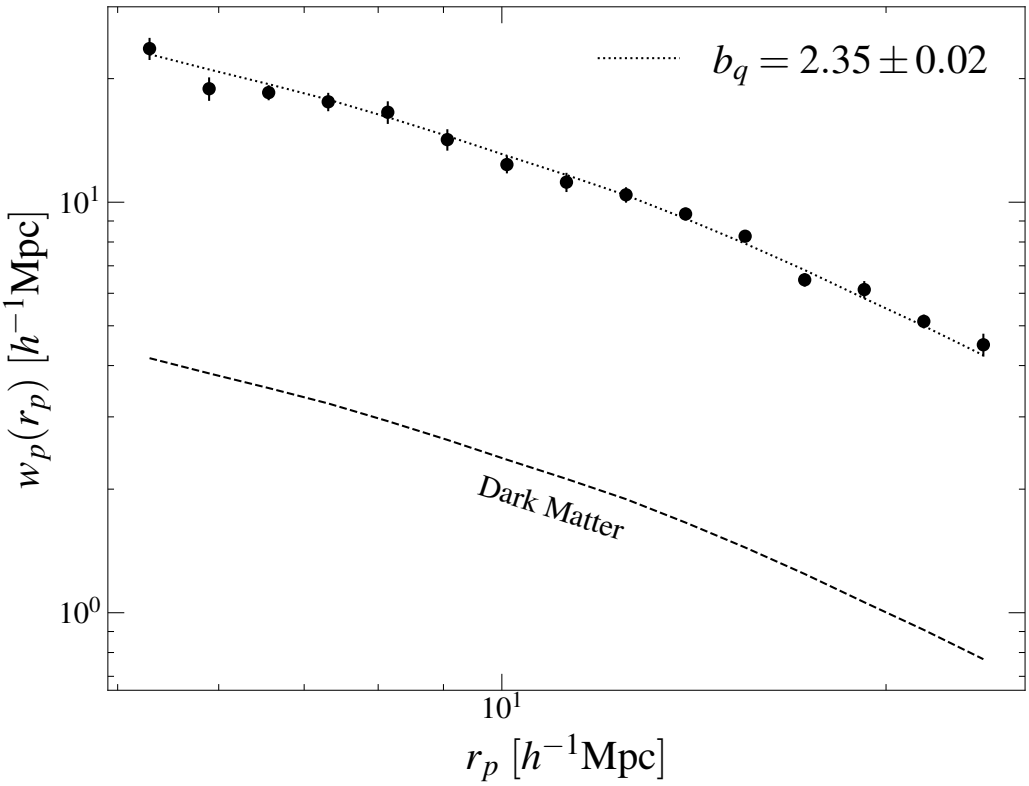

Figure 2.6: The projected autocorrelation function of eBOSS quasars (black circles), along with the modeled dark matter projected correlation function (dashed line) and the best model fit (dotted line). The dark matter model fits to the autocorrelations of LRGs and ELGs are of similar quality.

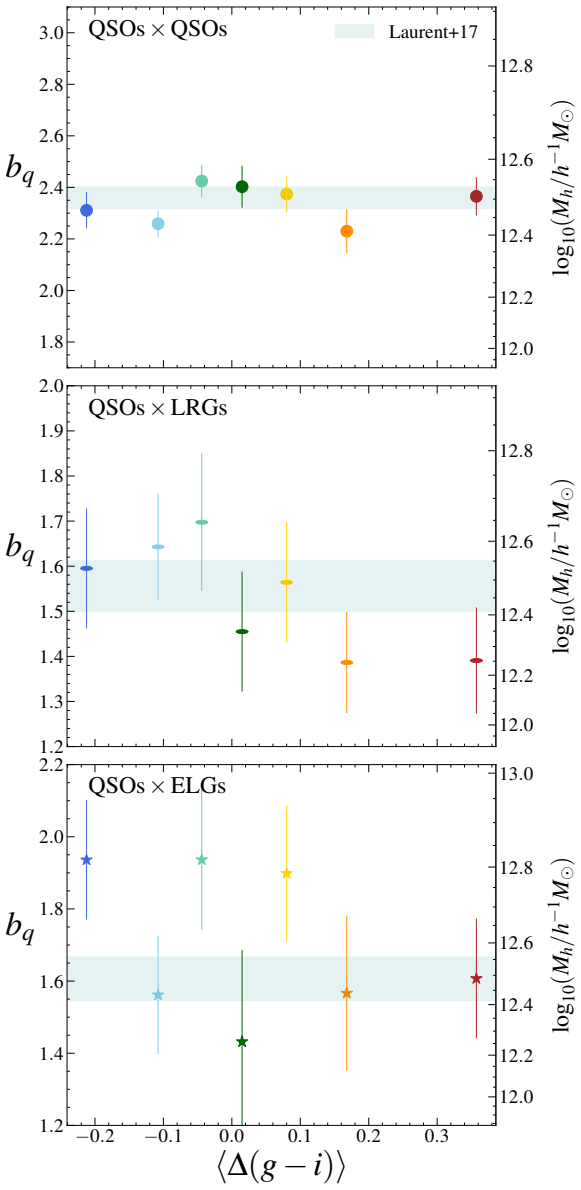

Figure 2.7: The measured bias as a function of median color offset for the three correlation function analyses. Top panel: the results from the cross-correlation of quasars with the entirety of the eBOSS quasar sample. Middle panel: the cross-correlation of quasars with LRGs. Bottom panel: the cross-correlation of quasars with ELGs. The circles, ellipses and stars show the best-fit bias from cross-correlations with quasars, LRGs, and ELGs. All three panels demonstrate a constant halo bias with quasar color within the uncertainties, appearing to scatter about bias estimates of optically-selected quasars from the literature (shown as teal bands representing 1σ bounds from Laurent et al. 2017). The secondary ordinate shows characteristic halo masses for the measured biases, demonstrating that quasars occupy similarly massive halos regardless of color and effective redshift.

of quasars' host halos and demonstrates a similar result. We display a summary of all of the above bias measurements in Figure 2.9, which shows the measured bias of the bluest and reddest bins of quasars as probed by the cross-correlations and the lensing analysis. It is apparent that our measured values are consistent with those found in the literature for optically-selected quasars. This figure also restates that the reddest and bluest optical quasars sample the underlying matter distribution in a statistically indistinguishable manner at high significance. This is in concordance with the findings of Shen et al. (2009), which found no difference in the clustering of quasars of different $g-i$ color, albeit at lower significance. Thus, the primary driver of optical quasar color is likely unrelated to the large-scale environment, which we further discuss in Section 2.6.

In Figure 2.10, we display the median radio-loudness as a function of quasar color for the eBOSS sample, defined in Eq. 2.2, where the median radio luminosity is derived from stacks of 1.4 GHz FIRST images. A clear relationship appears between quasar color and radio-loudness, similar to the trend found by White et al. (2007). Namely, the radio-loudness increases towards redder systems, except for a possible reversal in the bluest bin. This uptick may be understood as evidence that some proportion of quasars' colors stem from their orientation with the line of sight, as a blue face-on accretion disk with a jet traveling towards the observer will be relativistically beamed and radio-enhanced. However, the bulk trend of increasing radio-loudness in redder quasars is not immediately accounted for by an orientation-based mechanism, and bolsters the finding of excess FIRST detection fractions in optically red quasars (Klindt et al. 2019). As a check, we also test the FIRST-detection fraction as a function of color, finding that the reddest bin are detected ~ 2.5 times more often than quasars belonging to the bluest bin, consistent with the findings of (Klindt et al. 2019; Fawcett et al. 2020; Rosario et al. 2020). We conclude that the same differences in radio-properties highlighted by these studies appear in our sample, buttressing the result that orientation is not driving quasars' colors.

$\langle\Delta(g-i)\rangle$	(a) QSOs × QSOs		(b) QSOs × LRGs	
	b_q	M_h $10^{12}\ h^{-1}\ M_\odot$	b_q	M_h $10^{12}\ h^{-1}\ M_\odot$
-0.22	2.31±0.07	$2.95\pm^{0.35}_{0.32}$	1.60±0.13	$3.40\pm^{1.37}_{1.11}$
-0.11	2.26±0.05	$2.71\pm^{0.24}_{0.23}$	1.64±0.12	$3.86\pm^{1.27}_{1.07}$
-0.05	2.42±0.06	$3.52\pm^{0.33}_{0.32}$	1.70±0.15	$4.42\pm^{1.81}_{1.47}$
0.02	2.40±0.08	$3.41\pm^{0.44}_{0.41}$	1.45±0.13	$2.23\pm^{1.10}_{0.86}$
0.07	2.37±0.07	$3.26\pm^{0.36}_{0.34}$	1.56±0.13	$3.11\pm^{1.31}_{1.06}$
0.17	2.23±0.08	$2.58\pm^{0.39}_{0.36}$	1.39±0.11	$1.76\pm^{0.80}_{0.63}$
0.36	2.37±0.07	$3.22\pm^{0.38}_{0.36}$	1.39±0.12	$1.79\pm^{0.85}_{0.67}$

(c) QSOs × ELGs

b_q	M_h $10^{12}\ h^{-1}\ M_\odot$
1.94±0.17	$6.56\pm^{2.29}_{1.93}$
1.56±0.16	$2.72\pm^{1.46}_{1.12}$
1.94±0.19	$6.56\pm^{2.72}_{2.22}$
1.43±0.25	$1.81\pm^{1.99}_{1.20}$
1.90±0.19	$6.09\pm^{2.55}_{2.07}$
1.57±0.22	$2.76\pm^{2.00}_{1.41}$
1.61±0.17	$3.09\pm^{1.58}_{1.22}$

Table 2.2

A TABULATION OF THE BEST DARK MATTER MODEL FIT PARAMETERS TO THE CROSS-CORRELATIONS OF EBOSS QUASARS WITH OTHER EBOSS QUASARS (A), LRGS (B), AND ELGS (C), INCLUDING THE HALO BIAS AS WELL AS THE CHARACTERISTIC HALO MASS. THESE RESULTS ARE REPRESENTED VISUALLY IN FIGURE 2.7.

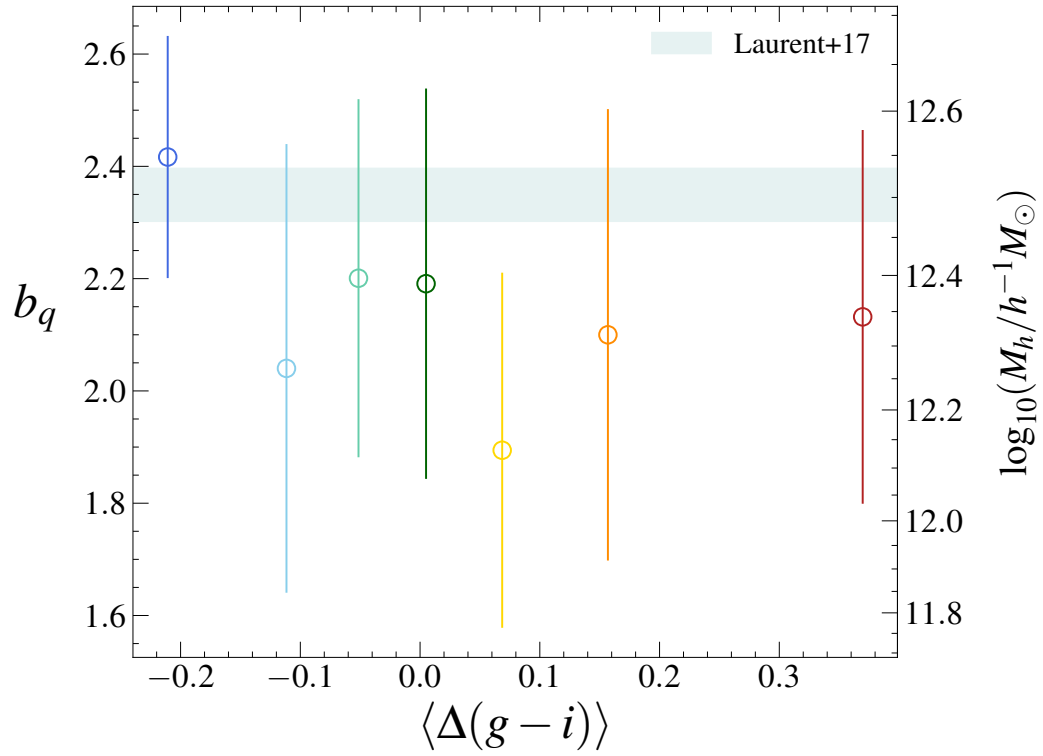

Figure 2.8: The measured bias (left ordinate) and implied characteristic host halo mass (right ordinate) as a function of median quasar color offset as measured through the Planck CMB lensing analysis of XDQSO photometric quasar candidates. Similarly to the results of the cross-correlation functions of eBOSS quasars (Figure 2.7), this analysis does not reveal any trend of halo bias with quasar color.

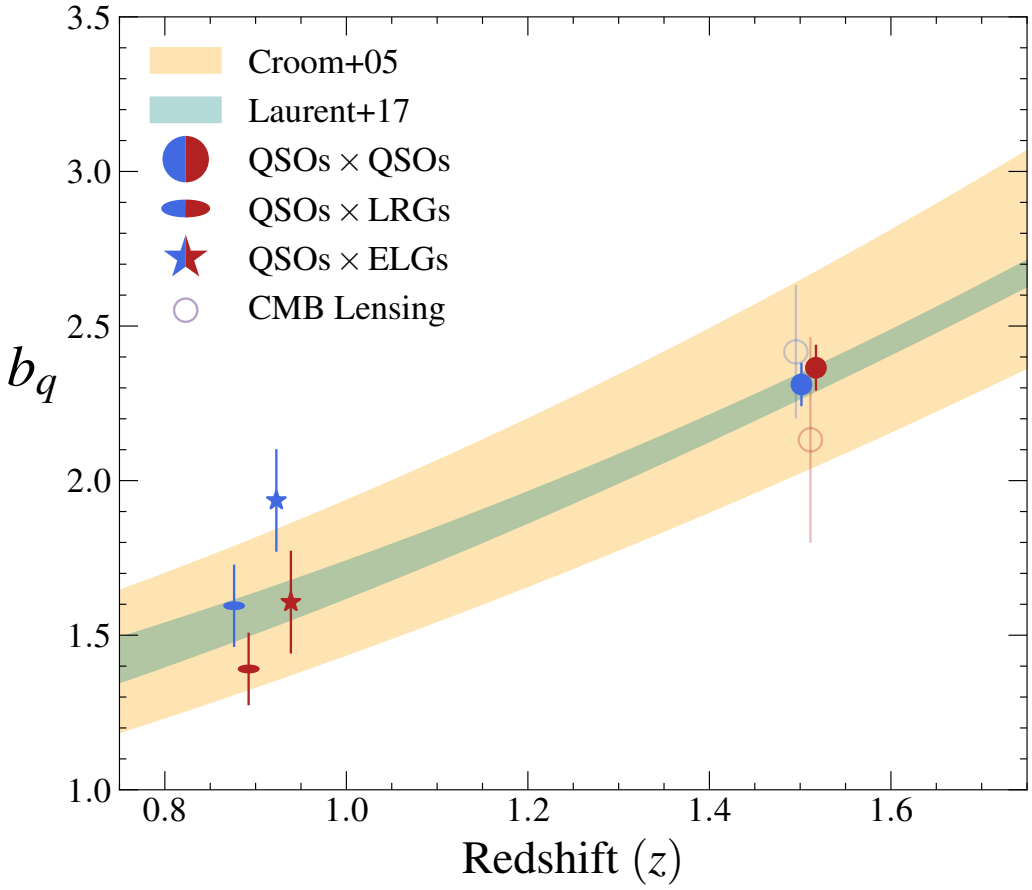

Figure 2.9: The bias of the reddest $\sim 14\%$ and bluest $\sim 14\%$ of quasars as probed by all of the above analyses at different effective redshifts. Circular markers, stars, and ellipses show results from cross-correlations with eBOSS quasars, ELGs, and LRGs, respectively. Open transparent circles display the results from the CMB lensing analysis of XDQSO photometric quasar candidates. All markers have been shifted slightly from their true effective redshifts for clarity. The parameterizations of optically-selected quasar bias with redshift from Croom et al. (2005) and Laurent et al. (2017) are shown with their 1σ bounds in orange and teal, respectively. The bias of the reddest and bluest quasars we measure in this work are consistent with the parameterizations from the literature, as well as with each other across $0.8 < z < 2.2$. Thus, optically red and blue quasars appear to occupy similar dark matter environments.

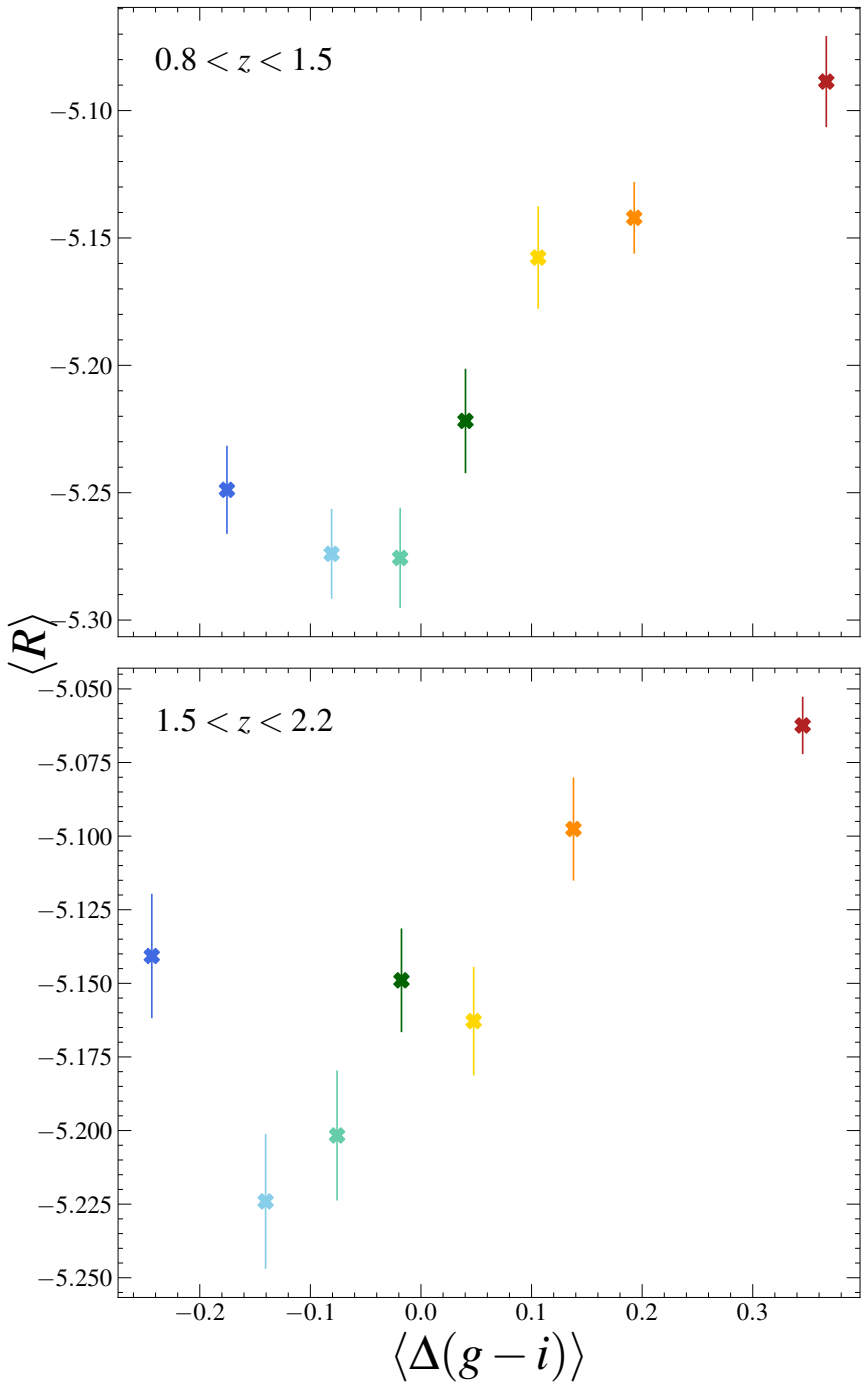

Figure 2.10: The median radio-loudness (Eq. 2.2) of the eBOSS quasar sample as a function of median color offset (from the Vanden Berk et al. 2001, template, Eq. 2.1). We compute this in two redshift bins, $0.8 < z < 1.5$ (top panel) and $1.5 < z < 2.2$ (bottom panel). These values were determined by median stacking FIRST images at the positions of quasars. The median radio-loudness clearly increases as quasars appear redder. The exception is the bluest bin, which may be interpreted as relativistic beaming. These values are all well below the classical radio-loud threshold of R = -4.24, converted from Klindt et al. (2019) using the Type 1 quasar template of Hickox et al. (2017).

2.6 Discussion

2.6.1 Interpretation

Robustly interpreting the clustering strength of different samples of quasars requires considering all other properties which might scale with halo bias, including redshift, luminosity, and also host galaxy properties like stellar mass and star formation rate (SFR) (Mendez et al. 2016). In this work, we have ensured that the quasar samples of different colors are matched in redshift and 1.5 μm luminosity by applying weighting schemes. We also note that Calistro Rivera et al. (2021) recently performed ultraviolet through far-infrared SED modeling on SDSS red and blue quasars to investigate their multiwavelength properties, finding that they have statistically identical distributions of stellar mass and SFR. Thus, we argue that we can interpret the clustering strength as a function of quasar color measured here as being uncontaminated by other differences between the samples.

We do not detect any trends of halo bias with optical quasar color from CMB lensing measurements of XDQSO quasars, nor from cross-correlations of eBOSS quasars with three independent tracers, robustly demonstrating that quasars exhibiting different optical colors reside in similar dark matter halo environments. This finding alone is consistent with the unified model of AGN, which attributes color to the chance inclination of the accretion disk/torus along the line-of-sight and would thus predict no trend of color with halo bias. However, recent studies have found that optically red quasars exhibit an excess of compact radio emission near the radio-quiet/radio-loud threshold, a result in direct tension with the unified model (Klindt et al. 2019; Rosario et al. 2020; Fawcett et al. 2020). According to the unified model, an inclined and red quasar would be less likely to exhibit excess radio emission owing to a lack of relativistic beaming toward the observer, and would would appear more extended due to jets propagating perpendicularly to the line-of-sight. In this work, we have confirmed the excess radio-loudness in red quasars through stacking of FIRST data. With a suite of analyses now demonstrating similar findings from a range of

different quasar samples and radio datasets, the differences between red and blue quasars cannot be primarily attributed to either orientation or large-scale environment.

We therefore suggest that the differences between red and blue quasars are generated on nuclear-galactic scales (roughly between the black hole's gravitational sphere of influence and the galaxy's extent, $10 \lesssim r \lesssim 10^4$ pc). This result is consistent with the finding that the excess radio-emission from red quasars is primarily driven by compact radio sources rather than extended classical radio-loud sources (Klindt et al. 2019). Rosario et al. (2021) recently studied a sample of red and blue quasars with high-resolution e-MERLIN imaging, confirming that red quasars exhibit excess radio emission only when the radio source is contained within the extent of the host galaxy. This interpretation is also supported by the work of Calistro Rivera et al. (2021), which found that SDSS red/blue quasars exhibit strikingly similar SEDs from the far-infrared to the ultraviolet aside from an excess of $2-5$ μm emission correlated with optical reddening. This excess was found to coincide with broad [OIII] lines associated with outflows, suggesting that winds of hot dust on nuclear scales may be the dominant source of optical quasar reddening. This could be consistent with the discovery of dust structures along the polar axis of some local AGN (Hönig et al. 2013; López-Gonzaga et al. 2016; Asmus et al. 2016), which might be associated with a quasar-driven wind (e.g., Schartmann et al. 2014).

It may be considered interesting that we do not detect trends of halo bias with color, given that red quasars are more often detected in FIRST than blue quasars, and FIRST-detected quasars at $0.3 < z < 2.3$ have been shown to inhabit more massive dark matter halos ($M_h \sim 5 \times 10^{13}\ h^{-1} M_\odot$) than their radio-quiet counterparts ($M_h \sim 2 \times 10^{12}\ h^{-1} M_\odot$, Retana-Montenegro & Röttgering 2017). This may indicate that the halo bias dichotomy in radio-loud/radio-quiet quasars only emerges when examining classical extended radio-loud systems. This idea is consistent with the finding that the differences in radio properties between red and blue systems appear to vanish when considering only systems with extended

radio sources (Klindt et al. 2019; Rosario et al. 2020; Fawcett et al. 2020; Rosario et al. 2021). It should be noted however that only $\sim 20\%$ of red quasars are FIRST-detected (Klindt et al. 2019), which would dilute any signal that may be present.

If quasars' observed optical colors are largely the result of different levels of dust extinction, it may also be considered surprising that we fail to detect trends of clustering with color given that obscured (Type 2) quasars have been shown to inhabit more massive halo environments than their unobscured (Type 1) counterparts (Hickox et al. 2011; Donoso et al. 2014; DiPompeo et al. 2014, 2015b, 2016a, 2017a; Powell et al. 2018), which likely indicates a large-scale evolutionary component in the obscuration mechanism (DiPompeo et al. 2017b; Whalen et al. 2020). In order to elucidate this, we repeat our clustering analysis as a function of color using $r-W2$ optical-infrared colors, which have been commonly utilized in the literature as a tracer of AGN obscuration (e.g., Hickox et al. 2007), rather than $g-i$ colors. We thus perform another cross-correlation analysis of the quasars in each bin with the entire eBOSS quasar sample, and derive bias values and characteristic halo masses exactly as described before. We show the resulting halo masses in Figure 2.11, along with the estimates of the characteristic halo masses of obscured and unobscured infrared-selected AGN from DiPompeo et al. (2017b), which utilized both angular clustering and CMB lensing measurements.

It is clear that eBOSS quasars occupy a small range of host halo mass regardless of optical-mid-infrared color, and that these masses are consistent with the halos surrounding unobscured infrared-selected AGN. It should be noted that the DiPompeo et al. (2017a) results are from a sample with a wider redshift range $0 < z < 3$, though quasar host halo mass appears to scale weakly with redshift. Given that the colors of infrared-selected quasars appear to be linked to their large-scale environments while optically-selected quasars' colors are not, this may suggest that the location of the obscuring material differs between Type 1 and Type 2 quasars. If obscured AGN and red quasars are simply different stages

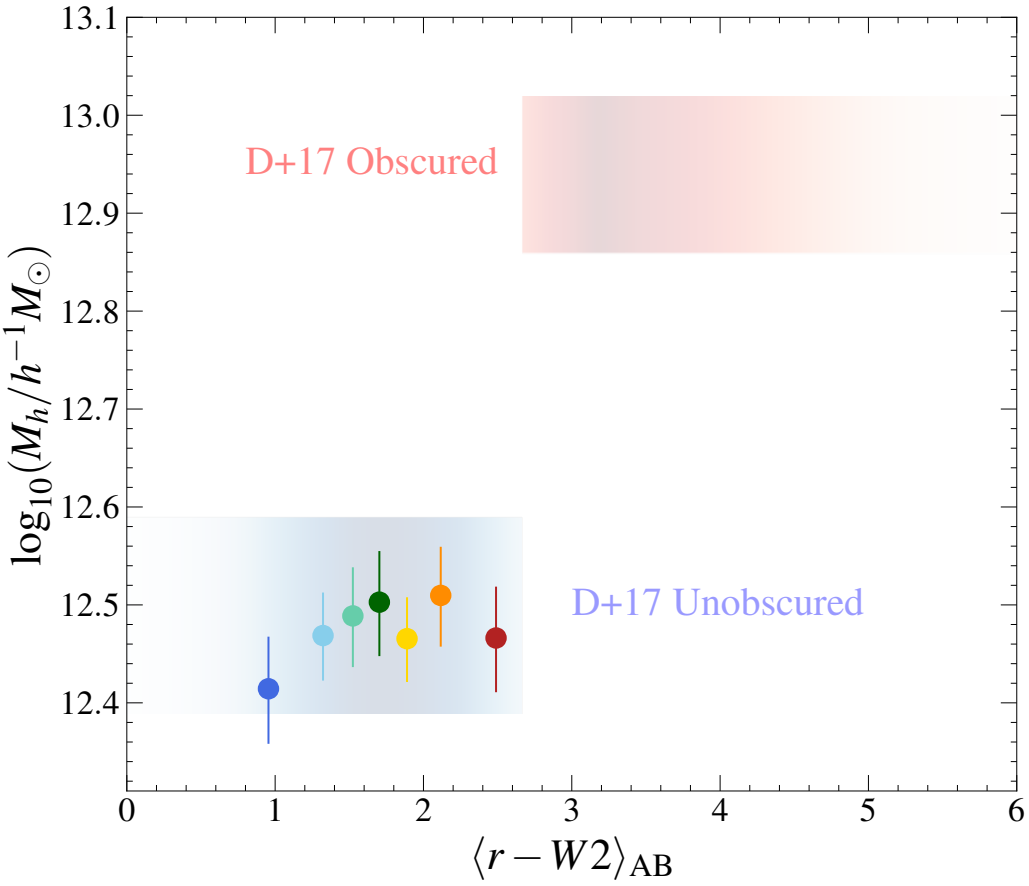

Figure 2.11: A comparison of the derived characteristic halo masses as a function of median $r-W2$ color of the eBOSS quasar sample (colored circles) and infrared-selected obscured/unobscured AGN from DiPompeo et al. (2017a) (red/blue gradients, respectively). The vertical span of the gradients represent the 68% confidence interval of the halo masses, while the opacity of the gradient represents the relative number of sources at a given $r-W2$ color in the DiPompeo et al. (2017a) sample. To derive halo masses for the eBOSS samples, we have binned the eBOSS sample in an identical manner as in Section 2.3.1, but have used $r-W2$ rather than $g-i$ colors, and then cross-correlated each bin with the entire eBOSS quasar sample. This demonstrates that optically-selected quasars occupy similar halo environments regardless of optical-infrared color, and that their characteristic halo mass is in excellent agreement with infrared-selected unobscured AGN.

in a continuous evolutionary sequence, this halo mass difference could alternatively indicate that the obscured phase lasts longer than the transition from obscured to unboscured, with red quasars representing a very brief intermediary stage (DiPompeo et al. 2017b; Whalen et al. 2020). Future work is required to test the halo bias across the full spectrum of color/obscuration to determine the level of extinction at which the halo bias "turns over" between Type 1 and 2 quasars, as well as control for redshift, luminosity, and host galaxy properties. This should constrain the nature of the obscuring material surrounding quasars as well as evolutionary models connecting star formation and AGN activity.

2.6.2 Systematics

Hot, ionized intra-cluster gas can bias CMB lensing measurements by distorting the CMB's spectrum via inverse-Compton scattering, known as the thermal Sunyaev-Zel'dovich (tSZ, Sunyaev & Zeldovich 1972) effect. We test for a possible tSZ bias to our lensing results by performing a stack in the Planck 2018 SMICA temperature map (Planck Collaboration et al. 2020c) at the same positions of quasars. We do not detect any small-scale temperature variation across the color bins expected for tSZ contamination. We do however detect a large-scale ($\sim 30^{\circ}$) temperature excess for the blue quasar sample at $\sim 1.5\sigma$. We examine the sky density of our red and blue quasar samples, and find an excess of blue quasars relative to red in the southern galactic hemisphere. This appears to be due to deeper SDSS imaging, which happens to overlap with a large-scale CMB warm spot, generating the large-scale temperature excess observed for blue quasars. This large-scale fluctuation should not affect our estimates of lensing, but we test this by removing all quasars from our sample which lie within HEALPix pixels where the ratio of red to blue quasars differs by more than 1σ from the median, and reperform the lensing analysis with this refined sample. Our results are unchanged within the uncertainties, indicating that this temperature fluctuation is not producing dominant systematic errors. Indeed, Geach & Peacock (2017) showed that lensing estimates are unbiased in the Planck maps even for galaxy clusters with strong

tSZ signals.

2.7 Conclusions

We have probed the dark matter halo environments surrounding SDSS quasars as a function of optical color to test whether red and blue quasars occupy different large-scale environments. We have achieved this by utilizing two independent measurements, the two-point cross-correlation functions of eBOSS quasars with three different eBOSS LSS tracers as well as the gravitational lensing of CMB photons around XDQSO quasar candidates. We do not detect any trends of halo bias with color, implying that red and blue quasars trace the underlying matter distribution in a similar manner and occupy dark matter halos of similar characteristic mass, $M_h \sim 3 \times 10^{12}\ h^{-1} M_\odot$. We have also corroborated recent findings demonstrating fundamental differences in the radio properties of red and blue quasars through a stacking analysis of FIRST images, strengthening the conclusion that accretion disk orientation is not the dominant driver of quasar color. These results together appear to suggest that the observed differences arise on nuclear-galactic scales, consistent with recent high-resolution radio-imaging results demonstrating that radio-properties differ between red and blue quasars only when the radio source lies within the host galaxy's extent (Rosario et al. 2021). We suggest that all of these observations are consistent with a model whereby red quasars' colors are generated by reddening through a nuclear dusty wind launched by the quasar system. Finally, we have shown that optically-selected quasars occupy similar halos across varying $r-W2$ optical-infrared colors, which may help constrain models of AGN and galaxy co-evolution and elucidate the relationship between reddened quasars and heavily obscured AGN.

Chapter 3

Host Dark Matter Halos of Wide-field Infrared Survey Explorer-selected Obscured and Unobscured Quasars: Evidence for Evolution

by Grayson C. Petter (Dartmouth College), Ryan C. Hickox (Dartmouth College), David M. Alexander (Durham University), Adam D. Myers (University of Wyoming), James E. Geach (University of Hertfordshire), Kelly E. Whalen (Dartmouth College), Carolina P. Andonie (Durham University)

This work first appeared in The Astrophysical Journal 2020, volume 946, page 27.

3.1 Introduction

Quasars are the most luminous class of Active Galactic Nuclei (AGNs), manifestations of accretion onto supermassive black holes (SMBHs) at the centers of galaxies. Since

their discovery (Schmidt 1963), AGNs and quasars have been classified into a taxonomy according to their observed multiwavelength properties (Padovani et al. 2017). However, it remains unclear whether or not many of the observed differences between quasar subpopulations reflect intrinsic or cosmologically relevant features of black hole growth.

The class of quasars which are "obscured" has been increasingly recognized as important to characterize. These systems often lack the optical and soft X-ray signatures of unobscured quasars, implying that significant columns of dust and gas lie between the sub-pc region surrounding the black hole and our line of sight. A majority of the AGN activity in the Universe is obscured (e.g., Ueda et al. 2014; Aird et al. 2015; Buchner et al. 2015; Ananna et al. 2019), which implies that traditional AGN surveys in the optical or X-rays have potentially led to a biased picture of SMBH growth. Some of the most pressing questions in extragalactic astrophysics require a complete census of AGN activity extending to even the most heavily obscured systems, including understanding the physical structure of AGN systems, probing the cosmic history of black hole growth, and uncovering the role that AGNs play in galactic evolution (Hickox & Alexander 2018). Crucial to these endeavors is understanding the nature of the obscuring material and whether obscured quasars represent intrinsically different systems to their unobscured counterparts.

The canonical view of the AGN structure is termed the "unified model" (Antonucci 1993; Urry & Padovani 1995; Elvis 2000; Netzer 2015; Ramos Almeida & Ricci 2017), which posits that most AGN systems consist of a similar axisymmetric structure of material surrounding the SMBH, but appear to differ because of our chance line-of-sight toward this structure. This model implies that a quasar will appear obscured when our viewing angle happens to nearly coincide with the plane of the dusty "torus". In this view, obscured and unobscured quasars represent intrinsically similar objects and are thus expected to occupy similar environments, even when accounting for potential relationships between accretion rate and torus covering factor (Whalen et al. 2020).

Alternatively, evolutionary models suggest that AGNs vary in their observed properties over the course of their lifetimes through interaction with their broader environments. In this scenario, obscured quasar activity could represent a specific evolutionary stage of SMBH growth. One such model posits that rapid star formation and SMBH growth are triggered during major galaxy mergers which funnel gas to the nuclear region of the merger remnant. This process is expected to generate quasar activity obscured by galactic or circumnuclear star-forming gas, before this activity heats and expels the gas to unveil an unobscured quasar (e.g., Sanders et al. 1988a; Hopkins et al. 2005, 2006, 2008; Alexander & Hickox 2012; Hickox & Alexander 2018). A number of recent studies have found evidence for AGN obscuration taking place on galactic scales or correlated with star formation activity (e.g., Chen et al. 2015; Buchner et al. 2017; Ricci et al. 2017a; Circosta et al. 2019; Li et al. 2021; Yan et al. 2021; Andonie et al. 2022; Gilli et al. 2022; Juneau et al. 2022), which may be connected to this evolutionary scenario.

A natural test of these models lies in measuring how quasars populate the large-scale structure (LSS) of the Universe, or the manner in which they occupy dark matter halos. This is because the properties of a quasar-hosting halo could not feasibly be connected with the obscuring torus' orientation along our particular line-of-sight, while a connection between halo properties and galaxy evolution is expected. Thus, this work aims to probe the nature of infrared-selected obscured and unobscured quasars by estimating the host dark matter halo properties of each class.

The connection between halo properties and obscuration is contested. Some studies find that obscured systems occupy more massive halos (Hickox et al. 2011; Elyiv et al. 2012; Donoso et al. 2014; DiPompeo et al. 2014, 2015b, 2016a, 2017a; Powell et al. 2018), some find the opposite (Cappelluti et al. 2010; Allevato et al. 2014), and others find no trend (Coil et al. 2009; Gilli et al. 2009; Ebrero et al. 2009; Mountrichas & Georgakakis 2012; Geach et al. 2013; Mendez et al. 2016; Jiang et al. 2016; Koutoulidis et al. 2018; Krumpe et al.

2018). These studies, though, vary in their selection, statistical power, analysis technique, obscuration definition, and sample distributions across luminosity and redshift.

In this work, we utilize the largest sample of mid-infrared-selected quasars to date in order to put precise constraints on the host halo properties of obscured and unobscured quasars at $\langle z \rangle = 1.4$. We probe these properties first by measuring angular autocorrelation functions, and interpreting these signals in a halo occupation distribution (HOD) framework. We find that obscured quasars occupy significantly more massive halos than unobscured quasars on average, but we do not detect a difference in the fraction that are satellites. Next, we test the halo properties with an entirely independent method, the cross-correlation of quasar positions with *Planck*'s map of the gravitational lensing of the CMB. We interpret the lensing signals with a linearly-biased model, and again find that obscured quasars occupy significantly more massive halos. The implied effective halo masses from the clustering and lensing analyses are in excellent agreement, which we show implies our results are robust against uncertainties in the obscured quasar redshift distribution. We interpret these results as favoring an evolutionary explanation for the obscuration of at least some quasars.

Throughout this work, we adopt a "Planck 2018" CMB+BAO Λ-CDM concordance cosmology (Planck Collaboration et al. 2020a), with $h = H_0/100$ km s^{-1}Mpc$^{-1} = 0.6766$, $\Omega_m = 0.3111$, $\Omega_\Lambda = 0.6888$, $\sigma_8 = 0.8102$, and $n_s = 0.9665$.

