



EIGER. VI. The Correlation Function, Host Halo Mass, and Duty Cycle of Luminous Quasars at $z \gtrsim 6$

Anna-Christina Eilers^{1,2}, Ruari Mackenzie³, Elia Pizzati⁴, Jorryt Matthee^{3,5}, Joseph F. Hennawi^{4,6}, Haowen Zhang⁷, Rongmon Bordoloi⁸, Daichi Kashino⁹, Simon J. Lilly³, Rohan P. Naidu^{2,12}, Robert A. Simcoe^{1,2}, Minghao Yue²,

Carlos S. Frenk¹⁰, John C. Helly¹⁰, Matthieu Schaller^{4,11}, and Joop Schaye⁴

¹ Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; eilers@mit.edu

² MIT Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³ Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, Zürich, 8093, Switzerland

⁴ Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

⁵ Institute of Science and Technology Austria (ISTA), Am Campus 1, 3400 Klosterneuburg, Austria

⁶ Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA

⁷ Department of Astronomy, University of Arizona, Tucson, AZ 85721, USA

⁸ Department of Physics, North Carolina State University, Raleigh, NC 27695, USA

⁹ National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

¹⁰ Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham, DH1 3LE, UK

¹¹ Lorentz Institute for Theoretical Physics, Leiden University, PO Box 9506, NL-2300 RA Leiden, The Netherlands

Received 2024 February 28; revised 2024 September 3; accepted 2024 September 3; published 2024 October 17

Abstract

We expect luminous ($M_{1450} \lesssim -26.5$) high-redshift quasars to trace the highest-density peaks in the early Universe. Here, we present observations of four $z \gtrsim 6$ quasar fields using JWST/NIRCam in the imaging and wide-field slitless spectroscopy mode and report a wide range in the number of detected [O III]-emitting galaxies in the quasars' environments, ranging between a density enhancement of $\delta \approx 65$ within a 2 cMpc radius—one of the largest protoclusters during the Epoch of Reionization discovered to date—to a density contrast consistent with zero, indicating the presence of a UV-luminous quasar in a region comparable to the average density of the Universe. By measuring the two-point cross-correlation function of quasars and their surrounding galaxies, as well as the galaxy autocorrelation function, we infer a correlation length of quasars at $\langle z \rangle = 6.25$ of $r_0^{\text{QQ}} = 22.0^{+3.0}_{-2.9}$ cMpc h^{-1} , while we obtain a correlation length of the [O III]-emitting galaxies of $r_0^{\text{GG}} = 4.1 \pm 0.3$ cMpc h^{-1} . By comparing the correlation functions to dark-matter-only simulations we estimate the minimum mass of the quasars' host dark matter halos to be $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.43^{+0.13}_{-0.15}$ (and $\log_{10}(M_{\text{halo,min}}^{\text{[OIII]}}/M_{\odot}) = 10.56^{+0.05}_{-0.03}$ for the [O III] emitters), indicating that (a) luminous quasars do not necessarily reside within the most overdense regions in the early Universe, and that (b) the UV-luminous duty cycle of quasar activity at these redshifts is $f_{\text{duty}} \ll 1$. Such short quasar activity timescales challenge our understanding of early supermassive black hole growth and provide evidence for highly dust-obscured growth phases or episodic, radiatively inefficient accretion rates.

Unified Astronomy Thesaurus concepts: Quasars (1319); Supermassive black holes (1663); Clustering (1908); Astrostatistics techniques (1886); Galaxy dark matter halos (1880); High-redshift galaxy clusters (2007); Large-scale structure of the universe (902)

1. Introduction

The existence of supermassive black holes (SMBHs) with $\sim 10^9 - 10^{10} M_{\odot}$ harbored in the center of luminous quasars at redshifts $z \gtrsim 6$ when the Universe was less than a billion years old, poses significant challenges to our understanding of black hole growth (e.g., C. Mazzucchelli et al. 2017b; E. Bañados et al. 2018; J. Yang et al. 2020; F. Wang et al. 2021; M. Yue et al. 2024; L. Bigwood et al. 2024). In the standard picture of SMBH growth, the rest mass energy of the accreted matter is divided between radiation and black hole growth via the so-called radiative efficiency ϵ , implying that the emission of quasar light is concurrent with the growth of the black hole.

Every massive galaxy is thought to have undergone a luminous quasar phase, and thus high-redshift quasars represent the progenitors of the dormant SMBH population found in the centers of nearly all local bulge-dominated galaxies (T. A. Small & R. D. Blandford 1992; Q. Yu & S. Tremaine 2002). This fundamental link between high-redshift quasars and local galaxies is supported by the tight correlations between the masses of SMBHs and various galaxy properties, such as the stellar velocity dispersion or bulge mass (e.g., J. Magorrian et al. 1998; K. Gebhardt et al. 2000; L. Ferrarese & D. Merritt 2000; N. Häring & H.-W. Rix 2004).

In the current paradigm, black holes grow exponentially starting from an initial black hole seed M_{seed} during the lifetime t_{Q} of the quasar, which is the time that galaxies spent in the luminous quasar state, i.e.,

$$M_{\text{BH}}(t_{\text{Q}}) = M_{\text{seed}} \exp(t_{\text{Q}}/t_{\text{S}}). \quad (1)$$

The growth occurs on a characteristic timescale t_{S} known as the e -folding time, or “Salpeter” time (E. E. Salpeter 1964),

¹² NHFP Hubble Fellow.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

which is approximately

$$t_S \approx 0.45 \left(\frac{\epsilon}{1 - \epsilon} \right) \left(\frac{L}{L_{\text{Edd}}} \right)^{-1} \text{Gyr}, \quad (2)$$

where L and L_{Edd} denote the bolometric and Eddington luminosity of the quasar, respectively. Assuming a thin accretion disk around the black hole, the radiative efficiency ϵ ranges between $\sim 6\%$ and 34% , the exact value depending on the spin of the black hole (N. I. Shakura & R. A. Sunyaev 1973; K. S. Thorne 1974).

Thus, long timescales of quasar activity—comparable to the Hubble time t_H —are required with nearly continuous Eddington-limited accretion, in order to explain the growth of the observed SMBHs at high redshift (e.g., T. Tanaka & Z. Haiman 2009; M. C. Begelman & M. Volonteri 2017). It follows that the fraction of cosmic time that galaxies shine as luminous quasars, known as the duty cycle $f_{\text{duty}} = t_Q/t_H(z)$, is expected to be approximately unity for quasars in the early Universe.

This theoretical picture of SMBH formation predicts that early SMBHs form and grow in the centers of rare and massive dark matter halos, which constitute the highest-density peaks of the cosmological density field (G. Efstathiou & M. J. Rees 1988; S. Cole & N. Kaiser 1989; A. Nusser & J. Silk 1993; S. G. Djorgovski et al. 2003; M. Volonteri & M. J. Rees 2006; D. Sijacki et al. 2009; T. Costa et al. 2014; M. Trebitsch et al. 2019; T. Costa 2024). Thus, luminous quasars are expected to reside in massive overdensities at early cosmic times, which implies that one should find a high number of companion galaxies in their vicinity according to our Λ CDM cosmological model (e.g., J. A. Munoz & A. Loeb 2008; J. L. Tinker et al. 2010). The regions around some high-redshift quasars are predicted to constitute proto-clusters, progenitors of the galaxy clusters that populate the local Universe, albeit the growth of these structures across cosmic time sensitively depends on a variety of environmental factors (e.g., R. E. Angulo et al. 2012).

However, to date, the search for the expected abundance of companion galaxies around $z \gtrsim 6$ quasars has led to inconclusive results. While some studies find strong galaxy enhancements in quasar fields using mostly narrowband filters to capture the Ly α emission of galaxies in the quasars' environments (e.g., S. Kim et al. 2009; L. Morselli et al. 2014), others do not (e.g., E. Bañados et al. 2013; C. Simpson et al. 2014; C. Mazzucchelli et al. 2017a). The first, recently published observations with JWST/NIRCam WFSS of two high-redshift quasar fields revealed an abundance of [O III]-emitting galaxies in the quasars' environment, indicating that these particular quasars lie indeed in highly overdense regions in the Universe (D. Kashino et al. 2023; F. Wang et al. 2023). It is unclear whether the conflicting results arise from the limited sensitivity, velocity offsets between the narrowband filter and the quasars' redshifts, or a too small field of view of some of the observations (e.g., R. A. Overzier 2016), or alternatively, at least some quasars may inhabit less massive hosts at $z \gtrsim 6$ than expected.

The mass of the dark matter halos hosting these luminous quasars can be inferred by means of their clustering properties. The two-point correlation function is a powerful tool to probe the connection between the observed light and host dark matter halos (e.g., P. S. Osmer 1981; T. Shanks & B. J. Boyle 1994;

S. M. Croom et al. 2001, 2005; Y. Shen et al. 2007). At low and intermediate redshifts of $1 \lesssim z \lesssim 4$ the two-point auto-correlation function has been studied using several thousands of quasars distributed over large areas on the sky. Across this wide redshift range, most quasar clustering studies find similar masses of the dark matter halos that host luminous quasars of a few times $10^{12} M_\odot$ (e.g., C. Porciani et al. 2004; S. M. Croom et al. 2005; A. L. Coil et al. 2007; A. D. Myers et al. 2007; Y. Shen et al. 2007; N. Padmanabhan et al. 2009; N. P. Ross et al. 2009; R. Trainor & C. C. Steidel 2013; S. Eftekharzadeh et al. 2015; W. He et al. 2018). This implies that quasars at intermediate redshifts are highly biased tracers of the underlying density field and reside within some of the most massive dark matter halos at these cosmic times. At $z \sim 1$, however, the dark matter halo mass is close to the characteristic mass in the Press–Schechter mass function (W. H. Press & P. Schechter 1974), indicating that quasars in the local Universe are largely unbiased (S. M. Croom et al. 2005). Thus, we observe an overall increase in the quasars' bias factor with redshift, i.e., the ratio of the host halo mass for a typical quasar to the mean halo mass at the same epoch rises with increasing redshift.

In turn, this implies an increase in the quasars' duty cycle from $\lesssim 1\%$ at low redshifts, i.e., $z \lesssim 1$ (S. M. Croom et al. 2005; N. Padmanabhan et al. 2009; P. Laurent et al. 2017), to a few tens of percent at $z \sim 4$ (Y. Shen et al. 2007; E. Pizzati et al. 2024; but see S. Eftekharzadeh et al. (2015) who report a lower duty cycle at similar redshifts). Thus, the quasar activity timescales at these redshifts are expected to be relatively long, i.e., $t_Q > 10^8$ yr, therefore providing sufficient time to grow the observed SMBHs.

At redshifts $z > 6$, however, this approach to determine the host dark matter halo masses by measuring the quasars' autocorrelation function becomes unfeasible as the space density of luminous ($M_{1450} < -27$) quasars drops dramatically to less than 1 Gpc^{-3} (F. Wang et al. 2019; J.-T. Schindler et al. 2023). Using a much fainter ($M_{1450} \gtrsim -25$), but more numerous population of quasars, a first attempt of the autocorrelation measurement at $z \sim 6$ has recently been reported by the Subaru High- z Exploration of Low-luminosity Quasars (SHELLQs) Collaboration (J. Arita et al. 2023), indicating host dark matter halo masses similar or slightly higher than those found at lower redshifts, but uncertainties are large. However, while autocorrelation measurements of quasars at these high redshifts remain challenging, the complementary quasar–galaxy cross-correlation measurement is still viable. Assuming that both quasars and galaxies trace the same underlying dark matter density distribution, their respective autocorrelation functions determine the cross-correlation function between these two classes of objects.

During the last year, JWST has revolutionized our ability to study the early Universe, enabling us to reveal and spectroscopically confirm faint galaxies down to $m_{\text{AB}} \sim 29$ mag within the environment of high-redshift quasars. Using NIRCam in imaging and WFSS mode, JWST recently obtained large mosaic observations in five quasar fields at $6 \lesssim z \lesssim 7$ as part of the Emission line galaxies and Intergalactic Gas in the Epoch of Reionization (EIGER) Collaboration (D. Kashino et al. 2023). The filter for the *grism* spectra was tuned to spectroscopically identify galaxies in the quasar fields by means of their [O III] emission line doublet $\lambda\lambda 4960, 5008$. Using this data set, we present in this paper for the first time the quasar–galaxy cross-correlation measurement at $z \gtrsim 6$. Combined with the

autocorrelation of the [O III]-emitting galaxies, we infer the clustering properties of these high-redshift luminous quasars, and determine their host dark matter halo mass as well as their duty cycle.

