



# Subaru High- $z$ Exploration of Low-luminosity Quasars (SHELLQs). XVIII. The Dark Matter Halo Mass of Quasars at $z \sim 6$

Junya Arita<sup>1</sup> , Nobunari Kashikawa<sup>1,2</sup> , Yoshiki Matsuoka<sup>3</sup> , Wanqiu He<sup>4</sup> , Kei Ito<sup>1</sup> , Yongming Liang<sup>5</sup> , Rikako Ishimoto<sup>1</sup> , Takehiro Yoshioka<sup>1</sup> , Yoshihiro Takeda<sup>1</sup> , Kazushi Iwasawa<sup>6,7</sup> , Masafusa Onoue<sup>8,9</sup> , Yoshiki Toba<sup>3,4,10</sup> , and Masatoshi Imanishi<sup>4</sup>

<sup>1</sup> Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan; [jarita@astron.s.u-tokyo.ac.jp](mailto:jarita@astron.s.u-tokyo.ac.jp)

<sup>2</sup> Research Center for the Early Universe, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>3</sup> Research Center for Space and Cosmic Evolution, Ehime University, 2-5 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

<sup>4</sup> National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>5</sup> Institute for Cosmic Ray Research, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan

<sup>6</sup> Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès, 1, E-08028 Barcelona, Spain

<sup>7</sup> ICREA, Pg. Lluís Companys 23, E-08010 Barcelona, Spain

<sup>8</sup> Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China

<sup>9</sup> Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), The University of Tokyo, Chiba 277-8583, Japan

<sup>10</sup> Academia Sinica Institute of Astronomy and Astrophysics, 11F of Astronomy-Mathematics Building, AS/NTU, No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan

Received 2023 April 14; revised 2023 July 3; accepted 2023 July 3; published 2023 September 8

## Abstract

We present, for the first time, dark matter halo (DMH) mass measurement of quasars at  $z \sim 6$  based on a clustering analysis of 107 quasars. Spectroscopically identified quasars are homogeneously extracted from the Hyper Suprime-Cam Strategic Survey Program wide layer over  $891 \text{ deg}^2$ . We evaluate the clustering strength by three different autocorrelation functions: projected correlation function, angular correlation function, and redshift-space correlation function. The DMH mass of quasars at  $z \sim 6$  is evaluated as  $5.0^{+7.4}_{-4.0} \times 10^{12} h^{-1} M_{\odot}$  with the bias parameter  $b = 20.8 \pm 8.7$  by the projected correlation function. The other two estimators agree with these values; though, each uncertainty is large. The DMH mass of quasars is found to be nearly constant  $\sim 10^{12.5} h^{-1} M_{\odot}$  throughout cosmic time, suggesting that there is a characteristic DMH mass where quasars are always activated. As a result, quasars appear in the most massive halos at  $z \sim 6$ , but in less extreme halos thereafter. The DMH mass does not appear to exceed the upper limit of  $10^{13} h^{-1} M_{\odot}$ , which suggests that most quasars reside in DMHs with  $M_{\text{halo}} < 10^{13} h^{-1} M_{\odot}$  across most of the cosmic time. Our results supporting a significant increasing bias with redshift are consistent with the bias evolution model with inefficient active galactic nucleus feedback at  $z \sim 6$ . The duty cycle ( $f_{\text{duty}}$ ) is estimated as  $0.019 \pm 0.008$  by assuming that DMHs in some mass interval can host a quasar. The average stellar mass is evaluated from stellar-to-halo mass ratio as  $M_* = 6.5^{+9.6}_{-5.2} \times 10^{10} h^{-1} M_{\odot}$ , which is found to be consistent with [C II] observational results.