3.2 Data

3.2.1 WISE Quasar Sample

Though heavily obscured quasars can be challenging to distinguish from normal galaxies in the optical waveband, they can be recovered simply via their red mid-infrared colors (Lacy et al. 2004; Stern et al. 2005, 2012; Donley et al. 2012; Assef et al. 2013) which trace the reprocessed emission from the dusty torus. The Wide-field Infrared Survey Explorer

(*WISE;* Wright et al. 2010) has thus revolutionized obscured quasar selection by mapping the entire sky in four mid-infrared bands centered at (named) 3.6 (W1), 4.5 (W2), 12 (W3), and 22 μm (W4). These data currently provide the only diagnostic capable of producing highly reliable samples of millions of both obscured and unobscured quasars over the entire extragalactic sky.

We therefore elect to use the Assef et al. (2018) 90% reliable (R90) criterion on *WISE* photometry to select our parent sample of quasars. Rather than adopting the publicly released catalog associated with Assef et al. (2018) which used the criterion to select candidates from the AllWISE catalog (Cutri et al. 2021), we apply the criterion to the newer CatWISE 2020 (Eisenhardt et al. 2020; Marocco et al. 2020a; Marocco et al. 2021a) data release. This catalog is generated from *WISE* observations taken between 2010 and 2018, incorporating six times more exposures in the W1 and W2 channels than were used in generating the AllWISE catalog. This deeper imaging enables more precise measurements of photometric colors and therefore more reliable selection of quasars. Crucially, the deeper imaging will also enable a uniformly complete selection across the sky to fainter fluxes, greatly simplifying the reconstruction of the selection function necessary to perform a clustering measurement. We thus query the IRSA[1] database and select all objects in the CatWISE catalog satisfying the R90 criterion, adopting **mpro** Vega magnitudes. We also apply a magnitude limit in the W2 channel such that the selection is $> 99\%$ complete across the entire sky, and a bright-end cut to exclude infrared stars:

$$\begin{cases} \mathrm{W1} - \mathrm{W2} > 0.65 \times \exp[0.153 \times (\mathrm{W2} - 13.86)^2], \\ 9 < \mathrm{W2} < 16.5 \end{cases} \tag{3.1}$$

As the Assef et al. (2018) criterion was calibrated in the extragalactic Boötes field where Galactic objects are relatively rare, we must apply masks to remove regions which likely

[1] https://irsa.ipac.caltech.edu

suffer from higher contamination rates. For this, we follow the procedure described by Assef et al. (2018), though we create multi-order coverage (MOC) maps (Fernique et al. 2014) to accomplish this rather than remove sources geometrically. In particular, we mask regions within 10° of the Galactic plane and within 30° of the Galactic center. We also mask regions occupied by Galactic Planetary Nebulae (Acker et al. 1992), H_{II} regions (Anderson et al. 2014), star-forming regions (Lynds 1962, 1965), and resolved nearby galaxies in the Catalog and Atlas of the Local Volume Galaxies (LVG; Karachentsev et al. 2013) or the 2MASS extended source catalog (XSC; Skrutskie et al. 2006). We refer the reader to Assef et al. (2018) for more detail in this masking procedure. We will refer to this mask as the "contamination mask" throughout the remainder of this work.

One of the advantages of the CatWISE 2020 catalog over previous *WISE* releases is the order of magnitude improvement in detecting proper motions owing to the longer time baseline. This information allows the removal of stellar or solar system object contaminants from our quasar candidate sample. We thus remove sources with measured motions $> 0.25''/\mathrm{yr}$, which are able to be detected at $> 5\sigma$ across the sky down to our flux limit. This cut removes 1.5% of objects from the masked catalog.

Optical Data: Binning by Obscuration

The distribution of optical to mid-infrared colors of infrared-selected quasars appears bi-modal, as rest-frame UV-optical emission is easily extincted by dust, while near-infrared emission is less so. Therefore, a simple color cut can be used to classify obscured systems. Quasars with optical-infrared colors $r-\mathrm{W2} > 3.1$ [AB] typically show X-ray absorption corresponding to absorbing column densities of $N_{\mathrm{H}} > 10^{22}$ (Hickox et al. 2007). In order to classify our sources as "obscured" or "unobscured," we sought an optical survey deep enough to detect most of the quasars and as wide as possible to maintain a large sample size. We therefore utilize *r*-band optical data from the ninth release (DR9) of the Dark Energy Spectroscopic Instrument Legacy Imaging Survey (DESI-LS; Dey et al. 2019).

This survey covers $\sim 20,000$ deg^2 of extragalactic sky to $r_{\rm AB} \sim 24$, and the unprecedented combination of depth and area ensures that we are able to estimate the degree of obscuration for a uniform sample consisting of greater than one million quasars for the first time. We match our parent sample of *WISE* quasars to the DESI-LS catalog with a matching radius of $2''$ (Assef et al. 2013) using NOIRLab's Astro Data Lab[2] tools. The optical photometry is corrected for Galactic reddening using the map of Schlegel et al. (1998). We retain the $\sim 8\%$ of *WISE* quasar candidates within the DESI-LS footprint lacking optical counterparts and assign lower limits to their $r-W2$ color indices based on the local r-band imaging depth. To estimate the color distribution of sources undetected in DESI-LS, we match the non-detections to the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP) "DEEP" catalog (Aihara et al. 2018) in the areas where the surveys overlap. 96% of the DESI-LS non-detections are detected in HSC-DEEP, implying a negligible contamination rate of spurious WISE sources in our sample.

The resulting optical-infrared color distribution of our quasar sample is shown in Figure 3.1. It is clear that *WISE* mid-infrared selection reveals a population with a bimodal optical-infrared color distribution, and thus uncovers both obscured and unobscured sources in roughly equal proportion. We adopt the criterion of Hickox et al. (2007) ($r-$W2 [AB] ¿ 3.1) to separate the sample into obscured and unobscured subsets. This results in 52% of the sample being classified as unobscured, and 48% as obscured.

3.2.2 Sample Purity

In order to robustly interpret clustering measurements of quasar candidates, it is important to optimize and quantify the purity of the quasar selection. Mid-infrared quasar selection can be contaminated by low-redshift star-forming galaxies (SFGs) as well as by luminous high-redshift galaxies. Barrows et al. (2021) showed that the most significant source of sample contamination in the Assef et al. (2018) R90 catalog is from $z = 0.2-0.3$ galax-

[2]https://datalab.noirlab.edu

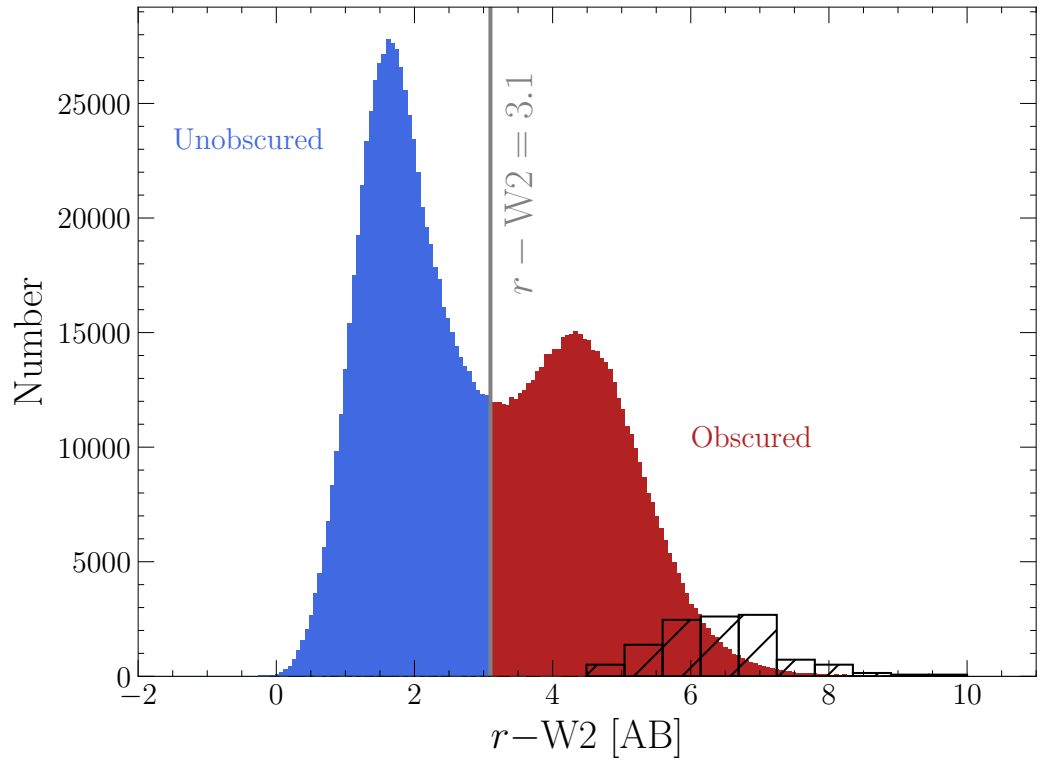

Figure 3.1: The optical-infrared ($r-$W2) color distribution of 1.4 million *WISE*-selected quasar candidates. The distribution is clearly bimodal, reflecting that infrared quasar selection recovers both obscured and unobscured populations. We classify quasars with $r-$W2 ¿ 3.1 [AB] as obscured, displaying them in red, and show unobscured quasars in blue. We note that all masking described in §3.2.1 and contaminant removal (§3.2.2) has been applied before producing this distribution. We also show the color distribution of the 8% of sources not detected in DESI-LS using black hatched boxes, obtained by matching against the HSC-SSP DEEP catalog.

ies with high specific star formation rates. Messias et al. (2012) developed a criterion to discriminate between low redshift star-forming galaxies and AGN using a near-infrared to mid-infrared color, $K-[4.5]$. In this work, we show that an alternative three-band diagnostic can effectively isolate low-redshift star-forming contaminants, as these objects are redder at $r-W2$ versus $z-W2$ than quasars. This was discovered by observing that optically-bright candidates appear bimodal in this color space, with objects redder at $r-W2$ compared to $z-W2$ having photometric redshifts $z<0.7$ (Barrows et al. 2021) and appearing resolved in DESI-LS imaging.

We display our candidates in this color space in Figure 3.2. Using a separate color, we show candidates which are highly resolved in the optical imaging, defined as being significantly better fit with a round exponential profile (REX) than as a point source ($\chi^2_{\rm REX}-\chi^2_{\rm PSF}>1000$). We also show predicted colors of optical quasars using the Type-1 template of Hickox et al. (2017), and of the prototypical starburst galaxy M82 from Polletta et al. (2007). It is clear that the optically-bright end of the R90 quasar selection criterion includes low-redshift galaxy-dominated contaminants. We use a simple intersection of two lines to excise sources in the SFG region of color space, which removes 12% of our full sample. We note that the quasar color distribution shown in Figure 3.1 was produced after culling these galaxy contaminants.

We briefly consider the reliability of this new catalog by matching it with the photometric redshift catalog produced in the Boötes field (Duncan et al. 2021). This catalog contains AGN diagnostic flags for every source denoting whether it appears in the 'Milliquas' compilation of spectroscopically confirmed quasars from the literature (Flesch 2019), exhibits mid-infrared *Spitzer*-IRAC colors indicative of AGN activity (Donley et al. 2012), or coincides with a luminous X-ray source (Kenter et al. 2005). 95% of our sample sources pass at least one of these criteria, while only 75% of sources in the Assef et al. (2018) R90 catalog do. We attribute this reliability increase to the more precise CatWISE photometric colors

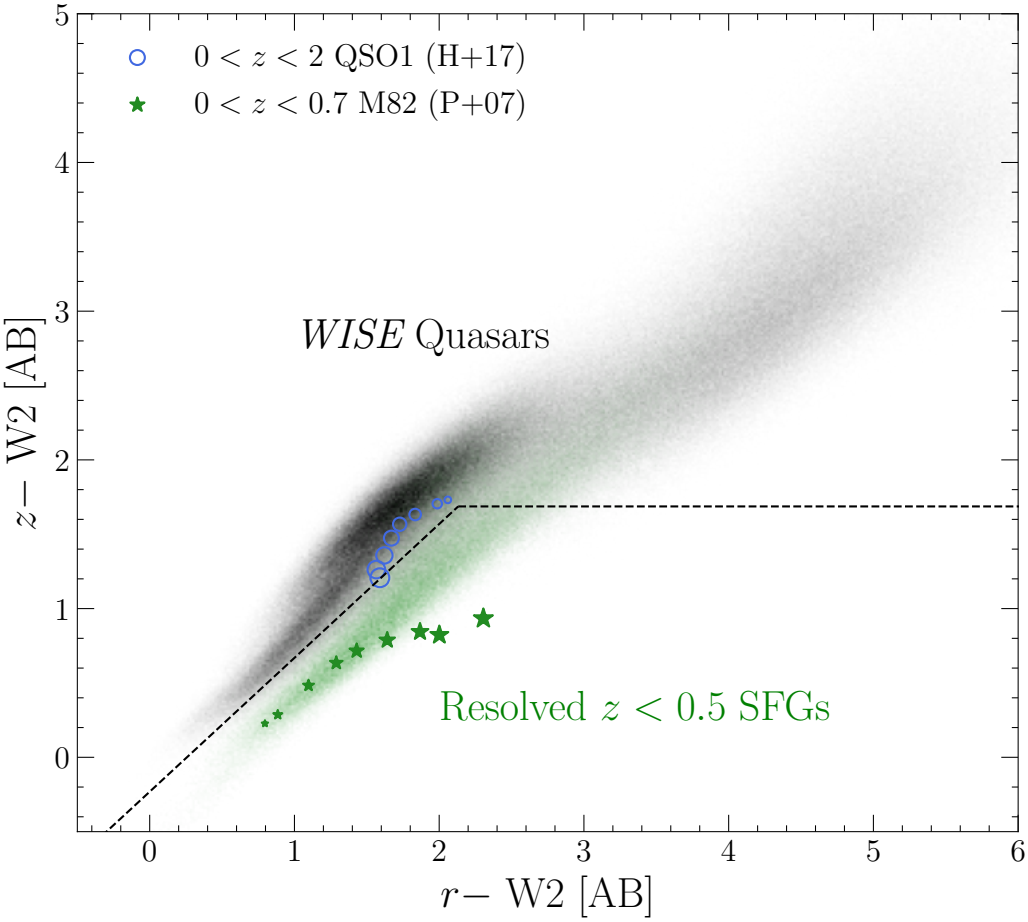

Figure 3.2: The distribution of Assef et al. (2018) R90 infrared quasar candidates in an optical (r), near-infrared (z), mid-infrared (W2) color space. We show the density of candidates in greyscale, and in green we show the density of highly resolved sources in the optical band representing low-redshift galaxy contaminants. Unobscured candidates (r-W2 ¡ 3.1) appear bimodal roughly along a line of constant $r-z$, which differentiates low-redshift SFG contaminants from quasars. Template colors of Type-1 quasars (Hickox et al. 2017) at $0<z<2$ and the starburst M82 at $0<z<0.7$ (Polletta et al. 2007) are shown with blue open circles and green stars, respectively, and increasing marker size represents linear steps towards increasing redshift. We show our criterion to remove low-z star-forming galaxy contaminants with a black dashed line.

as well as our removal of galaxy contaminants. Our catalog is thus expected to be highly reliable and any systematics introduced into our clustering measurements from non-quasar contamination will be subdominant.

3.2.3 Redshift Information

Interpreting the angular clustering or CMB lensing signal for a sample of quasars requires knowledge of the sample's redshift distribution. Targeted quasar redshift surveys are currently limited to optically-bright relatively unobscured quasars (e.g., Lyke et al. 2020; Alexander et al. 2022; Chaussidon et al. 2022), and individual blind redshift surveys lack either the breadth or depth to constrain the redshift distribution of our full sample. In order to obtain a nearly-complete and statistically representative estimate of the redshifts of our sample quasars, we therefore look to the well-characterized Boötes and COSMOS (Scoville et al. 2007) fields.

We select the portion of our sample falling within the footprint of the AGN and Galaxy Evolution Survey (AGES, Kochanek et al. 2012) in Boötes or the zCOSMOS (Lilly et al. 2007) footprint in the COSMOS field. We then performed a literature search for photometric and spectroscopic redshifts of the objects in these fields by matching them against various publicly-available redshift catalogs with a $2''$ search radius. We prioritize spectroscopic over photometric measurements, and adopt the spectroscopic estimate with the highest quality flag when available. We also reject photometric redshift estimates of $z > 3$ as unrealistic and likely the result of poor fits. In Boötes, we take photometric redshifts from Chung et al. (2014), and then Duncan et al. (2021), with the latter taking priority. In COSMOS, we adopt photometric redshifts from Delvecchio et al. (2017) and then Marchesi et al. (2016), which in turn depend on analyses from Salvato et al. (2009, 2011), Ilbert et al. (2009), and Laigle et al. (2016). The surveys providing the spectroscopic redshifts include AGES (Kochanek et al. 2012), SDSS DR16 (Ahumada et al. 2020), zCOSMOS (Scoville et al. 2007), PRIMUS (Coil et al. 2011; Cool et al. 2013), IDEOS (Hernán-Caballero et al.

2016), FMOS-COSMOS (Kashino et al. 2019), DEIMOS10K (Hasinger et al. 2018), and the studies of Prescott et al. (2006), Trump et al. (2009), Bussmann et al. (2012), Casey et al. (2012), Lacy et al. (2013), Assef et al. (2013), Comparat et al. (2015), Kartaltepe et al. (2015), Onodera et al. (2015), and Schulze et al. (2018). Overall, we find that 100% and 97% of unobscured and obscured quasars of the 899 total sources in these fields have a redshift measurement, respectively. We thus treat the redshift distributions as representative of their respective samples.

We show the resulting redshift distributions of the obscured and unobscured samples in Figure 3.3. Incorporating both photometric and spectroscopic redshifts, the distributions appear similar between the samples, implying their clustering properties can be appropriately compared. We verify this by measuring a p-value of $p = 0.15$ in a KS test, failing to reject the null hypothesis that the two samples are drawn from the same distribution. However, only photometric redshifts are available for approximately half of the obscured sample. Therefore, we will test the robustness of our clustering results against possible systematics in the obscured population's redshift distribution in §3.6.

3.2.4 Planck Lensing Convergence Map

The CMB radiation from $z \approx 1090$ has been gravitationally lensed by the intervening large-scale structure (LSS), and thus encodes information about the dark matter halos that host galaxies across cosmic time. Therefore, measuring the cross-correlation between the lensing convergence and the angular overdensity of quasars independently constrains the sample's host halo properties. In this work, we utilize the 2018 release of the CMB lensing convergence map (Planck Collaboration et al. 2020b) provided by data from the *Planck* satellite to probe the halo properties of quasars as a function of obscuration. We adopt the map generated using the minimum-variance (MV) estimator, which used both temperature and polarization data to reconstruct the lensing convergence. We also make use of simulated noise maps provided with the data release to estimate uncertainties. To generate

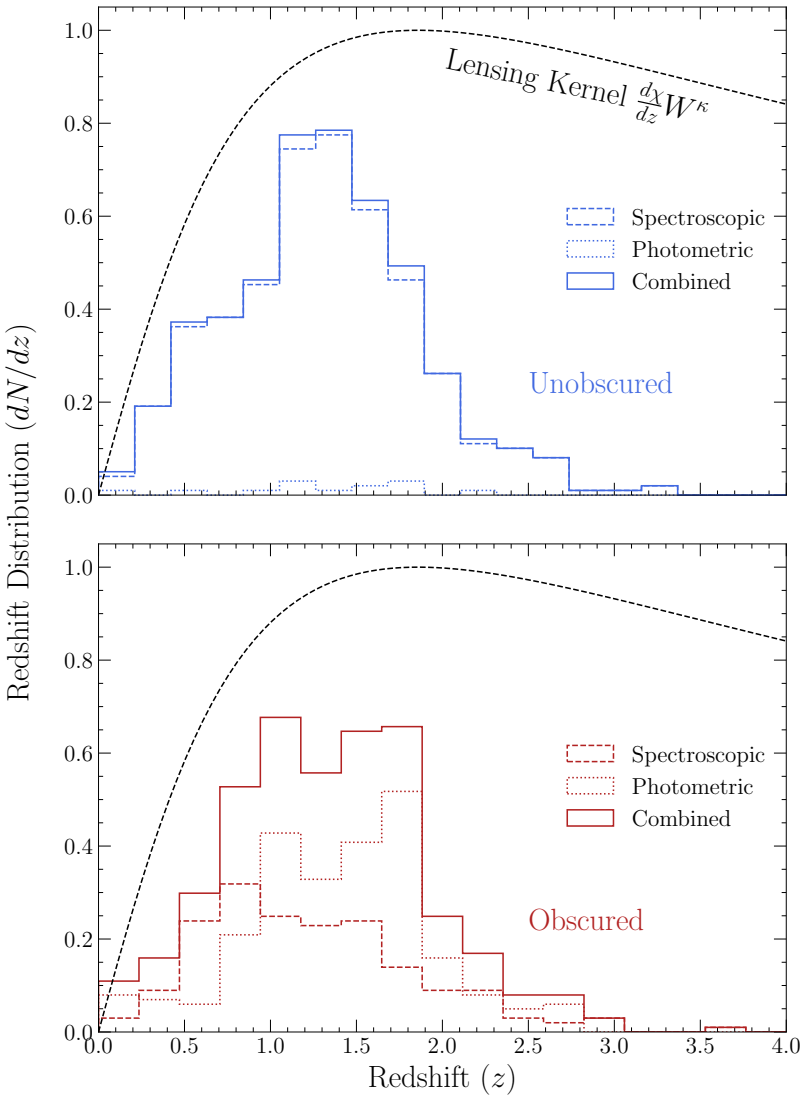

Figure 3.3: The redshift distributions (dN/dz) of the unobscured (top panel in blue) and obscured (bottom panel in red) quasar samples obtained by cross-matching to surveys in the Boötes and COSMOS fields. The samples are similarly distributed in redshift (a KS test p-value of $p = 0.15$ fails to reject the null hypothesis that the samples are drawn from the same distribution), implying that we can appropriately compare their clustering properties. Redshift distribution information stemming from spectroscopic redshifts is shown with a dashed line, a dotted line shows information from photometric redshifts, and a solid line shows the combinations. The unobscured distribution is derived mostly from spectroscopic redshifts, while we rely on photometric redshifts for half of the obscured population. We also display the shape of the CMB lensing kernel (Eq. 3.16) with a dashed black line, showing that the LSS surrounding our samples should efficiently lens CMB photons.

the lensing and noise maps, we transform the provided $\kappa_{\ell m}$ coefficients at $\ell < 2048$ into NSIDE=1024 resolution `HEALPix` maps.

3.2.5 Ancillary Photometry

Finally, we estimate the bolometric luminosity distributions of obscured and unobscured quasars to compute space densities and occupation statistics in §3.6.3. We use the rest-frame 6 μm band as a bolometric luminosity tracer which is minimally affected by obscuration. As 90% our sample falls within the Donley et al. (2012) AGN wedge, we suggest this wavelength is also primarily tracing the torus luminosity and is uncontaminated by star formation activity. For checking the location of these sources in *Spitzer* color-color space, we utilize mid-infrared *Spitzer* photometry for our sources from the *Spitzer* Deep, Wide-field Survey (SDWFS; Ashby et al. 2009) as collated by the *Herschel* Extragalactic Legacy Project (HELP, Shirley et al. 2019, 2021). We display the 6 μm luminosity distributions for obscured and unobscured quasars in Figure 3.4.

3.3 Measurements

3.3.1 Masking

Clustering measurements are inherently sensitive to systematics arising from a non-uniform selection function across the sky. Therefore, in addition to the contamination mask generated following Assef et al. (2018), we create several more masks to excise regions likely to introduce systematics into our measurements. As we permit optically-undetected quasars in our sample, we must carefully characterize the imaging footprint of DESI-LS to avoid classifying quasars as obscured simply because they lack deep optical imaging. Fortunately, DESI-LS DR9 provides random catalogs (Myers et al. 2022)[3] populated within the imaging footprint, with an associated local depth for each random point and each band.

[3] https://portal.nersc.gov/cfs/cosmo/data/legacysurvey/dr9/randoms/

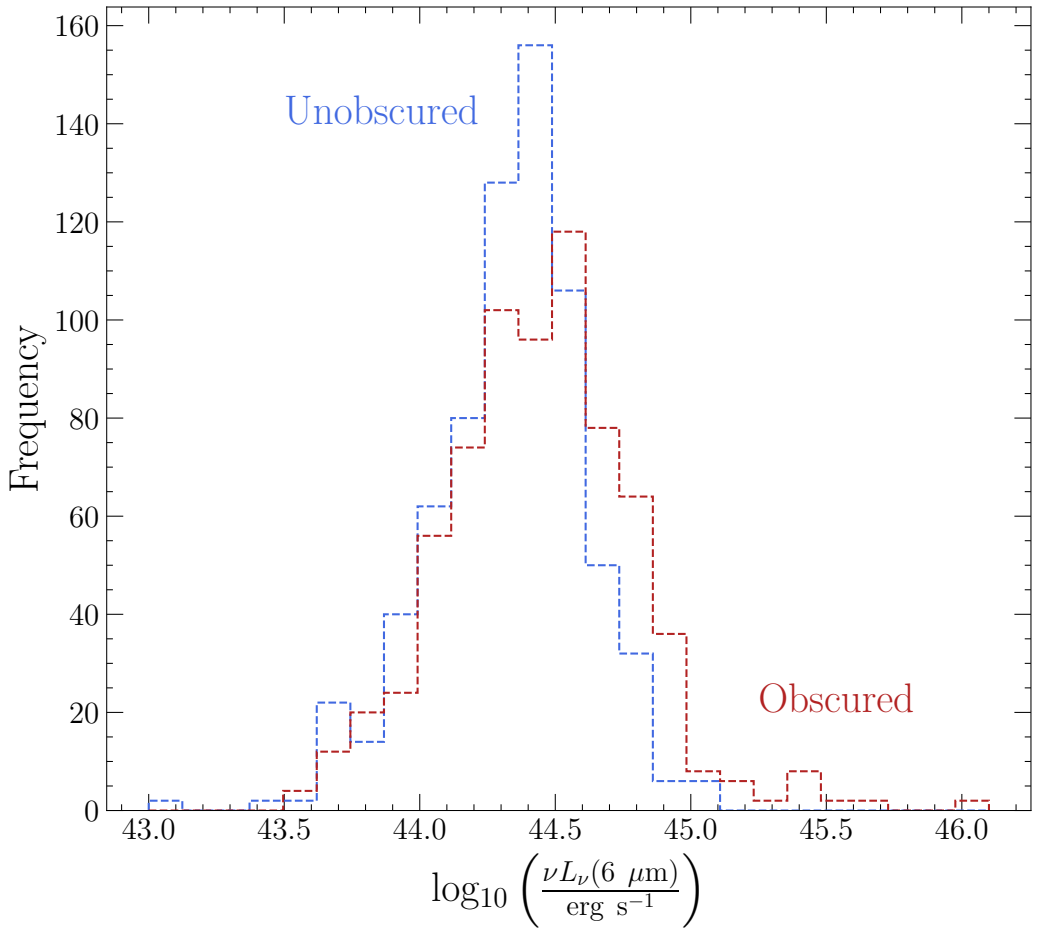

Figure 3.4: The 6 μm luminosity distributions for the obscured and unobscured quasar samples computed using SDWFS mid-infrared photometry. The samples are similarly distributed in luminosity, with a modest tail towards higher luminosities in the obscured sample. We utilize these distributions along with the clustering measurements in this work to study the occupation statistics of each subset in §3.6.3.

We take the median of the r-band depth in `HEALPix` cells, and mask cells where the 5-σ depth is $r < 23$. Next, we create a "reddening mask", excising regions with a high degree of Galactic reddening, $E(B-V) > 0.1$ (Schlegel et al. 1998). All of these masks are generated with `HEALPix` at NSIDE=1024 resolution.

We must also remove regions with severe *WISE* imaging artifacts, such as diffraction spikes, saturated pixels, latents and ghosts. However, these regions are often localized to scales of a few arcseconds, so masking the parent NSIDE=1024 pixel (corresponding to a $3.4'$ resolution) of every flagged region would cause us to remove an unacceptably large fraction of the sky from our analysis. Instead, we use the UnWISE bitmasks (Meisner et al. 2019) to remove sources in the data and random catalogs overlapping with flagged pixels. This process, as well as the contamination masking procedure (§ 3.2.1) effectively masks a fraction of each parent pixel, which we must correct for when generating the quasar overdensity map to cross-correlate with the CMB lensing map. Thus, we compute the fractional area lost within each NSIDE=1024 pixel, naming this the "lost-area" mask. We further apply a boolean mask to remove pixels in which the lost area is greater than 20% (e.g., Krolewski et al. 2020).

After applying the aforementioned masks, we visually inspect the sky distribution of our sample using `TOPCAT` (Taylor 2005) to check for remnant artifacts not captured by the masks. We find a small number of remaining regions which suffer from egregious bright star artifacts, where the quasar candidate sky distribution mirrors the pattern of large-scale (> 1 degree) diffraction spikes. As our candidate quasars are selected on their very red mid-infrared colors, these regions appear to be affected by the reddest and brightest infrared stars in the sky. We therefore find that using stars selected in the W3 band allows us to effectively mask these regions. We query the AllWISE catalog for sources brighter than W3 [Vega] < -1.5 (of which there are only ~ 200 sources in the extragalactic sky), and mask regions around them. We find that using a mask diameter in degrees equal to

the absolute value of the W3 Vega magnitude effectively removes the affected regions, as brighter stars tend to affect larger areas. We note that this choice represents a convenient rather than optimal solution. Finally, we observe an increase of the density of quasar candidates without optical counterparts within a few degrees of the Ecliptic poles, where the *WISE* imaging is deepest. We thus conservatively apply a mask within 10° of each Ecliptic pole.

We summarize the sample selection and masking as follows. The Assef et al. (2018) R90 selection of Eq. 3.1 curates a sample of $\sim$2.3 million *WISE*-detected sources within the $\sim 20,000$ deg^2 DESI-LS footprint. We report the percent reduction for each masking step in reference to this original sample as many sources are masked by multiple masking steps. 5% of the original sources are excised by the contamination mask generated following Assef et al. (2018) (§3.2.1). 10% of the original DESI-LS footprint is masked by our optical imaging depth requirement, while 6% of the area is vetoed by our reddening mask. The UnWISE bitmask flagging procedure removes 6% of sources, and the lost-area mask removes 3%. The bright infrared star mask removes 1% of the footprint, while the ecliptic pole mask removes 2%. After removing the 12% of likely contaminants with galaxy-like colors (§3.2.2) and the 1% of stellar interlopers with large proper motions (§3.2.1), we are left with a sample of $\sim$ 1.4 million quasar candidates over an effective area of $\sim$15,000 deg^2. We display the final combined mask atop the sky density of quasar candidates as a visual representation of this summary in Figure 3.5.

We apply all above masks to both the data and random catalogs required to perform a clustering measurement. We also mask the *Planck* lensing map with the mask provided alongside the Planck Collaboration et al. (2020b) data release.

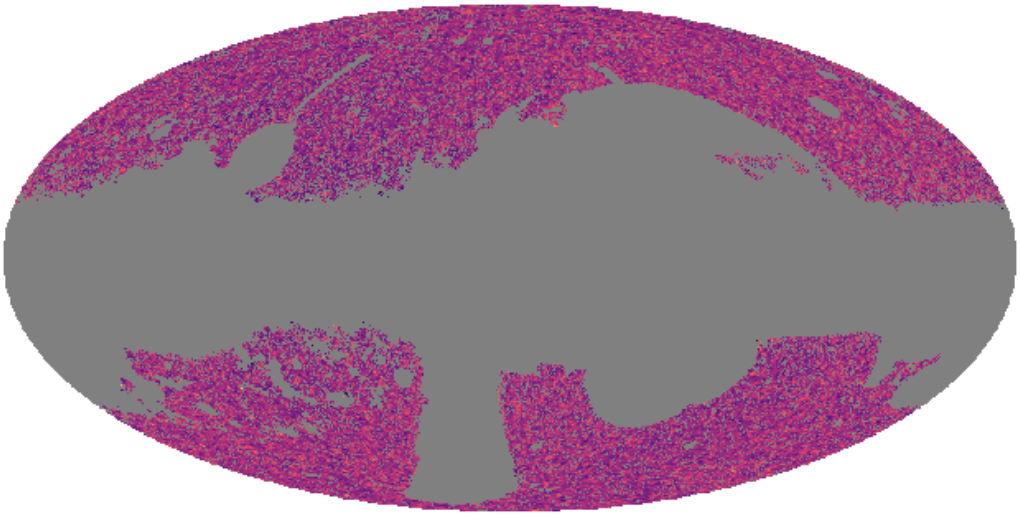

Figure 3.5: The sky density of ~ 1.4 million *WISE* quasar candidates over an effective footprint of $\sim 15,000$ deg^2, shown in Galactic coordinates. The density map has been smoothed with a $30'$ FWHM Gaussian for visual clarity. The combined mask, shown in gray and summarized in §3.3.1 shows significant complexity on both large and small scales, resulting in a quasar sample with sufficiently reduced angular systematics to perform clustering measurements.

3.3.2 Weighting Scheme

After applying the aforementioned masks to the data and random catalogs, we develop weights to further correct for large-scale angular systematics in the density of quasars (e.g., DiPompeo et al. 2017a). These weights will be used in both the lensing and the clustering analyses.

We might expect our sample density to vary with Galactic latitude due to increasing stellar density near the Galactic plane, or Ecliptic coordinates reflecting the *WISE* scanning pattern. Furthermore, by using optical data to bin the sample into obscured and unobscured objects, we might expect each subsample's density to vary with optical imaging depth. We measure the quasar density as a function of each of these variables and find the variation is largest for Ecliptic latitude and logarithmic optical imaging depth. We thus generate two-dimensional histograms of each sample along these dimensions and assign weights as the ratio of data points to random points in each bin. We have subsampled the random

catalog using the assigned weights such that the randoms match the data, and use these random catalogs in the correlation function measurements. Alternatively, for the CMB lensing cross-correlation, we assign the inverse of the weights to the data points rather than the randoms, using these weights when constructing the quasar overdensity field.

3.3.3 Angular Correlation Functions

Without redshift information for each of the objects in our quasar sample, we are limited to calculating angular correlation functions, which measure the clustering projected onto the surface of the sky. The two-point angular autocorrelation function $w(\theta)$ is defined as the excess probability — above that expected from an unclustered Poisson distribution — of finding a pair of objects at angular separation θ (Peebles 1980). We estimate the angular correlation function in this work using the Landy & Szalay (1993) estimator:

$$w(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{DD(\theta)}, \tag{3.2}$$

where *DD*, *DR*, and *RR* are weighted counts, normalized by number density, as a function of separation, for data-data pairs, data-random pairs, and random-random pairs, respectively.

We measure the angular correlation function using `Corrfunc` (Sinha & Garrison 2020) in 15 logarithmically-spaced bins between scales of $10^{-2.5} < \theta < 10^{-0.25}$ degrees. This range probes both regimes in which the one-halo term dominates, as well as the two-halo term, while avoiding the $\sim 6''$ resolution limit of *WISE* and also the regime at large scales where the Limber approximation begins to break down at the $> 10\%$ level, which is $\gtrsim 1^\circ$ for our sample's redshift distribution (Simon 2007).

To estimate uncertainties on the correlation functions, we perform a bootstrap resampling of the data (e.g., Efron 1982; Norberg et al. 2009). We divide the sample footprint into 30 equally-sized patches using the *k*-means clustering algorithm within the `scikit-learn`

package (Pedregosa et al. 2012). We then randomly draw from these patches with replacement, and recalculate a correlation function with the data and random points within these patches. This process is repeated 500 times and the variance across realizations is taken to be the variance of our measurement. We find that this error estimation agrees well with analytic Poisson errors in all cases on these scales (e.g., Myers et al. 2006).

3.3.4 Lensing Cross-Correlations

We also calculate the cross-power spectra $C_\ell^{\kappa q}$ of the *WISE* quasar overdensity fields with the CMB lensing convergence field measured by the *Planck* satellite. First, a quasar overdensity map is produced at the same resolution as the lensing map by performing a weighted count of sources in each cell, where the weights correct for the angular systematics discussed in §3.3.2. We correct this density field for effects of sub-pixel masking by dividing the density by the fractional-area-lost map. Finally, we convert this into a fractional overdensity map:

$$\delta_q = \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle} \tag{3.3}$$

We estimate the cross spectrum between the two maps using the psuedo-C_ℓ (e.g., Peebles 1973) algorithm `MASTER` (Hivon et al. 2002), as implemented in the `NaMaster` package (Alonso et al. 2019). This algorithm allows for fast and nearly-unbiased estimation of angular power spectra in the presence of sky masks. We measure the quasar-lensing cross-spectrum in 10 logarithmically-spaced bins of angular multipole moment (ℓ) from $100 < \ell < 1000$. When fitting models to the observed spectrum, we bin the theoretical curves using the same mode-coupling matrix and bandpower binning scheme as used for the data, though in all figures we show the unbinned model spectra for simplicity.

To estimate uncertainties on the cross-spectra, we utilize simulated noise maps released as part of the Planck lensing data product. After generating 60 noise maps in the same manner as the data, we measure the cross-spectrum between the quasar overdensity field and each

of the noise maps, taking the variance as our uncertainty estimate.

3.4 Modeling

3.4.1 Modeling Correlation Functions

We model the correlation functions in a halo occupation distribution (HOD) framework (e.g., Berlind & Weinberg 2002). To first order, the HOD $\langle N(M)\rangle$ is the mean number of quasars belonging to halos of mass M, decomposed into contributions from quasars at the centers of halos, $\langle N_c(M)\rangle$, and secondary or 'satellite" quasars belonging to the same halos $\langle N_s(M)\rangle$:

$$\langle N\rangle = \langle N_c\rangle + \langle N_s\rangle \tag{3.4}$$

The average number density of quasars is then an integral of the HOD over the halo mass function $dn/dM(M,z)$, for which we adopt the model of Tinker et al. (2008):

$$n_q(z) = \int dM \langle N\rangle \frac{dn}{dM} \tag{3.5}$$

We adopt the HOD model developed by Zheng et al. (2007) and Zehavi et al. (2011), which models the occupation of central objects as a softened step function:

$$\langle N_c\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log_{10}(M/M_{\mathrm{min}})}{\sigma_{\log_{10}M}}\right)\right] \tag{3.6}$$

where **erf** is the error function, M_{min} is the characteristic minimum halo mass required to host a quasar, and $\sigma_{\log_{10}M}$ is the softening parameter. The satellite HOD is given by:

$$\langle N_s \rangle = \Theta(M - M_0) \left(\frac{M - M_0}{M_1} \right)^{\alpha} \tag{3.7}$$

where Θ is the Heaviside step function, M_0 is the minimum mass to host a satellite quasar, and M_1 is the mass at which the term transitions to the power-law form. We note that while the an HOD model of this form has often been used in halo occupation studies of AGNs and quasars (e.g., Chatterjee et al. 2012; Richardson et al. 2012; Georgakakis et al. 2019; Alam et al. 2020), and thus we adopt this model for the sake of comparison to results from the literature.

This HOD model has five free parameters, which our present data will not be able to simultaneously constrain. We thus simplify the model as follows. We set $M_0 = M_{\rm min}$, such that the minimum halo mass required to host a satellite is the same as that required to host a central quasar. We fix the relatively unimportant parameter governing the softness of the step function to $\sigma_{\log_{10}M} = 0.4$ dex. Finally, we fix $M_1 = 12 \times M_{\rm min}$, motivated by the AGN HOD simulation results of Georgakakis et al. (2019) relying on the empirical accretion rate distributions of Georgakakis et al. (2017) and Aird et al. (2018) for luminous AGN. We note that fixing M_1 to be the same for obscured and unobscured quasars may not be valid. However, M_1 and α are degenerate with one another, roughly driving the relative strength of the one-halo term and thus the fraction of quasars which are satellites. Breaking the degeneracy between α and M_1 is not possible with present data. Importantly however, breaking this degeneracy in Eq. 3.7 is not critical to our analysis or interpretation, as we are primarily concerned with deriving the satellite fraction via $\langle N_s \rangle$. Thus, the two free parameters in our model are $M_{\rm min}$ and α. These parameters roughly govern the two-halo term and the one-halo term, respectively.