In Section 2, we will first describe the JWST/NIRCam observations of the high-redshift quasar fields analyzed in this study. We will briefly comment in Section 3 on the large diversity in the observed abundance of galaxies that we observe in the environments of the quasars, before measuring the two-point cross-correlation function of the quasars and their surrounding galaxies, as well as the galaxy–galaxy autocorrelation function in Section 4. In Section 5, we infer the characteristic mass of the quasars’ host dark matter halos based on their clustering properties and estimate their duty cycle in Section 6. We discuss the implications of our results in Section 7 before summarizing in Section 8.

Throughout this work, we adopt a flat Λ CDM cosmology with $H_0 = 67.66 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3097$, and $\Omega_\Lambda = 0.6889$ (Planck Collaboration et al. 2020).

2. JWST/NIRCam Observations of High-redshift Quasar Fields

The EIGER Collaboration is observing a total of six $z \gtrsim 6$ quasar fields utilizing NIRCam in WFSS and the imaging mode (ID: #1243). Of these six fields, the observations of five quasar fields were completed at the time of writing, and their analysis is presented in this work. Four out of five fields are used to study the quasar–galaxy clustering (as in one field our setup does not cover the [O III] emission at the quasar’s redshift), while all five quasar fields are used when measuring the galaxy autocorrelation (see Section 4).

2.1. Data Reduction

An overview of the observing program as well as a detailed description of the data reduction can be found in D. Kashino et al. (2023) and J. Matthee et al. (2023). We will briefly summarize the setup of the program and the main steps of the data reduction here.

The NIRCam imaging and WFSS observations of the five quasar fields were conducted between 2022 August 22 and 2023 June 2. The observations consist of four visits per quasar field forming a 2×2 overlapping mosaic pattern, which provides an area of approximately $3' \times 6'$ per field centered on the quasar. The central $40'' \times 40''$ region around the quasar is covered by all four mosaic tiles.

The WFSS observations use the *grism* R in the long-wavelength (LW) channel combined with the F356W filter. This yields spectra covering the wavelengths between $3.14.0 \mu\text{m}$ dispersed along the detector rows with a spectral resolution of $R \approx 1500$. The two NIRCam modules A and B disperse the spectra in opposite directions. These reversed *grism* spectra facilitate significantly the identification of the source objects in the direct images, and enable us to remove contaminating lines from other sources.

Our observations employ a three-point primary dither using INTRAMODULEX and a four-point subpixel dither, which yields a total of 12 exposures per visit and module for the direct images in the two short-wavelength (SW) filters (F115W and F200W), and 24 WFSS images. By co-adding the multiple exposures, the total exposure time in a single visit amounts to 4380 s per filter for the SW imaging and 8760 s for WFSS. We

obtained additional direct imaging in the LW channel F356W filter after the WFSS observations, with F200W in the SW channel. Given the small-scale dithering and larger-scale mosaic pattern, the total exposure time varies across the field, ranging from 1.5 ks at the edges sampled by only a single primary dither up to 13 ks in the center of each field for SW imaging, and from 2.9 ks up to 35 ks for WFSS (D. Kashino et al. 2023).

The WFSS data were reduced with a combination of the `jwst` pipeline (version 1.9.4) and additional custom processing steps. We first process the raw exposure files with the `Detector1` step from the pipeline and use the `Spec2` step to assign the astrometric solution for each image. The data are flat-fielded by applying the `Image2` step and sky background variations as well as $1/f$ noise are removed using the median counts in each column. We construct an emission line image by applying a running median filter along each row to remove any continuum emission (see D. Kashino et al. 2023, for details).

2.2. Identification of [O III]-emitting Galaxies

The spectral range covered by the filter F356W of $3.1 \leq \lambda \leq 4.0 \mu\text{m}$ enables us to identify [O III] emitters within the redshift range $5.33 \lesssim z \lesssim 6.96$. This implies that we can study the [O III]-emitting galaxies in the environment of four of the targeted quasars, while for one quasar, J1120+0641 at $z = 7.0848$, the [O III] emission lines lie outside of the observed redshift range. The galaxies observed in this field, however, will be included in the measurement of the galaxy–galaxy autocorrelation function (Section 4.2).

We employ two complementary methods to identify [O III]-emitting galaxies from the emission line images. The first approach starts with the co-added emission line image and searches for combinations of emission lines that could plausibly arise from doublets ([O III] $\lambda\lambda$ 4960, 5008) or triplets ([O III] $\lambda\lambda$ 4960, 5008 + H β λ 4861) using `SExtractor`. To evaluate the validity of these detections, we then try to identify a source galaxy in the F356W direct image using the *grism* model, which enables us to estimate the location on the sky for a given position in the *grism* image.

As a second method to identify [O III] emitters, we employ an inverse approach and use the catalog of sources identified in the F356W images as a starting point. For each source, we extract a 2D spectrum from the median-filtered *grism* images and run `SExtractor` to identify emission lines. [O III] emitters are identified if emission line doublets or triplets are detected at $\geq 3\sigma$ close to the expected locus of the spectral trace of the object. The list of detected [O III] emitters from both approaches are afterward combined and cross-validated. We estimate the systematic uncertainty on the derived redshifts to be about 100 km s^{-1} due to the inaccuracy of the *grism* model (D. Kashino et al. 2023).

Note that our identification algorithm for [O III] emitters has been fine-tuned and improved since the publication of the first results from the EIGER Collaboration (D. Kashino et al. 2023; J. Matthee et al. 2023). This results in an increased number of discovered [O III]-emitting galaxies. Within the field around quasar J0100+2802 we now discover a total of 180 [O III]-emitting galaxy systems (213 individual galaxies; see Section 2.3 for details on how we define a group) compared to the previously reported 117. Our total sample of individual [O III] emitters in all five quasar fields amounts to 861 (D. Kashino et al. 2024, in preparation).

Table 1
Overview of the Number of [O III]-emitting Galaxies in the Quasar Fields

Quasar Field	R.A. (hms)	Decl. (dms)	z	M_{1450}	$\log_{10}(M_{\text{BH}}/M_{\odot})$	N_{tot} ($\geq 10^{42}$ erg s $^{-1}$)	N_{gal} ($ \Delta v \leq v_{\text{max}}$)	V (cMpc 3)	δ ($r < 2$ cMpc)
J0148+0600	01:48:37.64	+06:00:20.06	5.99 ^a	−26.99	9.89 ^{+0.05} _{−0.06}	113	47	2469	65 ± 15
J1030+0524	10:30:27.10	+05:24:55.00	6.308 ^b	−27.39	9.19 ± 0.01	112	2	2631	3 ± 4
J0100+2802	01:00:13.02	+28:02:25.84	6.3270 ^a	−27.62	10.06 ± 0.01	121	24	2554	29 ± 10
J1148+5251	11:48:16.64	+52:51:50.30	6.4189 ^a	−29.14	9.64 ± 0.01	110	19	2650	16 ± 8
J1120+0641	11:20:01.48	+06:41:24.30	7.0848 ^a	−26.63	9.08 ± 0.03	66

Note. The columns denote the quasar field, the quasars’ coordinates R.A. and decl., the quasars’ redshifts (determined by means of *a*: [C II] or *b*: Mg II emission lines), and absolute magnitude at 1450 Å in the rest frame, the masses of the quasars’ SMBHs determined via the H β emission line (M. Yue et al. 2024) without any systematic uncertainties of ~ 0.5 dex, the number of [O III]-emitting galaxy systems above a luminosity limit of $L_{[\text{O III}],5008} = 10^{42}$ erg s $^{-1}$ detected in the whole quasar field, the number of these galaxies within $|\Delta v| \leq 1000$ km s $^{-1}$ of the quasar’s redshift, the total observed survey volume V around the quasar within ± 1000 km s $^{-1}$, and the observed overdensity δ within a radius of 2 cMpc around the quasar.

2.3. Groups of [O III]-emitting Galaxies

When analyzing the distribution of [O III]-emitting galaxy pairs as a function of their angular separation, we find a significant excess in the number of “clumps,” i.e., pairs or multiples, on very small scales at separations $\lesssim 2''$, which corresponds to approximately 12 kpc at $z = 6$ (see Figure 4 in J. Matthee et al. 2023). All of these [O III]-emitting systems within $2''$ separation are similarly closely separated in redshift ($|\Delta v| \lesssim 600$ km s $^{-1}$), suggesting that these components are physically associated with each other. Consequently, we merge all [O III] emitters within an angular separation $< 2''$ and $|\Delta v| < 1000$ km s $^{-1}$ together into a single system. The motivation for this is to mitigate uncertainties in the identification of individual objects that can arise due to the choice of deblending parameters used in SExtractor.

This merges about $\sim 13\%$ of the detected individual [O III] emitters and results in a total of 750 [O III]-emitting systems discovered in the five quasar fields. Since the completeness of our detections drops significantly below a luminosity limit of $L_{[\text{O III}],5008} < 10^{42}$ erg s $^{-1}$ (see Section 2.4), we restrict ourselves in this study to the 533 systems more luminous than this threshold (see Table 1). Furthermore, we restrict our analysis to systems that have been found using the second method to identify [O III] emitters starting from the source catalog identified in the F356W image (see Section 2.2), since the completeness of our observations is assessed using this method (see Section 2.4). This step only removes 11 objects, with a total of 522 systems remaining for our analysis.

Figure 1 shows our observations for four of the quasar fields, for which the quasars’ environment can be studied with our setup. For each field, we show in the top panels the NIRCcam images (F115W, F200W, and F356W) indicating the location of the quasar in the center as well as all [O III]-emitting galaxies within the quasar’s environment, i.e., within $|\Delta v| \leq 1000$ km s $^{-1}$ from the quasar’s systemic redshift. The bottom panels show the quasar spectra observed with the ground-based spectrographs FIRE (R. A. Simcoe et al. 2008) on the Magellan Telescopes, X-Shooter (J. Vernet et al. 2011) on the Very Large Telescope (VLT), and MOSFIRE (I. S. McLean et al. 2012) and the Echellette Spectrograph and Imager (ESI; A. I. Sheinis et al. 2002) on the Keck Telescopes (see A.-C. Eilers et al. 2023; D. Ďurovčíková et al. 2024, for details on the ground-based quasar spectra), as well as all detected [O III]-emitting systems in the quasar field as a function of the redshift and projected distance from the quasar.

2.4. Modeling the Completeness of Our Observations and Creating Random Galaxy Catalogs

In order to measure the clustering in the quasar fields as well as the excess in galaxies in the quasars’ environments compared to a blank field, we need to construct a set of “random” galaxy catalogs. These catalogs contain a large number of sources with similar properties as the observed galaxies but are randomly distributed across the survey volume.

Our selection function is not trivial due to the four-point mosaic, the different sensitivities of NIRCcam’s modules A and B, as well as the wavelength dependence of the noise over our field of view. To this end, we inject random sources into our survey volume and forward-model their spectra through our survey coverage and completeness. Their [O III] luminosities are randomly drawn from the [O III] luminosity distribution function (J. Matthee et al. 2023, with updated model parameters from R. Mackenzie et al. 2024, in preparation) and a minimum luminosity of $L_{[\text{O III}],5008} = 10^{42}$ erg s $^{-1}$ in the [O III] emission line at 5008 Å. Note that the [O III] luminosity function is estimated over the entire redshift range, but masking all galaxies at the quasars’ redshifts, such that no biases due to potential overdensities in the quasars’ environments arise. The spatial positions and redshifts are randomly chosen within the observed survey volume. We then check the spatial and spectral coverage of this source in our survey and ensure that both emission lines of the [O III] doublet, i.e., at 4960 and 5008 Å, would have been observed in our setup. Since both our methods to identify [O III]-emitting galaxies are based on the observability of the line doublet (see Section 2.2), we would miss all sources in our catalog for which only a single line would be observed.