With HOD parameters specified, the derived halo properties of interest can be expressed. These include the effective bias:

$$b_{\rm eff} = \int dz \frac{1}{n_q} \frac{dN}{dz} \int dM \frac{dn}{dM} \langle N \rangle b_h(M,z) \tag{3.8}$$

and the satellite fraction:

$$f_{\rm sat} = \int dz \frac{1}{n_q} \frac{dN}{dz} \int dM \frac{dn}{dM} \langle N_s \rangle \tag{3.9}$$

The effective halo mass $M_{\rm eff}$ is computed by solving for the mass which would result in the effective bias (Eq. 3.8) when integrated over a sample's redshift distribution.

For a given HOD, the power spectrum of the quasar overdensity can be written as the sum of power contributed by pairs of quasars within the same halos (the one-halo term, $P_{1h}(k,z)$) and pairs of quasars between distinct halos (the two-halo term, $P_{2h}(k,z)$).

The one-halo term can be decomposed into pairs between satellites and pairs between satellites and central quasars. The satellites are prescribed to follow the density profile of halos. The function $\tilde{u}(k,M,z)$ is the Fourier transform of the dark matter density profile, for which we adopt the "NFW" (Navarro et al. 1997) model. We use the analytic solution of this profile's transform given by Scoccimarro et al. (2001):

$$P_{1h} = \frac{1}{n_q^2} \int dM \frac{dn}{dM} [2\tilde{u} \langle N_c N_s \rangle + \tilde{u}^2 \langle N_s (N_s - 1) \rangle] \tag{3.10}$$

We assume the satellites follow Poisson statistics such that $\langle N_s(N_s-1)\rangle = \langle N_s \rangle^2$. Setting $M_0 = M_{\rm min}$ (see above) also imposes the "central condition" (Murray et al. 2021), which stipulates that a halo can host satellite quasars only if it hosts a central quasar, such that $\langle N_c N_s \rangle = \langle N_s \rangle$.

The halo density profile transform $\tilde{u}$ approaches unity as $k \to 0$, such that the one-halo power spectrum approaches a constant at large scales, adding spurious power which can dominate over the two-halo term. We suppress this unphysical power by modifying the one-

halo power spectrum with the ad-hoc correction of Mead et al. (2021), with a characteristic damping scale $k^* = 10^{-2}$ h/Mpc:

$$P_{1h} \rightarrow P_{1h} \times \frac{(k/k^*)^4}{1+(k/k^*)^4} \tag{3.11}$$

The two-halo term is given by the matter power spectrum $P_m(k,z)$ multiplied by a bias factor:

$$P_{2h} = P_m \left[\int dM \frac{dn}{dM} b_h(M,z) \left(\langle N_c \rangle + \tilde{u} \langle N_s \rangle \right) \right]^2 \tag{3.12}$$

A known problem in the halo model framework is the underprediction of power compared to N-body simulations in the "quasi-linear regime," the transition scales between which the one and two halo terms dominate (e.g., Fedeli et al. 2014; Mead et al. 2015). This arises due to a natural breakdown of linear perturbation theory as well as halo exclusion effects, though self-consistently modeling these effects is an active area of research. Here, we adopt the empirical function of Mead et al. (2015):

$$P_{qq} = \left[(P_{1h})^\beta + (P_{2h})^\beta \right]^{1/\beta}, \tag{3.13}$$

which remedies this effect by smoothing the power spectrum in the transition region. We adopt the best-fit value of $\beta = 0.719$ from Mead et al. (2021).

With a HOD power spectrum specified, we can compute an angular correlation function over the redshift distribution of our sample using the Limber (1953) approximation (e.g., Peebles 1980; Peacock 1991; DiPompeo et al. 2017a):

$$w(\theta) = \int dz \frac{H(z)}{c} \left(\frac{dN}{dz} \right)^2 \int \frac{dk\, k}{2\pi} P_{qq}\, J_0[k\theta\chi(z)] \tag{3.14}$$

where J_0 is the zeroth-order Bessel function of the first kind. The k-space integral of Eq. 3.14 is simply a Hankel transform of the power spectrum, for which we utilize the `FFTLog` algorithm (Talman 1978; Hamilton 2000).

Thus, given two free HOD parameters of $M_{\rm min}$ and α along with a source redshift distribution, a model angular correlation function may be calculated. We fit the observed correlation functions using a Markov chain Monte Carlo (MCMC) method, as implemented in the `emcee` package (Foreman-Mackey et al. 2013).

3.4.2 Modeling the Lensing Cross-Spectrum

We also model the cross-spectrum between matter overdensity and CMB lensing convergence. The quasar overdensity and lensing fields are both projections of three-dimensional density fields onto the plane of the sky, and thus the cross-spectrum is given by a line-of-sight integral of the matter power spectrum $P(k,z)$ over the two respective projection kernels under the Limber (1953) approximation with the first order correction ($l \rightarrow l+1/2$; LoVerde & Afshordi 2008):

$$C_\ell^{\kappa q} = c \int dz \, \frac{W^\kappa(z) W^q(z)}{\chi^2(z) H(z)} P\left(k = \frac{\ell + 1/2}{\chi(z)}, \, z\right) \tag{3.15}$$

Here, c is the speed of light, $\chi(z)$ is the comoving distance, and $H(z)$ is the Hubble parameter. We generate the matter power spectra using the analytic form of Eisenstein & Hu (1998). The CMB lensing kernel, a measure of the efficiency of lensing by structure (when multiplied by $d\chi/dz = c/H(z)$) as a function of redshift is given by (e.g., Cooray & Hu 2000; Song et al. 2003):

$$W^\kappa(z) = \frac{3}{2} \Omega_{m,0} \left(\frac{H_0}{c}\right)^2 (1+z) \chi(z) \frac{\chi_{\rm CMB} - \chi(z)}{\chi_{\rm CMB}} \tag{3.16}$$

The quasar overdensity kernel is in turn given by:

$$W^q(z) = b_h \frac{H(z)}{c} \frac{dN}{dz} \tag{3.17}$$

where dN/dz is the normalized redshift distribution of the lenses and b_h is the linear halo bias.

We have elected to fit the lensing spectra as linearly biased with respect to dark matter as we have found that our present data are unable to constrain the two-parameter HOD model introduced in the previous section. In particular, the Planck lensing map is noisiest at the small scales (large ℓ-modes) where the one-halo term dominates. Therefore, we instead fit the lensing spectra by assuming an effective halo mass as the single free parameter and inserting the mass-bias relation of Tinker et al. (2010) into Eq. 3.17.

3.5 Results

The results of the angular autocorrelation measurements are displayed in Figure 3.6. It is apparent that obscured quasars cluster significantly more strongly than their unobscured counterparts (across a similar redshift range), implying that obscured quasars are more biased tracers of matter and occupy more massive dark matter halos. The data for both samples are very well fit by our two parameter HOD model. However, the relative strength of the one-halo term is similar for the two populations, implying similar satellite fractions.

The posterior distributions of the HOD parameters from the MCMC fits to the correlation functions are displayed in Figure 3.7. It is clear that obscured quasars occupy their host halos in a signifcantly different manner than unobscured quasars. The minimum mass required to host an obscured quasar is higher, implying both larger effective biases and host halo masses. In particular, obscured quasars appear to occupy effective halos of $\sim 10^{12.9} h^{-1} M_\odot$, while unobscured quasars occupy halos of $\sim 10^{12.6} h^{-1} M_\odot$. However, we do

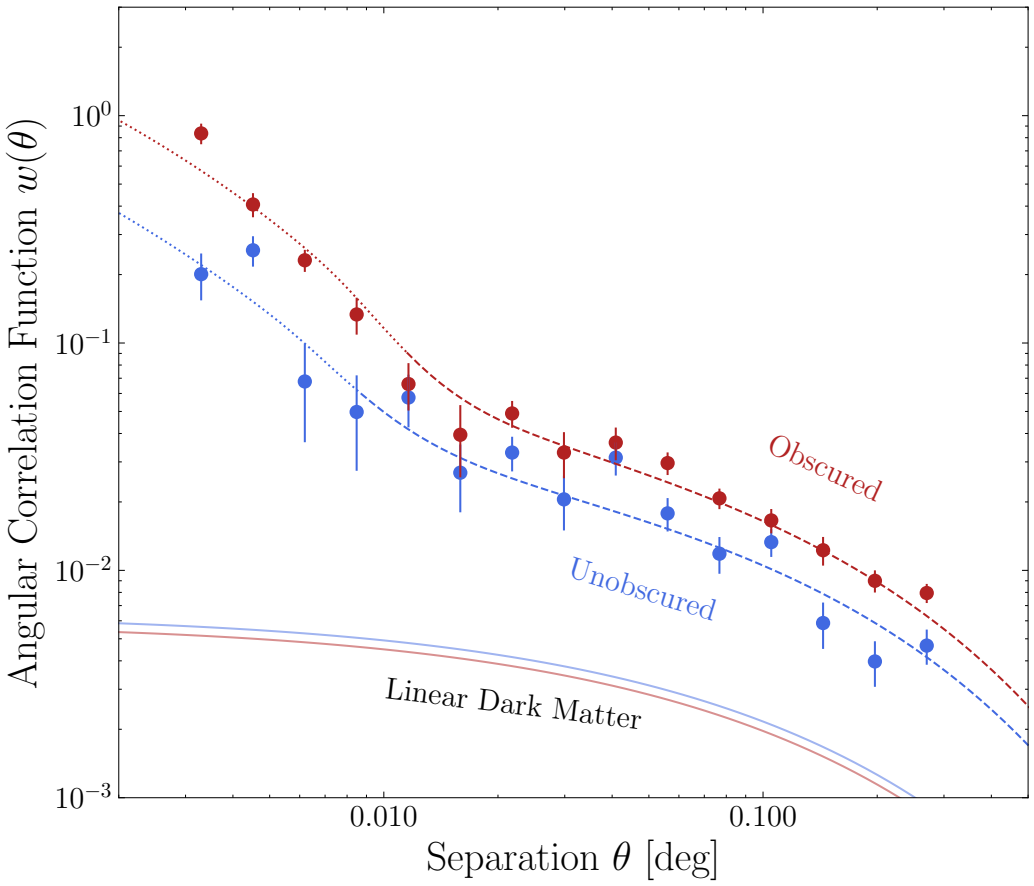

Figure 3.6: The angular autocorrelation functions of obscured and unobscured quasars are shown with red and blue markers, respectively. The model dark matter correlation functions for the corresponding redshift distributions are shown with solid lines. Finally, the best HOD model fits are shown with dashed/dotted lines. The dotted line indicates where the one-halo term dominates, while the dashed line indicates two-halo term domination. Obscured quasars cluster more strongly than their unobscured counterparts, implying they are a more biased tracer of matter and occupy more massive halos.

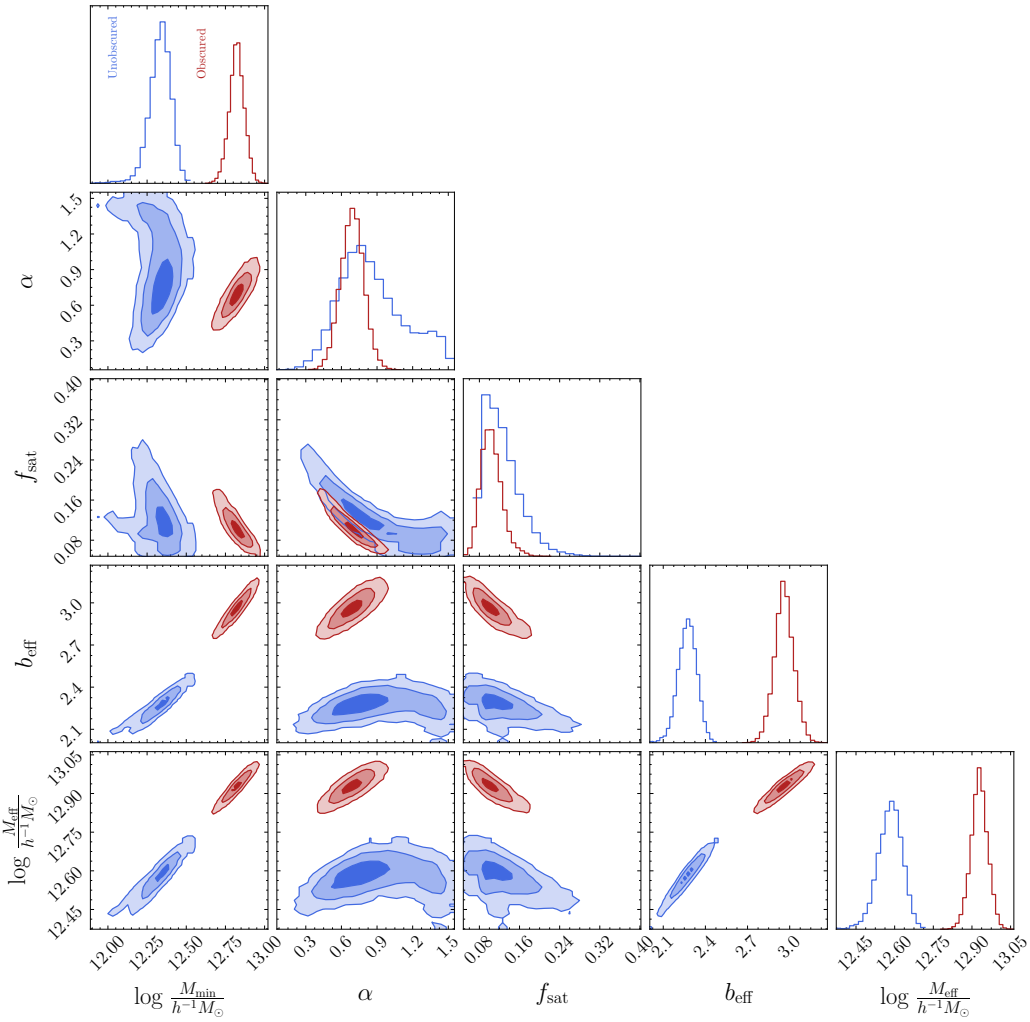

Figure 3.7: The posterior distributions of our two-parameter HOD model from MCMC fits to the angular autocorrelation functions of obscured and unobscured quasars, shown with red and blue contours, respectively. It is clear that the host halo properties differ significantly, with obscured quasars requiring a larger minimum host halo mass, implying that the effective mass and bias are also larger. However, we do not detect a difference in the one-halo term index α, or the satellite fraction $f_{\rm sat}$.

not detect a difference in the satellite power law index α, nor the derived satellite fraction, finding that $\sim 5-20\%$ of obscured and unobscured quasars are satellites within their halos. A non-detection of a difference in the satellite fraction is in contrast to the study of Mitra et al. (2018). We thus find from angular clustering measurements that obscured quasars occupy more massive halos than unobscured systems, but are not more likely to be satellites within their halos.

We also investigate the host halo properties of quasars as a function of obscuration using an independent technique, by calculating cross-correlations of quasar overdensities with the CMB lensing convergence map measured by *Planck*. We display the result of this in Figure 3.8. In agreement with the clustering analysis, we find that obscured quasars are a significantly more biased tracer of dark matter, again implying more massive host halo environments. We are unable to constrain the one-halo term, and thus satellite fraction with present data, as the *Planck* CMB lensing map is noisiest at small scales, as discussed in §3.4.2. Therefore, we fit the observed spectra with linearly-biased dark matter models, which provide good fits to the data.

We summarize our results in Figure 3.9 and tabulate them in Table 3.1. This shows most strikingly that the implied effective halo masses from clustering and CMB lensing analyses are in excellent agreement, with both showing that obscured quasars occupy significantly more massive effective halos. We show that infrared-selected unobscured quasars occupy similar effective halos as optically-selected spectroscopic Type-1 quasars from eBOSS (e.g., Laurent et al. 2017). We also show that our results are consistent with those of DiPompeo et al. (2017a), demonstrating that the enhanced clustering of *WISE*-selected obscured quasars persists both with improved precision and at higher redshifts.

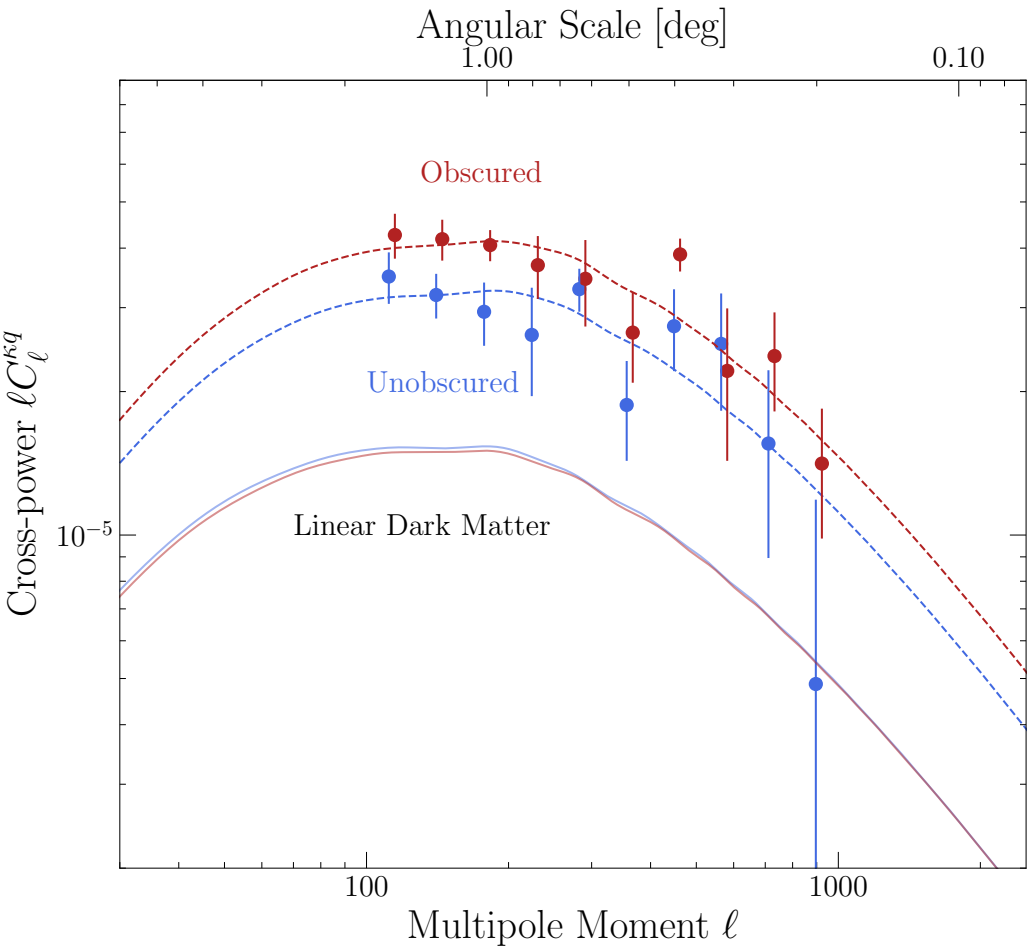

Figure 3.8: The cross-power spectra (multiplied by ℓ to reduce dynamic range) between obscured/unobscured quasar overdensity and Planck CMB lensing convergence κ are shown with red and blue markers. The model spectra of linear dark matter for the corresponding redshift distributions are shown using solid lines, while the best model fits are shown with dashed lines. Obscured quasar density correlates more strongly with lensing convergence, implying that obscured quasars are more biased tracers of matter and occupy more massive halos.

Table 3.1. Obscured and unobscured quasar bias and halo masses

Analysis	Sample	b_q	$\log_{10}(M_h/h^{-1}M_\odot)$
Clustering			
	Unobscured	2.27 ± 0.06	$12.58^{+0.04}_{-0.05}$
	Obscured	2.96 ± 0.07	12.93 ± 0.03
Lensing			
	Unobscured	2.29 ± 0.18	12.60 ± 0.09
	Obscured	3.02 ± 0.19	12.96 ± 0.06

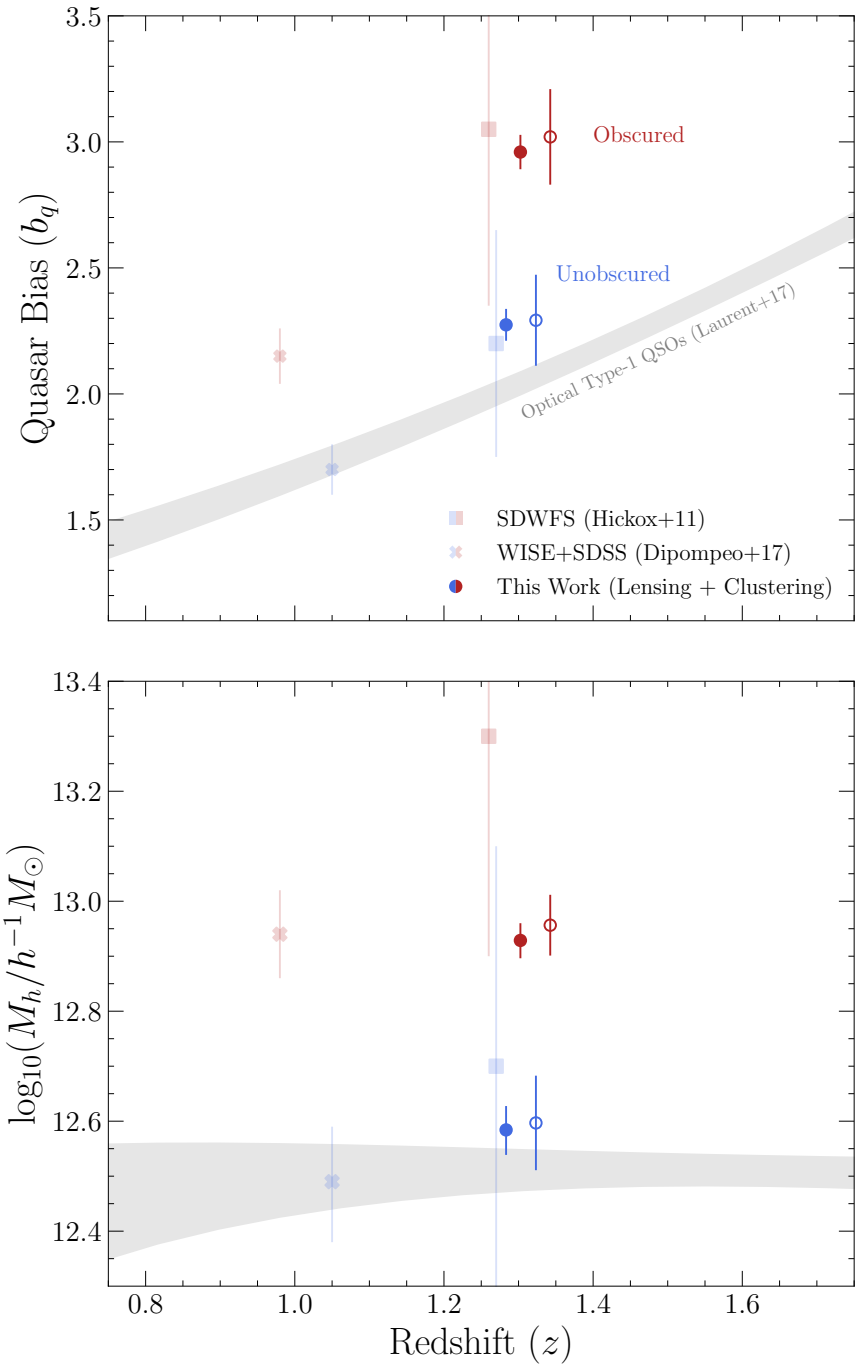

Figure 3.9: The bias (top panel) and effective halo mass (bottom panel) of obscured and unobscured quasars. We show the results for obscured quasars in red, and unobscured quasars in blue. Our angular clustering results are shown with closed circles, while the CMB lensing results are displayed with open circles, offset horizontally for visual clarity. We also show clustering measurement results of infrared quasars at similar effective redshifts from (Hickox et al. 2011) and (DiPompeo et al. 2017a), as well as for SDSS Type-1 quasars (Laurent et al. 2017).

3.6 Discussion

In §2.5, we have shown that *WISE*-selected obscured quasars occupy significantly more massive halos than unobscured quasars do. In this section, we will discuss possible systematics in these measurements and speculate on interpretations of this observed halo occupation difference.

3.6.1 Redshift Systematics

Inferring the host halo properties of a sample from angular statistics is inherently sensitive to systematics in the sample's redshift distribution, with the uncertainty in the latter oftentimes dominating over the statistical uncertainty of the clustering measurement (e.g., Coil 2013). As approximately half of our estimation of the obscured quasar redshift distribution relies on photometric redshifts (Figure 3.3), it is important to quantify whether our result could be explained by photometric redshift systematics, as was suggested by the authors of Mendez et al. (2016) to explain the similar results of Donoso et al. (2014) and DiPompeo et al. (2014).

We test this possibility by exploring the potential obscured quasar redshift distributions which would resolve the observed effective halo mass difference between obscured and unobscured quasars while reproducing our measurements. This was achieved by replacing the observed photometric redshift distribution of obscured quasars with many simulated distributions and refitting the data. First, the obscured quasar photometric redshift distribution of Figure 3.3 was fit as a normal distribution, finding a mean and dispersion of $\mu_{\rm obs} = 1.4$, $\sigma_{\rm obs} = 0.5$. We then vary these parameters, shifting the mean by $-1 < \mu_{\rm sim} - \mu_{\rm obs} < 3$ and varying the dispersion between $0.1 < \sigma_{\rm sim}/\sigma_{\rm obs} < 1.5$. Redshifts are randomly drawn from this new distribution, and only values $z > 0$ are retained. Distributions with $> 10\%$ of the cumulative distribution function below zero are rejected outright. These simulated redshifts are finally recombined with the observed spectroscopic redshifts. We then refit

the obscured quasar clustering and lensing signals (on scales dominated by the two-halo term) and determine which combinations of the redshift distribution parameters imply an effective halo mass consistent with that measured for unobscured quasars. The combinations which resolve the tension are explored in an MCMC analysis, for which we display the results in Figure 3.10.

It is clear that the dependence of clustering and CMB lensing measurements on redshift systematics are independent, with clustering being primarily sensitive to the distribution width and the lensing sensitive to the mean. We observe that the only possible configuration which would resolve the halo mass tension while reproducing both the clustering and lensing measurements would be if the obscured quasars without spectroscopic redshifts were distributed around $z \approx 3.5$, which we regard as infeasible. The photometric redshifts stem mostly from Duncan et al. (2021), which reports a typical photometric redshift scatter of 7% and an outlier fraction of $\sim 20\%$ for AGNs. In order to resolve the observed clustering differences, the true outlier fraction for obscured quasars would need to be nearly 100% if these systems were truly narrowly distributed at $z \approx 3.5$. Therefore, we argue that our measurements of halo mass differences are robust against systematics from uncertainties in the redshift distribution. This test highlights the power of combining clustering and lensing measurements to mitigate redshift systematics.

3.6.2 Interpretation of Clustering Difference

A significant observed difference in the host halo properties of matched samples of obscured and unobscured quasars appears to negate the simplest unified model which attributes obscuration solely to viewing angle. An interesting alternative in the context of galaxy evolution is an evolutionary model which posits that obscured and unobscured quasars represent different phases in a coevolutionary scheme between galaxies and their nuclear black holes. DiPompeo et al. (2017b) proposed a simple evolutionary model in which black hole growth lags behind dark matter halo growth (e.g., Alexander et al. 2008;

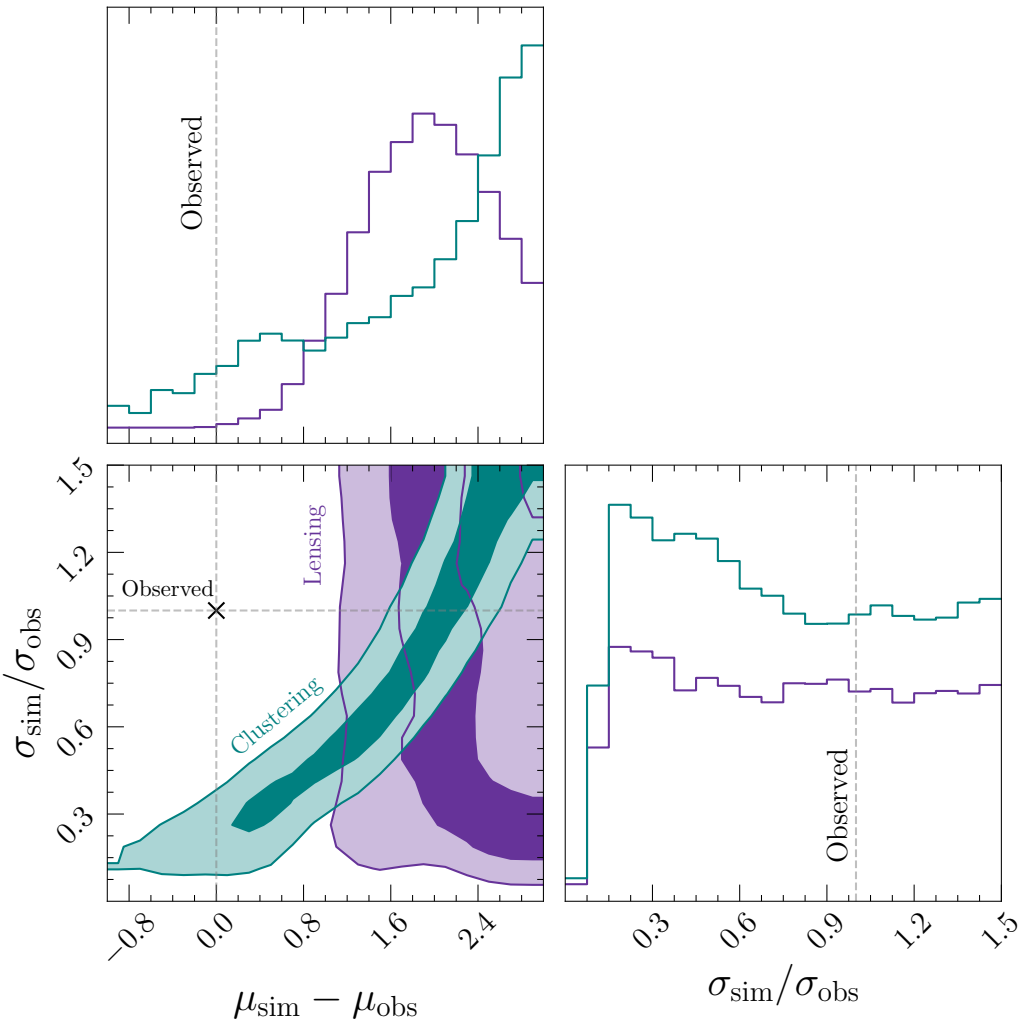

Figure 3.10: A corner plot demonstrating a test for systematics in our effective halo mass results due to photometric redshift distribution uncertainties. We have approximated the measured photometric redshift distribution of obscured quasars as a Gaussian, then shifted the mean μ and dispersion σ from their observed values until the implied effective halo mass for obscured quasars matches that of unobscured quasars. We show the 68% and 95% confidence contours for the redshift distribution properties required to resolve the halo mass discrepancy for the angular clustering and CMB lensing measurements in teal and purple, respectively. In order to resolve the halo mass tension while reproducing both the clustering and lensing signals, the half of our obscured quasar sample in deep fields without spectroscopic redshifts would need to be distributed around $z \approx 3.5$ instead of the value derived from photometric redshifts of $z = 1.4$, which we regard as implausible. Our result is thus robust against redshift systematics.

Woo et al. 2008; Kormendy & Ho 2013). If obscured quasars represent an early evolutionary phase of rapid SMBH growth and thus their black holes are systematically undermassive with respect to their halos (they have not yet "caught up" to the SMBH halo relation), then the obscured subset of a luminosity-limited quasar sample (which is in turn black hole mass-limited for a fixed Eddington ratio distribution) will systematically occupy more massive halos.

However, it is also possible that such a difference could arise from non-evolutionary effects. The modeling work of Whalen et al. (2020) explored whether the observed clustering difference between obscured and unobscured quasars reported by DiPompeo et al. (2017a) could be explained by such non-evolutionary scaling relations. In particular, they studied whether the results could be understood within the "radiation-regulated unification" schema of Ricci et al. (2017b), where an observed relationship between AGN Eddington ratio and torus covering factor is interpreted to imply that the radiation pressure from black holes accreting closer to their limits can expel toroidal dust and expose more unobscured sight lines to a potential observer. Therefore, one would expect a reduced incidence of obscured quasars represented at high Eddington ratios, which has been confirmed at low redshift by Ananna et al. (2022). This implies that at fixed luminosity, obscured quasar activity will preferentially be generated by more massive black holes accreting at lower Eddington ratios, which will thus lie in more massive halos according to black hole-halo mass relations (e.g., Kormendy & Ho 2013). Such a model could potentially explain the observed host halo difference without necessitating an evolutionary scheme.

Whalen et al. (2020) also studied whether the observed clustering difference could be explained by scaling relations between host galaxy stellar mass and ISM column density (Pannella et al. 2009; Buchner et al. 2017; Whitaker et al. 2017). As more massive galaxies are observed to contain larger columns of obscuring gas, obscured quasars may be preferentially selected in massive galaxies and thus halos without regard to an evolutionary

process.

The modeling work of Whalen et al. (2020) showed that the observed halo mass difference of DiPompeo et al. (2017a) cannot be explained either through radiation-regulated unification nor through scaling relations with galaxy mass while simultaneously reproducing the observed fraction of obscured quasars. With the updated effective host halo masses presented in this work, which are consistent with but roughly twice as precise as the results of DiPompeo et al. (2017a), these arguments become even more stringent. We thus favor an evolutionary explanation of the host halo mass difference observed in this work.

However, accurate comparison of the clustering properties between populations of quasars requires controlling across properties such as redshift, quasar luminosity, and host galaxy properties such as stellar mass and star formation rate (SFR). Indeed, several works suggest that AGN clustering properties depend only on the host galaxy property distributions revealed by different AGN selection techniques (e.g., Mendez et al. 2016; Yang et al. 2018; Powell et al. 2018, 2020; Krishnan et al. 2020; Aird & Coil 2021). In this work, we have shown that *WISE* obscured and unobscured quasars exhibit similar redshift and bolometric luminosity distributions. However, a potential concern is that obscured quasars may appear to occupy more massive halos because of a selection effect in which they are detected in host galaxies differing from the hosts of unobscured systems. Future work on modeling the full spectral energy distributions of *WISE* quasars is required to test whether the observed clustering difference may be understood in terms of host galaxy properties. However, Andonie et al. (2022) recently performed panchromatic SED modeling of infrared-selected quasars in the COSMOS field, finding no difference in the stellar mass distribution for obscured and unobscured quasars.

3.6.3 Non-torus-obscured Population

Although a robust trend between quasar obscuration and halo properties rules out the possibility that obscuration is *always* driven by viewing angle with the torus, it does not exorcise the role of nuclear obscuration. Instead, we expect a fraction of obscured quasars to appear obscured due to orientation effects. As the clustering of torus-obscured objects should match that of unobscured systems, the measured clustering of the obscured population in fact represents a lower limit for the population of systems obscured by evolutionary processes, as noted by DiPompeo et al. (2016b). The true clustering of this "non-torus-obscured" (NTO) population then depends on the NTO fraction ($f_{\rm NTO}$), the proportion of obscured quasars which are obscured by evolutionary effects as opposed to orientation.

We thus investigate the halo properties required of NTO quasars in order to produce the observed clustering as a function of NTO fraction. We adopt the method of DiPompeo et al. (2016b) to compute the implied host halo properties for NTO quasars using our updated measurements. Assuming similar redshift distributions, the bias of the NTO population is given by:

$$b_{\rm NTO} = \frac{b_{\rm ob} - b_{\rm unob}}{f_{\rm NTO}} + b_{\rm unob} \tag{3.18}$$

Therefore, the smaller the fraction of obscured quasars belonging to the NTO class, the more biased and rare halos NTO quasars must occupy to drive the observed clustering enhancement of all obscured quasars. For a physical interpretation, we use this NTO bias to estimate the minimum host halo mass, halo occupation fraction, and characteristic lifetime of NTO quasars as a function of NTO fraction. We convert between this NTO bias and a minimum halo mass required to host a quasar as a function of the NTO fraction $f_{\rm NTO}$ using:

$$b(M > M_{\rm min}) = \frac{\int_{M_{\rm min}}^{\infty} dM \frac{dn}{dM} b(M)}{\int_{M_{\rm min}}^{\infty} dM \frac{dn}{dM}}. \tag{3.19}$$

We note that this minimum mass differs slightly from that defined in our HOD modeling

which includes a softening of the step function. Given a minimum halo mass, we are able to compute an occupation fraction ($f_{\rm occ}$, the number fraction of halos more massive than a threshold mass which host a given type of quasar) by comparing the space density of quasars with the density of halos more massive than the minimum mass. To estimate the space density of our quasar samples, we integrate the Shen et al. (2020) quasar luminosity function (QLF) over the observed 6μm luminosity distributions (§3.2.5), also integrating over the sample's redshift distribution. Next, we divide the resulting density by two in order to isolate the density of the obscured population (which comprises roughly half of the total, §3.2). For NTO quasars we finally multiply the density by the NTO fraction. Next, we compute the space density of halos as a function of minimum mass by performing mass and redshift integrals over the halo mass function of Tinker et al. (2008).

Finally, we provide estimates of the quasar lifetime, the characteristic time a SMBH is expected to be observable as a quasar of a given phase. Assuming that every halo hosts one SMBH, the quasar lifetime is simply proportional to the halo occupation fraction (Haiman & Hui 2001; Martini & Weinberg 2001). A majority of our sample quasars lie at $z \approx 0.5-2$, corresponding to a span of ~ 5 Gyr of cosmic time. An approximate estimate of the quasar lifetime is thus given by this factor multiplied by the occupation fraction.

We show the results of these calculations in Figure 3.11. It is apparent that if NTO quasars make up a small proportion of the obscured population, their host halos must be very massive in order to drive the clustering of the full obscured population to the observed value. Conversely, we can use this relation to put constraints on the minimum NTO fraction required by our measurements in order to avoid unphysically large host halo masses for these objects. At sufficiently small NTO fractions ($\lesssim 15\%$), the implied host halos of NTO quasars become massive and thus rare enough that the corresponding occupation fraction exceeds unity. This forbidden region is shown in Figure 3.11 with a gray hatched region. This appears to be a compelling line of evidence that a non-negligible ($\gtrsim 15\%$) fraction

of obscured quasars are obscured by a mechanism aside from orientation, such as galactic or circumnuclear dust. We also note that our quasar lifetime analysis suggests that if NTO quasars represent a distinct evolutionary phase in the evolution of SMBH growth, the duration of the obscured phase appears greater than the unobscured phase by a factor $\gtrsim 3$.

3.6.4 Possible Connection to Dust-obscured Galaxies

By matching our quasar catalog to SDWFS 24 μm photometry, we find that $\sim 30\%$ of the objects in the the obscured quasar sample would be classified as dust-obscured galaxies (DOGs, Dey et al. 2008) with $F_{24}/F_r > 1000$. DOGs are ultra-luminous infrared galaxies (ULIRGs) often accompanied by quasar activity at $z = 1-2$, where the optical emission from both stars and SMBH accretion appears heavily attenuated, implying strong galactic-scale absorption. These objects thus are expected to be good candidates for galaxies caught in the violent post-merger evolutionary stage expected to produce obscured quasar activity and star formation in the models of Sanders et al. (1988a) and Hopkins et al. (2008). The DOGs in our sample are expected to be quasar-dominated given that their infrared colors satisfy the Donley et al. (2012) criterion of power-law infrared spectra. We observe that the DOG fraction increases in our obscured quasar sample towards systems with redder $r-$W2 colors. To test whether the enhanced clustering of obscured quasars observed in this work could be driven by the reddest tail of sources with significant overlap with the DOG population, we split the obscured sample into two further subsets and again perform each clustering and lensing analysis as previously presented. We split the obscured population in two using a cut of $r-$W2 = 4.5, which corresponds to the peak of the obscured quasar distribution (Figure 3.1). The subset redder than this cut consists of $\sim 50\%$ DOGs. We do not split the unobscured sample, as Petter et al. (2022) showed that unobscured spectroscopic quasars' clustering is not connected with their $r-$W2 colors. We display the effective halo masses from these analyses as a function of $r-$W2 color in Figure 3.12.

Interestingly, the effective halo mass of *WISE* quasars appears to continue increasing to-

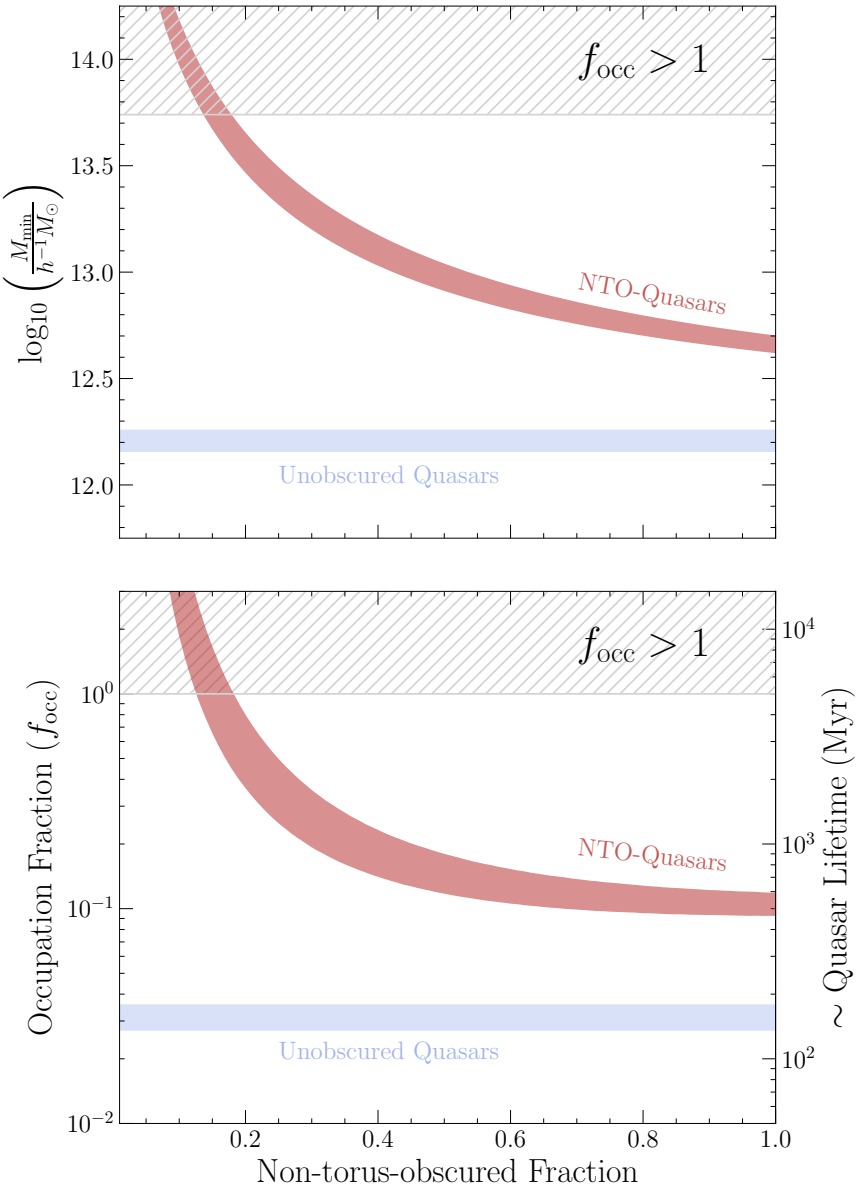

Figure 3.11: Top panel: the implied minimum host halo mass for "non-torus-obscured" (NTO) and unobscured quasars as a function of the NTO fraction of our sample. Bottom panel: the implied occupation fraction of NTO and unobscured quasars, or the fraction of halos more massive than the corresponding minimum mass which host a quasar of a given type. The right axis shows a characteristic lifetime of quasars in a given phase corresponding to the occupation fraction. If only a small fraction of obscured quasars belong to the NTO population, the bias and host halo masses of these systems must be very large in order to drive the observed clustering of the entire obscured population. Conversely, this constrains the NTO fraction to be $\gtrsim 15\%$ in order to avoid unphysically massive host halos, where the quasar occupation fraction exceeds unity (see §3.6.3).

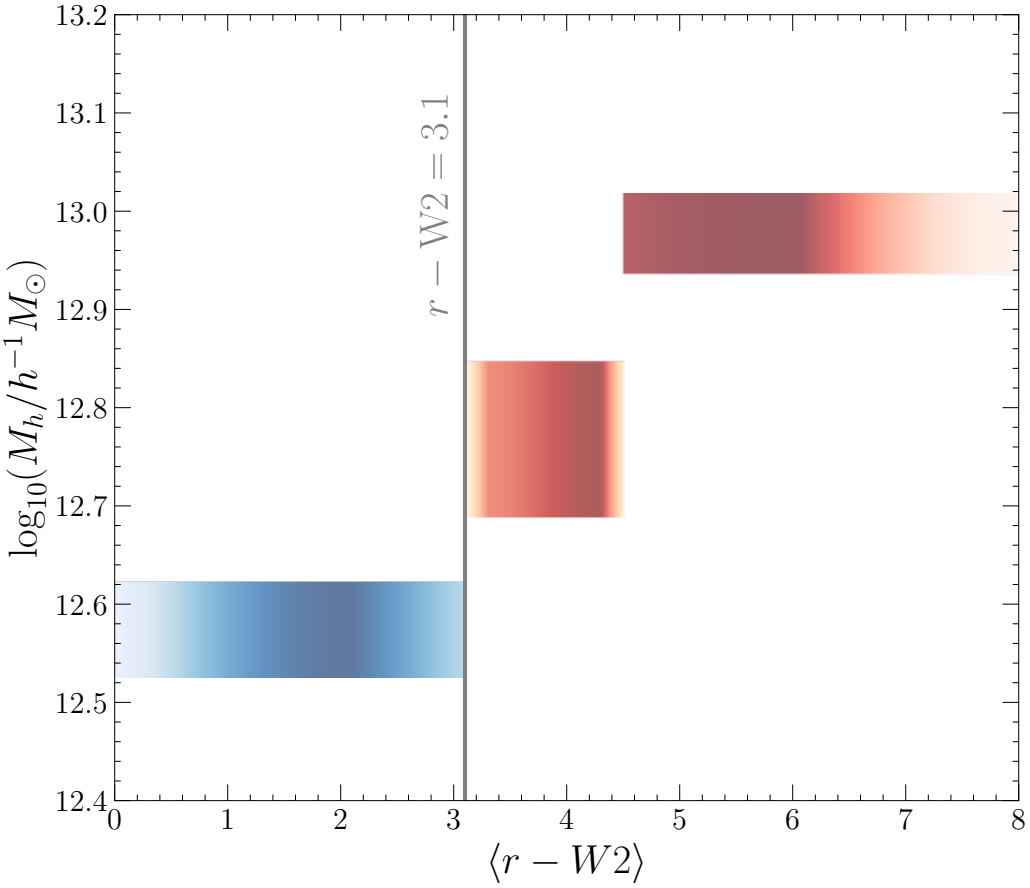

Figure 3.12: A further exploration of the effective host halo masses of quasars as a function of obscuration. We have now split the obscured sample into two subsets using a color cut (r-W2 [AB] = 4.5). We show the effective halo masses of the samples split by optical-infrared color using gradients, where the vertical span represents the 68% confidence interval, and the opacity of the gradient represents the relative number of sources of a given color. The effective halo mass appears to continue increasing for obscured quasars towards redder colors, suggesting the enhanced clustering may be dominantly driven by the reddest and most obscured subset.

wards more extreme red optical-infrared colors. We find that the results from clustering and lensing again agree, which indicates this trend is robust. The reddest sample appears to occupy halos of effective mass $\sim 10^{13.0} h^{-1} M_{\odot}$. Meanwhile, the study of Toba et al. (2017) measured the clustering of *WISE*-selected bright DOGs (which tend to be quasar-dominated), finding that they occupy very massive effective halos of $\sim 10^{13.6} h^{-1} M_{\odot}$. Given a majority of the reddest sample would be selected as DOGs, we speculate that the observed enhanced clustering of obscured infrared-selected quasars may be driven by the reddest and most heavily obscured subset, which are often DOGs and expected to be obscured by host-galaxy material. However, we note that we perform a full HOD fit of each subsample, but do not detect an excess of small scale clustering in the reddest bin as might be expected for a sample dominated by merger systems.

3.6.5 Comparison to Literature Results

It is important to understand the origin of the wide array of seemingly conflicting results on the clustering of obscured and unobscured AGNs in the literature. Assessing the literature, we broadly find that studies detecting enhanced clustering of obscured quasars tend to be selected in the infrared (Hickox et al. 2011; Donoso et al. 2014; DiPompeo et al. 2014, 2015b, 2016a, 2017a, this work) or ultra-hard X-ray (Powell et al. 2018), while studies which do not detect a difference tend to use softer X-ray selection (e.g., with *Chandra* or *XMM-Newton*) over relatively smaller fields (Coil et al. 2009; Gilli et al. 2009; Ebrero et al. 2009; Mountrichas & Georgakakis 2012; Koutoulidis et al. 2018), though there are exceptions (Geach et al. 2013; Mendez et al. 2016; Krumpe et al. 2018).

We posit that the diversity of results may be understood by one of two selection effects. In one scenario, quasar selection at certain wavelengths may reveal obscured and unobscured populations which are not matched in host galaxy properties that correlate with dark matter halo mass. In this scenario, it is possible, for instance, that infrared quasar selection may uncover obscured quasars which occupy more massive galaxies than unobscured quasars,

which would produce a difference in observed halo mass given that galaxy stellar mass correlates with halo mass. However, the modeling work of Whalen et al. (2020) appears to point against this scenario.

Alternatively, it is possible that the effective halo mass of a sample of quasars is driven by an extreme subset of that sample, which are not equally recoverable in all wavebands. In this context, we suggest that the enhanced clustering of obscured quasars may be driven by the most heavily obscured subset, which are able to be selected at infrared and hard X-ray wavelengths but not with soft X-rays. We have shown in Figure 3.12 that the clustering of obscured quasars continues to increase towards redder optical-infrared colors and greater overlap with the DOG population, possibly supporting this view. Quasar-dominated DOGs have been shown to often be obscured by Compton-thin or Compton-thick columns (e.g., Fiore et al. 2008; Lanzuisi et al. 2009; Stern et al. 2014; Piconcelli et al. 2015; Del Moro et al. 2016; Assef et al. 2016; Vito et al. 2018), which would prevent their detection in typical soft X-ray surveys. We thus speculate that the seemingly conflicting results in the literature regarding the host halos of obscured and unobscured quasars may be explained if the enhanced clustering of infrared-selected obscured quasars is driven by deeply-buried systems missed by soft X-ray surveys.

3.7 Conclusions

In this paper, we have selected ~ 1.4 million quasars over $\sim 15,000$ deg^2 of sky in the mid-infrared with *WISE* and split them into obscured and unobscured samples using optical-infrared colors. We have then probed their host halo properties by measuring their angular clustering statistics as well as the typical gravitational deflection of CMB photons passing through their surrounding large-scale structures. We have interpreted the clustering signals within an HOD framework, finding that the minimum and effective halo masses of obscured quasars are significantly higher than their unobscured counterparts, with obscured quasars

occupying effective halos of $\sim 10^{12.9} h^{-1} M_{\odot}$, while unobscured quasars occupy effective halos of $\sim 10^{12.6} h^{-1} M_{\odot}$. We however do not detect a difference in the one-halo term, finding that $\sim 5-20\%$ of both obscured and unobscured quasars are satellites within their halos. Interpreting the CMB lensing signals with a linearly biased halo model, we find excellent agreement with the clustering results. We showed that this agreement confirms the halo mass difference is robust against uncertainty in the obscured sample's redshift distribution. We discuss interpretations of the observed clustering difference, and favor an evolutionary explanation for the obscuration of at least some quasars. Finally, we detect a hint that the enhanced clustering is driven by the systems with the most extremely red optical-infrared colors, which have a significant overlap (¿50%) with the Dust-obscured Galaxy (DOG) population.

With this work, we have now estimated the halo properties of infrared quasars across the majority of the extragalactic sky. We are therefore likely approaching the limit of the statistical power available to probe the halos of purely *WISE*-selected quasars through angular autocorrelations. Consequently, further investigation of the halo properties of these quasars must utilize other techniques. While the *Planck* map of CMB lensing will remain the only all-sky dataset for the near future, ground-based experiments such as the Atacama Cosmology Telescope, South Pole Telescope, and soon the Simons Observatory and CMB-S4 will improve the lensing precision on smaller scales (e.g., Omori et al. 2017; Darwish et al. 2021), and may allow direct measurements of the quasar one-halo term. The weak lensing of background galaxies can also be used to study AGN halo properties (e.g., Leauthaud et al. 2015; Luo et al. 2022). Upcoming missions including *Euclid* and the Rubin Observatory will measure the weak gravitational lensing of galaxies at $z > 2$ over a wide area, probing the large-scale structure around the peak of quasar activity at $z = 1-2$. These missions will also detect and measure photometric redshifts for galaxies at $z = 1-2$, enabling cross-clustering measurements to increase the signal-to-noise ratio for the inherently rare population of bright quasars. Finally, spectroscopic surveys of quasars would allow spatial

clustering measurements and more precise obscuration criteria, but extending spectroscopy to the most heavily obscured systems will prove challenging due to their inherent optical faintness.

In order to interpret the clustering difference between the obscured and unobscured quasars observed in this work, a comparison of the host galaxy properties through full panchromatic SED modeling (e.g., Andonie et al. 2022) is required in future work. Such a characterization, combined with the results of this work, will provide a probe of the full halo-galaxy-SMBH connection, testing models of AGN structure and of galaxy-SMBH coevolution.

Chapter 4

Environments of Luminous Low-frequency Radio Galaxies Since Cosmic noon: Jet-mode Feedback Dominates in Groups

by Grayson C. Petter (Dartmouth College), Ryan C. Hickox (Dartmouth College), Leah K. Morabito (Durham University), David M. Alexander (Durham University)

This manuscript is under review with the Astrophysical Journal at the time of writing.

4.1 Introduction

Galaxies form in the standard Λ-cold dark matter cosmology when baryons cool onto collapsed dark matter halos (White & Rees 1978; White & Frenk 1991). However, galaxy formation models which incorporate only the first-order baryonic physics such as gravitational shock-heating and radiative cooling fail to match a wide array of observations of the

most massive systems, galaxy groups and clusters. In particular, the efficiency of gas cooling and in turn galaxy growth is observed to be poorer than these models predict in more massive systems (Balogh et al. 2001; Benson et al. 2003; Borgani et al. 2004; Somerville et al. 2008; Benson 2010; Donahue & Voit 2022).

The advent of the ROSAT, *Chandra* and *XMM-Newton* X-ray telescopes revealed an unexpected abundance of hot gas in clusters and groups. The cooling time of X-ray gas cores in clusters, groups, and elliptical galaxies is observed to be much shorter than the ages of the systems (Edge et al. 1992; Cavagnolo et al. 2008; Dunn & Fabian 2008), implying that thermal energy must be resupplied by an internal mechanism (Tabor & Binney 1993; Binney & Tabor 1995).

It is now recognized that a supermassive black hole (SMBH) lies at the center of every massive galaxy (Magorrian et al. 1998; Kormendy & Ho 2013), and that accretion onto this object triggers active galactic nucleus (AGN) activity which plays an integral role in shaping the host galaxy's growth and evolution through feedback processes (Best et al. 2006; McNamara & Nulsen 2007; Cattaneo et al. 2009; McNamara & Nulsen 2012; Fabian 2012; Alexander & Hickox 2012; Heckman & Best 2014; Hardcastle & Croston 2020). This picture was cemented theoretically when cosmological semi-analytic and hydrodynamic simulations including AGN feedback were able to address the cooling flow problem and reproduce observed local stellar mass functions (Croton et al. 2006; Bower et al. 2006; McCarthy et al. 2010).

This AGN feedback can take one of two forms, jet/radio-mode where relativistic jets are launched by inefficient accretion, and quasar/wind-mode where radiatively-driven winds are spawned by efficient accretion. The discovery of X-ray cavities coincident with radio jets (Boehringer et al. 1993; Carilli et al. 1994; Fabian et al. 2000, 2006) provided evidence for jet-mode feedback in the aforementioned groups and clusters. These systems serve as laboratories with which to estimate the jet energy deposition (McNamara et al. 2000; Chu-

razov et al. 2000; Bîrzan et al. 2004, 2008; Cavagnolo et al. 2010; Timmerman et al. 2022), showing that the mechanical power can greatly exceed that of the synchrotron luminosity, enough to inflate and heat the atmospheres of groups and clusters.

Alternatively, radiation pressure from efficient accretion can drive quasar/wind-mode feedback (Silk & Rees 1998; Elvis 2000; Veilleux et al. 2005; Di Matteo et al. 2005; Springel et al. 2005; Hopkins et al. 2006; Thacker et al. 2006; Laha et al. 2021), evidenced by observed gaseous outflows coincident with quasar activity (Alexander et al. 2010; Rupke & Veilleux 2011; Harrison et al. 2012; Liu et al. 2013; Harrison et al. 2014; Greene et al. 2014; Brusa et al. 2015; Zakamska et al. 2016; Fiore et al. 2017; Perrotta et al. 2019). However to date, the complete picture of the relative role that jet-mode and wind-mode feedback play is elusive. Radio galaxies are thought to be preferentially triggered by an advection-dominated accretion flow (Narayan & Yi 1994), when the SMBH is fed by condensing hot gas in a massive halo. Meanwhile, quasar activity is generated by a thin accretion disk (Shakura & Sunyaev 1973), expected to be fed by cold gas streams in lower-mass halos (Kereš et al. 2009). This suggests that jet-mode and wind-mode feedback may be expected to occur in different environments.

Modern hydrodynamic and semi-analytic galaxy evolution models routinely implement feedback from AGNs and match a wide array of observables (Gonzalez-Perez et al. 2014; Vogelsberger et al. 2014; Schaye et al. 2015; Somerville & Davé 2015; Dubois et al. 2016; Croton et al. 2016; Kaviraj et al. 2017; Springel et al. 2018; Davé et al. 2019; Rennehan et al. 2023). However, feedback effects typically take place on smaller scales than the resolution elements of simulations, such that models must assume "sub-grid" prescriptions. For a complete understanding of AGN feedback, observations are required to constrain which flavor of feedback dominates, and in which environments. Therefore in this work, we study the environments in which luminous low-frequency radio galaxies occur over the past $\sim$ 10 Gyr.

Radio galaxies at $z < 1$ have been found to cluster strongly, residing in halos of masses $\sim 10^{13.2} - 10^{13.5} h^{-1} M_{\odot}$ (Magliocchetti et al. 2004; Mandelbaum et al. 2009; Hickox et al. 2009), implying they typically deposit their energy into massive group environments. It is important to study the environments radio galaxies occupied during "cosmic noon" ($z \sim 2$), when the cosmic star formation efficiency began to decline likely in part due to AGN feedback (Madau & Dickinson 2014). To date, relatively few studies have explored the clustering of radio galaxies at $z > 1$, which have been confined to small fields, and conducted at GHz frequencies (Lindsay et al. 2014; Magliocchetti et al. 2017; Hale et al. 2018). Recent advances in low-frequency radio surveys such as with the LOw-Frequency ARray (LOFAR; van Haarlem et al. 2013) now offer a new window on radio galaxy clustering. The depth and area curate an unprecedentedly large radio AGN sample, and the low frequency observations reveal the old electron populations and thus likely better trace the long-term energy input of jets into their surroundings.

In this work, we study the host halo environments of $\sim 150,000$ luminous radio galaxies ($L_{150\ \mathrm{MHz}} \gtrsim 10^{25.25}$ W/Hz) selected at low frequencies using the LOFAR telescope with tomographic clustering and lensing measurements out to $z < 2$. We find that luminous radio galaxies occupy massive halo environments characteristic of galaxy groups ($M_h \gtrsim 10^{13} h^{-1} M_{\odot}$). We estimate the duty cycle of these systems is of order $\sim 10\%$, and has been increasing with cosmic time since $z \sim 2$, interestingly meanwhile the duty cycle of quasars has been declining. We finally estimate the average kinetic heating power released by these systems into halos as a function of mass, finding that jet-mode heating dominates over quasar/wind-mode in group-scale halos at $z < 2$.

Throughout this work, we adopt a "Planck 2018" CMB+BAO Λ-CDM concordance cosmology (Planck Collaboration et al. 2020a), with $h = H_0/100\ \mathrm{km\ s^{-1} Mpc^{-1}} = 0.6766$, $\Omega_m = 0.3111$, $\Omega_\Lambda = 0.6888$, $\sigma_8 = 0.8102$, and $n_s = 0.9665$. We adopt the convention where synchrotron spectra are parameterized as $S_\nu \propto \nu^\alpha$, and perform K-corrections throughout

using an assumed spectral index for radio galaxies of $\alpha = -0.7$. Any magnitudes are presented in the Vega system, and any logarithms are in base ten unless otherwise stated.

4.2 Data

In this work, we study the clustering properties of radio galaxies selected at 150 MHz using LOFAR data. We use Wide-field Survey Explorer (*WISE;* Wright et al. 2010) infrared and DESI Legacy Imaging Survey (DLIS; Dey et al. 2019) optical counterparts for redshift information of the radio sources, allowing study of the time-evolution of their clustering. Finally, we tomographically cross-correlate their positions with spectroscopic galaxy samples from the Sloan Digital Sky Survey (SDSS; York et al. 2000) for additional resolution in the time-evolution.

4.2.1 The LOFAR Two-metre Sky Survey

The LOFAR is a sensitive interferometric telescope currently revolutionizing surveys of the radio sky. In particular, the pioneering LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017) is the first to produce deep (RMS $\sim 100\ \mu$Jy beam^{-1}) and high-resolution ($6''$) maps at low frequencies (~ 150 MHz) over a wide area, and represents an order of magnitude sensitivity improvement for typical sources over the Faint Images of the Radio Sky at Twenty-centimeters (FIRST) survey (Becker et al. 1995; Helfand et al. 2015), conducted at 1.4 GHz. The second data release of LoTSS-wide (DR2; Shimwell et al. 2022) covers 5634 deg^2 and catalogs 4,396,228 sources.

This is a premier dataset to study radio galaxy clustering (Siewert et al. 2020; Tiwari et al. 2022; Hale et al. 2023; Nakoneczny et al. 2023) for a variety of reasons. The detection depth surpasses the switch point of ~ 1.5 mJy below which star-forming galaxies and radio-quiet AGNs become the dominant populations (Best et al. 2023), and therefore a deeper survey would not select a larger sample of radio galaxies on the basis of flux alone. The

wide area is useful for statistical power in clustering measurements due to the inherent rarity of luminous radio galaxies, and the survey footprint significantly overlaps with that of the SDSS, enabling cross-correlations with spectroscopic samples. The high angular resolution allows accurate identification of optical/infrared counterparts (e.g., Williams et al. 2019; Hardcastle et al. 2019, 2023) for host galaxy property and redshift estimation. Finally, low-frequency observations are well-suited to tracing the environments in which feedback may be taking place, as they are more sensitive to steep-spectrum, lobe-dominated sources which better trace the long-term integrated power input into their surroundings than GHz frequencies, which are more sensitive to core-dominated or beamed emission.

We adopt the LoTSS DR2 source catalog presented in Hardcastle et al. (2023, hereafter H23), which has attempted to associate double-lobed duplicate detections into single sources, and match to optical (DLIS; Dey et al. 2019) and/or infrared (unWISE; Schlafly et al. 2019) counterparts when possible. The association results from a combination of likelihood-ratio cross-match methods, machine learning methods, and visual-inspections by astronomers and citizen scientists. Sources with optical counterparts in the DLIS are enhanced by photometric redshift estimates (Duncan 2022), which we leverage in our study. The samples we construct using optical counterpart photometric redshifts at $z < 1$ in §4.3 are composed entirely of sources brighter than the > 8 mJy "bright" threshold referred to in H23, which received the highest priority of visual inspections. We thus are confident in the robustness and completeness of our photo-z-selected samples, but we refer the reader to H23 for a detailed discussion of the association techniques.

We choose to use only the LoTSS data in the northern Galactic hemisphere, as this region makes up the majority of the DR2 footprint (74%), has more uniform imaging depth, and better overlap with DLIS imaging and SDSS spectroscopic surveys.

4.2.2 eBOSS Quasar Sample

One of the goals of this work is to study radio galaxy clustering and its evolution at $z > 1$. We therefore utilize quasars with spectroscopic redshifts from SDSS surveys as biased matter tracers for cross-correlations with LoTSS-selected radio galaxies at $1 < z < 2$.

The Extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) was a spectroscopic survey designed to measure baryon acoustic oscillations in the distribution of three tracers, star-forming emission-line galaxies, luminous red galaxies, and quasars. The quasar sample tracing high-redshift structure consists of 343,708 uniformly-targeted (Myers et al. 2015) systems at $0.8 < z < 2.2$, representing the largest statistical sample of spectroscopic quasars to date. The eBOSS collaboration has produced large-scale structure catalogs including randoms and weights for measuring unbiased correlation functions (Ross et al. 2020). We utilize the versions presented in Rezaie et al. (2021), incorporating updated systematic weights.

4.2.3 LoTSS-deep Data

We make use of derived data products from the LoTSS-deep survey DR1 (Tasse et al. 2021; Sabater et al. 2021) to verify our sample selections and redshift distributions. This lies amongst the deepest (RMS ~ 20 μJy bm^{-1}) radio surveys to date, but covers more than an order of magnitude wider area than the comparable depth VLA-COSMOS survey (Smolčić et al. 2017), totaling 25 deg^2 with overlapping deep multiwavelength photometry in the Boötes, Lockman-hole, and ELAIS-N1 fields. This photometry allows for host identification (Kondapally et al. 2021), photometric redshift estimation (Duncan et al. 2021), source classification, and host galaxy property constraint through spectral energy decomposition (Best et al. 2023, hereafter B23), and the measurement of the radio AGN and star formation luminosity functions (Kondapally et al. 2022). The redshift information is crucial to our clustering and lensing analysis, and the luminosity functions will allow es-

timation of the occupation statistics when combined with clustering measurements. We update the catalogs with the new Dark Energy Spectroscopic Instrument (DESI) early release data (DESI Collaboration et al. 2023) for additional spectroscopic redshifts where available. We prioritize the data in the Boötes field unless otherwise stated, as this region has the best spectroscopic coverage primarily courtesy of the AGN and Galaxy Evolution Survey (AGES; Kochanek et al. 2012).

4.2.4 Planck Cosmic Microwave Background Lensing Map

Cosmic microwave background photons emitted during the recombination epoch have been gravitationally-lensed by the intervening structure, and provide a complimentary probe of high-redshift galaxies' host halo properties along with correlation functions. A number of high-resolution CMB experiments have now produced wide-area maps of the lensing convergence κ, a projected surface mass density tracing structure from $0.5 \lesssim z \lesssim 5$ (e.g., Planck Collaboration et al. 2020a; Omori et al. 2023; Madhavacheril et al. 2023). Only *Planck*'s all-sky survey overlaps with LoTSS DR2, and thus we use the *Planck* final release (PR4) lensing map (Carron et al. 2022) as an independent constraint on the host-halo properties of the LoTSS radio galaxies. We also make use of the provided simulated maps to estimate uncertainties. We produce maps from the combined temperature and polarization data (minimum-variance reconstruction) at NSIDE=1024 resolution using ℓ-modes ¡ 2048 for cross-correlation with radio galaxy overdensity maps.

4.2.5 CatWISE Infrared Photometry

Many luminous radio galaxies at higher redshifts ($z \gtrsim 1$) are hosted by galaxies too faint to be readily detected in the optical waveband at the DLIS depth. However, their hosts can be easier recovered at near-infrared wavelengths, as aged stellar populations exhibit a negative K-correction out to $z \sim 2$ at observed-frame 3-5 μm. Thus, we update the H23 catalog with deeper *WISE* data than was used in H23 to push the host galaxy detection to higher

redshift. *WISE* is a space telescope which has mapped the entire sky in four infrared bands, centered at 3.4, 4.6, 12 and 22 μm (named W1, W2, W3, and W4 respectively). *WISE* continues to operate in the two shorter-wavelength bands in its post-cryogenic phase (Mainzer et al. 2011), and the deeper resulting imaging continues to detect fainter and higher-redshift galaxies. The H23 LoTSS catalog has associated radio sources with unWISE counterparts when possible, and 69% of $S_{150\ \mathrm{MHz}} > 2$ mJy sources have a W1/W2 measurement. However, the CatWISE2020 (Marocco et al. 2020b; Marocco et al. 2021a) catalog probes deeper than unWISE, which facilitates detection of host galaxy counterparts of radio galaxies at higher-redshift. We match the H23 catalog with CatWISE2020 using a conservative match radius of 2.5″ (from the H23 counterpart if available, then the radio position), and let the CatWISE measurements supersede unWISE where applicable. This raises the fraction of $S_{150\ \mathrm{MHz}} > 2$ mJy systems with infrared counterparts to 79%. As expected, systems detected in CatWISE but not unWISE are found to have a redshift distribution peaking at $z \sim 2$ in the LoTSS-deep fields. We will leverage the feature that higher-redshift galaxies will appear both redder and fainter in W1/W2 space (Schlafly et al. 2019; Krolewski et al. 2020) to select $z > 1$ radio galaxies in §4.3.

4.3 Radio Galaxy Samples

We aim to study the environments of luminous radio galaxies across cosmic time, including at the relatively unexplored epoch of $z > 1$. However, the subset of LoTSS sources dominated by radio galaxies ($S_{150\ \mathrm{MHz}} \gtrsim 2$ mJy) exhibits a broad redshift distribution peaking at $z \sim 0.5$ with a long tail towards higher redshifts (Duncan et al. 2021; Alonso et al. 2021, B23). Thus, in this work we use optical/infrared counterparts to separate the radio galaxies into three broad redshift regimes. We then also utilize a *tomographic* technique whereby we cross-correlate radio sources with tracer populations of galaxies at known redshift for additional time-resolution in the study of their clustering.

The host galaxies of luminous radio galaxies are massive at $z \lesssim 1$ ($M_\star \sim 10^{11} M_\odot$; B23), and thus readily detectable in the optical waveband at the depth of the DLIS out to $z \lesssim 1$, where their photo-zs are reliable (Duncan 2022). At higher redshifts, host galaxies undetectable in the DLIS can be recovered at infrared wavebands probed by *WISE* (§4.2.5). We thus use optical counterparts to LoTSS sources augmented by photo-zs to select samples of radio galaxies at $z < 1$, and *WISE* counterparts to select systems at $z > 1$.

We match the nearly-pure AGN flux-limited ($S_{150\ \mathrm{MHz}} > 2.5$ mJy) subset of the H23 catalog with *WISE* counterparts to the LoTSS-deep catalog in the Boötes field for deep photometric or spectroscopic redshift information (Duncan et al. 2021), and display the resulting *WISE* color-magnitude diagram in Figure 4.1. Sources are colored by their broad redshift regime and galaxies belonging to redshift epochs cluster together in the diagram, demonstrating that *WISE* magnitudes and colors can be used to isolate radio galaxies at high redshift.

4.3.1 High-z Radio Galaxies (HzRGs)

To curate the $z > 1$ sample, we adopt the sliding color cut form of Schlafly et al. (2019) designed to select high redshift galaxies shown in Figure 4.1 with a red dashed line, along with a 90% W2 completeness cut (Marocco et al. 2021b). We also incorporate a photometric redshift cut of $z_{\mathrm{phot}} > 1$ for sources with an optical counterpart to cull low-redshift interlopers.

$$\begin{cases} S_{150\ \mathrm{MHz}} > 2.5\ \mathrm{mJy} \\ \mathrm{W1} - \mathrm{W2} > (17 - \mathrm{W2})/4 + 0.15 \\ \mathrm{W2} < 17.25 \\ z_{\mathrm{phot}} > 1 \mid z_{\mathrm{phot}} == \mathrm{Null} \end{cases} \tag{4.1}$$

This flux cut corresponds to a luminosity limit of $\sim 10^{25.25}$ W/Hz at $z \sim 1.5$. We refer to this sample of bright LoTSS sources with faint/red *WISE* counterparts as our "high-redshift

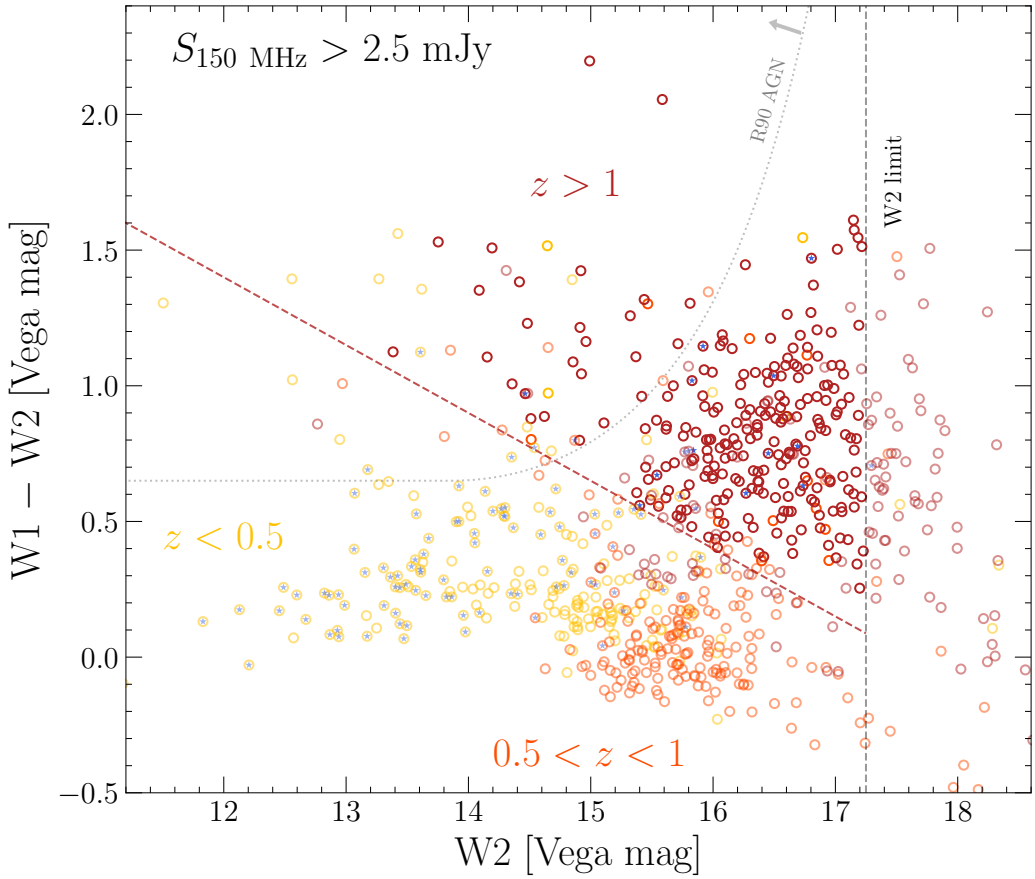

Figure 4.1: An infrared *WISE* diagram of LoTSS-bright ($S_{150\ \mathrm{MHz}} > 2.5$ mJy) galaxies in the Boötes field, colored by their broad redshift regime (Duncan et al. 2021, B23), with yellow, orange, and red corresponding to the redshift ranges $z \in [0, 0.5], [0.5, 1], [1, 3]$, respectively. Galaxies at different redshift epochs broadly cluster together in the diagram. We display our sliding color cut for selecting high-redshift $z > 1$ radio galaxies (Eq. 4.1) with a red dashed line, highlighting the sources satisfying this cut in bold. The Assef et al. (2018) R90 criterion for selecting radiative-mode quasars is drawn with a gray dotted line; most of these radio galaxies would not be selected as *WISE* quasars. Sources with radio emission consistent with their far-infrared star formation rates (B23) are highlighted with an additional internal blue star marker, and are mostly confined to low redshifts. This figure demonstrates that the infrared counterpart properties of LoTSS sources can be used to select a complete and reliable sample of radio galaxies at $1 \lesssim z \lesssim 2$.

radio galaxy" (HzRG; $z > 1$) sample throughout the rest of the text.

4.3.2 Intermediate-z Radio Galaxies (IzRGs)

At $z \lesssim 1$, reliable photo-zs enable construction of approximately radio luminosity-limited samples, as opposed to the flux-limited HzRG sample. For broad correspondence with the luminosity distribution probed by the HzRG sample, we select systems with $L_{150\ \mathrm{MHz}}(z_{\mathrm{phot}}) > 10^{25.25}$ W/Hz. As $S_{150\mathrm{MHz}} > 8$ mJy sources received the highest-priority of visual associations in H23, the above luminosity limit requires that we limit to $z_{\mathrm{phot}} \lesssim 0.8$ for uniform selection. Additionally, we use the converse of the *WISE* color cut for HzRGs to remove likely $z > 1$ interlopers. Together, these cuts form a selection for intermediate-z radio galaxies (IzRGs, $0.5 < z < 0.8$):

$$\begin{cases} 0.5 < z_{\mathrm{phot}} < 0.8 \\ L_{150\ \mathrm{MHz}}(z_{\mathrm{phot}}) > 10^{25.25}\ \mathrm{W/Hz} \\ \mathrm{W1} - \mathrm{W2} < (17 - \mathrm{W2})/4 + 0.15 \end{cases} \tag{4.2}$$

4.3.3 Low-z Radio Galaxies (LzRGs)

Finally, we extend our study once more to lower redshift with the same strategy presented for IzRGs, forming a selection for low-z radio galaxies (LzRGs; $0.25 < z < 0.5$):

$$\begin{cases} 0.25 < z_{\mathrm{phot}} < 0.5 \\ L_{150\ \mathrm{MHz}}(z_{\mathrm{phot}}) > 10^{25.25}\ \mathrm{W/Hz} \\ \mathrm{W1} - \mathrm{W2} < (17 - \mathrm{W2})/4 + 0.15 \end{cases} \tag{4.3}$$

4.3.4 Luminosity Distributions

We show the luminosity-redshift distributions of our samples in Figure 4.2. This demonstrates that our selections define samples with similar luminosity thresholds ($L_{150\ \mathrm{MHz}} \gtrsim 10^{25.25}$ W/Hz) from $0.25 \lesssim z \lesssim 2$. These systems are complete to luminosities ~ 1 dex below the break of the radio AGN luminosity function (Kondapally et al. 2023), implying that they are representative of the global heating output of jetted AGN. The top panel shows that the samples are dominated by LERGs at $z < 2$, with negligible contamination rates of radio-quiet AGN and SFGs.