To model the completeness of our line-selected sample into the random catalogs, we apply the completeness models that will be described in a forthcoming companion paper (R. Mackenzie et al. 2024, in preparation). In short, each field has an individual model, which consists of a noise cube and a completeness interpolating function. The noise cube represents the typical propagated error in a 2D stacked spectrum as a function of the position and wavelength. The function is constructed by injecting millions of mock sources into the 2D spectra and rerunning the [O III] emitter identification step. Thus, the measured completeness is not a step function at some limiting flux, but instead is a smooth function that we interpolate from the injection experiments. This function then translates the flux of the weaker of the two [O III] emission lines, i.e., 4960 Å, divided by the local error, to a completeness fraction. Dividing by the error essentially rescales the flux such that the

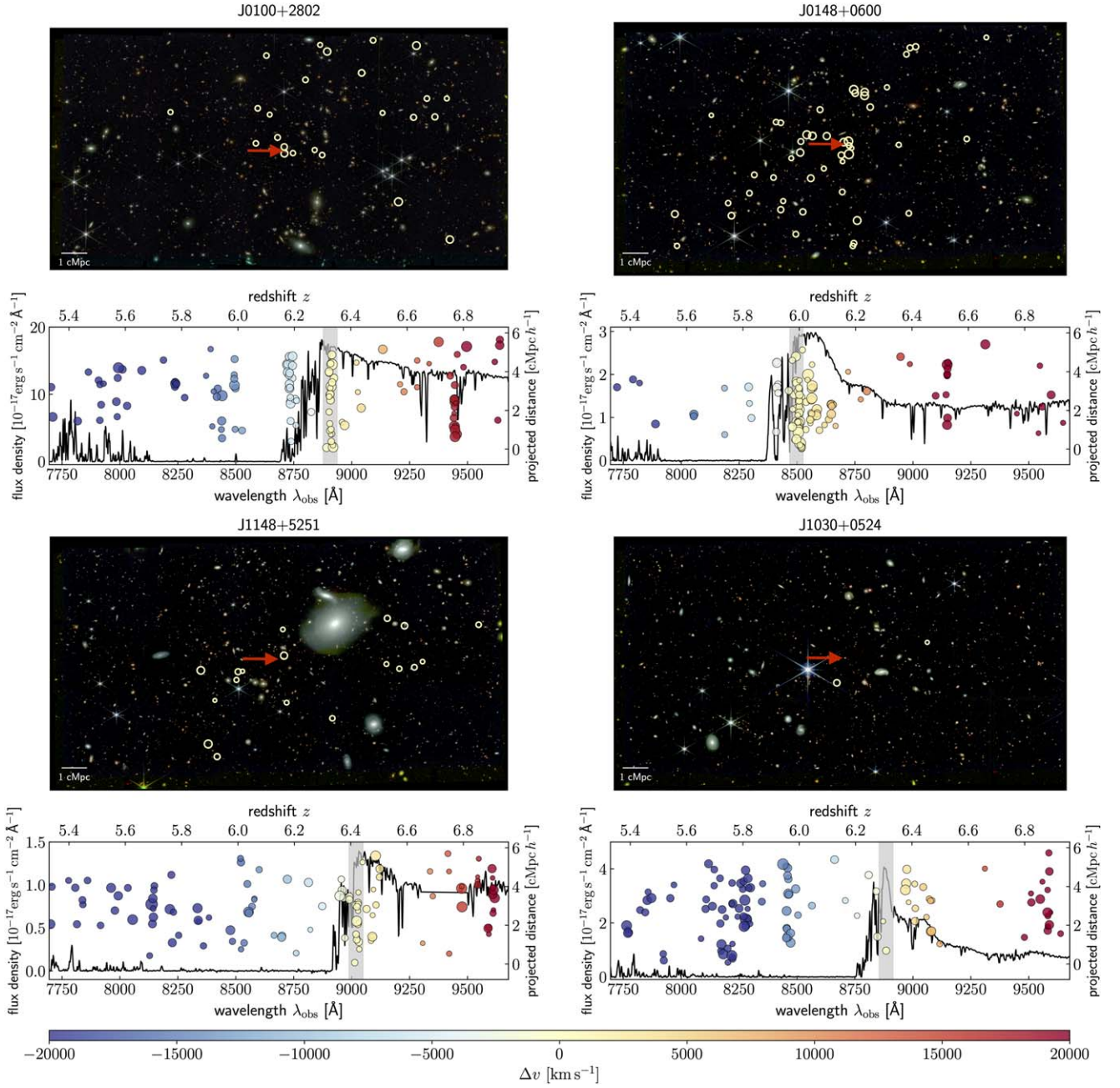


Figure 1. Observations of four $z \gtrsim 6$ quasar fields. The top panels show the RGB images constructed from NIRCcam imaging data (F115W, F200W, and F356W) of each quasar field. The bottom panels show the optical and near-IR spectra of the quasars observed with X-Shooter/VLT and FIRE/Magellan (J0100+2802, J0148+0600, J1030+0524), or MOSFIRE+ESI/Keck (J1148+5251). The colored data points show the redshifts and projected distances of all [O III]-emitting galaxies detected in the NIRCcam *grism* spectra above a luminosity of $L_{[\text{O III}],5008} = 10^{42} \text{ erg s}^{-1}$; their color corresponds to their velocity offset with respect to the quasars' redshifts and their size correlates with the logarithm of their luminosity, $\log_{10} L_{[\text{O III}],5008}$. The gray shaded area denotes the velocity interval of $|\Delta v| \leq v_{\text{max}}$ around the redshift of the quasar. All galaxies that fall within this velocity range are also circled in the respective RGB images to show the spatial distribution around the quasar (indicated by a red arrow).

completeness function is independent of the local sensitivity or redshift. To create the random galaxy catalogs we calculate these completeness maps of each EIGER field, and then draw random samples using the completeness as a probability threshold. Since our observations become highly incomplete for sources below a luminosity of $L_{[\text{O III}],5008} = 10^{42} \text{ erg s}^{-1}$, we only consider galaxies above this luminosity limit for our analysis, which corresponds to a line flux of for the weaker [O III] emission line of $f_{[\text{O III}],4960} = 0.7 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ at $z \approx 6.3$.

We estimate an average completeness of $\sim 41\%$ for galaxies with a luminosity of $L_{[\text{O III}],5008} = 10^{42} \text{ erg s}^{-1}$ at $6.0 < z < 6.4$, which increases to $\sim 98\%$ for $L_{[\text{O III}],5008} = 10^{43} \text{ erg s}^{-1}$. Note that these completeness estimates are averaged over the five quasar fields and the full spatial coverage of each mosaic, for regions closer to the quasar the completeness is higher. For more details on the completeness evaluation of our observations we refer the reader to (J. Matthee et al. 2023, R. Mackenzie et al. 2024, in preparation).

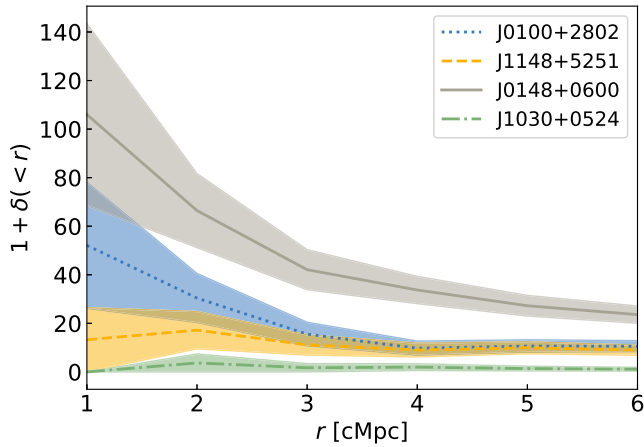


Figure 2. Density contrast of [O III] emitters around the four $z \gtrsim 6$ quasars analyzed in this work. The overdensity is measured in cylinders around the quasars with radius r and a depth of $|\Delta v| = 1000 \text{ km s}^{-1}$ around the quasars’ redshifts.

3. Large Diversity in Quasar Environments

Current SMBH formation models predict that luminous quasars in the high-redshift Universe trace the highest-density peaks and thus reside in a significantly overdense environment (e.g., D. Sijacki et al. 2009). To this end we estimate the density contrast, i.e., $\delta = N_{\text{gal}}/N_{\text{exp}} - 1$, by comparing the number of galaxies detected in close vicinity to the quasars N_{gal} , to the number of galaxies one would expect within the same observed volume N_{exp} in a blank “random” part of the Universe. In this case, N_{gal} denotes simply the number counts of [O III]-emitting galaxy systems with $|\Delta v| \leq 1000 \text{ km s}^{-1}$ from the quasars’ systemic redshift and within different radii. The expected number of objects N_{exp} is estimated by multiplying the galaxy density n_{gal} inferred from the [O III] luminosity function (J. Matthee et al. 2023, R. Mackenzie et al. 2024, in preparation) by the completeness map derived in Section 2.4 and by the observed volume V around the quasar, i.e., a cylinder with a given radius and depth of $\pm 1000 \text{ km s}^{-1}$. We estimate the uncertainties on the density contrast from Poisson noise on the galaxy counts.

Our results are shown as a function of the radius r for the four quasar fields in Figure 2 and suggest a strong diversity among the observed quasar fields. We find significant overdensities around three of the four observed quasars. In close vicinity of the quasars within $r < 2 \text{ cMpc}$, we find density enhancements of $\delta(r < 2 \text{ cMpc}) = 29 \pm 10$ and $\delta(r < 2 \text{ cMpc}) = 16 \pm 8$ around the quasars J0100+2802 and J1148+5251, respectively. With a total of $N_{\text{gal}} = 47$ galaxies observed in the complete observed volume around the quasar J0148+0600, this quasar resides in a spectacular protocluster within the Epoch of Reionization with an overdensity of $\delta(r < 2 \text{ cMpc}) = 65 \pm 15$.

Interestingly, one of the targeted quasar fields around J1030+0542 shows only two [O III]-emitting galaxies within the $|\Delta v| \leq 1000 \text{ km s}^{-1}$ velocity interval from the quasar and therefore no galaxy enhancement compared to a blank field, i.e., $\delta(r < 2 \text{ cMpc}) = 3 \pm 4$. We will discuss the possible implications of the highly diverse quasar environments in Section 7. All measurements and galaxy counts across the whole observed volume, as well as the density enhancements at $r < 2 \text{ cMpc}$, are listed in Table 1.

4. The Clustering Properties of High-redshift Quasars

In this section, we present measurements of the cross-correlation function between the quasars and the [O III]-emitting galaxies in their environment (Section 4.1), as well as of the autocorrelation function of [O III] emitters outside of the quasars’ environments (Section 4.2). By jointly fitting these two correlation functions we will infer the clustering properties of the luminous quasars themselves (Sections 4.3 and 4.4).

4.1. The Quasar–Galaxy Cross-correlation Function

We first measure the volume-averaged projected cross-correlation function between the quasars and the surrounding [O III]-emitting galaxies, i.e., $\chi_{\text{QG}}(R_{\text{min}}, R_{\text{max}})$, which is a dimensionless quantity defined as the real-space quasar–galaxy cross-correlation function $\xi_{\text{QG}}(R, Z)$ integrated over and normalized by a comoving volume V (e.g., J. F. Hennawi et al. 2006; C. García-Vergara et al. 2017), i.e.,

$$\chi_{\text{QG}}(R_{\text{min}}, R_{\text{max}}) = \frac{2}{V} \int_{R_{\text{min}}}^{R_{\text{max}}} \int_0^{Z_{\text{max}}} \xi_{\text{QG}}(R, Z) 2\pi R \, dR \, dZ. \quad (3)$$

We consider galaxies with a maximum velocity difference from the central quasar of $|\Delta v| \leq v_{\text{max}}$ with $v_{\text{max}} = 1000 \text{ km s}^{-1}$, and thus $Z_{\text{max}} = v_{\text{max}}(1+z)/H(z)$, where $H(z)$ denotes the Hubble parameter at redshift z . The integrated volume V is a cylindrical shell with inner and outer radii R_{min} and R_{max} , respectively, as well as a depth of Z_{max} . We integrate over the redshift range, in order to avoid the need to model redshift-space distortions and redshift uncertainties.

Following previous work (e.g., C. García-Vergara et al. 2017) we choose the estimator

$$\chi_{\text{QG}}(R_{\text{min}}, R_{\text{max}}) = \frac{\langle \text{QG} \rangle}{\langle \text{QR} \rangle} - 1, \quad (4)$$

where $\langle \text{QG} \rangle$ denotes the number of *detected* galaxies in radial bins around the quasar, while $\langle \text{QR} \rangle$ describes the number of *expected* galaxies within the same volume in a random “blank” field. Thus, in order to infer $\langle \text{QR} \rangle$, we calculate the number of pairs between the quasar and galaxies from our random catalog within the same radial bins and $|\Delta v| \leq v_{\text{max}}$, and normalize them to the number of galaxies we would have expected to find within the same observed cylindrical volume in a “blank” field. As before, we estimate the expected quasar–galaxy pair counts by means of the number density of galaxies inferred from the [O III] luminosity function multiplied by the observed volume in each shell and the completeness of our survey volume. Uncertainties on the cross-correlation measurements are estimated using Poisson noise on the detected galaxy counts.¹³ Our measurements for χ_{QG} are listed in Table 2 and shown in Figure 3.