4.3.5 Redshift Distributions

Crucial to interpretation of a clustering or lensing measurement is knowledge of a sample's redshift distribution. After curating the HzRG, IzRG, and LzRG samples using optical/infrared information, we estimate their redshift distributions using the LoTSS source-associated redshift catalogs of (Duncan et al. 2021, B23). We display the resulting redshift distributions in Figure 4.3. The proposed selections clearly curate samples in distinct redshift regimes, allowing the study of radio galaxy clustering evolution.

To generate the redshift distribution of HzRGs which relies heavily on photometric rather than spectroscopic estimates, we have incorporated the photo-z probability density functions. Even after this step, we observe that the redshift distribution features "peaks", notably at $z \sim 1.1$ and $z \sim 1.7$, as noted in Kondapally et al. (2022). These result from photometric redshift aliasing due to gaps in filter coverage, and are thus unphysical. We thus filter the HzRG redshift distribution with a linear Savitsky-Golay filter, and use this smoothed redshift distribution in all subsequent analyses, though this does not materially affect our results.

To confirm the validity of these cross-match redshift distributions, we derive independent estimates using the clustering redshift technique (Newman 2008; Ménard et al. 2013; Chi-

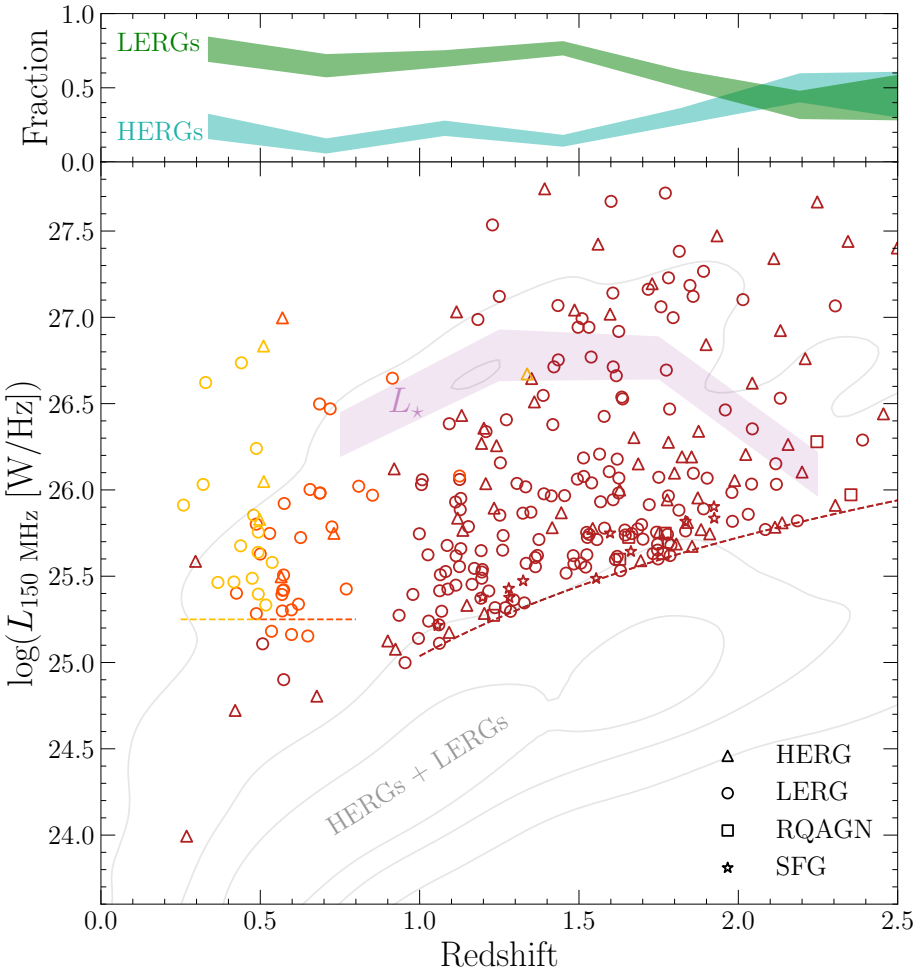

Figure 4.2: The luminosity-redshift distributions of our samples using redshift estimates from the Boötes LoTSS-deep field. The dashed yellow, orange, and red lines show the luminosity thresholds corresponding to the selections for LzRGs, IzRGs, and HzRGs, respectively. This demonstrates that we are able to curate samples with similar luminosity thresholds at three characteristic redshift epochs. Gray contours show the distribution of the full radio galaxy populations recovered in LoTSS-deep. Circular, triangular, square, and star markers denote classifications as LERGs, HERGs, RQAGN, and SFGs, respectively. The top panel shows the LERG and HERG fractions of our combined samples binned in redshift, demonstrating that our sample contains a mix of the two populations, with LERGs dominating at $z < 2$. The characteristic break luminosity of the radio galaxy luminosity function $L_\star$ (Kondapally et al. 2023) is shown with a purple band, demonstrating that our selections curate representative samples of radio galaxies.

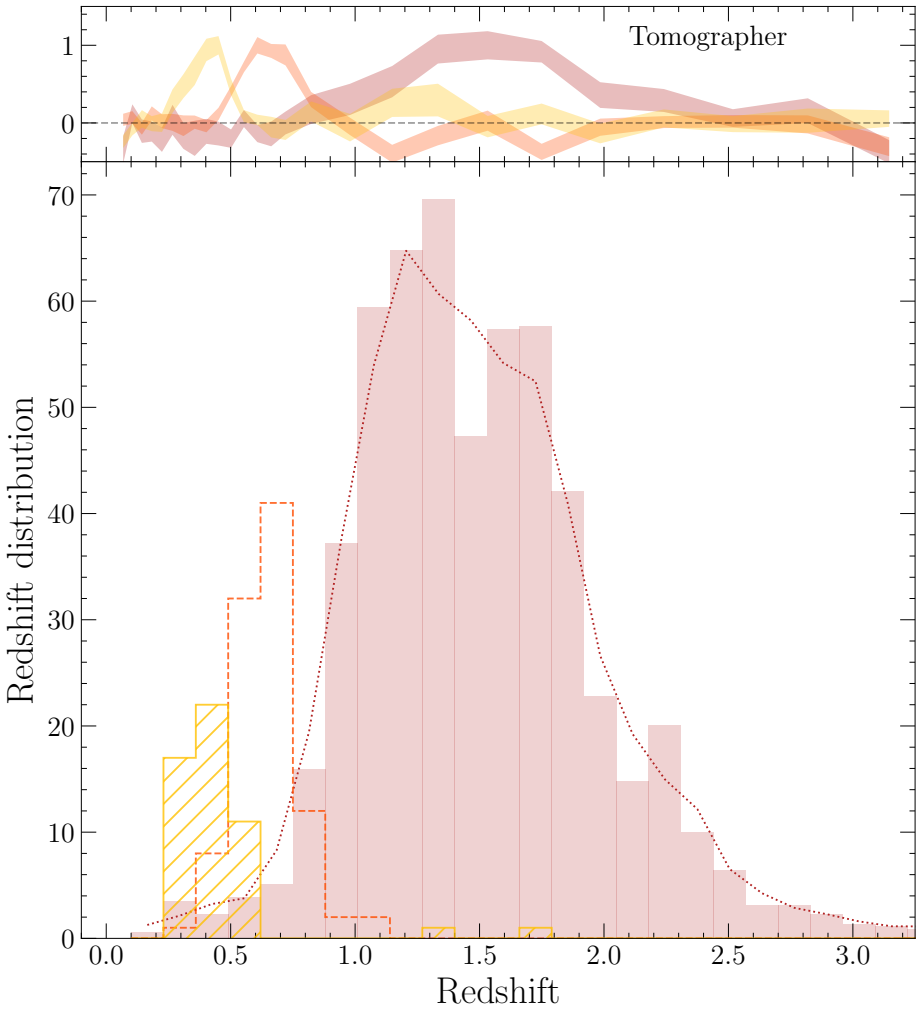

Figure 4.3: We show the redshift distributions of LzRGs, IzRGs, and HzRGs with yellow hatched, orange dashed, and filled red histograms, respectively. We also show a Savitsky-Golay filter fit to the HzRG distribution with a dotted red line to smooth over the artificial peaks in the distribution, which we use in our analysis. The top panel displays corresponding redshift distributions implied from a clustering-based method using the `Tomographer` assuming a fiducial redshift-bias dependence (see text). This qualitatively verifies the efficacy of the selection techniques and the validity of the cross-match redshift distributions. This figure demonstrates that our selections curate radio galaxies in three largely distinct redshift epochs.

ang & Ménard 2019), as implemented in the `Tomographer` web tool.[1] This yields a bias-scaled redshift distribution $b(z)dN/dz$, but untangling any bias evolution from the redshift distribution is degenerate with the goal of this work. Instead, we use this information as a qualitative check on our redshift distributions and assume that the bias evolves inversely with the cosmological growth factor ($b \propto 1/D(z)$; e.g., Alonso et al. 2021). We show the results of the clustering redshift technique in the top panel of Figure 4.3, confirming that our selections generate radio galaxy samples at three largely distinct redshift epochs. We do not use this clustering-based redshift distribution in any further analysis, rather this serves as a qualitative check that our sample selections are performing as intended.

4.3.6 Sample Properties

We thus have constructed samples of luminous ($L_{150\ \mathrm{MHz}} \gtrsim 10^{25.25}$ W/Hz) radio galaxies at low ($z = 0.4$), intermediate ($z = 0.6$), and high redshift ($z = 1.5$) using optical and infrared counterparts to LoTSS sources to probe the host halo environments of low-frequency radio galaxies over cosmic time. We note that this sample is comprised of the most luminous radio galaxies, with limits ~ 1 dex below the break luminosity $L_\star$. As the total kinetic power output of all radio galaxies is dominated by relatively-luminous sources (Kondapally et al. 2023), systems of the luminosities considered here contribute more than half of all the kinetic luminosity density of all radio AGNs. If radio jets are indeed important in shaping galaxy evolution, this sample should thus be significantly representative of the environments in which feedback is taking place. We note that recovering the remainder of fainter radio galaxies over wide swaths of sky will be challenging without corresponding far-infrared data to disentangle star formation processes. The sample has a low contamination rate by pure SFGs and radio quiet AGN, and is dominated by low-excitation radio galaxies (LERGs). As noted in Kondapally et al. (2022), this sample is also hosted by predominantly star-forming galaxies at high redshift, transitioning to being hosted almost

[1] http://tomographer.org

entirely by luminous passive galaxies at $z < 1$.

4.3.7 Masking

Clustering measurements are sensitive to systematics in the selection function of the given sample of galaxies. It is therefore paramount to characterize the angular selection function of the survey. We characterize the footprint of the LoTSS-wide DR2 using a multi-order coverage (MOC) map.[2] We then stitch together an estimate of the RMS imaging noise from the provided RMS map of each facet,[3] by projecting to a `NSIDE=4096` `HEALPix` map. Next, we downgrade the map to `NSIDE=128` using the median of the child pixels. Finally, we mask regions of the sky with local median RMS > 0.2 mJy beam^{-1}, totalling 1.5% of the LoTSS Galactic north footprint. Our samples of radio sources ($S_{150\ \mathrm{MHz}} > 2.5$ mJy) are thus expected to be highly complete.

Requiring optical/infrared counterparts to the radio sources from *WISE* and/or DLIS introduces additional considerations in the selection functions of our samples. We thus leverage a series of `HEALPix` maps[4] characterizing angular systematics generated by the DESI collaboration (Myers et al. 2023). First, we limit the footprint to that of the DLIS DR8 using the provided MOC. We mask regions of abnormally shallow *WISE* imaging depth (5σ W2 PSF depth < 17.25). Next, we limit the footprint to where DLIS made at least 3 observations in each of the g, r, and z bands, and the 5σ PSF z-band depth was at least $z[\mathrm{AB}] > 23$. We also construct a veto `NSIDE=4096` mask of regions with DLIS bitmask flags touching a bright star or large galaxy in the Siena Galaxy Atlas (Moustakas et al. 2023). Finally, we measure trends of each sample's density with Galactic/ecliptic coordinates, stellar density, and Galactic reddening. We observe significant deviations at reddening values $E(B-V) > 0.1$ and at stellar densities > 2000 deg^{-2}, and thus mask these regions. In

[2] https://hips.astron.nl/ASTRON/P/lotss_dr2_high/Moc.fits
[3] https://lofar-surveys.org/dr2_release.html
[4] https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/1.1.1/pixweight/main/resolve/dark/pixweight-1-dark.fits

total, 12.5% of the LoTSS-DR2 north Galactic footprint is masked by our procedure, leaving 3654 deg^2 of survey area. In the case of cross-correlations with SDSS spectroscopic quasars, we mask the HzRG sample with the eBOSS masks,[5] and correspondingly we mask the spectroscopic sample with the LoTSS mask described above.

4.4 Measurements

We perform a variety of measurements to study the halo environments of luminous low-frequency radio galaxies over cosmic time. All correlation function measurements are performed with the `Corrfunc` (Sinha & Garrison 2020) package.

4.4.1 Radio Galaxy Autocorrelations

First, we measure the angular autocorrelation functions of the luminous radio galaxy samples. The angular autocorrelation function is the excess probability above that of a Poisson distribution for detecting a pair of galaxies separated by an angle θ (Peebles 1980). We mask the data and random catalogs identically (§4.3.7), and then measure the angular clustering using the (Landy & Szalay 1993) estimator:

$$w(\theta) = \frac{DD - 2DR + RR}{DD}, \tag{4.4}$$

where DD, DR, and RR are normalized counts for data-data, data-random, and random-random pairs, respectively. We interpret the autocorrelations both with linear and full halo occupation distribution models in (§4.5). Therefore, we measure the autocorrelations at projected angular scales for the median redshift of the sample corresponding to $0.5 - 25\ h^{-1}$Mpc for the full halo occupation distribution constraint, but fit only over $5 - 25\ h^{-1}$Mpc when assuming linear models.

[5] https://data.sdss.org/sas/dr16/eboss/lss/catalogs/DR16/

To estimate uncertainties on all correlation functions, we split the survey footprint into 30 equal area patches using k-means clustering, then perform a bootstrap resampling of the patches (e.g., Efron 1982; Norberg et al. 2009), calculating a correlation function with the data and randoms from each draw. This process is repeated 500 times and the variance across realizations is taken to be the variance of our measurement.

We show the autocorrelation function measurements of our radio galaxy samples in Figure 4.4. Each sample's clustering is very well fit by a halo occupation distribution model (§4.5.2), allowing the inference of the host halo properties of radio galaxies.

4.4.2 Tomographic Cross-Correlations

To probe the redshift evolution of radio galaxy clustering at $z > 1$, we also cross-correlate the radio galaxies with samples of quasars at known redshift, using a tomographic technique. We thus cross correlate HzRGs with eBOSS quasars in redshift slices. After cleaning each sample with the mask of the other (§4.3.7), we bin the eBOSS sample into two slices, $1 < z < 1.5$, and $1.5 < z < 2$. We then measure angular cross-correlations of each slice with the HzRGs, using the Davis & Peebles (1983) estimator:

$$w(\theta) = \frac{D_1 D_2}{D_1 R_2} - 1, \tag{4.5}$$

where HzRGs and quasars make up the first and second samples, respectively. This estimator depends only on the quasar selection function. We use the provided quasar weights as specified by Rezaie et al. (2021), while the radio galaxies are weighted to unity. We show the results of these cross-correlations in Figure 4.5. The measurements are well-fit by linearly biased models, and their bias factors are larger than the bias of quasars being correlated against, implying they occupy more massive halos.

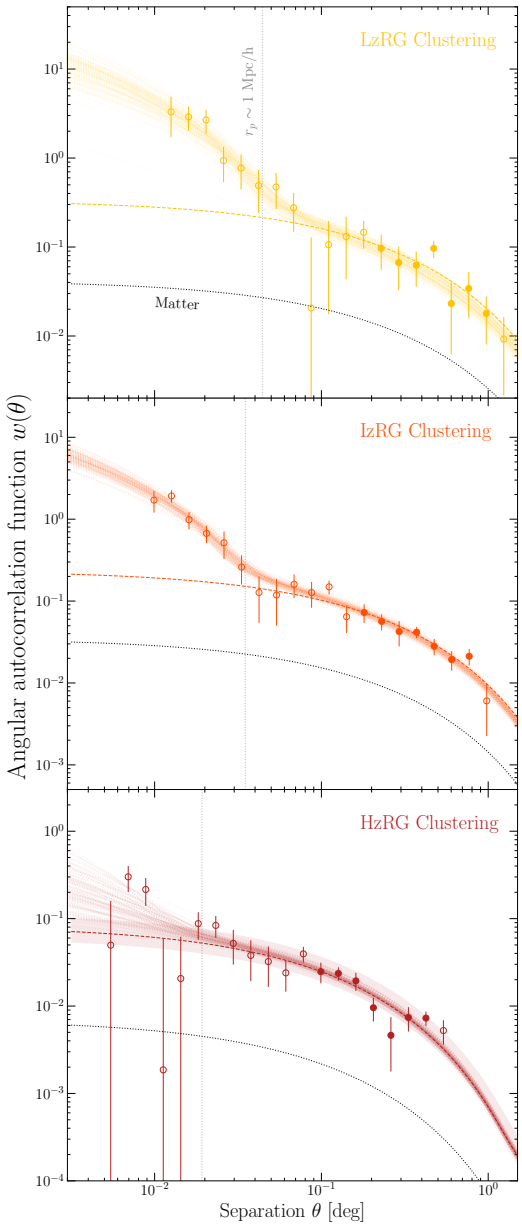

Figure 4.4: The autocorrelation function measurements of LzRGs (top panel in yellow), IzRGs (middle panel in orange) and HzRGs (bottom panel in red). Filled markers show the measurements on linear projected scales ($5-25\ h^{-1}$ Mpc) at the median redshift of the sample, while open markers represent other scales. The projected scale of ~ 1 Mpc/h delineating the boundary of the one and two-halo terms is shown with vertical gray lines. The best fit of a linear model to the linear scales is shown with a dashed colored line. We also show the best fit HOD models to the full range of scales using transparent lines drawn from the posterior distributions. We thus are able to constrain the host halo properties of luminous radio galaxies at three redshift epochs at $z < 2$.

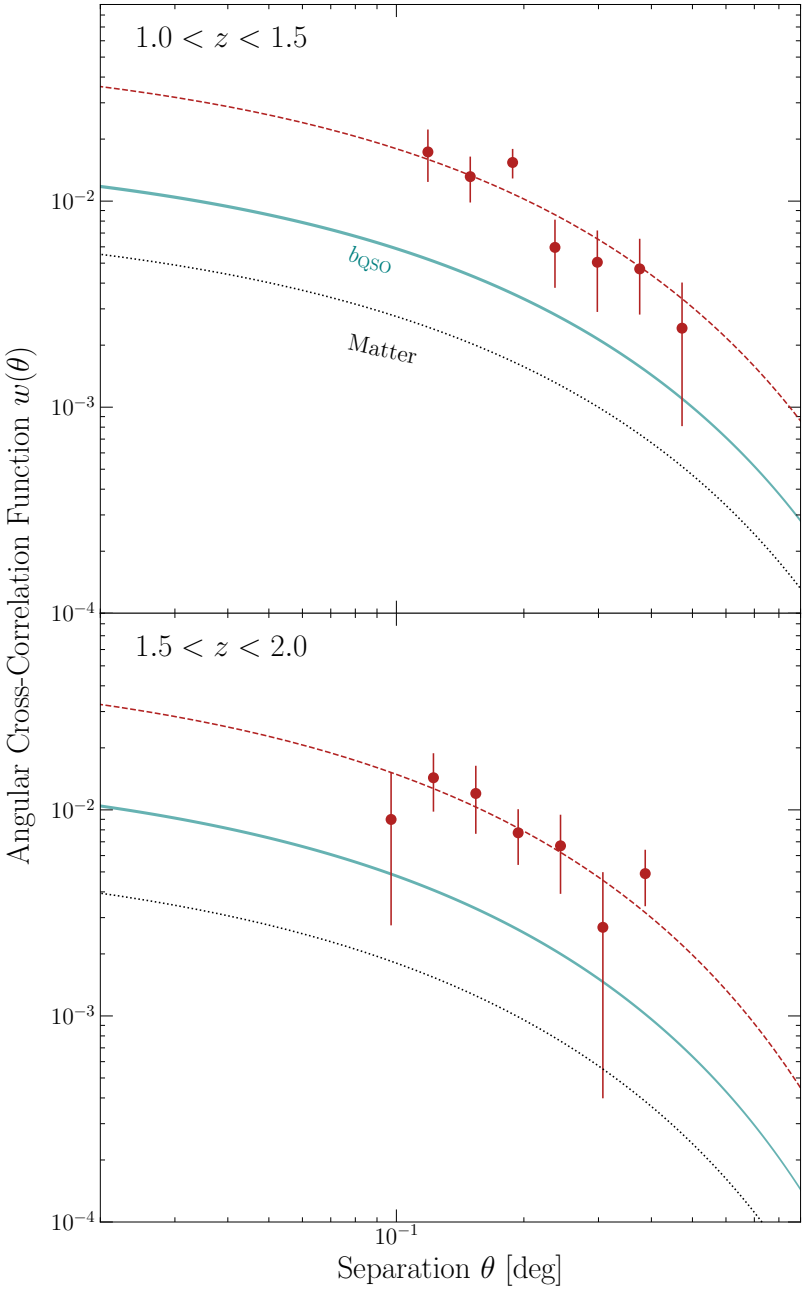

Figure 4.5: The tomographic angular cross correlation function measurements between HzRGs and eBOSS quasars in two redshift slices, displayed in the upper left corners. The linear matter model for the samples' redshift overlap is shown with a black dotted line, while the 1σ confidence interval of the quasar tracer bias is shown with teal bands. HzRGs are best fit with higher bias parameters than quasars, indicating they occupy more massive halos from $1 < z < 2$.

4.4.3 Matter Tracer Autocorrelations

In order to interpret the above cross-correlations, we must estimate the bias parameters of the tracer populations being correlated against. Therefore, we measure the projected spatial clustering of the eBOSS quasars and BOSS galaxies in each redshift slice. We first measure the two-dimensional correlation function $\xi(r_p, \pi)$ using the (Landy & Szalay 1993) estimator, before integrating over the line-of sight to the projected correlation function for mitigation of redshift space distortions (Davis & Peebles 1983; Kaiser 1987), up to a maximum line-of-sight separation $\pi_{\rm max} = 40\ {\rm h}^{-1}{\rm Mpc}$:

$$w_p(r_p) = 2\int_0^{\pi_{\rm max}} d\pi\, \xi(r_p, \pi). \tag{4.6}$$

4.4.4 Radio Galaxy CMB Lensing

Complimenting the autocorrelation measurement of HzRGs, we acquire an independent constraint on the halos hosting HzRGs using a cross correlation of their positions with the *Planck* CMB lensing convergence (κ) map (§4.2.4). First, we produce a fractional overdensity (δ) map of radio galaxies at the same NSIDE=1024 resolution as the lensing map:

$$\delta_{RG} = \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle}, \tag{4.7}$$

where ρ represents the galaxy counts-in-cells. We then must estimate the cross-power spectrum of this overdensity map with the lensing map. Harmonic analysis of maps which do not cover the entire sphere induces mode-coupling, such as for surveys with a limited footprint. However, a fast and nearly-optimal estimator to measure an unbiased psuedo-spectrum is implemented in the `NaMASTER` package (Hivon et al. 2002; Alonso et al. 2019). We mask the RG density map (§4.3.7), and mask the lensing map with the pro-

vided mask after apodizing the edges with a 1° FWHM Gaussian (Krolewski et al. 2020). With `NaMASTER`, we measure the cross-spectrum in 10 logarithmically-spaced bins from $200 < \ell < 2000$. For error estimation, we repeat the above process with 100 provided realizations of simulated lensing maps.

We show the resulting cross-spectrum of HzRG density with CMB lensing convergence in Figure 4.6. This spectrum is well-fit by a linearly-biased halo model (§4.5), allowing inference of the host halo properties in an independent manner from the clustering.

We perform a lensing analysis only of HzRGs for a number of reasons. Firstly, CMB lensing is most efficient at $z \sim 1-2$. Interpretation of the HzRG measurements also relies on constraint of the redshift distribution through photometric redshifts moreso than LzRGs and IzRGs, allowing a comparison of the HzRG lensing and clustering measurements to test for systematics in the redshift distribution (e.g., Petter et al. 2023). The LzRG sample is too sparse and at too low redshift to detect a significant signal. Finally, we observe that a cross-correlation between IzRG positions and CMB lensing is not well-fit by a linear model. This implies contamination in the CMB lensing map from the sources, which is consistent with the fact that this sample is significantly brighter in the radio band than HzRGs.

4.5 Modeling

To interpret the clustering measurements and constrain the environments in which radio galaxies release their energy, we model the measurements within a halo model framework (Seljak 2000; Cooray & Sheth 2002). Throughout this section, $\chi(z)$ is the comoving distance, k is the comoving wavenumber, and c is the speed of light. We assume that dark matter halos follow the "NFW" (Navarro et al. 1997) density profile, and the mass-concentration relation of Duffy et al. (2008), where the mass is defined within a radius containing 200 times the universal critical density. We adopt the halo mass function

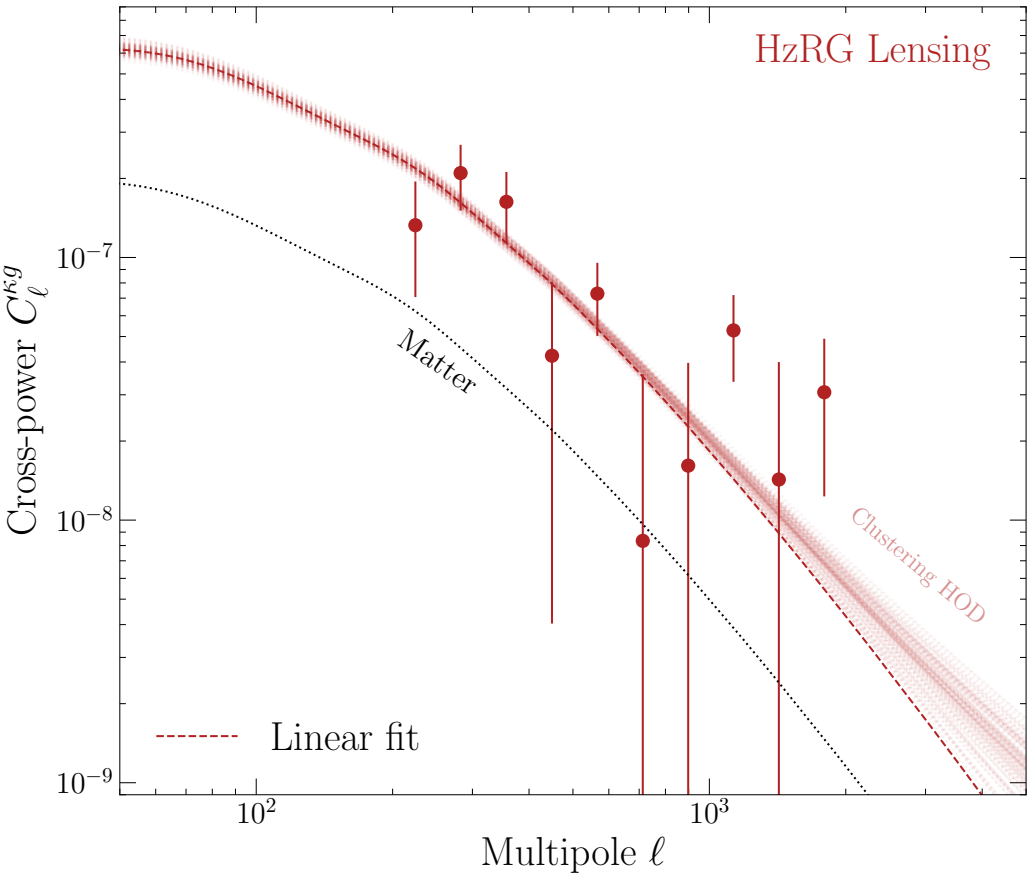

Figure 4.6: The cross correlation measurement of HzRG densities with *Planck* CMB lensing convergence κ. The matter spectrum for the redshift distribution is shown with a black dashed line, while the best linearly-biased fit is shown with a red dashed line. We also show the predictions of the lensing spectrum drawn from the HOD posterior fit to the auto-correlation clustering measurement (Fig. 4.4), demonstrating excellent agreement between the lensing and clustering analyses, and providing a systematic test of the HzRG redshift distribution.

$dn/dM(M,z)$ of Tinker et al. (2008), and the halo mass-bias relation $b(M,z)$ of Tinker et al. (2010). Finally, we utilize `CAMB` (Lewis et al. 2000) to compute all linear matter power spectra $P_{\rm mm}(k,z)$.

An angular correlation function between two matter tracers can be modeled using the Limber (1953) approximation (Peebles 1980; Peacock 1991):

$$w(\theta) = \int dz \frac{dz}{d\chi} \left(\frac{dN}{dz}_i \frac{dN}{dz}_j \right) \int \frac{dk\, k}{2\pi} P_{ij}\, J_0(k\theta\chi), \tag{4.8}$$

where J_0 is the zeroth-order Bessel function of the first kind, and $dN/dz_{i,j}$ are the normalized redshift distributions of the two tracers, identical in the case of an autocorrelation. For a linearly-biased model, the power spectrum is then given by bias factors multiplied by the matter power spectrum:

$$P_{ij}(k,z) = b_i b_j P_{\rm mm}(k,z). \tag{4.9}$$

We also model the spatial autocorrelations of eBOSS quasars to interpret the tomographic cross-correlation measurements. This involves deducing the tracer bias by modeling the projected correlation functions (§4.4.3). The spatial correlation function $\xi(r)$ is the Fourier transform of the galaxy/quasar auto spectrum P_{gg}. The projected correlation function is then an Abel transform of $\xi(r)$:

$$w_p(r_p) = 2 \int_{r_p}^{\infty} dr \, \frac{r\xi(r)}{\sqrt{r^2 - r_p^2}}. \tag{4.10}$$

4.5.1 Modeling CMB Lensing

We also model the cross-spectrum between radio galaxy overdensity and CMB lensing convergence. This is again given by a Limber (1953) integral over the lensing and galaxy projection kernels:

$$C_\ell^{\kappa g} = \int dz \frac{d\chi}{dz} \frac{W^\kappa W^g}{\chi^2} P_{\rm mm}\left(k = \frac{\ell + 1/2}{\chi},\, z\right), \tag{4.11}$$

where the CMB lensing kernel is given by (e.g., Cooray & Hu 2000):

$$W^\kappa(z) = \frac{3}{2}\Omega_{m,0}\left(\frac{H_0}{c}\right)^2 (1+z)\chi \frac{\chi_{\rm CMB} - \chi}{\chi_{\rm CMB}}, \tag{4.12}$$

and the linearly-biased galaxy overdensity kernel is:

$$W^g(z) = b\frac{dz}{d\chi}\frac{dN}{dz}. \tag{4.13}$$

The effective redshift at which we report the bias measurement is weighted by the lensing kernel (Modi et al. 2017; Krolewski et al. 2023):

$$z_{\rm eff} = \int dz \frac{d\chi}{dz}\frac{dN}{dz}\frac{z}{\chi^2} W^\kappa. \tag{4.14}$$

We do not consider magnification bias, as we measure the response of number counts of our only flux-limited sample (HzRGs) against limiting flux/magnitude ($s_\mu \equiv d\,\log_{10}N/dm = 0.412 \pm 0.001$), which is sufficiently near the value of $s_\mu = 0.4$ at which magnification bias is zero.

4.5.2 Halo Occupation Distribution Modeling

We also model the radio galaxy autocorrelation functions in a halo occupation distribution (HOD) framework (Scoccimarro et al. 2001; Berlind & Weinberg 2002) for a more complete interpretation of how these systems occupy their host halos than the linear bias model. This entails replacing the power P_{ij} in Eq. 4.8 with an HOD power spectrum.

An HOD power spectrum is the sum of the "1-halo" power arising from pairs of galaxies

in common halos, and the "2-halo" term from pairs between different halos. The HOD $\langle N(M)\rangle$ is the mean number of galaxies belonging to halos of mass M, decomposed into contributions from galaxies at the centers of halos, $\langle N_c(M)\rangle$, and secondary or 'satellite" galaxies belonging to the same halos $\langle N_s(M)\rangle$:

We adopt the HOD model of Zheng et al. (2007) and Zehavi et al. (2011), which implements the central occupation as a softened step function:

$$\langle N_c\rangle = \frac{1}{2}\left[1+\mathrm{erf}\left(\frac{\log_{10}(M/M_{\mathrm{min}})}{\sigma_{\log_{10}M}}\right)\right] \tag{4.15}$$

where **erf** is the Gauss error function, M_{min} is the minimum halo mass required to host a radio galaxy, and $\sigma_{\log_{10}M}$ is the softening parameter. The satellite HOD is given then as:

$$\langle N_s\rangle = \Theta(M-M_0)\left(\frac{M-M_0}{M_1}\right)^{\alpha} \tag{4.16}$$

where Θ is the Heaviside step function, M_0 is the minimum mass to host a satellite quasar, and M_1 is the mass at which the term transitions to the power-law form. We enforce $M_0 = M_{\mathrm{min}}$ to reduce the degrees of freedom, such that the minimum mass to host a central or satellite is identical. Our HOD models thus contain four free parameters, M_{min}, $\sigma_{\log_{10}M}$, M_1, and α.

We compute the HOD power spectra given a set of parameters using the `Core Cosmology Library` (Chisari et al. 2019) package. Finally, unphysical one-halo power at large scales is suppressed using the Mead et al. (2021) prescription, and a smoothing is applied to intermediate scales between the one and two-halo regimes (Mead et al. 2015). We refer the reader to Petter et al. (2023) for additional details in the HOD modeling procedure. The observed autocorrelation functions are fit using the `emcee` (Foreman-Mackey et al. 2013) sampler, using broad Gaussian (1σ) priors of $\log_{10}M_{\mathrm{min}} = 13\pm2$, $\sigma_{\log M} = 0.5\pm0.5$, $\log_{10}M_1 = 14\pm2$, and $\alpha = 1\pm0.5$.

4.6 Results

We can now interpret all of the observed autocorrelation functions, cross-correlations with CMB lensing, and cross-correlation functions with spectroscopic galaxy tracers in the halo model framework to constrain the environments in which radio galaxies release their energy.

4.6.1 Effective Halo Masses

We display the effective host halo mass results from all clustering measurements of luminous low-frequency radio galaxies at $z < 2$ in Fig. 4.7. These systems occupy massive group halos of $\sim 10^{13} - 10^{14}\ h^{-1}M_{\odot}$ for the past ~ 10 Gyr, with an increasing effective halo mass with cosmic time consistent with the mean growth rate of halos (Fakhouri et al. 2010). We conclude that luminous radio galaxies are typically hosted by group-scale halos across time, and that the occurrence of these systems is intimately connected to host halo mass (e.g.; Hatch et al. 2014). These observations support the paradigm where radiatively-inefficient accretion resulting in jets is driven by gas condensation from the hot halo, and may also reflect the fact that more massive halos contain more gas for jets to collide into and emit at low frequencies via shocks.

Though there are relatively few studies of radio galaxy environments at $z > 1$, our results at lower redshifts appear consistent with previous studies at $z < 1$. Clustering studies of radio galaxies at $z < 1$ reveal they exist in group environments (Mandelbaum et al. 2009; Hickox et al. 2009). This is consistent with studies of the galaxy overdensities surrounding $z < 1$ radio galaxies, which also show they are on average found in galaxy groups (Best 2004; Ineson et al. 2015; Ching et al. 2017; Croston et al. 2019).

At higher redshift ($z > 1$), we find similar agreement with previous results including clustering (Lindsay et al. 2014; Magliocchetti et al. 2017; Hale et al. 2018) and environmental (Hatch et al. 2014) studies, conducted at GHz frequencies. These results support the

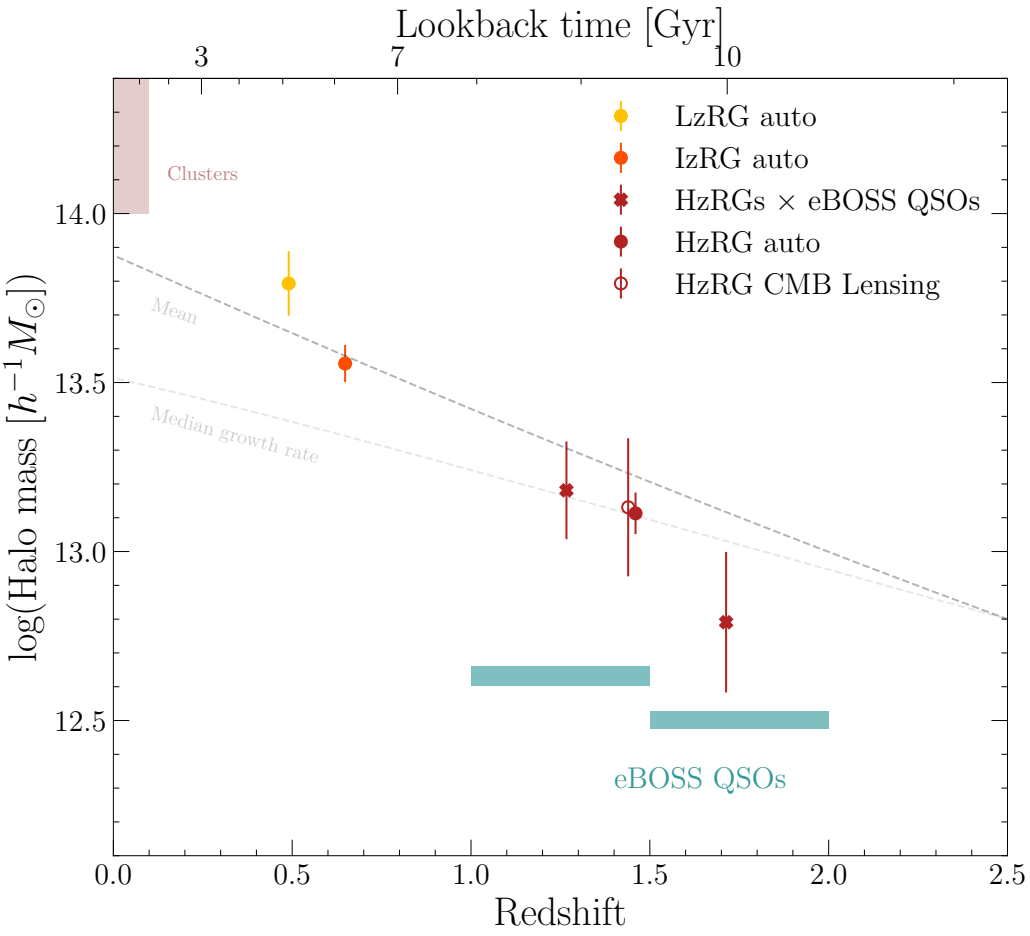

Figure 4.7: The effective halo masses of luminous low-frequency radio galaxies from $z < 2$. The red cross markers show the tomographic cross-correlations between eBOSS QSOs (halo masses shown with blue bands) and HzRGs. The solid circles in yellow, orange, and red show the linear fits to the autocorrelations of LzRGs, IzRGs and HzRGs, and the open circle shows the result from HzRG CMB lensing. Radio galaxies are strongly clustered across cosmic time, residing in group-scale halos. The median and mean halo growth rates from Fakhouri et al. (2010) are shown with dashed and dotted lines respectively, for a halo whose progenitor mass was $10^{12.8}h^{-1}M_\odot$ at z=2.5. This shows that the effective host halo mass of radio galaxies evolves at a rate consistent with the mean growth rate of halos. The halos hosting luminous low-frequency radio galaxies at cosmic noon are expected to evolve on average into massive groups or clusters by the present.

paradigm where radio galaxy activity must be intimately connected with the large-scale environment (Magliocchetti 2022).

4.6.2 Halo Occupation Distributions

For a more sophisticated interpretation of the manner in which radio galaxies occupy their host halos than a single effective host mass, we display the posterior distributions of HOD model fits to the autocorrelation functions of LzRGs, IzRGs and HzRGs in Figure 4.8. We observe that luminous radio galaxies broadly exhibit similar HODs at $z \lesssim 2$, hosted by halos $M_h \gtrsim 10^{13} h^{-1} M_\odot$. Interestingly, the similar effective satellite halo mass of $M_1 \sim 10^{14.25} h^{-1} M_\odot$ across redshift may imply a difference in the satellite fraction at $z > 1$ compared to $z < 1$, because of the fact that these extremely massive halos are increasingly rare at higher redshift. Indeed, we estimate the satellite fraction of HzRGs to be $f_{\rm sat} = 1^{+2}_{-0.1}\%$, while the satellite fractions of IzRGs are $f_{\rm sat} = 13^{+7}_{-4}\%$, and LzRGs are $f_{\rm sat} = 23^{+19}_{-11}\%$. We caution that this is a suggestive lead for future study, and tentatively highlight that this may be connected to the evolution of star-forming properties, FRI/FRII status, and accretion mode from high to low redshift.

We also display the confidence intervals of the HODs in Figure 4.9. This more clearly demonstrates the similarity of radio galaxies' HODs since $z \lesssim 2$. The amplitudes of the HODs are modulated by their respective duty cycles (§4.6.3). We also display the best fit HOD for optically-luminous quasars from early DESI data (Prada et al. 2023). This shows that radio galaxies occupy systematically more massive halos than quasars, which has implications about which environments jets and quasar winds provide their feedback energy into. We will incorporate this information in weighing the relative contribution of jets and winds as a function of environment in §4.6.4.

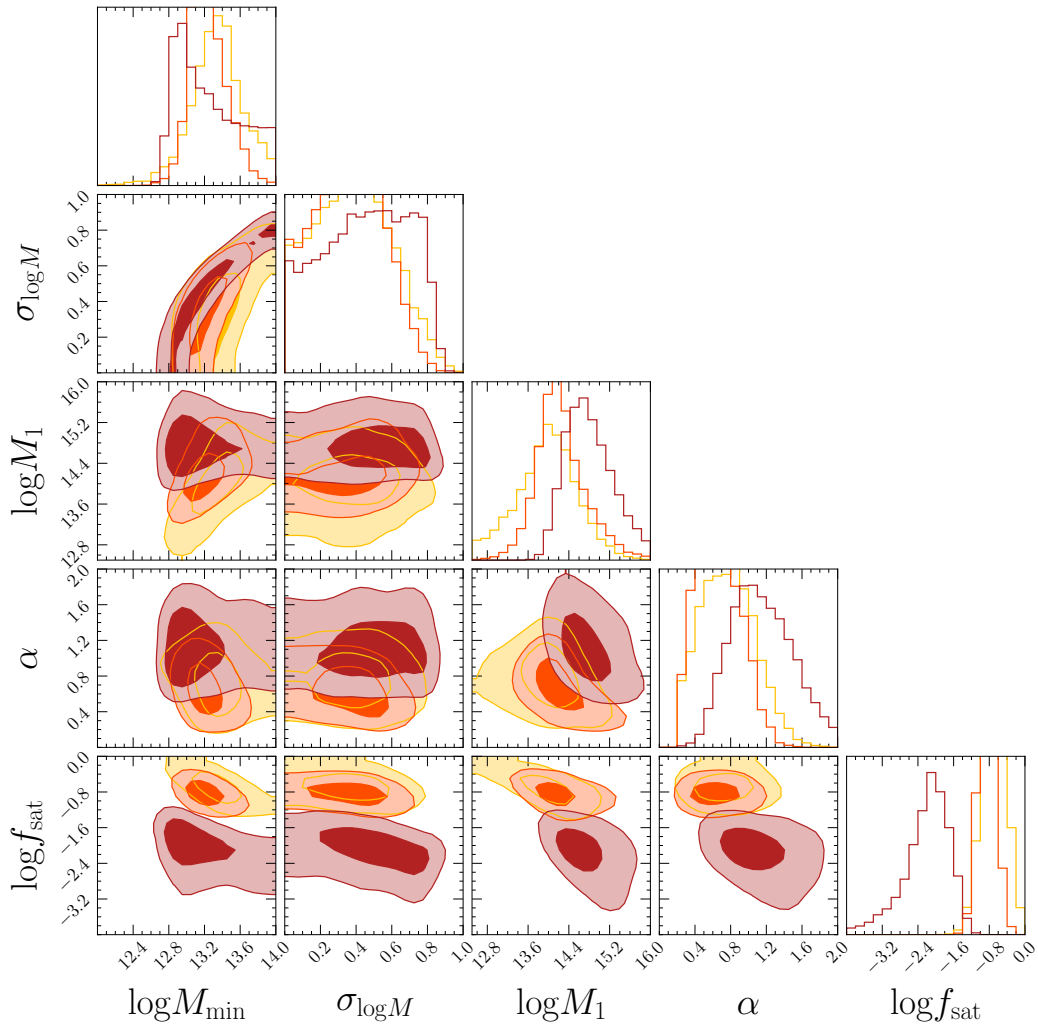

Figure 4.8: A corner plot of the HOD parameter posterior distributions (68 and 95% intervals) from fits to the autocorrelation functions of radio galaxies at $z \sim 0.4$ (yellow), ~ 0.6 (orange) and ~ 1.5 (red). Radio galaxies at $z < 2$ appear to have central occupations similar across time, hosted by halos more massive than $\gtrsim 10^{13} h^{-1} M_{\odot}$. However, the evolving halo mass function implies that radio galaxies appear to be satellites in their halos more often at $z < 1$ than at $z \sim 1.5$, at rates of $\sim 10\%$ and $\sim 1\%$, respectively.

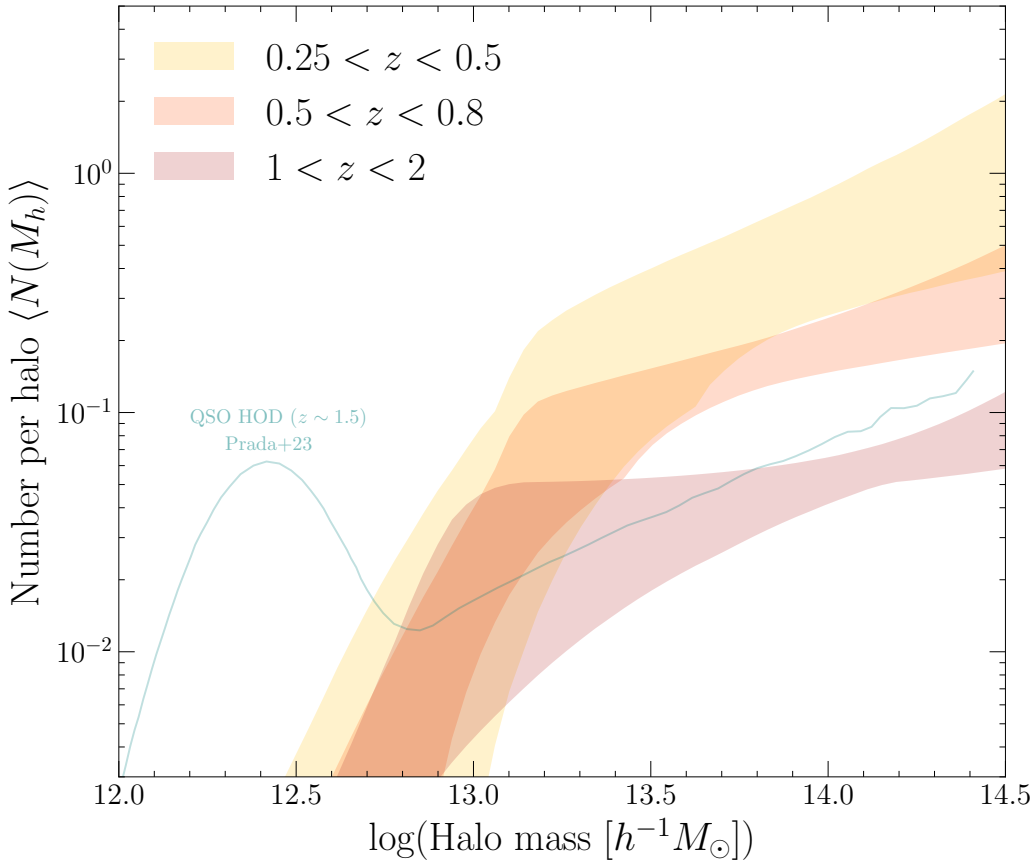

Figure 4.9: The halo occupation distributions (68% confidence intervals) inferred from fits to angular autocorrelations of LzRGs, IzRGs, and HzRGs. The populations are strongly clustered, hosted by halos $\gtrsim 10^{13} h^{-1} M_\odot$ across cosmic time. We also show the HOD of Type-1 quasars at $z \sim 1.5$ derived from early DESI data (Prada et al. 2023) in blue, demonstrating that radio galaxies occupy significantly more massive halos than quasars at $z \sim 1.5$.

4.6.3 Duty Cycle

A constraint on the properties of halos hosting radio galaxies can be combined with the radio AGN luminosity function (LF) to constrain the occupation fraction or duty cycle of these systems. A comparison of the number density of radio galaxies in the sample to the number density of halos massive enough to host them yields the fraction of said halos at a given time which host an observable radio AGN. Assuming that every sufficiently massive halo hosts a central galaxy with a SMBH, this occupation fraction is interpreted as a duty cycle $f_{\rm duty}$, or fraction of time that a SMBH hosted by these halos is observable as a radio galaxy in the survey (e.g., Haiman & Hui 2001; Martini & Weinberg 2001).

To estimate the radio galaxy duty cycle, we refit all clustering measurements (§4.4) with a minimum halo mass parameter $M_{\rm min}$ required to host a radio galaxy rather than the effective mass used before, using the relation between minimum mass and bias:

$$b(M > M_{\rm min}) = \frac{\int_{M_{\rm min}}^{\infty} dM \frac{dn}{dM} b(M)}{\int_{M_{\rm min}}^{\infty} dM \frac{dn}{dM}}. \tag{4.17}$$

We then estimate the space density of halos more massive than $M_{\rm min}$ by integrating the Tinker et al. (2008) halo mass function over mass and the considered redshift distribution. To estimate the space densities of the radio galaxies, we integrate the LoTSS-deep radio AGN LFs of Kondapally et al. (2022). We consider the best-fit LFs including all radio-excess AGNs, allowing for luminosity and density evolution, as presented in Kondapally et al. (2023). We account for the fact that our HzRG sample is limited by flux rather than luminosity by integrating above an evolving luminosity threshold with interpolation between LFs measured in different redshift epochs. At $z < 0.5$ where a best-fit LF is not provided, we interpolate between the LF measured at $0.5 < z < 1$, and the local ($z < 0.3$) LF from Sabater et al. (2019) and Kondapally et al. (2022). We adopt the local LF of Kondapally et al. (2022) at $L_{150\ \rm MHz} < 10^{26}$ W/Hz, and the LF of Sabater et al. (2019) above this limit, for reasons discussed in Kondapally et al. (2022), though choosing one or

the other does not change the interpretation of our results.

We display the resulting duty cycle estimates in Figure 4.10. Luminous low-frequency radio galaxies exhibit duty cycles of order $\sim 10\%$ since $z < 2$, mildly increasing over the past 10 Gyr. The characteristic observable lifetime is thus $\sim$1 Gyr, consistent with the estimate of Magliocchetti et al. (2017). We also display the implied duty cycle of eBOSS quasars from (Laurent et al. 2017), which rises and falls before and after cosmic noon. Thus, $z \sim 2$ appears to mark the beginning of the end of the "quasar epoch", and the emergence of more frequent radio galaxy activity.

4.6.4 Heating Power per Halo

Having constrained the host halo properties of luminous radio galaxies, we can now consider the energetics of the heating they supply to their environments. Here, we estimate the time-averaged kinetic heating power injected per host halo, and compare this to the same quantity for radiatively-driven quasar winds across cosmic time.

Kondapally et al. (2023) recently estimated the kinetic heating rate of radio galaxies as a function of 150 MHz luminosity - known as a specific heating rate $\Psi(L_{150}, z)$ - by convolving the luminosity functions of Kondapally et al. (2022) with a conversion between luminosity and kinetic jet power. They use the Heckman & Best (2014) relation calibrated using X-ray cluster cavities to trace the work done by jets observed at 1.4 GHz (Bîrzan et al. 2008; Cavagnolo et al. 2010). This heating rate thus specifies the spatially-averaged kinetic power released into a unit volume as a function of radio luminosity. Kondapally et al. (2023) finds that the total heating is dominated by relatively high-luminosity sources ($L_{150} \gtrsim 10^{25}$ W/Hz), similar to our sample. They find that the total heating is dominated by LERGs, and that the global kinetic heating dominates over radiative quasar-mode feedback since $z \lesssim 2$. However, as quasars and radio galaxies occupy halos differently, a more meaningful comparison is between the energy injected by jets versus winds within halos in

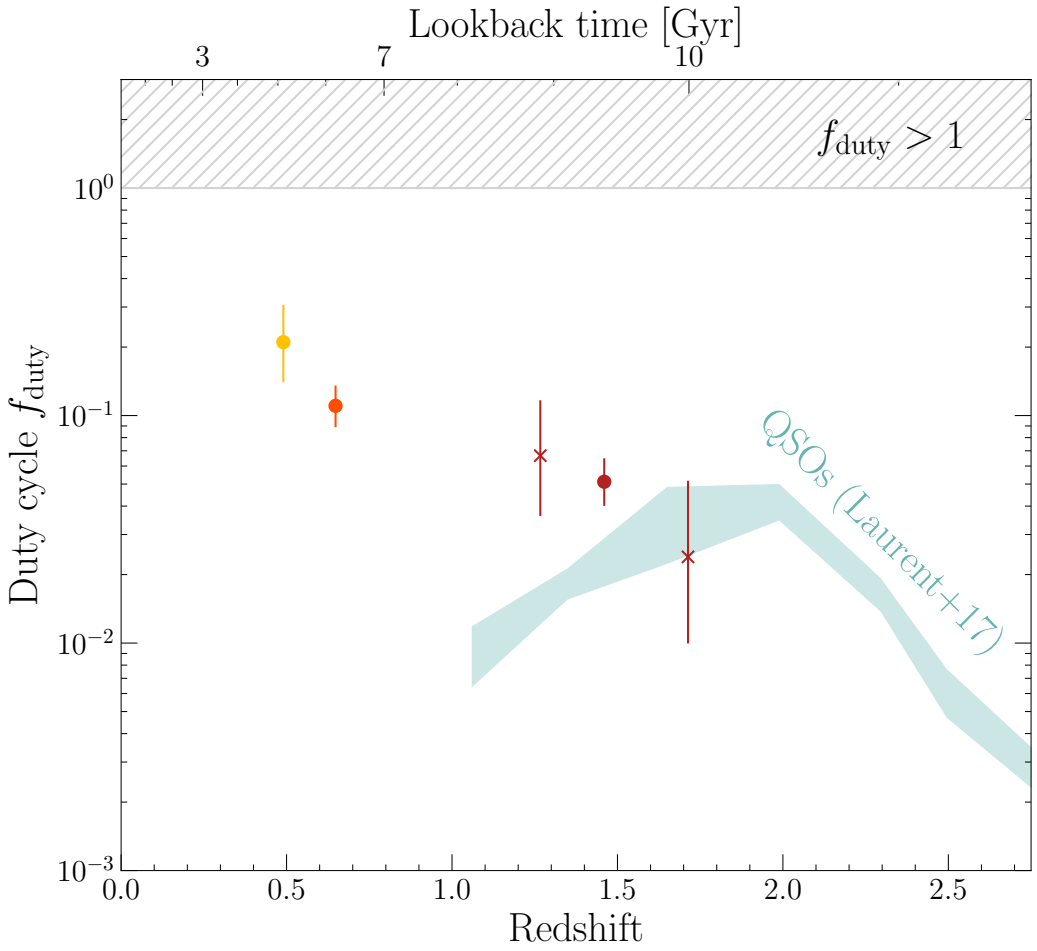

Figure 4.10: The duty cycle of luminous low-frequency radio galaxies over $z < 2$ implied by our clustering measurements and the luminosity functions of Kondapally et al. (2022) are shown in yellow, orange and red markers. These systems appear to be active for $\sim 10\%$ of the Hubble time, and rise from $\sim 5\%$ at $z \sim 1.5$ to $\sim 20\%$ at $z \sim 0.4$. We also show the 1σ confidence interval for the duty cycle of eBOSS quasars determined by Laurent et al. (2017) with a teal band. Interestingly, the luminous radio galaxy and quasar duty cycles appear to diverge after $z \sim 1.5$.

a certain mass range.

We integrate the Kondapally et al. (2023) specific heating rates (for all radio-excess AGNs) over luminosity using the flux or luminosity limit of our sample at each redshift, yielding a total kinetic heating power sourced from these jets per unit volume, a kinetic luminosity density. The volume and time-averaged kinetic power injected by jets per halo into halos within some mass interval is given by the division of the kinetic luminosity density with the number density of halos in the mass interval, then multiplied by the fraction of luminous radio galaxies which reside in halos within that mass interval, which depends on the HOD:

$$P_{\rm kin/halo}^{\rm Jets} = \frac{\int_{L_{\rm min}}^{\infty} dL \Psi(L,z)}{\int_{M_{\rm min}}^{M_{\rm max}} dM_h \frac{dN_h}{dM_h}} \times \frac{\int_{M_{\rm min}}^{M_{\rm max}} dM_h \langle N(M_h) \rangle \frac{dN_h}{dM_h}}{\int_{0}^{\infty} dM_h \langle N(M_h) \rangle \frac{dN_h}{dM_h}}. \tag{4.18}$$

We can now compare this energy output to that of radiative-mode quasars' kinetic contribution in the form of winds. Hopkins et al. (2007) estimated the cumulative bolometric power density $U_{\rm bol}$ released by radiative-mode AGN over time by integrating their AGN luminosity functions under the Soltan (1982) argument, assuming a 10% radiative efficiency. Only a fraction of this energy is ultimately transferred to kinetic energy in gaseous winds, found to vary considerably between studies ($0.0001\% - 1\%$; Fiore et al. 2017; Lacy et al. 2019; Lutz et al. 2020; Dall'Agnol de Oliveira et al. 2021; Kakkad et al. 2022; Massingill et al. 2023). Like in K23, here we adopt a characteristic range of $f_{\rm wind} = 0.1 - 0.5\%$. To compute the average quasar-driven wind power density, we multiply this factor by the bolometric energy density released per unit of cosmological time within the same redshift intervals probed by our radio galaxy samples. Analogous with Eq. 4.18, the power released by quasars into halos in a mass interval is:

$$P_{\rm kin/halo}^{\rm winds} = \frac{f_{\rm wind} U_{\rm bol}(z_1, z_2)}{\int_{M_{\rm min}}^{M_{\rm max}} dM_h \frac{dN_h}{dM_h}} \times \frac{\int_{45}^{\infty} dL \rho L}{\int_0^{\infty} dL \rho L} \times \frac{\int_{M_{\rm min}}^{M_{\rm max}} dM_h \langle N(M_h) \rangle \frac{dN_h}{dM_h}}{\int_0^{\infty} dM_h \langle N(M_h) \rangle \frac{dN_h}{dM_h}}. \tag{4.19}$$

The second term modulates the power by the fraction of the total AGN luminosity density contributed by quasars relative to all AGN. For this we adopt the best-fit AGN luminosity function of Shen et al. (2020).

The computation in Eq. 4.19 depends on the HOD of quasars, which we must assume for the purposes of this work. The effective masses of halos hosting Type-1 quasars appear remarkably consistent since $z \sim 6$ and across luminosity (Porciani et al. 2004; Croom et al. 2005; Porciani & Norberg 2006; Shen et al. 2007; Myers et al. 2007; da Ângela et al. 2008; Padmanabhan et al. 2009; Eftekharzadeh et al. 2015; Laurent et al. 2017; Geach et al. 2019; Timlin et al. 2018; He et al. 2018; Petter et al. 2022; Arita et al. 2023; Yuan et al. 2023; Prada et al. 2023; Eilers et al. 2024), at a few times $10^{12} h^{-1} M_{\odot}$. However, the full quasar HOD (Richardson et al. 2012; Eftekharzadeh et al. 2019) and its possible evolution or luminosity dependence has remained difficult to study. Here, we adopt the quasar HOD of Prada et al. (2023), assuming that it is valid across cosmic time and at all luminosities above $L_{\rm bol} \gtrsim 10^{45}$ erg/s, though we acknowledge our following interpretation may depend on the validity of this assumption.

Dividing equations 4.18 and 4.19 gives the ratio of kinetic power injected by luminous jets versus luminous quasar winds into halos within a given mass range. We estimate this ratio in ten mass intervals from $10^{12} - 10^{13.75} h^{-1} M_{\odot}$, and display the result in Figure 4.11. This demonstrates that the feedback energy injected globally by jets dominates in halos $M_h \gtrsim 10^{13} h^{-1} M_{\odot}$, while quasar winds may dominate in lower-mass halos. This can be understood as a consequence of the result of Kondapally et al. (2023) that jet-heating dom-

inates the total heating budget at $z \lesssim 2$ across all environments, unified with our clustering constraints showing that radio galaxies are typically found in halos $\gtrsim 10^{13} h^{-1} M_{\odot}$, while quasars are found in lower-mass halos (Fig. 4.9).

We raise the caveat that considering only Type-1 / unobscured quasars may provide an incomplete picture of radiatively-driven winds. Obscured quasars are canonically obscured by orientation with the line of sight, but a growing body of evidence may instead favor an evolutionary model of obscuration (Sanders et al. 1988b; Alexander & Hickox 2012; Hickox & Alexander 2018), where some quasars are obscured by galactic or circumnuclear gas during an evolutionary phase. Obscured / reddened quasars are sometimes observed to be more strongly clustered than unobscured systems (Hickox et al. 2011; Donoso et al. 2014; DiPompeo et al. 2017a; Petter et al. 2023), occupy galaxies forming stars more vigorously (Chen et al. 2015; Andonie et al. 2022), and show enhanced radio emission (Klindt et al. 2019), possibly tracing winds (Rosario et al. 2021). The high clustering amplitude and duty cycle of $\sim 10\%$ (Petter et al. 2023) are similar to the radio galaxies studied in this work, suggesting a possible stronger influence on the halos considered here compared to unobscured quasars. However, further constraints on obscured quasars' clustering evolution, their luminosity functions, and their radiation-wind coupling efficiency will be required to make meaningful constraints on their contribution to kinetic feedback.

We also note that by necessity, we have selected a heterogeneous radio galaxy population, containing a mix of LERGs and HERGs, FRI and FRII systems, which are hosted by a mix of star-forming and quiescent galaxies. Croston et al. (2019) finds that FRI systems at $z < 0.4$ occupy systematically richer environments than FRII analogues. Binning samples by FR class may become possible with upcoming sub-arcsecond LOFAR imaging (Sweijen et al. 2022).

We note that the conversion rate between observed radio luminosity and mechanical power is subject to significant systematic uncertainties (Willott et al. 1999; Hardcastle & Krause

2013; Hardcastle 2018). Therefore, further calibration of such a relation at 150 MHz and at high redshift, along with more precise clustering measurements are required to uncover which form of feedback dominates in massive halos at high redshift.

4.7 Conclusions

In this work, we select luminous ($L_{150\,\mathrm{MHz}} \gtrsim 10^{25.25}$ W/Hz) low-frequency radio galaxies, bin them into redshift epochs using optical and infrared counterparts, and study the evolution of their clustering properties from $z < 2$. We summarize our results as follows:

- These systems are strongly clustered at $z < 2$, hosted group-scale halos ($10^{13} - 10^{14} h^{-1} M_{\odot}$). The effective halo mass thus increases with cosmic time at a rate consistent with the mean growth rate of halos. The minimum halo mass of $\sim 10^{13} h^{-1} M_{\odot}$ derived from HOD modeling is consistent with a model where the triggering of low-frequency radio galaxy activity is strongly linked to the condensation of gas from a hot halo.

- The coevolution of the radio galaxy clustering and luminosity function implies that the duty cycle of these systems is of order $\sim 10\%$ for the past 10 Gyr, mildly increasing during this epoch where the quasar duty cycle has been declining. The characteristic observable lifetime of these systems is thus order $\sim$1 Gyr.

- The kinetic power injected into group halos ($M_h \gtrsim 10^{13} h^{-1} M_{\odot}$) appears to be dominated by luminous radio galaxies rather than radiatively-efficient quasar wind power since $z < 2$. Radio galaxies are thus expected to be more efficient drivers of feedback in groups for at least the last $\sim$ 10 Gyr. Improved constraints on the relation from radio luminosity to kinetic power, the fraction of quasar luminosity coupled to winds, the HODs of radio galaxies and quasars, and any possible evolution with luminosity to confirm this result.

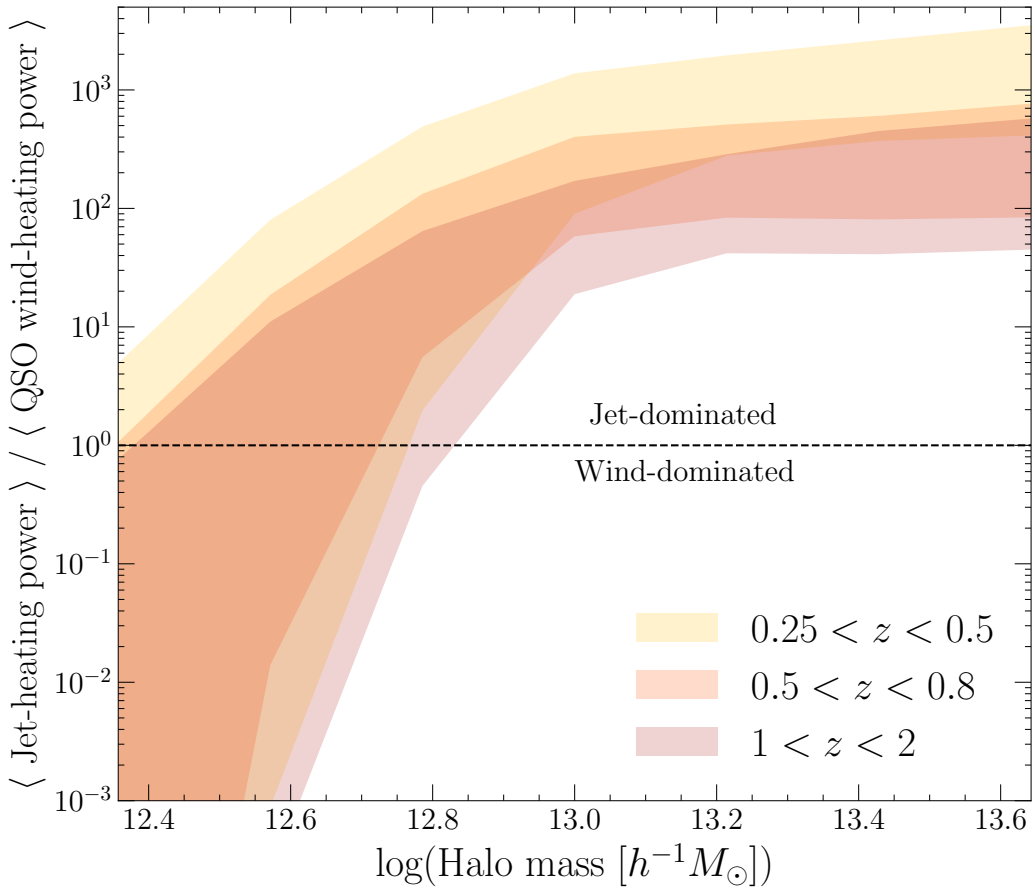

Figure 4.11: The ratio of the kinetic heating power injected into halos within a mass interval of 0.25 dex by luminous radio galaxies compared with quasar winds. This comparison results from combining the radio galaxy HODs presented in this work, the jet heating rates of K23, the bolometric quasar power densities of Hopkins et al. (2007), the quasar HOD of Prada et al. (2023) and assuming a wind-coupling range $f_{\rm wind} = 0.1 - 0.5\%$ according to equations 4.18 and 4.19. Luminous jets appear to dominate the global heating in environments $\gtrsim 10^{13} h^{-1} M_\odot$ at $z < 2$. We note that this computation assumes that the HODs of radio galaxies and quasars are luminosity independent, and the quasar HOD is also redshift-independent. In this picture, jet-mode feedback appears to dominate the global kinetic input into group and cluster-scale halos, while quasar winds may dominate at lower masses.

The upcoming WEAVE-LOFAR spectroscopic follow up survey of LoTSS-detected sources (Smith et al. 2016) will improve precision on the clustering properties of radio galaxies, mitigate against any systematics in the radio galaxy redshift distribution, and possibly support separately studying clustering for LERGs and HERGs. Furthermore, upcoming sub-arcsecond imaging (Sweijen et al. 2022) of the LoTSS-wide area using the International LOFAR baselines should allow purer selection of radio galaxies to fainter fluxes using brightness temperature criteria (Morabito et al. 2022). Higher-resolution imaging should also allow for more precise cross-matching at alternate wavelengths, and enable clustering studies as a function of FRI/II status. Finally, future X-ray missions such as *Athena* (Nandra et al. 2013) will detect the gas in groups and clusters out to $z \sim 2$ heated in part by the systems studied in this work.

This work shows that jet-mode feedback appears to be the dominant source of heating in massive environments at $z < 2$. However, further constraints on the evolution of radio galaxy clustering with redshift and luminosity, and the conversions between radio luminosity and mechanical power are required to perform a census of jet-mode feedback, and uncover the complete picture of how supermassive black hole accretion shapes galaxy growth across time.

Chapter 5

Future Work and Conclusions

This thesis has focused on selecting AGN of different observational classes in wide-field surveys, and putting precise constraints on their host halo properties. With these measurements, I was able to test unified versus evolutionary models of AGN obscuration and constrain the halo mass scales at which jet-mode feedback dominates over quasar wind-mode feedback.

Now, a number of upcoming surveys have begun collecting data which will further push studies of the AGN-halo connection into the precision era. The Dark Energy Spectroscopic Instrument (DESI) will measure spectroscopic redshifts for ~ 3 million optical quasars, representing an order of magnitude improvement over eBOSS. Likewise, the eROSITA (Merloni et al. 2012) final depth survey should detect ~ 3 million X-ray AGN. *Euclid* (Laureijs et al. 2011), *Roman* (Green et al. 2012), and the Rubin Observatory (Ivezić et al. 2019) will detect unprecedentedly large galaxy samples out to $z \sim 3$ which could be used for cross-correlation studies with AGN. Finally, arcminute-resolution maps of CMB with the Atacama Cosmology Telescope (Swetz et al. 2011) and South Pole Telescope (Carlstrom et al. 2011) will improve CMB lensing constraints, and also permit Sunyaev-Zel'dovich measurements around AGN to search for evidence of feedback. These datasets, combined

with novel modeling techniques have the potential to significantly advance our understanding of AGN feedback and structure.

In this Chapter, I will outline three potential avenues to further advance our understanding of the AGN-halo connection, and finally summarize and conclude.

5.1 Future Work

5.1.1 Quasar Clustering as a Function of Luminosity

Quasars' clustering amplitude has long been expected to scale with their bolometric luminosity, given that a more massive black hole is capable of powering a more luminous quasar according to the Eddington limit scaling ($L_{\rm Edd} \propto M_{\rm BH}$). As there is a tight correlation between black hole and galaxy bulge mass (Kormendy & Ho 2013), and also between bulge mass and halo mass (Behroozi et al. 2019), transitively there should be a connection between halo and central black hole mass, and possibly thus halo mass and quasar luminosity.

This matter is relevant to understanding which environments quasars release their feedback energy into. In Chapter 4, I assumed that quasars' host halo mass is independent of luminosity when comparing the jet-to-wind power injection as a function of halo mass. However, as higher luminosity quasars inject more energy, they dominate the total heating energy budget even though they are far less abundant than fainter AGN. If quasars' host halo properties indeed do correlate with luminosity, the trend of jet-to-wind power as a function of halo mass presented in Figure 4.11 will need to be revised.

However, many studies have found a surprising lack of evidence for a strong trend of host halo mass with quasar luminosity (Myers et al. 2007; Shen et al. 2009; Geach et al. 2019). This has generally been interpreted to imply that quasar activity is highly variable and sporadic compared to the timescales of galaxy evolution, such that a black hole of a given

mass at a given time is accreting at a stochastic fraction of the Eddington limit (e.g., Hickox et al. 2014). This effect would imply that a given luminosity-limited sample of quasars contains a mixture of systems powered by black holes of a wide range of masses, accreting at a similarly wide range of Eddington rates. Thus, any $M_{\rm BH} - M_{\rm halo}$ relation becomes smeared out when binning quasars by luminosity. However, this effect is expected to abate at the highest end of the luminosity function because the Eddington limit sets a floor of how massive a black hole must be at minimum to power a quasar of a given luminosity. The most luminous quasars beyond the break of the luminosity function releasing $\sim 10^{47}$ erg/s corresponds to a black hole mass of $\sim 10^9 M_\odot$ if it is accreting at the Eddington limit, and even more massive if it is accreting at $< \lambda_{\rm Edd}$. Thus, quasars with these extreme luminosities essentially *must* be powered by the most massive black holes. In other words, the mass range of the black hole engine powering quasars must narrow to the most massive subset at the highest luminosities. Therefore, we should expect the most luminous quasars to cluster very strongly. However, this effect has thus far been difficult to detect. The number density of systems this luminous is very small, meaning autocorrelation function measurements will prove challenging.

An ideal sample to address this topic has recently been released in the "Quaia" catalog (Storey-Fisher et al. 2024). Though the *Gaia* satellite is a mission primarily to map stars in the Milky Way, its sensitivity to point sources and all-sky coverage makes it useful in selecting luminous unobscured quasars. Its multiple observations of each part of the sky can detect the natural variability in quasars, and its incredible sensitivity to proper motions and parallaxes can cull stellar contaminants. Finally, *Gaia* collects low resolution UV-IR spectra, which can detect quasar emission lines and estimate spectroscopic redshifts. Gaia Collaboration et al. (2023) used all of these data to select samples of quasar candidates with a variety of machine-learning algorithms, with about $\sim 50\%$ reliability. Storey-Fisher et al. (2024) then combined this catalog with further proper motion cuts and *WISE* colors to select an extremely pure sample of ~ 1.3 million quasars over three quarters of the sky.

Because this sample has no pre-selection cuts like previous spectroscopic surveys, Quaia likely includes essentially *all* of the most luminous Type-1 quasars outside of the Galactic plane out to $z \sim 3$. Because clustering measurements are limited by number counts, and we only have one realization of the universe to observe, we likely will not obtain larger samples for increased autocorrelation signal-to-noise for these types of quasars going forward.

I have taken the Quaia catalog, and matched to the SDSS DR16 catalog of spectroscopic quasars (Lyke et al. 2020), for which black hole masses and bolometric luminosities have been estimated (Wu & Shen 2022). I then trained a neural network on the common objects between the catalogs to enable prediction of a bolometric luminosity based on the Gaia magnitudes, colors, and redshift estimates. This model can then be used to predict luminosities for all objects in the Quaia catalog. The most luminous systems as a function of redshift can then be selected, and their clustering can be studied and compared with lower luminosity systems. At lower redshifts ($z \sim 1$), a cross correlation measurement with more numerous galaxy systems such as luminous red galaxies will improve S/N to better constrain the clustering of the inherently sparse luminous quasars. These measurements will reveal whether the quasars releasing the most radiative energy into their surroundings occupy different environments than their less-luminous counterparts.

5.1.2 A More Sophisticated Model of the AGN-halo Connection

We have now put percent-level constraints on the effective masses of halos that host optical, infrared, and radio AGN. Optical quasars appear to occupy halos of $\sim 10^{12.5} h^{-1} M_{\odot}$, and luminous radio AGN occupy $\gtrsim 10^{13} h^{-1} M_{\odot}$ halos over the past ~ 10 Gyr, suggesting that halo mass is a crucial parameter in triggering AGN activity. However, a more sophisticated model of the AGN-halo connection than a single effective mass model is required to understand the details of this connection. In Chapters 3 & 4, we have modeled AGN clustering with a halo occupation distribution (HOD) formalism, constraining the number of AGN per halo as a function of mass, and the *satellite fraction* (number of AGN which are secondary

galaxies orbiting the central galaxy of the halo) to the $\sim 10\%$ level. Now, we are poised to constrain truly realistic models of how AGN occupy their host halos for the first time, incorporating halo properties beyond mass, the positions and velocities of the AGN inside the halo, and the manner in which AGN sample the global galaxy population.

Recent advances in wide-field multiplexed spectroscopic redshift surveys, in the resolution and volume of N-body cosmological simulations, and in novel modeling techniques are beginning to permit such a second-order exploration of the connection between galaxies and the underlying matter distribution. The `AbacusHOD` tool (Yuan et al. 2022) has been developed to extend studies of the halo-galaxy connection beyond the standard HOD formalism, modeling galaxy clustering by populating mock galaxies into halos identified in the state-of-the-art `AbacusSummit` N-body simulations (Maksimova et al. 2021), allowing far more customization than analytic models. This flexible technique can then predict clustering signatures dependent on secondary halo parameters at fixed mass (known as *secondary/assembly bias*), such as halo concentration, formation age, spin, and shape. If AGN exhibit any secondary biases, the environments they release energy into will have been previously incorrectly estimated, and the secondary halo properties correlated with AGN activity may encode clues of how AGN are fueled. This modeling approach also permits relaxing the assumption that galaxies identically follow the velocity or density fields of the underlying halos' dark matter particles, thus probing any present *velocity bias*. If AGN exhibit a velocity or density bias, especially relative to the inactive host galaxy population, it may also reveal how the active BHs are being fueled.

Some of these effects manifest themselves in the two-dimensional or *anisotropic* correlation function. Redshift-space distortions (RSDs) arising from peculiar motions of galaxies with respect to the Hubble flow result in a perturbed correlation function only along the line-of-sight dimension (Kaiser 1987). While RSDs can be an inconvenience to be integrated over in projected clustering studies (as in my previous work), they imprint valuable

information about the motions and distributions of galaxies in their halos onto the full 2D clustering. Our approach to realistically model the AGN-halo connection will thus rely on measuring the anisotropic clustering of AGN and cross-correlations with other galaxy tracers.