¹³ Note that in reality, the uncertainties on the correlation measurements arise from a combination of Poisson counting errors as well as cosmic variance. The latter errors are, however, challenging to estimate and require extensive cosmological simulations, and thus we approximate the uncertainties by the Poisson errors alone.

Table 2
Quasar–Galaxy Cross-correlation and Galaxy–Galaxy Autocorrelation Measurements at $\langle z \rangle = 6.25$

R_{\min} (cMpc h^{-1})	R_{\max} (cMpc h^{-1})	$\langle QG \rangle$	$\langle QR \rangle$	$\langle GG \rangle$	$\langle GR \rangle$	$\langle RR \rangle$	χ_{QG}	χ_{GG}
0.06	0.10	2	3.3×10^{-3}	3	0.10	0.11	604.2 ± 427.9	26.0 ± 17.2
0.10	0.18	3	1.4×10^{-2}	6	0.33	0.34	210.8 ± 122.3	15.8 ± 8.0
0.18	0.31	3	4.3×10^{-2}	16	1.02	1.05	68.3 ± 40.0	13.2 ± 4.2
0.31	0.56	4	0.14	29	3.16	3.24	27.0 ± 14.0	7.0 ± 2.0
0.56	1.00	7	0.44	66	9.48	9.65	15.0 ± 6.1	4.9 ± 1.1
1.00	1.77	21	1.25	117	26.92	27.29	15.8 ± 3.7	2.3 ± 0.5
1.77	3.15	30	3.69	203	67.07	68.39	7.1 ± 1.5	1.0 ± 0.3
3.15	5.60	21	4.55	260	113.43	116.50	3.6 ± 1.0	0.3 ± 0.2

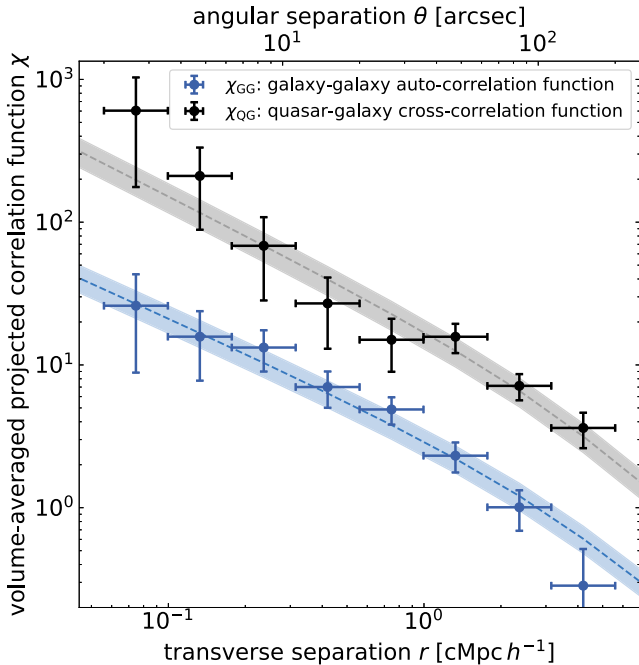


Figure 3. The volume-averaged projected quasar–galaxy cross-correlation function is shown as black data points, while the blue data points show the galaxy–galaxy autocorrelation function at an average redshift of $\langle z \rangle = 6.25$. The dashed lines and shaded areas show the best power-law fits with 2σ uncertainties when jointly fitting both sets of measurements while keeping the slopes $\gamma_{GG} = 1.8$ and $\gamma_{QG} = 2.0$ fixed (see Section 4.4).

4.2. Galaxy–Galaxy Autocorrelation Function

Analogous to Equation (3), we calculate the volume-averaged projected autocorrelation function of galaxies $\chi_{GG}(R_{\min}, R_{\max})$ by integrating the real-space galaxy–galaxy autocorrelation function $\xi_{GG}(R, Z)$ in shells of volume V . We consider all galaxies within the redshift interval $5.95 \leq z \leq 6.55$, which is approximately centered around the mean redshift of the quasars, $\langle z \rangle \approx 6.25$. Again, galaxies within the quasars’ environments are excluded from the autocorrelation measurement. We then calculate pairs of galaxies separated by $|\Delta v| \leq v_{\max}$.

We choose the Landy–Szalay estimator (S. D. Landy & A. S. Szalay 1993), i.e.,

$$\chi_{GG}(R_{\min}, R_{\max}) = \frac{\langle GG \rangle - 2\langle GR \rangle + \langle RR \rangle}{\langle RR \rangle} - 1, \quad (5)$$

where $\langle GG \rangle$, $\langle GR \rangle$, and $\langle RR \rangle$ denote the number of galaxy pairs, pairs of detected galaxies, and galaxies in the random

catalog, as well as pairs of galaxies within the random catalog, respectively. The terms $\langle GR \rangle$ and $\langle RR \rangle$ are normalized such that the number of random galaxies would match the number of detected galaxies in each quasar field within the considered redshift interval. Uncertainties on the measurements are approximated by the Poisson noise on the number of detected galaxies. The measurements of the volume-averaged projected galaxy–galaxy autocorrelation function are shown in Figure 3 and listed in Table 2.

4.3. Fitting the Cross- and Autocorrelation Functions

The volume-averaged projected correlation functions χ_{QG} and χ_{GG} represent integrals of the real-space correlation functions ξ_{QG} and ξ_{GG} as shown in Equation (3). We parameterize the real-space cross- and autocorrelation functions by a power law, i.e.,

$$\xi(r) = (r/r_0)^{-\gamma}, \quad (6)$$

where $r = \sqrt{R^2 + Z^2}$. Thus, the real-space correlation functions are governed each by two parameters, i.e., their slopes γ , denoted as γ_{QG} and γ_{GG} for the cross- and autocorrelation, respectively, and their scale lengths r_0 , denoted as r_0^{QG} and r_0^{GG} .

We use a Markov Chain Monte Carlo algorithm (D. Foreman-Mackey et al. 2013) with a Gaussian likelihood function and uniform priors for the correlation scale length of $r_0 \in [1, 30]$ cMpc h^{-1} and the slope $\gamma \in [1, 3]$ to fit our measurements. The median and 16th and 84th percentiles of the posterior distribution define our best estimate.

With both the slope and scale length as free model parameters, we obtain $r_0^{QG} = 12.4_{-2.3}^{+3.8}$ cMpc h^{-1} and $\gamma_{QG} = 1.6_{-0.3}^{+0.2}$ as the best fits for the quasar–galaxy cross-correlation function. When fixing the slope to $\gamma_{QG} = 2.0$ to enable a comparison to other studies (see Section 7), we obtain $r_0^{QG} = 9.1_{-0.6}^{+0.5}$ cMpc h^{-1} .

Similarly, for the galaxy–galaxy autocorrelation, we obtain $r_0^{GG} = 3.9 \pm 0.4$ cMpc h^{-1} and $\gamma_{GG} = 1.9 \pm 0.1$ when fitting for both scale length and slope. When fixing the slope to $\gamma_{GG} = 1.8$ as in previous studies (e.g., C. García-Vergara et al. 2017), we obtain $r_0^{GG} = 4.1 \pm 0.3$ cMpc h^{-1} .

4.4. Jointly Fitting Both Correlation Functions

We can now attempt to study the clustering properties of the quasars themselves and infer the real-space quasar–quasar autocorrelation function ξ_{QQ} by means of ξ_{QG} and ξ_{GG} . To this end, we follow a deterministic bias model (see, e.g., S. M. Croom et al. 2001; C. García-Vergara et al. 2017) and denote the density contrast of galaxies and quasars as δ_Q and

δ_G . Thus, the cross-correlation between quasars and galaxies is defined as $\xi_{QG} = \langle \delta_Q \delta_G \rangle$. Assuming that galaxies and quasars trace the same underlying dark matter distribution, the galaxy and quasar density contrast can be expressed as $\delta_G = b_G(\delta_{DM})\delta_{DM}$ and $\delta_Q = b_Q(\delta_{DM})\delta_{DM}$, respectively, where $b_G(\delta_{DM})$ and $b_Q(\delta_{DM})$ denote the galaxy and quasar bias, which are (possibly nonlinear) functions of the dark matter density contrast δ_{DM} . If we describe δ_Q and δ_G as two stochastic processes, their cross-correlation coefficient ρ can be expressed as

$$\rho = \frac{\langle \delta_Q \delta_G \rangle}{\sqrt{\langle \delta_Q \delta_Q \rangle \langle \delta_G \delta_G \rangle}}. \quad (7)$$

However, given the deterministic relations for δ_Q and δ_G as defined above, it follows that $\rho = 1$. Thus, the quasar–quasar autocorrelation ξ_{QQ} can be expressed as a combination of the other two correlation functions, ξ_{QG} and ξ_{GG} , i.e.,

$$\xi_{QQ} = \xi_{QG}^2 / \xi_{GG}. \quad (8)$$

Analogously to Equation (6), we parameterize all auto- and cross-correlation functions with a power law but now perform a joint fit to both the auto- and cross-correlation measurements. To this end, we fit ξ_{GG} as a function of γ_{GG} and r_0^{GG} , as well as ξ_{QG} parameterized by γ_{QQ} , γ_{GG} , r_0^{QQ} and r_0^{GG} using Equation (8).

To enable a comparison to previous work (see Section 7.1), we keep the slope of the two autocorrelation functions fixed, i.e., $\gamma_{GG} = 1.8$ (C. García-Vergara et al. 2017), and $\gamma_{QQ} = 2.0$ (Y. Shen et al. 2007; S. Eftekharzadeh et al. 2015), and only fit the two remaining free parameters, i.e., the scale lengths r_0^{GG} and r_0^{QQ} . Our best estimates determined by the median and 16th and 84th percentile of the posterior probability distributions for the scale lengths of the galaxy autocorrelation function at $\langle z \rangle \approx 6.25$ is $r_0^{GG} = 4.1 \pm 0.3 \text{ cMpc } h^{-1}$, while we obtain $r_0^{QQ} = 22.0_{-2.9}^{+3.0} \text{ cMpc } h^{-1}$ for the scale length of the quasars’ autocorrelation function. These estimates for the power laws and our measurements are shown in Figure 3.

We also attempt to jointly fit all four model parameters, i.e., two slopes and two correlation scale lengths. We caution that the parameters are strongly degenerate and approach the prior boundaries, but the marginalized posteriors suggest $r_0^{GG} = 4.0_{-0.3}^{+0.4} \text{ cMpc } h^{-1}$, $r_0^{QQ} = 23.7_{-5.6}^{+4.3} \text{ cMpc } h^{-1}$, $\gamma_{GG} = 1.9 \pm 0.1$, and $\gamma_{QQ} = 1.9 \pm 0.2$.

5. Mass Estimates of the Quasars’ Host Dark Matter Halos

In this section, we will estimate the mass of the dark matter halos that harbor luminous quasars at $z \gtrsim 6$ based on the quasars’ clustering properties. Our cosmological model Λ CDM provides a direct link between the clustering strength of a population of objects and the characteristic mass of their host dark matter halos, since more massive halos have a higher clustering bias (e.g., N. Kaiser 1984; H. J. Mo & S. D. M. White 1996; J. L. Tinker et al. 2010).

We will present three different approaches to constrain the host dark matter halo masses, all based on dark-matter-only simulations. We use N -body cosmological simulations to leverage the accurate representation of the halo clustering properties that they provide. Accurately modeling the distribution of halos in the Universe is particularly important at the small scales probed by our observations (i.e., at distances of $r \lesssim 10 \text{ cMpc } h^{-1}$), as these scales are highly nonlinear in

overdense environments. Alternative approaches based on nonlinear analytical extensions of linear theory to small scales (such as the “halo model” framework, e.g., A. Cooray & R. Sheth 2002) would likely return results that are not accurate enough for our purposes.

We first use a new N -body simulation FLAMINGO-10K, which is part of the FLAMINGO project (J. Schaye et al. 2023, see Section 5.1), and assume simple “step function” models for the halo occupation distributions (HODs; e.g., M. White et al. 2008) of quasars and galaxies. We then also compare to two empirical models for galaxies and SMBHs based on the *Uchuu* N -body simulation (T. Ishiyama et al. 2021), namely, the UNIVERSEMACHINE (P. Behroozi et al. 2019, see Section 5.2), as well as the TRINITY model (H. Zhang et al. 2023, see Section 5.3).

5.1. Host Dark Matter Halo Masses Based on the FLAMINGO-10K Cosmological Simulation

As a first approach, we use the N -body cosmological simulation FLAMINGO-10K to directly infer the host halo masses of quasars and [O III] emitters by matching the observed quasar–galaxy cross-correlation and galaxy–galaxy autocorrelation measurements with the clustering properties of dark matter halos. FLAMINGO-10K is a new dark-matter-only simulation part of the FLAMINGO suite (J. Schaye et al. 2023), which makes use of the open-source code SWIFT (M. Schaller et al. 2024), a highly parallel gravity and smoothed particle hydrodynamics solver. The simulation was run with 10080^3 cold dark matter particles and 5600^3 neutrino particles and has the same volume as the FLAMINGO flagship hydrodynamical run (2.8 cGpc on a side) but 8 times more dark matter and neutrino particles, i.e., with a dark matter particle mass of $8.4 \times 10^8 M_\odot$. The run adopts the “ $3 \times 2pt + \text{all}$ ” cosmology from T. M. C. Abbott et al. (2022; $\Omega_m = 0.306$, $\Omega_b = 0.0486$, $\sigma_8 = 0.807$, $H_0 = 68.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $n_s = 0.967$), with a summed neutrino mass of 0.06 eV. Based on this simulation, we can model the clustering properties of any subset of $z \approx 6$ dark matter halos in the mass range $M = 10^{10.5} - 10^{13} M_\odot$. Our simulation covers a total of $\approx 10^8$ dark matter halos over this mass range, 60, 857 (947, 4) of which above a halo mass of $10^{12} M_\odot$ ($10^{12.5} M_\odot$, $10^{13} M_\odot$). Thus, the number of lower-mass halos hosting [O III] emitters is very large, but we need to ensure that our simulations include sufficient high-mass halos to limit the effects of cosmic variance. Assuming that ~ 10 – 20 of such halos are sufficient, the cross-correlation function estimated from the simulations will be independent of sampling effects due to cosmic variance for halo masses up to $\sim 10^{12.8-12.9} M_\odot$. For details on the modeling procedure and the simulation, we refer the reader to E. Pizzati et al. (2024).

Assuming a prescription for how quasars and [O III] emitters populate dark matter halos, we model the quasar–galaxy cross-correlation, χ_{QG} , as well as the galaxy–galaxy autocorrelation, χ_{GG} . Here, we assume a “step function” HOD model for both quasars and [O III] emitters, where halos are populated by quasars ([O III] emitters) only above a minimum threshold mass $M_{\text{halo,min}}$ ($M_{\text{halo,min}}^{[\text{O III}]}$). The two minimum host halo masses, $M_{\text{halo,min}}$ and $M_{\text{halo,min}}^{[\text{O III}]}$, are free parameters in our model, which we infer by jointly fitting the observed correlation functions, χ_{QG} and χ_{GG} . We assume that these two correlation functions are independent, and write down a 2D Gaussian likelihood for the parameters $M_{\text{halo,min}}$ and $M_{\text{halo,min}}^{[\text{O III}]}$. By computing the median and 16th–84th percentiles of the marginalized likelihood distributions, we infer the minimum host dark matter

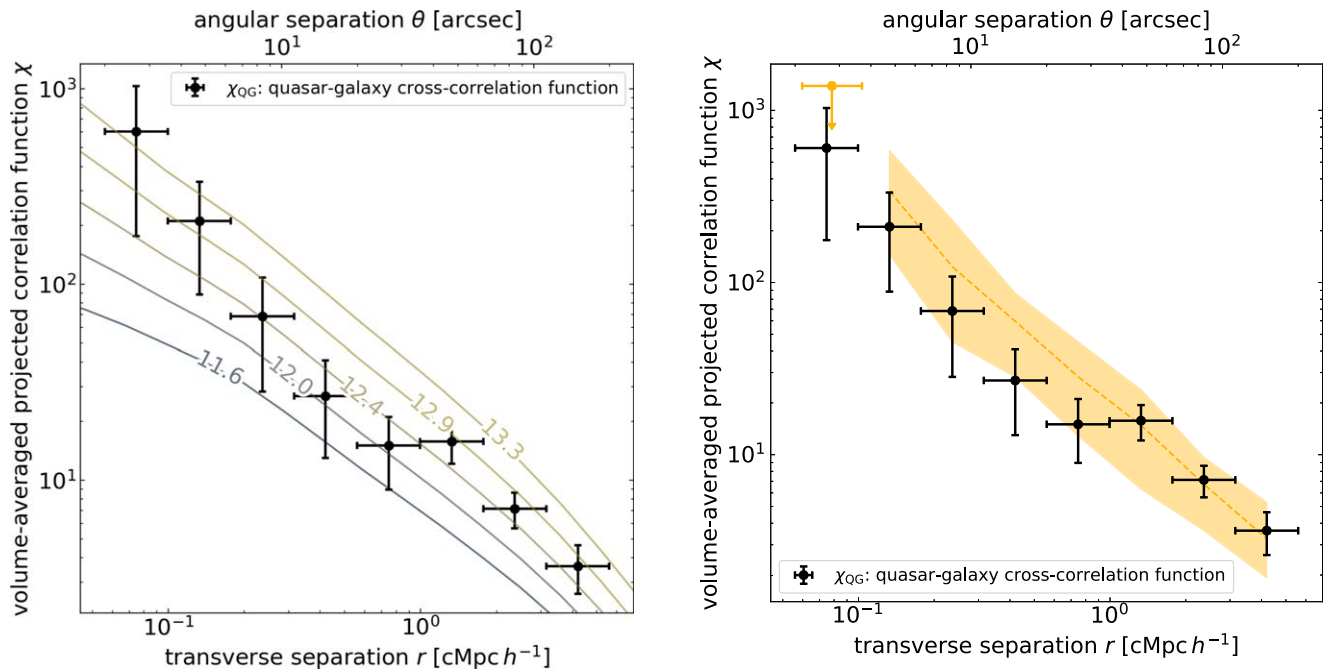


Figure 4. Left: comparison of the quasar–galaxy cross-correlation measurement to models from the FLAMINGO-10K simulation for different minimum quasar host dark matter halo masses, $M_{\text{halo,min}}$. By comparing the correlation functions of dark matter halos from the simulations to the measured quasar–galaxy cross-correlation and galaxy–galaxy autocorrelation, we obtain a best estimate for the quasars’ minimum host halo mass of $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.43^{+0.13}_{-0.15}$. Right: comparison between the measurements of the quasar–galaxy cross-correlation with predictions (median: dashed yellow curve; 16th–84th percentile: shaded region) for the observed quasar fields from the TRINITY model, which implies an average host dark matter halo mass of $\log_{10}(M_{\text{halo}}/M_{\odot}) = 12.14^{+0.24}_{-0.26}$.

halo mass of quasars to be $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.43^{+0.13}_{-0.15}$, while we find that [O III] emitters are hosted in halos of mass $\log_{10}(M_{\text{halo,min}}^{\text{[O III]}}/M_{\odot}) = 10.56^{+0.05}_{-0.03}$. In a companion paper (E. Pizzati et al. 2024), we model the full mass distribution of the dark matter halos hosting quasars and galaxies by jointly fitting the clustering properties as well as the luminosity functions of quasars and [O III] emitters, and find results that are in excellent agreement with the model presented here.

In the left panel of Figure 4, we compare the predictions from the FLAMINGO-10K simulation with the observed quasar–galaxy cross-correlation measurements, for different values of the minimum halo mass for the quasar hosts, $M_{\text{halo,min}}$. The cross-correlation terms are obtained by assuming a minimum host halo mass for [O III] emitters equal to the median value determined above, i.e., $\log_{10}(M_{\text{halo,min}}^{\text{[O III]}}/M_{\odot}) = 10.56$. The best fit between the simulated and observed cross- and autocorrelation functions is obtained with a minimum quasar host halo mass of $\log_{10}(M_{\text{halo,min}}/M_{\odot}) \approx 12.43$.

Note that the slope of the cross-correlation functions at very small scales, i.e., $r \lesssim 200 \text{ cMpc } h^{-1}$, appears to differ between the observations and the simulations. This could either be due to small number statistics given our current data set, as we only observe a total of five galaxies for the innermost two data points, or alternatively, uncertainties in the modeling procedure could cause the discrepancy. For instance, the identification of subhalos becomes increasingly challenging at such small separations to a massive structure, and could thus alternate the shape of the correlation functions at small scales. However, our conclusions are not affected by the two innermost data points, as the uncertainties on these measurements are high and thus do not influence the best fit significantly.

Additional differences at the smallest scales arise since we do not apply the same merging procedure to the subhalos in the simulations that we apply to the detected [O III] emitters as

described in Section 2.3. Merging the subhalos would lead to a slightly shallower slope in the simulated cross-correlation function at the smallest scales of $\lesssim 3''$, but since the uncertainties on our measurements on the smallest scales are large due to the low number of detected galaxies, this has a negligible effect on our estimates. When applying the same merging procedure to subhalos in the *Uchuu* simulations (see Section 5.2) as a check, we found that $\sim 2\%$ of the subhalos were affected compared to the $\sim 13\%$ of [O III] emitters affected by merging, which indicates that the [O III] emitters show mostly clumpy substructure within an individual subhalo.

In the following, we will briefly discuss two other methods to determine the host dark matter halo masses from the quasars’ clustering properties to ensure consistency between different simulations and models, but moving forward, we will consider the measurement obtained here as our best estimate for the minimum host dark matter halo mass of $z \gtrsim 6$ quasars.

5.2. Host Dark Matter Halo Masses Based on the UNIVERSEMACHINE

A second method to determine the dark matter halo masses of the observed quasars is based on finding analogous systems to the observed quasar environments using the UNIVERSEMACHINE (P. Behroozi et al. 2019) applied to the *Uchuu* *N*-body simulations. The details of this approach will be presented in detail in a forthcoming companion paper (R. Mackenzie et al. 2024, in preparation). In short, by comparing the number of simulated [O III] emitters around halos of different masses to the number of companion galaxies observed in the quasar fields, a distribution of possible host dark matter halo masses can be determined for each observed quasar.

We first construct a catalog of simulated environments around each halo and select all halos above $10^{10} h^{-1} M_{\odot}$ (at $z = 6.35$). In

total, this simulation contains $\sim 10^{8.5}$ such halos, approximately 1.3×10^6 (24, 525, 1, 860, 85, 8) of which are above a mass threshold of $10^{11.5} M_\odot$ ($10^{12.1} M_\odot$, $10^{12.4} M_\odot$, $10^{12.7} M_\odot$, $10^{12.9} M_\odot$). Each halo has its own realization where the halo position becomes the center of a mock observation (i.e., at the quasar position). The surrounding galaxies are then forward modeled through our completeness function and survey geometry. By calculating the joint distribution of the number of detected neighboring galaxies versus halo mass, and applying the Bayes equation with a flat prior on halo mass, one can estimate the mass probability distribution for an observed number of [O III]-emitting neighbors. The resulting most probable dark matter halo masses range between $10^{10.85} M_\odot \lesssim M_{\text{halo}} \lesssim 10^{12.85} M_\odot$ with a median of $\langle M_{\text{halo}} \rangle = 10^{12.4 \pm 0.5} M_\odot$, where the wide range of host halo masses is largely driven by the cosmic variance observed between the different quasar fields. The estimate for the minimum host dark matter halo masses using constraints from the UNIVERSEMACHINE is well consistent with the estimate based on the FLAMINGO-10K simulations.

5.3. Host Dark Matter Halo Masses Based on TRINITY

Finally, we compare the observed χ_{QG} with the prediction from the TRINITY model (H. Zhang et al. 2023, 2024) of the halo–galaxy–SMBH connection, to infer the typical host dark matter halo mass of the quasars. Briefly, TRINITY reconstructs empirical galaxy and black hole mass growth rate distributions as functions of the halo mass from redshifts $z = 0–10$, and forward models observables such as galaxy mass functions and quasar luminosity functions. SMBH properties such as radiative efficiencies, Eddington ratio distributions, and duty cycles are parameterized and constrained by observational data. For details on the TRINITY implementation, we refer the reader to H. Zhang et al. (2023).