Interestingly, a study of the DESI early release quasars (DESI Collaboration et al. 2023) suggests that they exhibit a strong velocity bias, implying they are orbiting their host halos more even more quickly than if all quasars were satellite galaxies in their host halos (Yuan et al. 2023). However, this analysis has not accounted for the fact that quasar redshifts are typically more uncertain than galaxy redshifts even when observed with high-resolution spectrographs owing to their broad emission lines, and the fact that high-ionization lines can appear blueshifted from the galaxy's systemic redshift in the presence of an outflow which will artificially mimic the effect of velocity bias. I have thus been developing a fork of `AbacusHOD` to account for the effect of quasar redshift errors on inferred HOD parameters by manually perturbing the mock quasars' velocities in the simulation according to the distribution of redshift errors intrinsic to the dataset. I have estimated the 2D clustering of the eBOSS quasar dataset (with updated systemic redshift estimates, Wu & Shen 2022) in five redshift bins from $0.9 < z < 3.2$, on scales larger than the $\sim 1'$ fiber collision limit of the eBOSS spectrograph. I then modeled these signals with `AbacusHOD`, using a model where quasars are assumed to be centrals in their halos, but allowed to have a velocity bias, where a maximum velocity bias of 1 means that the quasars are assigned random velocity kicks proportional to the host halo particle velocity dispersion. I tentatively find significant evidence for a velocity bias at all redshifts probed, even after accounting for the effect of redshift errors. However, further exploration of alternative models which might fit this data are required. The model constraining power would also be significantly enhanced with redshift-space clustering measurements on one-halo scales, which will be achieved by future DESI data releases, as DESI avoids the fiber collision problem by making multiple passes of each sky region. If a strong velocity/density bias indeed exists for quasars, it may

imply that quasars are typically recently accreted onto their host halo or involved frequently in violent mergers, which would be a key insight in explaining how AGN are fueled.

5.1.3 Measuring Sunyaev-Zel'dovich Pressure Around AGN and Galaxies

Now that we have put precise constraints on the masses of halos which host different varieties of AGN and begun estimate the typical budget of kinetic power injected by these systems into their surrounding intragroup/intracluster medium (ICM), a crucial test of AGN feedback models is to check whether this energy is leaving an observable imprint in the ICM surrounding AGN, and galaxies which may have hosted AGN activity in the past.

In addition to clustering, another observational tool to probe AGN feedback is now entering the precision era, a result of wide-field arcminute-resolution maps of the CMB intensity from experiments such as the Atacama Cosmology Telescope (ACT, Swetz et al. 2011). These telescopes can detect the characteristic distortion in the CMB's spectrum induced by inverse-Compton scattering off hot electrons along the line-of-sight at high significance, known as the thermal Sunyaev-Zel'dovich (SZ, Sunyaev & Zeldovich 1972) effect. Current SZ maps are noise-dominated except near the most massive galaxy clusters, but the average pressure in lower-mass systems can be recovered by *stacking* the maps at the positions of many targets, similar to the lensing stacks I performed in Chapter 2. Measuring the SZ effect directly traces the pressure of gas in the ICM, which may contain imprints of AGN feedback.

Gas accreted onto halos is gravitationally shock-heated, with more massive halos capable of heating to higher temperatures. Purely gravitational heating should thus result in a *self-similar* gas pressure profile across halos of varying mass, entailing a power-law relation between halo mass and contained thermal energy (Kaiser 1986; Komatsu & Seljak 2001). Nongravitational heating such as AGN feedback can thus manifest itself as a deviation

from self-similarity, either by contributing an excess of thermal energy to the CGM for a given halo mass, or by radially redistributing the gas pressure (Battaglia et al. 2012). With precision measurements of the halo masses that galaxies and AGN occupy, and with spatially-resolved measurements of the pressure in the surrounding medium, we are poised to perform strong tests of AGN feedback by searching for these effects.

The atmospheres surrounding powerful AGN are natural environments to search for evidence of feedback. However, the baseline SZ maps can be contaminated by AGN emission at microwave frequencies, stemming from synchrotron-emitting jets and far-infrared-emitting dust from star formation (correlated with AGN activity). A majority of SZ stacking analyses have thus probed samples of mostly passive, inactive galaxies which are microwave-quiet and yield an unbiased measure of the CGM pressure (e.g., Greco et al. 2015). These measurements attempt to detect the statistical contribution of historical AGN feedback in the global galaxy population. Alternatively, residual SZ effects can be recovered even in the presence of AGN/galaxy contamination, through stacked spectral energy distribution (SED) decomposition (Gralla et al. 2014; Crichton et al. 2016; Soergel et al. 2017; Hall et al. 2019). Rather than stacking the baseline SZ maps at the positions of AGN, one can stack flux maps from infrared to radio frequencies, and then model the observed stacked SEDs as a combination of emission from AGN, the CMB, and a spectral distortion from the SZ effect, allowing an unbiased recovery of the pressure. Wide-area surveys in the radio (e.g., ASKAP RACS, VLASS), the far-infrared (*Herschel*/AKARI) and microwave bands (the ACT single-frequency maps) enable these stacking analyses. SED-based detections of SZ pressure around quasars has so far been around the 4σ level, with tentative detections of excess thermal energy. The upcoming quasar surveys along with new CMB data with ACT DR6 should firmly detect or rule out such an excess of energy.

These methods permit a search for evidence of AGN feedback through SZ measurements with a multi-pronged approach, leveraging both stacked SEDs in the presence of microwave

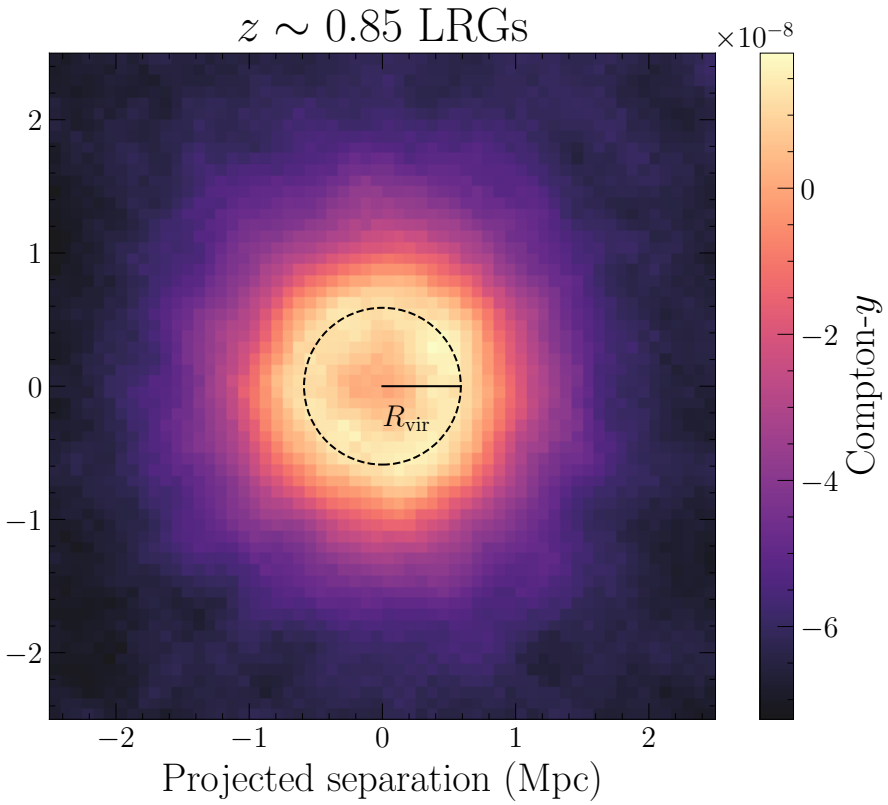

Figure 5.1: Stack of the ACT DR6 tSZ map at the positions of 8 million photometric LRGs at $0.7 < z < 1$. The stacked profile is negatively biased near the center, where radio emission from AGN likely is contaminating the signal. The virial radius of the effective host halo ($M_h = 10^{13.3} M_\odot/h$) is shown with a dashed line. SZ pressure is detected at high significance outside the virial radius, possibly probing the gas evacuated from galaxy groups by AGN feedback.

AGN contamination, and baseline SZ stacks otherwise.

The work presented in Chapter 4 has shown that AGN hosting powerful jets occupy galaxy groups across cosmic time and inject $\sim 10^{59}$ erg / Gyr into the intra-group medium, but this injection is expected to be dwarfed by the gravitational heating energy ($\sim 10^{61}$ erg) characteristic of these systems and thus difficult to detect in a spatially-integrated SZ measurement. However, jetted AGN may be expected to leave an imprint on the CGM by lifting gas and causing the pressure profile to deviate from the self-similar profile. Indeed, X-ray surveys of local galaxy groups find large decrements in the baryon fraction compared to clusters (Eckert et al. 2021), suggesting the gas has been heated by AGN feedback and lifted beyond the virial radius where it would fall below the surface brightness limits of X-ray telescopes. SZ measurements are ideal to search for this hot extended gas. The 1.6' ACT beam probes scales of $\lesssim R_{\rm vir}$ for massive groups across redshift, enabling studies of the gas pressure within and surrounding these systems' halos.

I show a preliminary stack of the ACT DR6 SZ map (Coulton et al. 2024) for a set of $0.7 < z < 1$ DESI-like photometric luminous red galaxies (Zhou et al. 2023) in Figure 5.1. LRGs occupy groups and are representative hosts for powerful jetted AGN, likely having undergone jet-mode feedback in the past, if not presently. Significant SZ pressure appears to be detected outside the virial radius, tentatively probing the gas evacuated from group environments. With a baseline stack in an annuli outside $R_{\rm vir}$ along with a SED decomposition measurement within the central beam to account for radio contamination should enable a strong test of the self-similar pressure profile in groups for systems with current and past feedback episodes.

5.2 Conclusions

Accreting supermassive black holes at the centers of galaxies play a significant role in regulating galaxy formation and evolution. The energy released during accretion onto these

objects in the form of radiation and mechanical jets appears capable of heating the surrounding gaseous medium on large scales, ultimately quenching star-formation and growth of the host galaxy. Certain observational classes of AGN may represent evolutionary stages in a galaxy-AGN coevolutionary scheme, rather than resulting from an orientation effect along the line-of-sight. This thesis has argued that probing how AGN populate the large-scale structure is a particularly robust test of these models.

In this thesis, I have utilized public survey data spanning the from the optical to radio band to select AGN samples of varying observational classes, leveraging large swaths of sky and probing a large fraction of the Hubble volume to enable unprecedentedly precise constraints on the properties of halos that host these systems. Now, the accelerating proliferation of large-scale structure data over the coming decade is set to enable precision tests of models for AGN feedback and phenomenology, furthering our understanding of the cosmic role played by the simultaneously exotic and ubiquitous supermassive black holes lurking at the centers of galaxies.

Appendix A

Python Packages

A.1 HaloModelPy

Many codes exist in the literature to predict observables in the halo model (e.g., Nishimichi et al. 2019; Chisari et al. 2019; Murray et al. 2021). However, most are optimized for cosmological analysis, whereas this thesis concerned probing host halo properties of AGN in a fixed fiducial cosmology. Additionally, most codes are not centered around estimating host halo parameters for a given galaxy sample, rather simply predicting the matter correlation function and letting the user fit e.g. for a halo bias.

To facilitate exploratory research in the connection between galaxies and their host halo properties, I developed `HaloModelPy`, a `Python` package to quickly compute observables in the halo model, including auto and cross-correlation functions in angular, spatial, projected, and redshift space, as well as lensing cross-powers or stacked profiles. It is optimized for an observational astrophysics use case, given that we fix the cosmology upon initialization, and prioritize real-space correlation functions rather than power spectra. We also always compute observables integrated over a given redshift distribution while many other codes predict spatial correlation functions at a single input redshift, as delta function redshift

distributions are not representative of realistic galaxy surveys. I also implement a number of methods to parameterize the galaxy-halo connection with redshift, including a constant bias model, a constant effective halo mass, a constant minimum required mass, and the Zheng et al. (2007) HOD model.

An example of the variety of models is shown in Figure A.1. I predict the spatial correlation function for a model with an effective bias, an effective host halo mass, a minimum halo mass, and a HOD. This demonstrates the simplicity of modeling a galaxy correlation function according to a number of different parameterizations for the halo-galaxy connection. I additionally show the modeling of the cross-power spectrum between the galaxy overdensity field and CMB lensing convergence in Figure A.2. If an user has both lensing and clustering measurements for a given sample, they can quickly fit both with the same model and test for agreement, as I did in Chapters 2, 3, and 4.

We transform directly from the matter power spectrum to observables such as the angular correlation function, whereas most routines (such as CCL) require an intermediate step of transforming the matter power to an angular power, then Fourier transforming to real-space. Our method reduces execution time, and we have implemented all transforms in fast vectorized `numpy` code. The optimization is useful when e.g., fitting an HOD to an observed correlation function with a Markov-Chain Monte Carlo (MCMC) technique requiring tens of thousands of model prediction executions.

Finally, we provide a set of fitting routines to rapidly infer host halo properties. Simply pass a dictionary containing a measured correlation function, a redshift distribution of the sample, clustering uncertainties or a covariance matrix, and a set of scale bins, and `HaloModelPy` will quickly fit for an effective bias, host halo mass, and minimum halo mass. A routine is also included to explore HOD parameter spaces with `emcee`. By passing a dictionary containing clustering measurements on one and two-halo scales, a redshift distribution, and a set of HOD priors, `HaloModelPy` will return a chain representing the

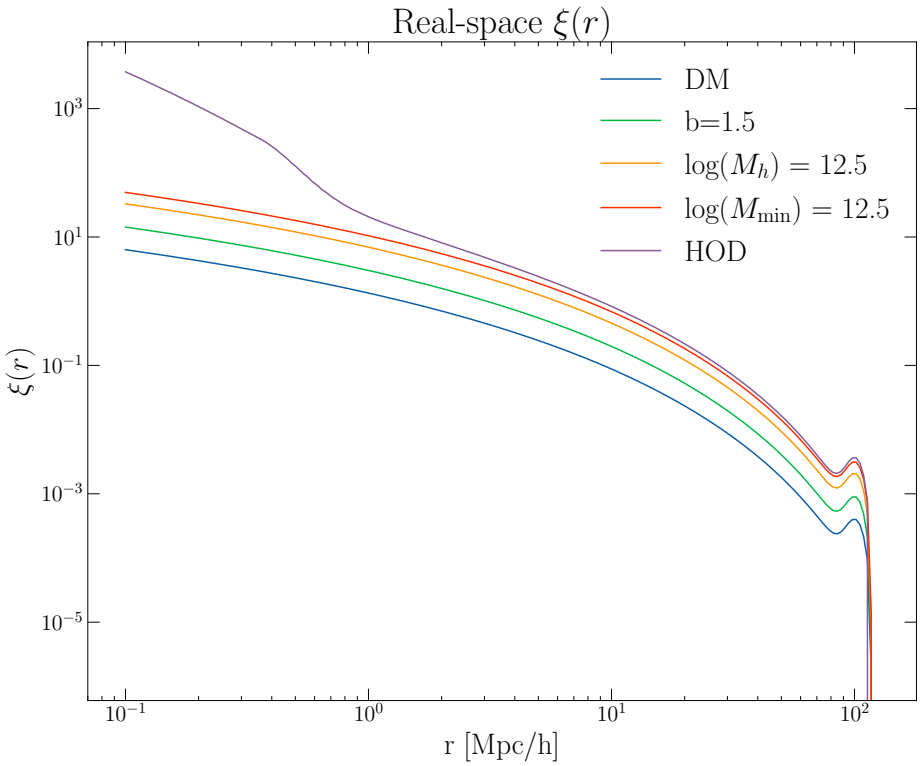

Figure A.1: A showcase of the different models implemented into `HaloModelPy` to model the halo-galaxy connection. These include a linearly biased model, a model biased according to a constant effective halo mass, a constant minimum mass, and a HOD model describing the central and satellite occupations.

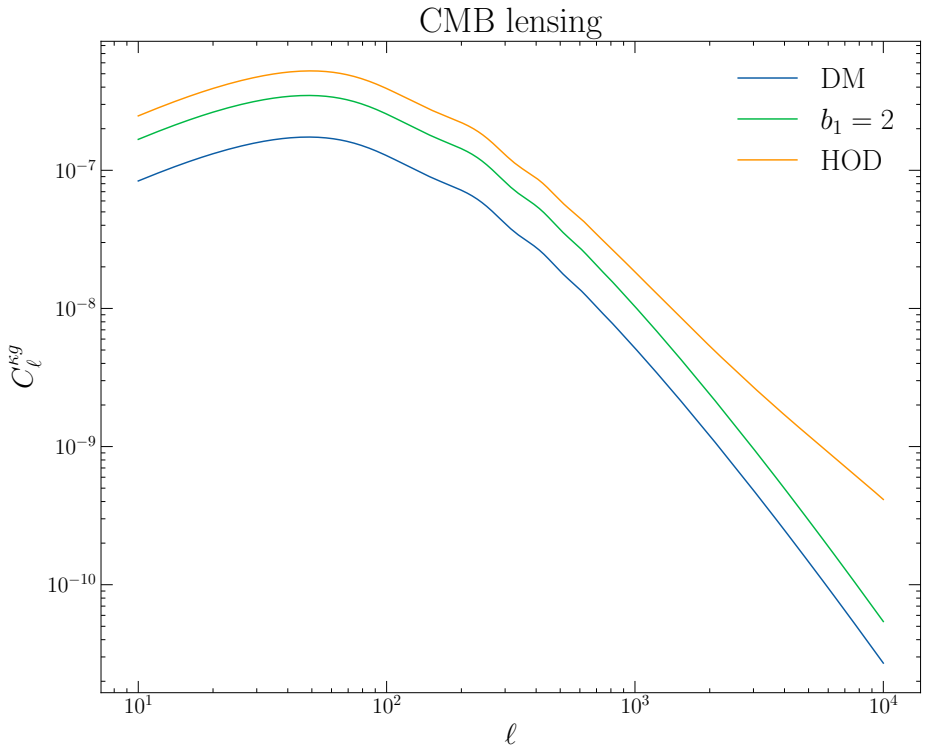

Figure A.2: A showcase of modeling the halo-galaxy connection as probed by the lensing of the CMB. We show the prediction for the cross-power spectrum between galaxy overdensity and CMB lensing convergence for a linearly-biased model with respect to dark matter, as well as an HOD.

distribution of parameter combinations which best fit the data.

`HaloModelPy` is maintained at https://github.com/gpetter/HaloModelPy. A further set of example usage cases can be explored in the Jupyter notebook at https://github.com/gpetter/HaloModelPy/blob/master/notebooks/showcase.ipynb. This package can be installed by creating a new python environment with the requisite packages. With anaconda:

```
conda create -n <env name> numpy scipy astropy pyccl
```

Then, using pip to install from the GitHub repository:

```
pip install git+https://github.com/gpetter/HaloModelPy.git
```

A.2 Corrfunc Helper

Large-scale structure research has been greatly aided by a number of public highly optimized pair counting codes (e.g., Jarvis 2015; Sinha & Garrison 2020) to compute correlation functions. However, as clustering measurements tend to require careful consideration of selection functions, weighting schemes, estimators, and handling of uncertainties, these codes tend to present a moderate learning curve for an user to begin performing clustering measurements and then inferring physical parameters.

To accelerate large-scale structure research, I developed `Corrfunc helper`. Once an user has compiled a catalog of galaxies and a random catalog parameterizing the selection function of the sample, `Corrfunc helper` allows the user to compute a variety of clustering statistics in a single line of Python code. This is a wrapper of the highly-optimized `Corrfunc` package, which handles the pair-counting.

An user simply passes a catalog either in the form of an `Astropy` table, or a tuple of coordinates each for the data and random samples, a set of scale bins to measure the correlation function within, and optional weights, and `Corrfunc helper` will return a measured correlation function with uncertainties. `Corrfunc helper` can also compute cross-correlation

functions between two galaxy samples simply by passing two data catalogs, and a random catalog corresponding to either one or both of the data catalogs.

This package implements both the Landy & Szalay (1993) estimator included with `Corrfunc` as well as the Davis & Peebles (1983) estimator, which can be especially useful in the case of a cross-correlation, as it only depends on one random catalog and thus the selection function of the second sample is not required to be carefully characterized.

By passing a catalog containing only sky positions 'RA' and 'DEC', the package will yield an angular correlation function $w(\theta)$. However, if the sample has spectroscopic redshifts, additionally passing a comoving distance coordinate 'CHI' `Corrfunc helper` will result in the computation of the spatial projected correlation function $w_p(r_p)$ and the two-dimensional correlation function $\xi(r_p, \pi)$.

Optimally estimating uncertainties on clustering measurements is a field of research unto itself. As measuring a correlation function is fundamentally a pair-counting exercise of discrete tracers, the uncertainty on clustering measurements in the low signal-to-noise regime tend to be dominated by Poisson errors:

$$1 + \xi / \sqrt{DD}. \tag{A.1}$$

Ideally, estimating errors on clustering measurements would be accomplished by generating many mock galaxy datasets with indistinguishable power spectra from the data, and measuring the covariance matrices directly from these realizations. This is the approach taken by experiments where mitigating systematics is critical, such as the cosmology experiments eBOSS and DESI (e.g., Zhao et al. 2021). However, this typically relies on outputs of computationally expensive N-body simulations and depends on sophisticated modeling of galaxy formation.

In absence of such models, a more typical technique to account for uncertainties from

both shot noise and cosmic variance typically involves a form of directly manipulating the observed data by resampling it. By splitting a survey footprint into a number of patches containing similar numbers of sources, one captures the field-to-field error stemming from cosmic variance. To estimate uncertainties on the correlation function, the *bootstrap* or *jackknife* resampling techniques are often used. Bootstrap resampling refers to splitting the footprint into N patches, then randomly drawing a set of those patches *with replacement*, meaning that a single patch is allowed to be drawn multiple times (or not at all) in a single realization. Then, the objects in those patches are used to compute a correlation function, and the variance across many of these realizations is taken to be the clustering uncertainty. Jackknife resampling typically refers to splitting the footprint into patches, then randomly removing one patch from the dataset for a given realization.

We implement a bootstrap resampling technique to estimate uncertainties and covariance matrices on clustering measurements in `Corrfunc helper`, following Norberg et al. (2009). This work finds that a bootstrap technique successfully recovers the true eigenvectors of the covariance matrix, but the normalization can be overestimated by up to $\sim 40\%$. They find that *oversampling* patches can yield accurate uncertainty normalizations, meaning if the survey is split into N patches, a given realization will draw an integer multiple of N.

We implement an automatic algorithm to split the footprint into patches, leveraging the K-means clustering algorithm implemented in `scikit-learn`. The user then can specify a number of bootstrap resamplings to perform, as well as optionally a factor to oversample by (the default is unity).

This package can be easily combined with the functionality `HaloModelPy` to model and fit the observed correlation functions, such that an user with an observed galaxy catalog, random catalog, and redshift distribution can quickly estimate the correlation function, and then fit it to deduce the bias, effective halo mass, or HOD parameters, all in a matter of min-

utes. An example is given in Figure A.3, where I take the catalogs of eBOSS quasars and randoms provided by Ross et al. (2020), use `Corrfunc helper` to measure the projected correlation function, then `HaloModelPy` to model this signal and infer the effective bias and halo mass, all with three lines of code. These tools thus aim to be useful in facilitating exploratory research in the galaxy-halo connection.

In [1]:
```
from halomodelpy import redshift_helper, clustering_fit
from corrfunc_helper import twoPointCFs
```

In [2]:
```
qso = Table.read('eBOSS_QSO.fits') # eBOSS quasar catalog
rand = Table.read('eBOSS_QSO_randoms.fits') # random catalog
```

In [3]:
```
rps = np.logspace(0.5, 1.4, 10) # projected radii in Mpc/h
cf = twoPointCFs.autocorr_cat(rps, qso, rand, estimator='LS', pimax=40.,
                              nbootstrap=500)
```

In [4]:
```
dndz = redshift_helper.dndz_from_z_list(qso['Z'], nbins=30)
fit = clustering_fit.fit_pipeline(dndz, cf)
fit['plot']
```

Out[4]:

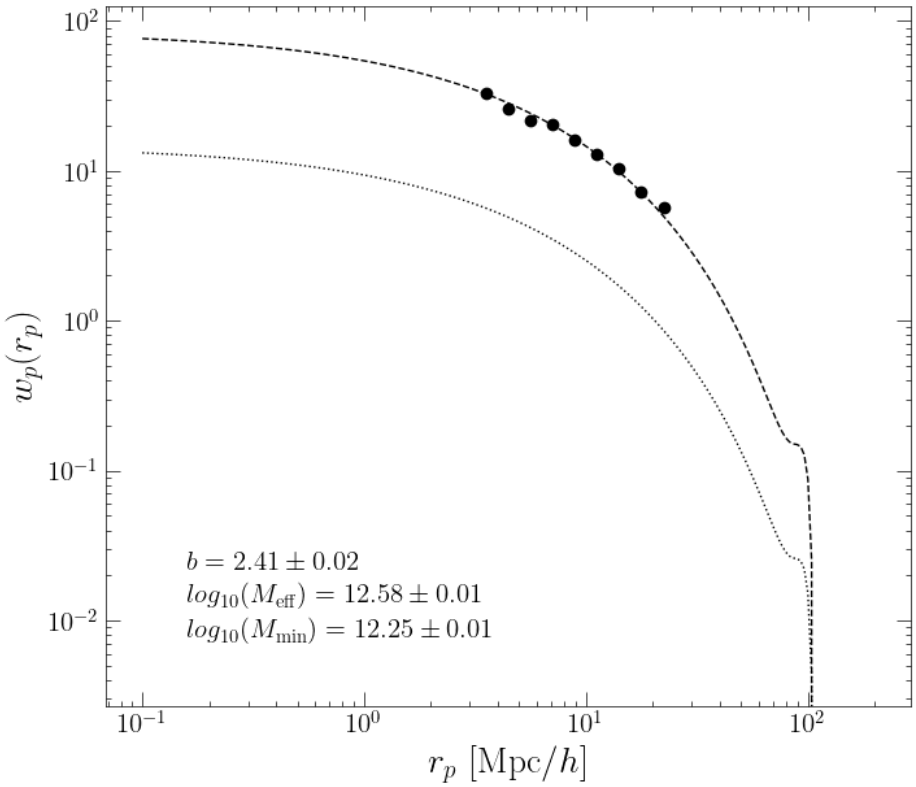

Figure A.3: An example simplified workflow leveraging both `Corrfunc helper` and `HaloModelPy` to both measure and then fit the clustering of quasars from the eBOSS.

Bibliography

Acker, A., Marcout, J., Ochsenbein, F., et al. 1992, The Strasbourg-ESO Catalogue of Galactic Planetary Nebulae. Parts I, II.

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3, doi: 10.3847/1538-4365/ab929e

Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S8, doi: 10.1093/pasj/psx081

Aird, J., & Coil, A. L. 2021, MNRAS, 502, 5962, doi: 10.1093/mnras/stab312

Aird, J., Coil, A. L., & Georgakakis, A. 2018, MNRAS, 474, 1225, doi: 10.1093/mnras/stx2700

Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892, doi: 10.1093/mnras/stv1062

Alam, S., Peacock, J. A., Kraljic, K., Ross, A. J., & Comparat, J. 2020, MNRAS, 497, 581, doi: 10.1093/mnras/staa1956

Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617, doi: 10.1093/mnras/stx721

Alexander, D. M., & Hickox, R. C. 2012, NewAR, 56, 93, doi: 10.1016/j.newar.2011.11.003

Alexander, D. M., Swinbank, A. M., Smail, I., McDermid, R., & Nesvadba, N. P. H. 2010, MNRAS, 402, 2211, doi: 10.1111/j.1365-2966.2009.16046.x

Alexander, D. M., Brandt, W. N., Smail, I., et al. 2008, AJ, 135, 1968, doi: 10.1088/0004-6256/135/5/1968

Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2022, arXiv e-prints, arXiv:2208.08517. https://arxiv.org/abs/2208.08517

Allevato, V., Finoguenov, A., Civano, F., et al. 2014, ApJ, 796, 4, doi: 10.1088/0004-637X/796/1/4

Alonso, D., Bellini, E., Hale, C., Jarvis, M. J., & Schwarz, D. J. 2021, MNRAS, 502, 876, doi: 10.1093/mnras/stab046

Alonso, D., Sanchez, J., Slosar, A., & LSST Dark Energy Science Collaboration. 2019, MNRAS, 484, 4127, doi: 10.1093/mnras/stz093

Ananna, T. T., Treister, E., Urry, C. M., et al. 2019, ApJ, 871, 240, doi: 10.3847/1538-4357/aafb77

Ananna, T. T., Weigel, A. K., Trakhtenbrot, B., et al. 2022, ApJS, 261, 9, doi: 10.3847/1538-4365/ac5b64

Anderson, L. D., Bania, T. M., Balser, D. S., et al. 2014, ApJS, 212, 1, doi: 10.1088/0067-0049/212/1/1

Andonie, C., Alexander, D. M., Rosario, D., et al. 2022, Monthly Notices of the Royal Astronomical Society, doi: 10.1093/mnras/stac2800

Andonie, C., Alexander, D. M., Rosario, D., et al. 2022, MNRAS, 517, 2577, doi: 10.1093/mnras/stac2800

Antonucci, R. 1993, ARA&A, 31, 473, doi: 10.1146/annurev.aa.31.090193.002353

Arita, J., Kashikawa, N., Matsuoka, Y., et al. 2023, arXiv e-prints, arXiv:2307.02531, doi: 10.48550/arXiv.2307.02531

Asgari, M., Mead, A. J., & Heymans, C. 2023, The Open Journal of Astrophysics, 6, 39, doi: 10.21105/astro.2303.08752

Ashby, M. L. N., Stern, D., Brodwin, M., et al. 2009, ApJ, 701, 428, doi: 10.1088/0004-637X/701/1/428

Asmus, D., Hönig, S. F., & Gandhi, P. 2016, ApJ, 822, 109, doi: 10.3847/0004-637X/822/2/109

Assef, R. J., Stern, D., Noirot, G., et al. 2018, ApJS, 234, 23, doi: 10.3847/1538-4365/aaa00a

Assef, R. J., Stern, D., Kochanek, C. S., et al. 2013, ApJ, 772, 26, doi: 10.1088/0004-637X/772/1/26

Assef, R. J., Walton, D. J., Brightman, M., et al. 2016, ApJ, 819, 111, doi: 10.3847/0004-637X/819/2/111

Balogh, M. L., Pearce, F. R., Bower, R. G., & Kay, S. T. 2001, MNRAS, 326, 1228, doi: 10.1111/j.1365-2966.2001.04667.x

Banerji, M., Alaghband-Zadeh, S., Hewett, P. C., & McMahon, R. G. 2015, MNRAS, 447, 3368, doi: 10.1093/mnras/stu2649

Barrows, R. S., Comerford, J. M., Stern, D., & Assef, R. J. 2021, ApJ, 922, 179, doi: 10.3847/1538-4357/ac1352

Battaglia, N., Bond, J. R., Pfrommer, C., & Sievers, J. L. 2012, ApJ, 758, 75, doi: 10.1088/0004-637X/758/2/75

Baxter, E. J., Keisler, R., Dodelson, S., et al. 2015, ApJ, 806, 247, doi: 10.1088/0004-637X/806/2/247

Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559, doi: 10.1086/176166

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143, doi: 10.1093/mnras/stz1182

Benson, A. J. 2010, PhR, 495, 33, doi: 10.1016/j.physrep.2010.06.001

Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, ApJ, 599, 38, doi: 10.1086/379160

Berlind, A. A., & Weinberg, D. H. 2002, ApJ, 575, 587, doi: 10.1086/341469

Best, P. N. 2004, MNRAS, 351, 70, doi: 10.1111/j.1365-2966.2004.07752.x

Best, P. N., Kaiser, C. R., Heckman, T. M., & Kauffmann, G. 2006, MNRAS, 368, L67, doi: 10.1111/j.1745-3933.2006.00159.x

Best, P. N., Kondapally, R., Williams, W. L., et al. 2023, MNRAS, 523, 1729, doi: 10.1093/mnras/stad1308

Binney, J., & Tabor, G. 1995, MNRAS, 276, 663, doi: 10.1093/mnras/276.2.663

Bîrzan, L., McNamara, B. R., Nulsen, P. E. J., Carilli, C. L., & Wise, M. W. 2008, ApJ, 686, 859, doi: 10.1086/591416

Bîrzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, ApJ, 607, 800, doi: 10.1086/383519

Boehringer, H., Voges, W., Fabian, A. C., Edge, A. C., & Neumann, D. M. 1993, MNRAS, 264, L25, doi: 10.1093/mnras/264.1.L25

Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078, doi: 10.1111/j.1365-2966.2004.07431.x

Bovy, J., Hennawi, J. F., Hogg, D. W., et al. 2011, ApJ, 729, 141, doi: 10.1088/0004-637X/729/2/141

Bovy, J., Myers, A. D., Hennawi, J. F., et al. 2012, ApJ, 749, 41, doi: 10.1088/0004-637X/749/1/41

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645, doi: 10.1111/j.1365-2966.2006.10519.x

Brusa, M., Bongiorno, A., Cresci, G., et al. 2015, MNRAS, 446, 2394, doi: 10.1093/mnras/stu2117

Buchner, J., Schulze, S., & Bauer, F. E. 2017, MNRAS, 464, 4545, doi: 10.1093/mnras/stw2423

Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, ApJ, 802, 89, doi: 10.1088/0004-637X/802/2/89

Bussmann, R. S., Dey, A., Armus, L., et al. 2012, ApJ, 744, 150, doi: 10.1088/0004-637X/744/2/150

Calistro Rivera, G., Alexander, D. M., Rosario, D. J., et al. 2021, A&A, 649, A102, doi: 10.1051/0004-6361/202040214

Canalizo, G., & Stockton, A. 2001, ApJ, 555, 719, doi: 10.1086/321520

Cappelluti, N., Ajello, M., Burlon, D., et al. 2010, ApJL, 716, L209, doi: 10.1088/2041-8205/716/2/L209

Carilli, C. L., Perley, R. A., & Harris, D. E. 1994, MNRAS, 270, 173, doi: 10.1093/mnras/270.1.173

Carlstrom, J. E., Ade, P. A. R., Aird, K. A., et al. 2011, PASP, 123, 568, doi: 10.1086/659879

Carron, J., Mirmelstein, M., & Lewis, A. 2022, JCAP, 2022, 039, doi: 10.1088/1475-7516/2022/09/039

Casey, C. M., Berta, S., Béthermin, M., et al. 2012, ApJ, 761, 140, doi: 10.1088/0004-637X/761/2/140

Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213, doi: 10.1038/nature08135

Cavagnolo, K. W., Donahue, M., Voit, G. M., & Sun, M. 2008, ApJL, 683, L107, doi: 10.1086/591665

Cavagnolo, K. W., McNamara, B. R., Nulsen, P. E. J., et al. 2010, ApJ, 720, 1066, doi: 10.1088/0004-637X/720/2/1066

Chatterjee, S., Degraf, C., Richardson, J., et al. 2012, MNRAS, 419, 2657, doi: 10.1111/j.1365-2966.2011.19917.x

Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2022, arXiv e-prints, arXiv:2208.08511. https://arxiv.org/abs/2208.08511

Chen, C.-T. J., Hickox, R. C., Alberts, S., et al. 2015, ApJ, 802, 50, doi: 10.1088/0004-637X/802/1/50

Chiang, Y.-K., & Ménard, B. 2019, ApJ, 870, 120, doi: 10.3847/1538-4357/aaf4f6

Ching, J. H. Y., Croom, S. M., Sadler, E. M., et al. 2017, MNRAS, 469, 4584, doi: 10.1093/mnras/stx1173

Chisari, N. E., Alonso, D., Krause, E., et al. 2019, ApJS, 242, 2, doi: 10.3847/1538-4365/ab1658

Chung, S. M., Kochanek, C. S., Assef, R., et al. 2014, ApJ, 790, 54, doi: 10.1088/0004-637X/790/1/54

Churazov, E., Forman, W., Jones, C., & Böhringer, H. 2000, A&A, 356, 788, doi: 10.48550/arXiv.astro-ph/0002375

Circosta, C., Vignali, C., Gilli, R., et al. 2019, A&A, 623, A172, doi: 10.1051/0004-6361/201834426

Coil, A. L. 2013, in Planets, Stars and Stellar Systems. Volume 6: Extragalactic Astronomy and Cosmology, ed. T. D. Oswalt & W. C. Keel, Vol. 6, 387, doi: 10.1007/978-94-007-5609-0_8

Coil, A. L., Hennawi, J. F., Newman, J. A., Cooper, M. C., & Davis, M. 2007, ApJ, 654, 115, doi: 10.1086/509099

Coil, A. L., Georgakakis, A., Newman, J. A., et al. 2009, ApJ, 701, 1484, doi: 10.1088/0004-637X/701/2/1484

Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ, 741, 8, doi: 10.1088/0004-637X/741/1/8

Comparat, J., Richard, J., Kneib, J.-P., et al. 2015, A&A, 575, A40, doi: 10.1051/0004-6361/201424767

Cool, R. J., Moustakas, J., Blanton, M. R., et al. 2013, ApJ, 767, 118, doi: 10.1088/0004-637X/767/2/118

Cooray, A., & Hu, W. 2000, ApJ, 534, 533, doi: 10.1086/308799

Cooray, A., & Sheth, R. 2002, PhR, 372, 1, doi: 10.1016/S0370-1573(02)00276-4

Coulton, W., Madhavacheril, M. S., Duivenvoorden, A. J., et al. 2024, PhRvD, 109, 063530, doi: 10.1103/PhysRevD.109.063530

Crichton, D., Gralla, M. B., Hall, K., et al. 2016, MNRAS, 458, 1478, doi: 10.1093/mnras/stw344

Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, MNRAS, 356, 415, doi: 10.1111/j.1365-2966.2004.08379.x

Croston, J. H., Hardcastle, M. J., Mingo, B., et al. 2019, A&A, 622, A10, doi: 10.1051/0004-6361/201834019

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11, doi: 10.1111/j.1365-2966.2005.09675.x

Croton, D. J., Stevens, A. R. H., Tonini, C., et al. 2016, ApJS, 222, 22, doi: 10.3847/0067-0049/222/2/22

Cutri, R. M., Wright, E. L., Conrow, T., et al. 2021, VizieR Online Data Catalog, II/328

da Ângela, J., Shanks, T., Croom, S. M., et al. 2008, MNRAS, 383, 565, doi: 10.1111/j.1365-2966.2007.12552.x

Dall'Agnol de Oliveira, B., Storchi-Bergmann, T., Kraemer, S. B., et al. 2021, MNRAS, 504, 3890, doi: 10.1093/mnras/stab1067

Darwish, O., Madhavacheril, M. S., Sherwin, B. D., et al. 2021, MNRAS, 500, 2250, doi: 10.1093/mnras/staa3438

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827, doi: 10.1093/mnras/stz937

Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465, doi: 10.1086/160884

Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44, doi: 10.3847/0004-6256/151/2/44

Del Moro, A., Alexander, D. M., Bauer, F. E., et al. 2016, MNRAS, 456, 2105, doi: 10.1093/mnras/stv2748

Delvecchio, I., Smolčić, V., Zamorani, G., et al. 2017, A&A, 602, A3, doi: 10.1051/0004-6361/201629367

DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2023, arXiv e-prints, arXiv:2306.06308, doi: 10.48550/arXiv.2306.06308

Dey, A., Soifer, B. T., Desai, V., et al. 2008, ApJ, 677, 943, doi: 10.1086/529516

Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168, doi: 10.3847/1538-3881/ab089d

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604, doi: 10.1038/nature03335

DiPompeo, M. A., Bovy, J., Myers, A. D., & Lang, D. 2015a, MNRAS, 452, 3124, doi: 10.1093/mnras/stv1562

DiPompeo, M. A., Hickox, R. C., Eftekharzadeh, S., & Myers, A. D. 2017a, MNRAS, 469, 4630, doi: 10.1093/mnras/stx1215

DiPompeo, M. A., Hickox, R. C., & Myers, A. D. 2016a, MNRAS, 456, 924, doi: 10.1093/mnras/stv2681