We utilize the TRINITY mock catalogs of galaxies and SMBHs, with positions, velocities, halo masses, galaxy masses, star formation rates, rest-frame UV luminosities, as well as black hole masses and bolometric luminosities (if hosting a SMBH) for each galaxy. These mock catalogs are generated also from the *Uchuu* N -body simulation (box size: $2 \text{ cGpc } h^{-1}$, mass resolution: $3.27 \times 10^8 M_\odot h^{-1}$, T. Ishiyama et al. 2021), and will be made publicly available in a forthcoming TRINITY publication (H. Zhang et al. 2024, in preparation). In order to calculate χ_{QG} and host halo mass distributions, we use all quasars in our $z = 6.35$ mock catalog and assign each mock quasar a weight to quantify its degree of similarity to the here analyzed quasars. In particular, the weight w for a simulated quasar with black hole mass and bolometric luminosity (M_{BH} , L_{bol}) is a sum of the log-normal probabilities given the observed mass and luminosities, $\{(M_{\text{BH},i}, L_{\text{bol},i})\}_{i=1}^{N=4}$, of the EIGER quasars, i.e.,

$$\begin{aligned}
 w &= \sum_{i=1}^{N=4} P[(M_{\text{BH}}, L_{\text{bol}}) | (M_{\text{BH},i}, L_{\text{bol},i})] \\
 &= \sum_{i=1}^{N=4} \frac{1}{2\pi\sigma_{L_{\text{bol},i}}\sigma_{M_{\text{BH},i}}} \\
 &\quad \times \exp\left[-\frac{(\log M_{\text{BH}} - \log M_{\text{BH},i})^2}{2\sigma_{M_{\text{BH},i}}^2}\right] \\
 &\quad \times \exp\left[-\frac{(\log L_{\text{bol}} - \log L_{\text{bol},i})^2}{2\sigma_{L_{\text{bol},i}}^2}\right], \quad (9)
 \end{aligned}$$

where $\sigma_{M_{\text{BH},i}}$ and $\sigma_{L_{\text{bol},i}}$ are the measurement uncertainties in the SMBH mass and luminosity of the i th quasar, respectively. For $\sigma_{M_{\text{BH},i}}$, we also add a scatter of 0.5 dex in quadrature to account for the random scatter in virial estimates of M_{BH} (R. J. McLure & M. J. Jarvis 2002; M. Vestergaard & B. M. Peterson 2006).

When calculating the predicted χ_{QG} we convert the galaxy UV luminosity, L_{UV} , into $L_{[\text{O III}],5008}$ by applying the scaling relation from J. Matthee et al. (2023), and include all mock galaxies with $L_{[\text{O III}],5008} \geq 10^{42} \text{ erg s}^{-1}$. To quantify χ_{QG} uncertainties from TRINITY, we simulate the observation 3000 times, i.e., we choose four mock quasars for every observation, where each quasar is randomly selected within the 1.5σ range of each observed quasar in the $(\log M_{\text{BH}}, \log L_{\text{bol}})$ space. We then calculate the 16th–84th percentile range of the measured χ_{QG} 's as the uncertainties from TRINITY.

As shown in the right panel of Figure 4, the resulting χ_{QG} from TRINITY are consistent with the observed values within the measurement uncertainties. This consistency allows us to infer the host dark matter halo masses of the quasars by examining the host halo mass distribution of the mock quasars, weighted by the weights defined in Equation (9). From this exercise, we obtain a median (with 16th–84th percentile) host halo mass of $\log_{10}(M_{\text{halo}}/M_\odot) = 12.14_{-0.26}^{+0.24}$.

6. Estimating the Quasars' Duty Cycle and Lifetime

Our cosmological model Λ CDM allows us to determine the number density of the quasars' host dark matter halos n_{hosts} from their characteristic mass determined in the previous section (e.g., H. J. Mo & S. D. M. White 1996; J. L. Tinker et al. 2010), which in turn enables constraints on the quasars' duty cycle and lifetime. As quasars temporarily subsample their hosts, their number density n_{Q} can be expressed as

$$n_{\text{Q}} \simeq \frac{t_{\text{Q}}}{t_{\text{H}}(z)} n_{\text{hosts}}, \quad (10)$$

where t_{Q} is the quasars' UV-luminous lifetime, and $t_{\text{H}}(z)$ denotes the Hubble time at redshift z , and hence $f_{\text{duty}} = t_{\text{Q}}/t_{\text{H}}(z)$ is the quasars' duty cycle at the given redshift (Z. Haiman & L. Hui 2001; P. Martini & D. H. Weinberg 2001). We will now determine these timescales of quasar activity, which are essential for understanding the concomitant growth phases of early SMBHs.

Using the FLAMINGO-10K simulation we find that the number density of halos with a mass larger than $\log_{10}(M_{\text{halo,min}}/M_\odot) = 12.43_{-0.15}^{+0.13}$ is $n_{\text{host}} = 5.3_{-3.9}^{+16.7} \times 10^{-8} \text{ cMpc}^{-3}$. For estimating the number density of luminous quasars we use the quasar luminosity function (QLF) at $\langle z \rangle \approx 6.25$ by J.-T. Schindler et al. (2023) and integrate it over the magnitude range $-26.5 > M_{1450} > -29.5$ bracketing the observed magnitudes of the analyzed quasars (see Table 1), which results in a quasar number density of $n_{\text{Q}} \approx 0.23 \text{ cGpc}^{-3}$. Using Equation (10) the two number densities determine a duty cycle of $f_{\text{duty}} = 0.4_{-0.3}^{+1.2}\%$, and hence a UV-luminous quasar lifetime of approximately $t_{\text{Q}} \sim 10^{6.5} \text{ yr}$ at $\langle z \rangle = 6.25$.

However, we caution that these timescale estimates depend sensitively on the number density of the host dark matter halos given the characteristic minimum halo mass derived in Section 5. Thus even relatively small differences in the host halo mass estimates of approximately ± 0.25 dex can lead to variations in the derived timescales of ± 1 dex, i.e., a host dark matter halo mass of $\log_{10}(M_{\text{halo,min}}/M_\odot) \approx 12.65$

($\log_{10}(M_{\text{halo,min}}/M_{\odot}) \approx 12.18$) with a number density of a factor of 10 lower (higher) than our best estimate would correspond to a UV-luminous quasar lifetime of $t_Q \sim 10^{7.5}$ yr ($t_Q \sim 10^{5.5}$ yr). Yet, even with this uncertainty, these inferred UV-luminous lifetime estimates represent a major departure from the expected long timescales of $t_Q \sim 10^9$ yr and $f_{\text{duty}} \sim 1$, required to grow the quasars' SMBHs. In order to match the abundance of host dark matter halos to the abundance of luminous quasars to ensure a duty cycle of unity, the required host dark matter halo mass would have to be $\sim 10^{13} M_{\odot}$.

7. Discussion

7.1. Comparison to Previous Studies

There is only one study to date measuring the autocorrelation function of $z \gtrsim 6$ quasars, which is the recent result by the SHELLQs Collaboration (J. Arita et al. 2023). The authors analyze the autocorrelation function of approximately 100 faint ($M_{1450} \gtrsim -25$) quasars at $z \sim 6$ distributed over 891 deg^2 . They probe scales between $r \sim 10\text{--}1000 \text{ cMpc } h^{-1}$, i.e., much larger scales than the scales of $r < 10 \text{ cMpc } h^{-1}$ we probe with our NIRCам WFSS data, and measure an autocorrelation scale length of $r_0^{\text{QQ}} = 23.7 \pm 11 \text{ cMpc } h^{-1}$, which is in excellent agreement to our inferred estimate of $r_0^{\text{QQ}} = 22.0_{-2.9}^{+3.0} \text{ cMpc } h^{-1}$. They infer a host dark matter halo mass using the linear halo model software HALOMOD (S. G. Murray et al. 2013, 2021) and obtain a significantly larger host dark matter halo mass estimate of $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.9_{-0.4}^{+0.4}$.

An independent approach to infer the host dark matter halos of a sample of luminous ($M_{1450} < -26.5$) quasars at $z \sim 6$ with similar properties to the ones analyzed in this study is presented in H. Chen et al. (2022). The authors use the flux transmission profile observed in the rest-frame UV quasar spectra to infer the density field around the quasars and obtain an estimate of the typical host dark matter halo masses of $\log_{10}(M_{\text{halo}}/M_{\odot}) = 12.5_{-0.7}^{+0.4}$, which is consistent with our results.

In Figure 5, we compare our results for the autocorrelation length, host dark matter halo mass, and duty cycle to these previous analyses at $z \sim 6$. We also add measurements at lower redshifts to study the redshift evolution of the parameters. We confirm previous findings that the dark matter halo mass of quasars appears to be approximately constant throughout cosmic time, suggesting that there is a characteristic halo mass of a few times $M_{\text{halo}} \sim 10^{12} M_{\odot}$ where quasars can be activated (e.g., S. M. Croom et al. 2005; J. Arita et al. 2023).

Interestingly, we know that galaxies typically quench at stellar masses $M_{\star} \sim 10^{11} M_{\odot}$ across a wide range of redshifts out to $z \sim 5$ (e.g., Y.-J. Peng et al. 2010; P. S. Behroozi et al. 2013; N. Caplar et al. 2015), corresponding to a halo mass $M_{\text{halo}} \sim 10^{12.0\text{--}12.5} M_{\odot}$, assuming an approximately constant stellar mass to halo mass ratio. Since there is considerable evidence that the quenching of galaxies is associated with, or possibly caused by, quasar activity (e.g., M. Cano-Díaz et al. 2012; N. Caplar et al. 2018), the identification of this same halo mass as the mass of the EIGER quasars' host halos at $z \sim 6$ may not be unexpected.

7.2. Overdensities around High-redshift Quasars?

The numerous previous inconclusive results aiming to determine the existence of an overdense environment around high-redshift quasars have cast doubt on whether these quasars indeed trace the highest-density peaks in our Universe. Our clustering measurement

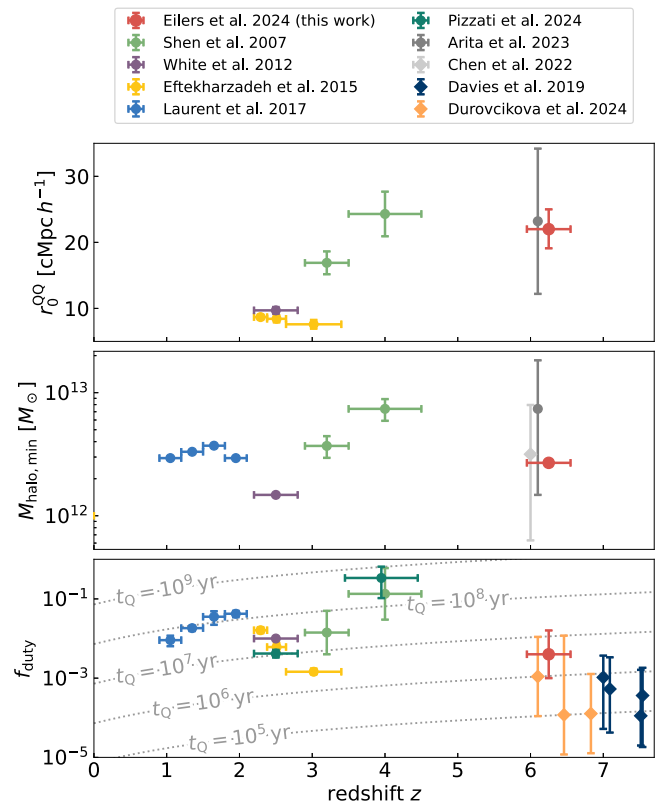


Figure 5. Redshift evolution of the quasars' autocorrelation scale length (top), the minimum host dark matter halo mass (middle), and the quasars' duty cycle (bottom). Our inferred scale length is well consistent with the recent autocorrelation measurement by J. Arita et al. (2023), but we infer a smaller host dark matter halo mass using dark-matter-only simulations (see Section 5). Our results indicate a very small duty cycle of quasars at $z \gtrsim 6$, which agrees with independent duty cycle measurements inferred from Ly α damping wings (diamond-shaped data points; F. B. Davies et al. 2019; D. Đurovčíková et al. 2024). The host dark matter halo mass estimate by H. Chen et al. (2022) is also inferred from the transmitted flux along quasar sightlines, while all other measurements are derived from quasar clustering studies (circular data points; Y. Shen et al. 2007; M. White et al. 2012; S. Eftekharzadeh et al. 2015; P. Laurent et al. 2017; J. Arita et al. 2023; E. Pizzati et al. 2024). The gray dotted lines in the bottom panel indicate different quasar lifetimes.

suggests that quasars reside in dark matter halos with a minimum mass of $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.43_{-0.15}^{+0.13}$, which implies that they do on average trace large overdensities but *not necessarily* the rarest and highest-density peaks.