DiPompeo, M. A., Hickox, R. C., Myers, A. D., & Geach, J. E. 2017b, MNRAS, 464, 3526, doi: 10.1093/mnras/stw2589

DiPompeo, M. A., Myers, A. D., Hickox, R. C., Geach, J. E., & Hainline, K. N. 2014, MNRAS, 442, 3443, doi: 10.1093/mnras/stu1115

DiPompeo, M. A., Myers, A. D., Hickox, R. C., et al. 2015b, MNRAS, 446, 3492, doi: 10.1093/mnras/stu2341

DiPompeo, M. A., Runnoe, J. C., Hickox, R. C., Myers, A. D., & Geach, J. E. 2016b, MNRAS, 460, 175, doi: 10.1093/mnras/stw986

Donahue, M., & Voit, G. M. 2022, PhR, 973, 1, doi: 10.1016/j.physrep.2022.04.005

Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142, doi: 10.1088/0004-637X/748/2/142

Donoso, E., Yan, L., Stern, D., & Assef, R. J. 2014, ApJ, 789, 44, doi: 10.1088/0004-637X/789/1/44

Dubois, Y., Peirani, S., Pichon, C., et al. 2016, MNRAS, 463, 3948, doi: 10.1093/mnras/stw2265

Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64, doi: 10.1111/j.1745-3933.2008.00537.x

Duncan, K. J. 2022, MNRAS, 512, 3662, doi: 10.1093/mnras/stac608

Duncan, K. J., Kondapally, R., Brown, M. J. I., et al. 2021, A&A, 648, A4, doi: 10.1051/0004-6361/202038809

Dunn, R. J. H., & Fabian, A. C. 2008, MNRAS, 385, 757, doi: 10.1111/j.1365-2966.2008.12898.x

Ebrero, J., Mateos, S., Stewart, G. C., Carrera, F. J., & Watson, M. G. 2009, A&A, 500, 749, doi: 10.1051/0004-6361/200911670

Eckert, D., Gaspari, M., Gastaldello, F., Le Brun, A. M. C., & O'Sullivan, E. 2021, Universe, 7, 142, doi: 10.3390/universe7050142

Edge, A. C., Stewart, G. C., & Fabian, A. C. 1992, MNRAS, 258, 177, doi: 10.1093/mnras/258.1.177

Efron, B. 1982, The Jackknife, the Bootstrap and other resampling plans

Eftekharzadeh, S., Myers, A. D., & Kourkchi, E. 2019, MNRAS, 486, 274, doi: 10.1093/mnras/stz770

Eftekharzadeh, S., Myers, A. D., White, M., et al. 2015, MNRAS, 453, 2779, doi: 10.1093/mnras/stv1763

Eilers, A.-C., Mackenzie, R., Pizzati, E., et al. 2024, arXiv e-prints, arXiv:2403.07986, doi: 10.48550/arXiv.2403.07986

Eisenhardt, P. R. M., Marocco, F., Fowler, J. W., et al. 2020, ApJS, 247, 69, doi: 10.3847/1538-4365/ab7f2a

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605, doi: 10.1086/305424

Elvis, M. 2000, ApJ, 545, 63, doi: 10.1086/317778

Elyiv, A., Clerc, N., Plionis, M., et al. 2012, A&A, 537, A131, doi: 10.1051/0004-6361/201117983

Fabian, A. C. 2012, ARA&A, 50, 455, doi: 10.1146/annurev-astro-081811-125521

Fabian, A. C., Sanders, J. S., Taylor, G. B., et al. 2006, MNRAS, 366, 417, doi: 10.1111/j.1365-2966.2005.09896.x

Fabian, A. C., Sanders, J. S., Ettori, S., et al. 2000, MNRAS, 318, L65, doi: 10.1046/j.1365-8711.2000.03904.x

Fakhouri, O., Ma, C.-P., & Boylan-Kolchin, M. 2010, MNRAS, 406, 2267, doi: 10.1111/j.1365-2966.2010.16859.x

Fawcett, V. A., Alexander, D. M., Rosario, D. J., et al. 2020, MNRAS, 494, 4802, doi: 10.1093/mnras/staa954

Fedeli, C., Semboloni, E., Velliscig, M., et al. 2014, JCAP, 2014, 028, doi: 10.1088/1475-7516/2014/08/028

Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23, doi: 10.1086/174036

Fernique, P., Boch, T., Donaldson, T., et al. 2014, MOC - HEALPix Multi-Order Coverage map Version 1.0, IVOA Recommendation 02 June 2014, doi: 10.5479/ADS/bib/2014ivoa.spec.0602F

Fiore, F., Grazian, A., Santini, P., et al. 2008, ApJ, 672, 94, doi: 10.1086/523348

Fiore, F., Feruglio, C., Shankar, F., et al. 2017, A&A, 601, A143, doi: 10.1051/0004-6361/201629478

Flesch, E. W. 2019, arXiv e-prints, arXiv:1912.05614. https://arxiv.org/abs/1912.05614

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

Gaia Collaboration, Bailer-Jones, C. A. L., Teyssier, D., et al. 2023, A&A, 674, A41, doi: 10.1051/0004-6361/202243232

Geach, J. E., & Peacock, J. A. 2017, Nature Astronomy, 1, 795, doi: 10.1038/s41550-017-0259-1

Geach, J. E., Peacock, J. A., Myers, A. D., et al. 2019, ApJ, 874, 85, doi: 10.3847/1538-4357/ab0894

Geach, J. E., Hickox, R. C., Bleem, L. E., et al. 2013, ApJL, 776, L41, doi: 10.1088/2041-8205/776/2/L41

Georgakakis, A., Aird, J., Schulze, A., et al. 2017, MNRAS, 471, 1976, doi: 10.1093/mnras/stx1602

Georgakakis, A., Comparat, J., Merloni, A., et al. 2019, MNRAS, 487, 275, doi: 10.1093/mnras/sty3454

Gilli, R., Zamorani, G., Miyaji, T., et al. 2009, A&A, 494, 33, doi: 10.1051/0004-6361:200810821

Gilli, R., Norman, C., Calura, F., et al. 2022, arXiv e-prints, arXiv:2206.03508. https://arxiv.org/abs/2206.03508

Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98, doi: 10.3847/1538-3881/aafc33

Glikman, E., Simmons, B., Mailly, M., et al. 2015, ApJ, 806, 218, doi: 10.1088/0004-637X/806/2/218

Glikman, E., Urrutia, T., Lacy, M., et al. 2012, ApJ, 757, 51, doi: 10.1088/0004-637X/757/1/51

Gonzalez-Perez, V., Lacey, C. G., Baugh, C. M., et al. 2014, MNRAS, 439, 264, doi: 10.1093/mnras/stt2410

Gralla, M. B., Crichton, D., Marriage, T. A., et al. 2014, MNRAS, 445, 460, doi: 10.1093/mnras/stu1592

Greco, J. P., Hill, J. C., Spergel, D. N., & Battaglia, N. 2015, ApJ, 808, 151, doi: 10.1088/0004-637X/808/2/151

Green, J., Schechter, P., Baltay, C., et al. 2012, arXiv e-prints, arXiv:1208.4012, doi: 10.48550/arXiv.1208.4012

Greene, J. E., Pooley, D., Zakamska, N. L., Comerford, J. M., & Sun, A.-L. 2014, ApJ, 788, 54, doi: 10.1088/0004-637X/788/1/54

Haiman, Z., & Hui, L. 2001, ApJ, 547, 27, doi: 10.1086/318330

Hale, C. L., Jarvis, M. J., Delvecchio, I., et al. 2018, MNRAS, 474, 4133, doi: 10.1093/mnras/stx2954

Hale, C. L., Schwarz, D. J., Best, P. N., et al. 2023, arXiv e-prints, arXiv:2310.07627. https://arxiv.org/abs/2310.07627

Hall, K. R., Zakamska, N. L., Addison, G. E., et al. 2019, MNRAS, 490, 2315, doi: 10.1093/mnras/stz2751

Hamilton, A. J. S. 2000, MNRAS, 312, 257, doi: 10.1046/j.1365-8711.2000.03071.x

Hardcastle, M. J. 2018, MNRAS, 475, 2768, doi: 10.1093/mnras/stx3358

Hardcastle, M. J., & Croston, J. H. 2020, NewAR, 88, 101539, doi: 10.1016/j.newar.2020.101539

Hardcastle, M. J., & Krause, M. G. H. 2013, MNRAS, 430, 174, doi: 10.1093/mnras/sts564

Hardcastle, M. J., Williams, W. L., Best, P. N., et al. 2019, A&A, 622, A12, doi: 10.1051/0004-6361/201833893

Hardcastle, M. J., Horton, M. A., Williams, W. L., et al. 2023, A&A, 678, A151, doi: 10.1051/0004-6361/202347333

Harrison, C. M., Alexander, D. M., Mullaney, J. R., & Swinbank, A. M. 2014, MNRAS, 441, 3306, doi: 10.1093/mnras/stu515

Harrison, C. M., Alexander, D. M., Swinbank, A. M., et al. 2012, MNRAS, 426, 1073, doi: 10.1111/j.1365-2966.2012.21723.x

Hasinger, G., Capak, P., Salvato, M., et al. 2018, ApJ, 858, 77, doi: 10.3847/1538-4357/aabacf

Hatch, N. A., Wylezalek, D., Kurk, J. D., et al. 2014, MNRAS, 445, 280, doi: 10.1093/mnras/stu1725

He, W., Akiyama, M., Bosch, J., et al. 2018, PASJ, 70, S33, doi: 10.1093/pasj/psx129

Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589, doi: 10.1146/annurev-astro-081913-035722

Helfand, D. J., White, R. L., & Becker, R. H. 2015, ApJ, 801, 26, doi: 10.1088/0004-637X/801/1/26

Hernán-Caballero, A., Spoon, H. W. W., Lebouteiller, V., Rupke, D. S. N., & Barry, D. P. 2016, MNRAS, 455, 1796, doi: 10.1093/mnras/stv2464

Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625, doi: 10.1146/annurev-astro-081817-051803

Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, ApJ, 782, 9, doi: 10.1088/0004-637X/782/1/9

Hickox, R. C., Myers, A. D., Greene, J. E., et al. 2017, ApJ, 849, 53, doi: 10.3847/1538-4357/aa8c77

Hickox, R. C., Jones, C., Forman, W. R., et al. 2007, ApJ, 671, 1365, doi: 10.1086/523082

—. 2009, ApJ, 696, 891, doi: 10.1088/0004-637X/696/1/891

Hickox, R. C., Myers, A. D., Brodwin, M., et al. 2011, ApJ, 731, 117, doi: 10.1088/0004-637X/731/2/117

Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2, doi: 10.1086/338126

Hönig, S. F., Kishimoto, M., Tristram, K. R. W., et al. 2013, ApJ, 771, 87, doi: 10.1088/0004-637X/771/2/87

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2005, ApJ, 630, 705, doi: 10.1086/432438

—. 2006, ApJS, 163, 1, doi: 10.1086/499298

Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, ApJS, 175, 356, doi: 10.1086/524362

Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, ApJ, 654, 731, doi: 10.1086/509629

Hopkins, P. F., Strauss, M. A., Hall, P. B., et al. 2004, AJ, 128, 1112, doi: 10.1086/423291

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236, doi: 10.1088/0004-637X/690/2/1236

Ineson, J., Croston, J. H., Hardcastle, M. J., et al. 2015, MNRAS, 453, 2682, doi: 10.1093/mnras/stv1807

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Jarvis, M. 2015, TreeCorr: Two-point correlation functions, Astrophysics Source Code Library, record ascl:1508.007

Jiang, N., Wang, H., Mo, H., et al. 2016, ApJ, 832, 111, doi: 10.3847/0004-637X/832/2/111

Juneau, S., Goulding, A. D., Banfield, J., et al. 2022, ApJ, 925, 203, doi: 10.3847/1538-4357/ac425f

Kaiser, N. 1986, MNRAS, 222, 323, doi: 10.1093/mnras/222.2.323

—. 1987, MNRAS, 227, 1, doi: 10.1093/mnras/227.1.1

Kakkad, D., Sani, E., Rojas, A. F., et al. 2022, MNRAS, 511, 2105, doi: 10.1093/mnras/stac103

Karachentsev, I. D., Makarov, D. I., & Kaisina, E. I. 2013, AJ, 145, 101, doi: 10.1088/0004-6256/145/4/101

Kartaltepe, J. S., Sanders, D. B., Silverman, J. D., et al. 2015, ApJL, 806, L35, doi: 10.1088/2041-8205/806/2/L35

Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, ApJS, 241, 10, doi: 10.3847/1538-4365/ab06c4

Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739, doi: 10.1093/mnras/stx126

Kenter, A., Murray, S. S., Forman, W. R., et al. 2005, ApJS, 161, 9, doi: 10.1086/444379

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160, doi: 10.1111/j.1365-2966.2009.14541.x

Kim, D., & Im, M. 2018, A&A, 610, A31, doi: 10.1051/0004-6361/201731963

Klindt, L., Alexander, D. M., Rosario, D. J., Lusso, E., & Fotopoulou, S. 2019, MNRAS, 488, 3109, doi: 10.1093/mnras/stz1771

Kochanek, C. S., Eisenstein, D. J., Cool, R. J., et al. 2012, ApJS, 200, 8, doi: 10.1088/0067-0049/200/1/8

Komatsu, E., & Seljak, U. 2001, MNRAS, 327, 1353, doi: 10.1046/j.1365-8711.2001.04838.x

Kondapally, R., Best, P. N., Hardcastle, M. J., et al. 2021, A&A, 648, A3, doi: 10.1051/0004-6361/202038813

Kondapally, R., Best, P. N., Cochrane, R. K., et al. 2022, MNRAS, 513, 3742, doi: 10.1093/mnras/stac1128

Kondapally, R., Best, P. N., Raouf, M., et al. 2023, MNRAS, 523, 5292, doi: 10.1093/mnras/stad1813

Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511, doi: 10.1146/annurev-astro-082708-101811

Koutoulidis, L., Georgantopoulos, I., Mountrichas, G., et al. 2018, MNRAS, 481, 3063, doi: 10.1093/mnras/sty2429

Krawczyk, C. M., Richards, G. T., Gallagher, S. C., et al. 2015, AJ, 149, 203, doi: 10.1088/0004-6256/149/6/203

Krishnan, C., Almaini, O., Hatch, N. A., et al. 2020, MNRAS, 494, 1693, doi: 10.1093/mnras/staa815

Krolewski, A., Ferraro, S., Schlafly, E. F., & White, M. 2020, JCAP, 2020, 047, doi: 10.1088/1475-7516/2020/05/047

Krolewski, A., Percival, W. J., Ferraro, S., et al. 2023, arXiv e-prints, arXiv:2305.07650, doi: 10.48550/arXiv.2305.07650

Krumpe, M., Miyaji, T., Coil, A. L., & Aceves, H. 2018, MNRAS, 474, 1773, doi: 10.1093/mnras/stx2705

Lacy, M., Storrie-Lombardi, L. J., Sajina, A., et al. 2004, ApJS, 154, 166, doi: 10.1086/422816

Lacy, M., Ridgway, S. E., Gates, E. L., et al. 2013, ApJS, 208, 24, doi: 10.1088/0067-0049/208/2/24

Lacy, M., Mason, B., Sarazin, C., et al. 2019, MNRAS, 483, L22, doi: 10.1093/mnrasl/sly215

Laha, S., Reynolds, C. S., Reeves, J., et al. 2021, Nature Astronomy, 5, 13, doi: 10.1038/s41550-020-01255-2

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24, doi: 10.3847/0067-0049/224/2/24

Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64, doi: 10.1086/172900

Lanzuisi, G., Piconcelli, E., Fiore, F., et al. 2009, A&A, 498, 67, doi: 10.1051/0004-6361/200811282

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193, doi: 10.48550/arXiv.1110.3193

Laurent, P., Eftekharzadeh, S., Le Goff, J.-M., et al. 2017, JCAP, 2017, 017, doi: 10.1088/1475-7516/2017/07/017

Leauthaud, A., J. Benson, A., Civano, F., et al. 2015, MNRAS, 446, 1874, doi: 10.1093/mnras/stu2210

Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473, doi: 10.1086/309179

Li, J., Silverman, J. D., Ding, X., et al. 2021, ApJ, 918, 22, doi: 10.3847/1538-4357/ac06a8

Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70, doi: 10.1086/516589

Limber, D. N. 1953, ApJ, 117, 134, doi: 10.1086/145672

Lindsay, S. N., Jarvis, M. J., & McAlpine, K. 2014, MNRAS, 440, 2322, doi: 10.1093/mnras/stu453

Liu, G., Zakamska, N. L., Greene, J. E., Nesvadba, N. P. H., & Liu, X. 2013, MNRAS, 436, 2576, doi: 10.1093/mnras/stt1755

López-Gonzaga, N., Burtscher, L., Tristram, K. R. W., Meisenheimer, K., & Schartmann, M. 2016, A&A, 591, A47, doi: 10.1051/0004-6361/201527590

LoVerde, M., & Afshordi, N. 2008, PhRvD, 78, 123506, doi: 10.1103/PhysRevD.78.123506

Ludlow, A. D., Bose, S., Angulo, R. E., et al. 2016, MNRAS, 460, 1214, doi: 10.1093/mnras/stw1046

Luo, W., Silverman, J. D., More, S., et al. 2022, arXiv e-prints, arXiv:2204.03817. https://arxiv.org/abs/2204.03817

Lutz, D., Sturm, E., Janssen, A., et al. 2020, A&A, 633, A134, doi: 10.1051/0004-6361/201936803

Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8, doi: 10.3847/1538-4365/aba623

Lynden-Bell, D. 1969, Nature, 223, 690, doi: 10.1038/223690a0

Lynds, B. T. 1962, ApJS, 7, 1, doi: 10.1086/190072

—. 1965, ApJS, 12, 163, doi: 10.1086/190123

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

Madhavacheril, M., Sehgal, N., Allison, R., et al. 2015, PhRvL, 114, 151302, doi: 10.1103/PhysRevLett.114.151302

Madhavacheril, M. S., Sifón, C., Battaglia, N., et al. 2020, ApJL, 903, L13, doi: 10.3847/2041-8213/abbccb

Madhavacheril, M. S., Qu, F. J., Sherwin, B. D., et al. 2023, arXiv e-prints, arXiv:2304.05203, doi: 10.48550/arXiv.2304.05203

Magliocchetti, M. 2022, A&A Rv, 30, 6, doi: 10.1007/s00159-022-00142-1

Magliocchetti, M., Popesso, P., Brusa, M., et al. 2017, MNRAS, 464, 3271, doi: 10.1093/mnras/stw2541

Magliocchetti, M., Maddox, S. J., Hawkins, E., et al. 2004, MNRAS, 350, 1485, doi: 10.1111/j.1365-2966.2004.07751.x

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285, doi: 10.1086/300353

Mainzer, A., Bauer, J., Grav, T., et al. 2011, ApJ, 731, 53, doi: 10.1088/0004-637X/731/1/53

Maksimova, N. A., Garrison, L. H., Eisenstein, D. J., et al. 2021, MNRAS, 508, 4017, doi: 10.1093/mnras/stab2484

Mandelbaum, R., Li, C., Kauffmann, G., & White, S. D. M. 2009, MNRAS, 393, 377, doi: 10.1111/j.1365-2966.2008.14235.x

Marchesi, S., Civano, F., Elvis, M., et al. 2016, ApJ, 817, 34, doi: 10.3847/0004-637X/817/1/34

Marocco, F., Eisenhardt, P. R. M., Fowler, J. W., et al. 2021a, ApJS, 253, 8, doi: 10.3847/1538-4365/abd805

—. 2021b, ApJS, 253, 8, doi: 10.3847/1538-4365/abd805

Marocco et al. 2020a, CatWISE2020 Catalog, IPAC, doi: 10.26131/IRSA551

—. 2020b, CatWISE2020 Catalog, IPAC, doi: 10.26131/IRSA551

Martini, P., & Weinberg, D. H. 2001, ApJ, 547, 12, doi: 10.1086/318331

Massingill, K., Mason, B., Lacy, M., et al. 2023, arXiv e-prints, arXiv:2310.09445. https://arxiv.org/abs/2310.09445

McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, MNRAS, 406, 822, doi: 10.1111/j.1365-2966.2010.16750.x

McNamara, B. R., & Nulsen, P. E. J. 2007, ARA&A, 45, 117, doi: 10.1146/annurev.astro.45.051806.110625

—. 2012, New Journal of Physics, 14, 055023, doi: 10.1088/1367-2630/14/5/055023

McNamara, B. R., Wise, M., Nulsen, P. E. J., et al. 2000, ApJL, 534, L135, doi: 10.1086/312662

Mead, A. J., Brieden, S., Tröster, T., & Heymans, C. 2021, MNRAS, 502, 1401, doi: 10.1093/mnras/stab082

Mead, A. J., Peacock, J. A., Heymans, C., Joudaki, S., & Heavens, A. F. 2015, MNRAS, 454, 1958, doi: 10.1093/mnras/stv2036

Meisner, A. M., Lang, D., Schlafly, E. F., & Schlegel, D. J. 2019, PASP, 131, 124504, doi: 10.1088/1538-3873/ab3df4

Ménard, B., Scranton, R., Schmidt, S., et al. 2013, arXiv e-prints, arXiv:1303.4722, doi: 10.48550/arXiv.1303.4722

Mendez, A. J., Coil, A. L., Aird, J., et al. 2016, ApJ, 821, 55, doi: 10.3847/0004-637X/821/1/55

Merloni, A., Predehl, P., Becker, W., et al. 2012, arXiv e-prints, arXiv:1209.3114, doi: 10.48550/arXiv.1209.3114

Messias, H., Afonso, J., Salvato, M., Mobasher, B., & Hopkins, A. M. 2012, ApJ, 754, 120, doi: 10.1088/0004-637X/754/2/120

Mitra, K., Chatterjee, S., DiPompeo, M. A., Myers, A. D., & Zheng, Z. 2018, MNRAS, 477, 45, doi: 10.1093/mnras/sty556

Modi, C., White, M., & Vlah, Z. 2017, JCAP, 2017, 009, doi: 10.1088/1475-7516/2017/08/009

Morabito, L. K., Sweijen, F., Radcliffe, J. F., et al. 2022, MNRAS, 515, 5758, doi: 10.1093/mnras/stac2129

Mountrichas, G., & Georgakakis, A. 2012, MNRAS, 420, 514, doi: 10.1111/j.1365-2966.2011.20059.x

Moustakas, J., Lang, D., Dey, A., et al. 2023, ApJS, 269, 3, doi: 10.3847/1538-4365/acfaa2

Murray, S. G., Diemer, B., Chen, Z., et al. 2021, Astronomy and Computing, 36, 100487, doi: 10.1016/j.ascom.2021.100487

Myers, A. D., Brunner, R. J., Nichol, R. C., et al. 2007, ApJ, 658, 85, doi: 10.1086/511519

Myers, A. D., Brunner, R. J., Richards, G. T., et al. 2006, ApJ, 638, 622, doi: 10.1086/499093

Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, ApJS, 221, 27, doi: 10.1088/0067-0049/221/2/27

Myers, A. D., Moustakas, J., Bailey, S., et al. 2022, arXiv e-prints, arXiv:2208.08518. https://arxiv.org/abs/2208.08518

—. 2023, AJ, 165, 50, doi: 10.3847/1538-3881/aca5f9

Nakoneczny, S. J., Alonso, D., Bilicki, M., et al. 2023, arXiv e-prints, arXiv:2310.07642. https://arxiv.org/abs/2310.07642

Nandra, K., Barret, D., Barcons, X., et al. 2013, arXiv e-prints, arXiv:1306.2307, doi: 10.48550/arXiv.1306.2307

Narayan, R., & Yi, I. 1994, ApJL, 428, L13, doi: 10.1086/187381

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493, doi: 10.1086/304888

Netzer, H. 2015, ARA&A, 53, 365, doi: 10.1146/annurev-astro-082214-122302

Newman, J. A. 2008, ApJ, 684, 88, doi: 10.1086/589982

Nishimichi, T., Takada, M., Takahashi, R., et al. 2019, ApJ, 884, 29, doi: 10.3847/1538-4357/ab3719

Norberg, P., Baugh, C. M., Gaztañaga, E., & Croton, D. J. 2009, MNRAS, 396, 19, doi: 10.1111/j.1365-2966.2009.14389.x

Oei, M. S. S. L., van Weeren, R. J., Hardcastle, M. J., et al. 2022, A&A, 660, A2, doi: 10.1051/0004-6361/202142778

Oguri, M., & Hamana, T. 2011, MNRAS, 414, 1851, doi: 10.1111/j.1365-2966.2011.18481.x

Omori, Y., Chown, R., Simard, G., et al. 2017, ApJ, 849, 124, doi: 10.3847/1538-4357/aa8d1d

Omori, Y., Baxter, E. J., Chang, C., et al. 2023, PhRvD, 107, 023529, doi: 10.1103/PhysRevD.107.023529

Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, ApJ, 808, 161, doi: 10.1088/0004-637X/808/2/161

Padmanabhan, N., White, M., Norberg, P., & Porciani, C. 2009, MNRAS, 397, 1862, doi: 10.1111/j.1365-2966.2008.14071.x

Padovani, P., Alexander, D. M., Assef, R. J., et al. 2017, A&A Rv, 25, 2, doi: 10.1007/s00159-017-0102-9

Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJL, 698, L116, doi: 10.1088/0004-637X/698/2/L116

Peacock, J. A. 1991, MNRAS, 253, 1P, doi: 10.1093/mnras/253.1.1P

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2012, arXiv e-prints, arXiv:1201.0490. https://arxiv.org/abs/1201.0490

Peebles, P. J. E. 1973, ApJ, 185, 413, doi: 10.1086/152431

—. 1980, The large-scale structure of the universe

Perrotta, S., Hamann, F., Zakamska, N. L., et al. 2019, MNRAS, 488, 4126, doi: 10.1093/mnras/stz1993

Petter, G. C., Hickox, R. C., Alexander, D. M., et al. 2023, ApJ, 946, 27, doi: 10.3847/1538-4357/acb7ef

—. 2022, ApJ, 927, 16, doi: 10.3847/1538-4357/ac4d31

Piconcelli, E., Vignali, C., Bianchi, S., et al. 2015, A&A, 574, L9, doi: 10.1051/0004-6361/201425324

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020a, A&A, 641, A6, doi: 10.1051/0004-6361/201833910

—. 2020b, A&A, 641, A8, doi: 10.1051/0004-6361/201833886

Planck Collaboration, Akrami, Y., Ashdown, M., et al. 2020c, A&A, 641, A4, doi: 10.1051/0004-6361/201833881

Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81, doi: 10.1086/518113

Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010, doi: 10.1111/j.1365-2966.2004.08408.x

Porciani, C., & Norberg, P. 2006, MNRAS, 371, 1824, doi: 10.1111/j.1365-2966.2006.10813.x

Powell, M. C., Urry, C. M., Cappelluti, N., et al. 2020, ApJ, 891, 41, doi: 10.3847/1538-4357/ab6e65

Powell, M. C., Cappelluti, N., Urry, C. M., et al. 2018, ApJ, 858, 110, doi: 10.3847/1538-4357/aabd7f

Prada, F., Ereza, J., Smith, A., et al. 2023, arXiv e-prints, arXiv:2306.06315, doi: 10.48550/arXiv.2306.06315

Prescott, M. K. M., Impey, C. D., Cool, R. J., & Scoville, N. Z. 2006, ApJ, 644, 100, doi: 10.1086/503325

Raichoor, A., de Mattia, A., Ross, A. J., et al. 2021, MNRAS, 500, 3254, doi: 10.1093/mnras/staa3336

Ramos Almeida, C., & Ricci, C. 2017, Nature Astronomy, 1, 679, doi: 10.1038/s41550-017-0232-z

Rennehan, D., Babul, A., Moa, B., & Davé, R. 2023, arXiv e-prints, arXiv:2309.15898. https://arxiv.org/abs/2309.15898

Retana-Montenegro, E., & Röttgering, H. J. A. 2017, A&A, 600, A97, doi: 10.1051/0004-6361/201526433

Rezaie, M., Ross, A. J., Seo, H.-J., et al. 2021, MNRAS, 506, 3439, doi: 10.1093/mnras/stab1730

Ricci, C., Bauer, F. E., Treister, E., et al. 2017a, MNRAS, 468, 1273, doi: 10.1093/mnras/stx173

Ricci, C., Trakhtenbrot, B., Koss, M. J., et al. 2017b, Nature, 549, 488, doi: 10.1038/nature23906

Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, AJ, 123, 2945, doi: 10.1086/340187

Richards, G. T., Hall, P. B., Vanden Berk, D. E., et al. 2003, AJ, 126, 1131, doi: 10.1086/377014

Richardson, J., Zheng, Z., Chatterjee, S., Nagai, D., & Shen, Y. 2012, ApJ, 755, 30, doi: 10.1088/0004-637X/755/1/30

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009, doi: 10.1086/300499

Rosario, D. J., Alexander, D. M., Moldon, J., et al. 2021, MNRAS, 505, 5283, doi: 10.1093/mnras/stab1653

Rosario, D. J., Fawcett, V. A., Klindt, L., et al. 2020, MNRAS, 494, 3061, doi: 10.1093/mnras/staa866

Rose, M., Tadhunter, C. N., Holt, J., & Rodríguez Zaurín, J. 2013, MNRAS, 432, 2150, doi: 10.1093/mnras/stt564

Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, MNRAS, 498, 2354, doi: 10.1093/mnras/staa2416

Rupke, D. S. N., & Veilleux, S. 2011, ApJL, 729, L27, doi: 10.1088/2041-8205/729/2/L27

Sabater, J., Best, P. N., Hardcastle, M. J., et al. 2019, A&A, 622, A17, doi: 10.1051/0004-6361/201833883

Sabater, J., Best, P. N., Tasse, C., et al. 2021, A&A, 648, A2, doi: 10.1051/0004-6361/202038828

Salpeter, E. E. 1964, ApJ, 140, 796, doi: 10.1086/147973

Salvato, M., Hasinger, G., Ilbert, O., et al. 2009, ApJ, 690, 1250, doi: 10.1088/0004-637X/690/2/1250

Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, ApJ, 742, 61, doi: 10.1088/0004-637X/742/2/61

Sanders, D. B., Soifer, B. T., Elias, J. H., et al. 1988a, ApJ, 325, 74, doi: 10.1086/165983

Sanders, D. B., Soifer, B. T., Elias, J. H., Neugebauer, G., & Matthews, K. 1988b, ApJL, 328, L35, doi: 10.1086/185155

Schartmann, M., Wada, K., Prieto, M. A., Burkert, A., & Tristram, K. R. W. 2014, MNRAS, 445, 3878, doi: 10.1093/mnras/stu2020

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521, doi: 10.1093/mnras/stu2058

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: 10.1088/0004-637X/737/2/103

Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30, doi: 10.3847/1538-4365/aafbea

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525, doi: 10.1086/305772

Schmidt, M. 1963, Nature, 197, 1040, doi: 10.1038/1971040a0

Schulze, A., Silverman, J. D., Kashino, D., et al. 2018, ApJS, 239, 22, doi: 10.3847/1538-4365/aae82f

Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, ApJ, 546, 20, doi: 10.1086/318261

Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1, doi: 10.1086/516585

Seljak, U. 2000, MNRAS, 318, 203, doi: 10.1046/j.1365-8711.2000.03715.x

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Shen, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2020, MNRAS, 495, 3252, doi: 10.1093/mnras/staa1381

Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, AJ, 133, 2222, doi: 10.1086/513517

Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, ApJ, 697, 1656, doi: 10.1088/0004-637X/697/2/1656

Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119, doi: 10.1046/j.1365-8711.1999.02692.x

Shimwell, T. W., Röttgering, H. J. A., Best, P. N., et al. 2017, A&A, 598, A104, doi: 10.1051/0004-6361/201629313

Shimwell, T. W., Hardcastle, M. J., Tasse, C., et al. 2022, A&A, 659, A1, doi: 10.1051/0004-6361/202142484

Shirley, R., Roehlly, Y., Hurley, P. D., et al. 2019, MNRAS, 490, 634, doi: 10.1093/mnras/stz2509

Shirley, R., Duncan, K., Campos Varillas, M. C., et al. 2021, MNRAS, 507, 129, doi: 10.1093/mnras/stab1526

Siewert, T. M., Hale, C., Bhardwaj, N., et al. 2020, A&A, 643, A100, doi: 10.1051/0004-6361/201936592

Silk, J., & Rees, M. J. 1998, A&A, 331, L1, doi: 10.48550/arXiv.astro-ph/9801013

Simon, P. 2007, A&A, 473, 711, doi: 10.1051/0004-6361:20066352

Sinha, M., & Garrison, L. H. 2020, MNRAS, 491, 3022, doi: 10.1093/mnras/stz3157

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

Smith, D. J. B., Best, P. N., Duncan, K. J., et al. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin, 271–280

Smolčić, V., Novak, M., Bondi, M., et al. 2017, A&A, 602, A1, doi: 10.1051/0004-6361/201628704

Soergel, B., Giannantonio, T., Efstathiou, G., Puchwein, E., & Sijacki, D. 2017, MNRAS, 468, 577, doi: 10.1093/mnras/stx492

Soltan, A. 1982, MNRAS, 200, 115, doi: 10.1093/mnras/200.1.115

Somerville, R. S., & Davé, R. 2015, ARA&A, 53, 51, doi: 10.1146/annurev-astro-082812-140951

Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, MNRAS, 391, 481, doi: 10.1111/j.1365-2966.2008.13805.x

Song, Y.-S., Cooray, A., Knox, L., & Zaldarriaga, M. 2003, ApJ, 590, 664, doi: 10.1086/375188

Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776, doi: 10.1111/j.1365-2966.2005.09238.x

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676, doi: 10.1093/mnras/stx3304

Stern, D., Eisenhardt, P., Gorjian, V., et al. 2005, ApJ, 631, 163, doi: 10.1086/432523

Stern, D., Assef, R. J., Benford, D. J., et al. 2012, ApJ, 753, 30, doi: 10.1088/0004-637X/753/1/30

Stern, D., Lansbury, G. B., Assef, R. J., et al. 2014, ApJ, 794, 102, doi: 10.1088/0004-637X/794/2/102

Storey-Fisher, K., Hogg, D. W., Rix, H.-W., et al. 2024, ApJ, 964, 69, doi: 10.3847/1538-4357/ad1328

Sunyaev, R. A., & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

Sweijen, F., van Weeren, R. J., Röttgering, H. J. A., et al. 2022, Nature Astronomy, 6, 350, doi: 10.1038/s41550-021-01573-z

Swetz, D. S., Ade, P. A. R., Amiri, M., et al. 2011, ApJS, 194, 41, doi: 10.1088/0067-0049/194/2/41

Tabor, G., & Binney, J. 1993, MNRAS, 263, 323, doi: 10.1093/mnras/263.2.323

Talman, J. D. 1978, Journal of Computational Physics, 29, 35, doi: 10.1016/0021-9991(78)90107-9

Tasse, C., Shimwell, T., Hardcastle, M. J., et al. 2021, A&A, 648, A1, doi: 10.1051/0004-6361/202038804

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Thacker, R. J., Scannapieco, E., & Couchman, H. M. P. 2006, ApJ, 653, 86, doi: 10.1086/508650

Timlin, J. D., Ross, N. P., Richards, G. T., et al. 2018, ApJ, 859, 20, doi: 10.3847/1538-4357/aab9ac

Timmerman, R., van Weeren, R. J., Botteon, A., et al. 2022, A&A, 668, A65, doi: 10.1051/0004-6361/202243936

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709, doi: 10.1086/591439

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878, doi: 10.1088/0004-637X/724/2/878

Tiwari, P., Zhao, R., Zheng, J., et al. 2022, ApJ, 928, 38, doi: 10.3847/1538-4357/ac5748

Toba, Y., Nagao, T., Kajisawa, M., et al. 2017, ApJ, 835, 36, doi: 10.3847/1538-4357/835/1/36

Trump, J. R., Impey, C. D., Elvis, M., et al. 2009, ApJ, 696, 1195, doi: 10.1088/0004-637X/696/2/1195

Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, ApJ, 786, 104, doi: 10.1088/0004-637X/786/2/104

Urrutia, T., Lacy, M., & Becker, R. H. 2008, ApJ, 674, 80, doi: 10.1086/523959

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803, doi: 10.1086/133630

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2, doi: 10.1051/0004-6361/201220873

Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549, doi: 10.1086/321167

Veilleux, S., Cecil, G., & Bland-Hawthorn, J. 2005, ARA&A, 43, 769, doi: 10.1146/annurev.astro.43.072103.150610

Vito, F., Brandt, W. N., Stern, D., et al. 2018, MNRAS, 474, 4528, doi: 10.1093/mnras/stx3120

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518, doi: 10.1093/mnras/stu1536

Webster, R. L., Francis, P. J., Petersont, B. A., Drinkwater, M. J., & Masci, F. J. 1995, Nature, 375, 469, doi: 10.1038/375469a0

Whalen, K. E., Hickox, R. C., DiPompeo, M. A., Richards, G. T., & Myers, A. D. 2020, ApJ, 888, 71, doi: 10.3847/1538-4357/ab5d31

Whitaker, K. E., Pope, A., Cybulski, R., et al. 2017, ApJ, 850, 208, doi: 10.3847/1538-4357/aa94ce

White, R. L., Helfand, D. J., Becker, R. H., Glikman, E., & de Vries, W. 2007, ApJ, 654, 99, doi: 10.1086/507700

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52, doi: 10.1086/170483

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341, doi: 10.1093/mnras/183.3.341

Wilkes, B. J., Schmidt, G. D., Cutri, R. M., et al. 2002, ApJL, 564, L65, doi: 10.1086/338908

Williams, W. L., Hardcastle, M. J., Best, P. N., et al. 2019, A&A, 622, A2, doi: 10.1051/0004-6361/201833564

Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 1999, MNRAS, 309, 1017, doi: 10.1046/j.1365-8711.1999.02907.x

Woo, J.-H., Treu, T., Malkan, M. A., & Blandford, R. D. 2008, ApJ, 681, 925, doi: 10.1086/588804

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868, doi: 10.1088/0004-6256/140/6/1868

Wu, Q., & Shen, Y. 2022, ApJS, 263, 42, doi: 10.3847/1538-4365/ac9ead

Yan, W., Hickox, R. C., Chen, C.-T. J., et al. 2021, ApJ, 914, 83, doi: 10.3847/1538-4357/abfaa0

Yang, G., Brandt, W. N., Darvish, B., et al. 2018, MNRAS, 480, 1022, doi: 10.1093/mnras/sty1910

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513

Yuan, S., Garrison, L. H., Hadzhiyska, B., Bose, S., & Eisenstein, D. J. 2022, MNRAS, 510, 3301, doi: 10.1093/mnras/stab3355

Yuan, S., Zhang, H., Ross, A. J., et al. 2023, arXiv e-prints, arXiv:2306.06314, doi: 10.48550/arXiv.2306.06314

Zakamska, N. L., Hamann, F., Pâris, I., et al. 2016, MNRAS, 459, 3144, doi: 10.1093/mnras/stw718

Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59, doi: 10.1088/0004-637X/736/1/59

Zhao, C., Chuang, C.-H., Bautista, J., et al. 2021, MNRAS, 503, 1149, doi: 10.1093/mnras/stab510

Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760, doi: 10.1086/521074

Zhou, R., Ferraro, S., White, M., et al. 2023, JCAP, 2023, 097, doi: 10.1088/1475-7516/2023/11/097

Zonca, A., Singer, L., Lenz, D., et al. 2019, Journal of Open Source Software, 4, 1298, doi: 10.21105/joss.01298