Our sample of four quasar fields has revealed a wide range of overdensities providing evidence for large cosmic variance. The observed density contrasts in the quasar fields range from $\delta(r < 2 \text{ cMpc}) = 65 \pm 15$ to $\delta(r < 2 \text{ cMpc}) = 3 \pm 4$. While J0148+0600 with 47 [O III]-emitting galaxies in its surroundings likely resides in one of the highest-density peaks, and traces one of the largest high-redshift protoclusters reported to date, J1030+0542 on the other hand appears to reside in an average density region with no enhancement compared to a random field (which is in contrast to the protocluster reported in this field by M. Mignoli et al. 2020).

It is interesting to note that the [C II] emission arising from the molecular gas of the host galaxy of J1030+0542 was not detected with Atacama Large Millimeter/submillimeter Array (ALMA), indicating that the far-IR (FIR) luminosity of this quasar is significantly fainter than for most other $z \sim 6$ quasars with similar rest-frame UV magnitudes (R. Decarli et al. 2018). This result, combined with the recent discoveries of large

overdensities around submillimeter-bright galaxies at $z > 5$ (T. Herard-Demanche et al. 2023; F. Sun et al. 2024), provides circumstantial evidence that FIR luminosities might correlate more strongly with dark matter halo mass than the UV luminosities of high-redshift objects do.

7.3. Implications for the Early Growth of SMBHs

As the emission of quasar light is concomitant with black hole accretion, the timescales of quasar activity have implications for the growth of SMBHs. Previous constraints on the quasars' duty cycle based on clustering studies at redshifts of $1 \lesssim z \lesssim 4$ suggest broadly an increase toward higher redshifts¹⁴ as shown in the bottom panel of Figure 5, implying relatively long UV-luminous quasar lifetimes of $t_Q \sim 10^7 - 10^8$ yr (e.g., P. Martini 2004).

Our standard black hole growth model requires such long quasar lifetimes in order to explain the growth of the observed billion solar mass black holes of high-redshift quasars, and consequently, high quasar duty cycles of $f_{\text{duty}} \sim 1$ have been postulated for quasars in the early Universe (e.g., P. Martini 2004; M. Volonteri 2012). However, the clustering analysis of $z \gtrsim 6$ quasars in this study surprisingly provides evidence for the contrary: we find a very small duty cycle, $f_{\text{duty}} \ll 1$, indicating that galaxies do not shine as UV-luminous quasars throughout the whole Hubble time at these redshifts, thus posing additional challenges to our understanding of early SMBH growth.

Interestingly, such low values are consistent with completely independent constraints on the UV-luminous duty cycle of quasars at similar redshifts derived from the extent of their proximity zones (e.g., A.-C. Eilers et al. 2017, 2018, 2020; I. T. Andika et al. 2020; K. A. Morey et al. 2021; S. Satyavolu et al. 2023b) and damping wing features (e.g., F. B. Davies et al. 2019; D. Āurovčřková et al. 2024) that are observed in the quasars' rest-frame UV spectra. These studies estimate the number of ionizing photons emitted by a quasar that would result in the observed flux transmission profile along the line of sight and thus enable constraints on their emission timescales for individual objects.¹⁵ These observations have previously suggested UV-luminous quasar lifetimes of only $t_Q \sim 10^6$ yr, which imply similarly low duty cycles of $f_{\text{duty}} \ll 1$ for high-redshift quasars (see also I. S. Khrykin et al. 2019, for a similar analysis of He II proximity zones at $z \sim 4$).

Thus, two independent arguments—one based on the quasars' clustering properties and the other based on the observed flux transmission profiles in their spectra—provide evidence for very short UV-luminous quasar activity timescales in the early Universe. This suggests that accretion onto the SMBHs occurs in highly radiatively inefficient phases with a radiative efficiency ϵ of less than the fiducial $\sim 10\%$ suggested

¹⁴ We note that the increase in the quasars' duty cycle with redshift is not found by S. Eftekharzadeh et al. (2015), who found significantly lower duty cycles than Y. Shen et al. (2007) at $z \sim 3-4$. However, their quasar sample using BOSS includes much fainter objects, which makes the selection of quasars at these redshifts as well as the completeness evaluation of their sample more challenging.

¹⁵ Note that for quasars that are embedded in an already highly ionized intergalactic medium (IGM), with a neutral gas fraction of $x_{\text{H I}} \sim 10^{-3}$ at $z \sim 6$, the proximity zone size reflects the lifetime of the last quasar episode and hence the duty cycle presents an upper limit on these episodic lifetime estimates. For quasars at $z \gtrsim 7$ that are embedded in a still fairly neutral IGM, $x_{\text{H I}} \gtrsim 0.1$, the observed damping wing signature in the proximity zone allows for an estimate of the integrated quasar lifetime, i.e., the duty cycle, due to the long recombination time (see, e.g., F. B. Davies et al. 2019).

by thin accretion disk models (e.g., F. B. Davies et al. 2019), or, alternatively, a significant fraction of the black hole growth happens in UV-obscured, dust-enshrouded environments (e.g., S. Satyavolu et al. 2023a).

The latter is a particularly intriguing scenario in light of the recent discovery of an abundance of faint and highly dust-reddened active galactic nuclei present at early cosmic times (J. Matthee et al. 2024; D. D. Kocevski et al. 2023; J. E. Greene et al. 2024; R. Maiolino et al. 2023; Y. Harikane et al. 2023), which might provide evidence for a highly obscured population of growing SMBHs. However, studies in the mid-infrared such as the recently published Systematic Mid-infrared Instrument Legacy Extragalactic Survey (J. Lyu et al. 2024) are necessary to identify completely dust-obscured growing SMBHs and securely determine the fraction of obscured quasars in the early Universe. If black hole growth in completely dust-enshrouded, UV-obscured environments is indeed responsible for the discrepancy between the theoretically expected long black hole growth timescales and the observed short UV-luminous growth phases, we would expect $\sim 100-1000$ dust-obscured quasars for every unobscured one, indicating the presence of a very large population of dust-obscured quasars in the early Universe that we yet have to detect. While this would imply a very high obscuration fraction of $>99\%$, cosmological simulations (M. Trebitsch et al. 2019; Y. Ni et al. 2020; F. Vito et al. 2022; J. S. Bennett et al. 2024), as well as a recent study using ALMA observations (R. Gilli et al. 2022), have shown that obscuration fraction of at least $\sim 80\%-90\%$ is expected for high-redshift quasars due to the presence of high-column density gas within the innermost regions of their host galaxies, resulting in low observable UV-luminous duty cycles.

8. Summary

In this paper, we present the first clustering measurement of luminous quasars at $z \gtrsim 6$ and [O III]-emitting galaxies in their environment in four quasar fields observed with JWST/NIRCam in the imaging and WFSS mode as part of the EIGER project. We identify 522 [O III]-emitting galaxy systems above a luminosity of $L_{[\text{O III}],5008} = 10^{42}$ erg s^{-1} in all five observed quasar fields, 92 of which reside in close vicinity, i.e., $|\Delta v| \leq 1000$ km s^{-1} , to the luminous quasars.

We find a large diversity in the observed density contrast between the different quasar fields, ranging from density enhancements of $\delta(r < 2 \text{ cMpc}) = 65 \pm 15$ —one of the most spectacular protoclusters within the Epoch of Reionization discovered to date encompassing 47 [O III]-emitting galaxies around the quasar J0148+0600—to $\delta(r < 2 \text{ cMpc}) = 3 \pm 4$, indicating no significant galaxy enhancement around the quasar J1030+0564. The other two fields around the superluminous quasar J0100+2802 and the quasar J1148+5251, show density contrasts of $\delta(r < 2 \text{ cMpc}) = 29 \pm 10$ and $\delta(r < 2 \text{ cMpc}) = 16 \pm 8$, respectively.

We present measurements of the volume-averaged projected galaxy-galaxy autocorrelation function and the quasar-galaxy cross-correlation function at an average redshift of $\langle z \rangle = 6.25$. We parameterize the real-space correlation functions with a power law and obtain estimates for the auto- and cross-correlation lengths of $r_0^{\text{GG}} = 4.1 \pm 0.3 \text{ cMpc } h^{-1}$ and $r_0^{\text{QG}} = 9.1^{+0.5}_{-0.6} \text{ cMpc } h^{-1}$ at $\langle z \rangle = 6.25$, respectively (keeping the slopes $\gamma_{\text{GG}} = 1.8$ and $\gamma_{\text{QG}} = 2.0$ fixed). When jointly fitting the two correlation functions to infer the scale length of

the quasar–quasar autocorrelation, we measure $r_0^{\text{QQ}} = 22.0_{-2.9}^{+3.0} \text{ cMpc } h^{-1}$ (with fixed slope $\gamma_{\text{QQ}} = 2.0$).

By comparing our measurements to dark-matter-only simulations and different empirical models, we find consistent estimates for the host dark matter halo mass: We determine a minimum host halo mass of $\log_{10}(M_{\text{halo,min}}/M_{\odot}) = 12.43_{-0.15}^{+0.13}$ as the best fit between our measurements and the cross-correlation function of dark matter halos in the FLAMINGO-10K simulation. When searching for analog systems to the quasar fields in UNIVERSEMACHINE we find that such quasars are hosted in dark matter halos with average masses of $\log_{10}(M_{\text{halo}}/M_{\odot}) = 12.4 \pm 0.5$. Lastly, when comparing the cross-correlation function to predictions from the TRINITY model, we find host dark matter halo masses with a median of $\log_{10}(M_{\text{halo}}/M_{\odot}) = 12.14_{-0.26}^{+0.24}$. All three estimates agree with each other within the uncertainties despite their very different approaches.

These results for the mass of the quasars’ host dark matter halos suggest that quasars do *not necessarily* reside in the rarest, highest-density peaks, and could explain why some previous studies did not find large overdensities of galaxies around high-redshift quasars. It also indicates a smaller than expected duty cycle of quasars in the early Universe, i.e., $f_{\text{duty}} \ll 1$. While the short UV-luminous duty cycles are consistent with independent results analyzing the flux transmission observed in rest-frame UV quasar spectra, the findings pose significant challenges for the growth of the early SMBHs which power the quasars’ emission. Such short UV-luminous quasar lifetimes imply that (a) either the enormous $\gtrsim 10^9 M_{\odot}$ black holes must have built up their mass extremely rapidly, e.g., via radiatively inefficient, “super-Eddington” accretion phases; or (b) a significant fraction ($\gtrsim 99\%$) of the black hole growth happens in dust-enshrouded and UV-obscured environments, indicating the existence of a vast population of obscured quasars in the early Universe.

Acknowledgments

The authors would like to thank the anonymous referee for the thoughtful comments, which significantly improved our manuscript, and Jan-Torge Schindler, Jiamu Huang, and Feige Wang for helpful discussions.

J.F.H. and E.P. acknowledge support from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program (grant agreement No. 885301). J.M. acknowledges support from the European Union (ERC, AGENTS, 101076224).

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The JWST data presented in this article were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. The specific observations analyzed are associated with program #1243 and can be accessed via doi:10.17909/m5mp-5v90.

This work used the DiRAC Memory Intensive service (Cosma8) at the University of Durham, which is part of the STFC DiRAC HPC Facility (www.dirac.ac.uk). Access to DiRAC resources was granted through a Directors Discretionary Time allocation in 2023/24, under the auspices of the UKRI-funded DiRAC Federation Project. The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/

1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.

We thank the Instituto de Astrofísica de Andalucía (IAA-CSIC), Centro de Supercomputación de Galicia (CESGA), and Spanish Academic and Research Network (RedIRIS) in Spain for hosting Uchuu DR1, DR2, and DR3 in the Skies & Universes site for cosmological simulations. The Uchuu simulations were carried out on the Aterui II supercomputer at the Center for Computational Astrophysics, CfCA, of the National Astronomical Observatory of Japan, and the K computer at the RIKEN Advanced Institute for Computational Science. The Uchuu Data Releases efforts have made use of the skunIAA_RedIRIS and skun6IAA computer facilities managed by the IAA-CSIC in Spain (MICINN EU-Feder grant EQC2018-004366-P).

Software: numpy (C. R. Harris et al. 2020), scipy (E. Jones et al. 2001), matplotlib (J. D. Hunter 2007), astropy (Astropy Collaboration et al. 2013, 2018, 2022), emcee (D. Foreman-Mackey et al. 2013).

ORCID iDs

Anna-Christina Eilers  <https://orcid.org/0000-0003-2895-6218>
 Ruari Mackenzie  <https://orcid.org/0000-0003-0417-385X>
 Elia Pizzati  <https://orcid.org/0000-0002-9712-0038>
 Jorryt Matthee  <https://orcid.org/0000-0003-2871-127X>
 Joseph F. Hennawi  <https://orcid.org/0000-0002-7054-4332>
 Haowen Zhang  <https://orcid.org/0000-0002-4321-3538>
 Rongmon Bordoloi  <https://orcid.org/0000-0002-3120-7173>
 Daichi Kashino  <https://orcid.org/0000-0001-9044-1747>
 Simon J. Lilly  <https://orcid.org/0000-0002-6423-3597>
 Rohan P. Naidu  <https://orcid.org/0000-0003-3997-5705>
 Robert A. Simcoe  <https://orcid.org/0000-0003-3769-9559>
 Minghao Yue  <https://orcid.org/0000-0002-5367-8021>
 Carlos S. Frenk  <https://orcid.org/0000-0002-2338-716X>
 Matthieu Schaller  <https://orcid.org/0000-0002-2395-4902>
 Joop Schaye  <https://orcid.org/0000-0002-0668-5560>

References

- Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2022, *PhRvD*, 105, 023520
 Andika, I. T., Jahnke, K., Onoue, M., et al. 2020, *ApJ*, 903, 34
 Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, *MNRAS*, 425, 2722
 Arita, J., Kashikawa, N., Matsuoka, Y., et al. 2023, *ApJ*, 954, 210
 Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *ApJ*, 935, 167
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
 Bañados, E., Venemans, B., Walter, F., et al. 2013, *ApJ*, 773, 178
 Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Natur*, 553, 473
 Begelman, M. C., & Volonteri, M. 2017, *MNRAS*, 464, 1102
 Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, 488, 3143
 Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJL*, 762, L31
 Bennett, J. S., Sijacki, D., Costa, T., Laporte, N., & Witten, C. 2024, *MNRAS*, 527, 1033
 Bigwood, L., Eilers, A.-C., & Simcoe, R. A. 2024, *MNRAS*, 529, 3511
 Cano-Díaz, M., Maiolino, R., Marconi, A., et al. 2012, *A&A*, 537, L8
 Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2015, *ApJ*, 811, 148
 Caplar, N., Lilly, S. J., & Trakhtenbrot, B. 2018, *ApJ*, 867, 148
 Chen, H., Eilers, A.-C., Bosman, S. E. I., et al. 2022, *ApJ*, 931, 29
 Coil, A. L., Hennawi, J. F., Newman, J. A., Cooper, M. C., & Davis, M. 2007, *ApJ*, 654, 115
 Cole, S., & Kaiser, N. 1989, *MNRAS*, 237, 1127

- Cooray, A., & Sheth, R. 2002, *PhR*, 372, 1
- Costa, T. 2024, *MNRAS*, 531, 930
- Costa, T., Sijacki, D., Trenti, M., & Haehnelt, M. G. 2014, *MNRAS*, 439, 2146
- Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, *MNRAS*, 356, 415
- Croom, S. M., Shanks, T., Boyle, B. J., et al. 2001, *MNRAS*, 325, 483
- Davies, F. B., Hennawi, J. F., & Eilers, A.-C. 2019, *ApJL*, 884, L19
- Decarli, R., Walter, F., Venemans, B. P., et al. 2018, *ApJ*, 854, 97
- Djorgovski, S. G., Stern, D., Mahabal, A. A., & Brunner, R. 2003, *ApJ*, 596, 67
- Đurovčiková, D., Eilers, A.-C., Chen, H., et al. 2024, *ApJ*, 969, 162
- Efstathiou, G., & Rees, M. J. 1988, *MNRAS*, 230, 5p
- Eftekharzadeh, S., Myers, A. D., White, M., et al. 2015, *MNRAS*, 453, 2779
- Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017, *ApJ*, 840, 24
- Eilers, A.-C., Hennawi, J. F., & Davies, F. B. 2018, *ApJ*, 867, 30
- Eilers, A.-C., Hennawi, J. F., Decarli, R., et al. 2020, *ApJ*, 900, 37
- Eilers, A.-C., Simcoe, R. A., Yue, M., et al. 2023, *ApJ*, 950, 68
- Ferrarese, L., & Merritt, D. 2000, *ApJL*, 539, L9
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- García-Vergara, C., Hennawi, J. F., Barrientos, L. F., & Rix, H.-W. 2017, *ApJ*, 848, 7
- Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJL*, 539, L13
- Gilli, R., Norman, C., Calura, F., et al. 2022, *A&A*, 666, A17
- Greene, J. E., Labbe, I., Goulding, A. D., et al. 2024, *ApJ*, 964, 39
- Haiman, Z., & Hui, L. 2001, *ApJ*, 547, 27
- Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, *ApJ*, 959, 39
- Håring, N., & Rix, H.-W. 2004, *ApJL*, 604, L89
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, 585, 357
- He, W., Akiyama, M., Bosch, J., et al. 2018, *PASJ*, 70, S33
- Hennawi, J. F., Strauss, M. A., Oguri, M., et al. 2006, *AJ*, 131, 1
- Herard-Demanche, T., Bouwens, R. J., Oesch, P. A., et al. 2023, arXiv:2309.04525
- Hunter, J. D. 2007, *CSE*, 9, 90
- Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, *MNRAS*, 506, 4210
- Jones, E., Oliphant, T., Peterson, P., et al., 2001 SciPy: Open Source Scientific Tools for Python
- Kaiser, N. 1984, *ApJL*, 284, L9
- Kashino, D., Lilly, S. J., Matthee, J., et al. 2023, *ApJ*, 950, 66
- Khrykin, I. S., Hennawi, J. F., & Worseck, G. 2019, *MNRAS*, 484, 3897
- Kim, S., Stiavelli, M., Trenti, M., et al. 2009, *ApJ*, 695, 809
- Kocevski, D. D., Onoue, M., Inayoshi, K., et al. 2023, *ApJL*, 954, L4
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Laurent, P., Eftekharzadeh, S., Le Goff, J.-M., et al. 2017, *JCAP*, 2017, 017
- Lyu, J., Alberts, S., Rieke, G. H., et al. 2024, *ApJ*, 966, 229
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285
- Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2023, arXiv:2308.01230
- Martini, P. 2004, *Coevolution of Black Holes and Galaxies*, 169 (Cambridge: Cambridge Univ. Press)
- Martini, P., & Weinberg, D. H. 2001, *ApJ*, 547, 12
- Matthee, J., Mackenzie, R., Simcoe, R. A., et al. 2023, *ApJ*, 950, 67
- Matthee, J., Naidu, R. P., Brammer, G., et al. 2024, *ApJ*, 963, 129
- Mazzucchelli, C., Bañados, E., Decarli, R., et al. 2017a, *ApJ*, 834, 83
- Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017b, *ApJ*, 849, 91
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, *Proc. SPIE*, 8446, 84460J
- McLure, R. J., & Jarvis, M. J. 2002, *MNRAS*, 337, 109
- Mignoli, M., Gilli, R., Decarli, R., et al. 2020, *A&A*, 642, L1
- Mo, H. J., & White, S. D. M. 1996, *MNRAS*, 282, 347
- Morey, K. A., Eilers, A.-C., Davies, F. B., Hennawi, J. F., & Simcoe, R. A. 2021, *ApJ*, 921, 88
- Morselli, L., Mignoli, M., Gilli, R., et al. 2014, *A&A*, 568, A1
- Munoz, J. A., & Loeb, A. 2008, *MNRAS*, 385, 2175
- Murray, S. G., Diemer, B., Chen, Z., et al. 2021, *A&C*, 36, 100487
- Murray, S. G., Power, C., & Robotham, A. S. G. 2013, *A&C*, 3, 23
- Myers, A. D., Brunner, R. J., Nichol, R. C., et al. 2007, *ApJ*, 658, 85
- Ni, Y., DiMatteo, T., Gilli, R., et al. 2020, *MNRAS*, 495, 2135
- Nusser, A., & Silk, J. 1993, *ApJL*, 411, L1
- Osmer, P. S. 1981, *ApJ*, 247, 762
- Overzier, R. A. 2016, *A&ARv*, 24, 14
- Padmanabhan, N., White, M., Norberg, P., & Porciani, C. 2009, *MNRAS*, 397, 1862
- Peng, Y.-J., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, 721, 193
- Pizzati, E., Hennawi, J. F., Schaye, J., & Schaller, M. 2024, *MNRAS*, 528, 4466
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A&A*, 641, A6
- Porciani, C., Magliocchetti, M., & Norberg, P. 2004, *MNRAS*, 355, 1010
- Press, W. H., & Schechter, P. 1974, *ApJ*, 187, 425
- Ross, N. P., Shen, Y., Strauss, M. A., et al. 2009, *ApJ*, 697, 1634
- Salpeter, E. E. 1964, *ApJ*, 140, 796
- Satyavolu, S., Eilers, A.-C., Kulkarni, G., et al. 2023b, *MNRAS*, 522, 4918
- Satyavolu, S., Kulkarni, G., Keating, L. C., & Haehnelt, M. G. 2023a, *MNRAS*, 521, 3108
- Schaller, M., Borrow, J., Draper, P. W., et al. 2024, *MNRAS*, 530, 2378
- Schaye, J., Kugel, R., Schaller, M., et al. 2023, *MNRAS*, 526, 4978
- Schindler, J.-T., Bañados, E., Connor, T., et al. 2023, *ApJ*, 943, 67
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, 500, 33
- Shanks, T., & Boyle, B. J. 1994, *MNRAS*, 271, 753
- Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, *PASP*, 114, 851
- Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, *AJ*, 133, 2222
- Sijacki, D., Springel, V., & Haehnelt, M. G. 2009, *MNRAS*, 400, 100
- Simcoe, R. A., Burgasser, A. J., Bernstein, R. A., et al. 2008, *Proc. SPIE*, 7014, 70140U
- Simpson, C., Mortlock, D., Warren, S., et al. 2014, *MNRAS*, 442, 3454
- Small, T. A., & Blandford, R. D. 1992, *MNRAS*, 259, 725
- Sun, F., Helton, J. M., Egami, E., et al. 2024, *ApJ*, 961, 69
- Tanaka, T., & Haiman, Z. 2009, *ApJ*, 696, 1798
- Thorne, K. S. 1974, *ApJ*, 191, 507
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *ApJ*, 724, 878
- Trainor, R., & Steidel, C. C. 2013, *ApJL*, 775, L3
- Trebtsch, M., Volonteri, M., & Dubois, Y. 2019, *MNRAS*, 487, 819
- Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, 536, A105
- Vestergaard, M., & Peterson, B. M. 2006, *ApJ*, 641, 689
- Vito, F., DiMascia, F., Gallerani, S., et al. 2022, *MNRAS*, 514, 1672
- Volonteri, M. 2012, *Sci*, 337, 544
- Volonteri, M., & Rees, M. J. 2006, *ApJ*, 650, 669
- Wang, F., Yang, J., Fan, X., et al. 2019, *ApJ*, 884, 30
- Wang, F., Yang, J., Fan, X., et al. 2021, *ApJL*, 907, L1
- Wang, F., Yang, J., Hennawi, J. F., et al. 2023, *ApJL*, 951, L4
- White, M., Martini, P., & Cohn, J. D. 2008, *MNRAS*, 390, 1179
- White, M., Myers, A. D., Ross, N. P., et al. 2012, *MNRAS*, 424, 933
- Yang, J., Wang, F., Fan, X., et al. 2020, *ApJL*, 897, L14
- Yu, Q., & Tremaine, S. 2002, *MNRAS*, 335, 965
- Yue, M., Eilers, A.-C., Simcoe, R. A., et al. 2024, *ApJ*, 966, 176
- Zhang, H., Behroozi, P., Volonteri, M., et al. 2023, *MNRAS*, 518, 2123
- Zhang, H., Behroozi, P., Volonteri, M., et al. 2024, *MNRAS*, 531, 4974