Unified Astronomy Thesaurus concepts: Large-scale structure of the universe (902); Quasars (1319); Supermassive black holes (1663)

## 1. Introduction

According to the current  $\Lambda$  cold dark matter (CDM) theory, the tiny density fluctuation of dark matter (DM) in the early universe grows and subsequently collapses into dark matter halos (DMHs). These halos continuously accrete and hierarchically merge to form high-mass DMHs. Galaxies are nurtured in the center of DMHs, and almost all the galaxies harbor a supermassive black hole (SMBH) in their centers (Kormendy & Richstone 1995). Quasars are believed to be powered by gas accretion onto SMBHs (Salpeter 1964; Lynden-Bell 1969) and outshine in the multiple wavelengths. Since quasars are one of the most luminous objects in the universe, they are observable even at  $z \gtrsim 7$  (e.g., Mortlock et al. 2011; Bañados et al. 2018; Matsuoka et al. 2019; Yang et al. 2020; Koptelova & Hwang 2022). Quasars are important objects to study the open questions in the early universe; however, it remains unclear how high- $z$  quasars are physically related to the underlying DMHs they inhabit.

One of the important questions is when and how the coevolution between galaxies and SMBHs manifested, i.e., the masses of which are correlated with those of their host galaxies. While this relationship in the local universe is well established (Magorrian et al. 1998; Tremaine et al. 2002; Kormendy & Ho 2013), it remains to be elucidated in the early universe. The parent DMH, which governs both the SMBH and the galaxy, holds the key to unveiling the underlying physical mechanism of their relationship. The gas accumulated by the gravitational potential of DMHs is consumed to form stars; thus, the relationship between stellar mass and DMH mass is quite natural (White & Rees 1978). It is believed that the gas further loses the angular momentum due to the radiation from active star formation in the DMH and flows into the central SMBH (Ferrarese 2002; Granato et al. 2004; Shankar et al. 2006) to grow more massive. Otherwise, the steady high density cold gas flow directly from the halo could be responsible for sustaining critical accretion rates leading to rapid growth of  $\sim 10^9 M_{\odot}$  black holes as early as  $z \sim 7$  (e.g., Di Matteo et al. 2012). Therefore, the mass of a galaxy DMH hosting an SMBH is crucial for understanding their coevolutionary growth.

The DMH mass is also a key physical quantity to understand the active galactic nucleus (AGN) feedback, which is thought



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

to play a significant role in regulating the star formation of the host galaxies (Kormendy & Ho 2013; Heckman & Best 2014), because it can constrain the duty cycle, the fraction of DMHs that host active quasars (e.g., Cole & Kaiser 1989; Haiman & Hui 2001; Martini & Weinberg 2001). Hopkins et al. (2007) showed that feedback efficiency will greatly change DMH mass evolution at high- $z$ . According to their model, feedback prevents gas accretion against gravity by radiation pressure, and works to stop SMBH growth and eventually defuses the quasar phase. Thus, if feedback is inefficient to stop the SMBH growth at high- $z$ , quasars will live in the highest-mass DMHs. Since the quasar activity has a huge impact on the host galaxy, unveiling the feedback efficiency helps to advance our understanding of the coevolution.

The clustering analysis is an effective method to estimate DMH mass. It quantifies the distribution of objects often through a two-point correlation function. The two-point correlation function  $\xi(r)$  is defined based on the probability  $dP$  that an object is observed in the volume element  $dV$  apart from the separation  $r$  from a given object (Totsuji & Kihara 1969; Peebles 1980);

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

where  $\bar{n}$  is the mean number density of the objects. Quasar host galaxies are believed to reside in the peak of the underlying DM density distribution (Dekel & Lahav 1999). Using the bias (linear bias) parameter  $b$ , the relation between two-point correlation functions of quasars  $\xi_Q(r)$  and that of DM  $\xi_{\text{DM}}(r)$  can be expressed as

$$\xi_Q(r) = b^2 \xi_{\text{DM}}(r). \quad (2)$$

The bias parameter has been modeled theoretically (e.g., Jing 1998; Sheth et al. 2001; Seljak & Warren 2004; Mandelbaum et al. 2005; Tinker et al. 2010), which gives insight into the DMH mass.

The previous initiative works (Porciani et al. 2004; Croom et al. 2005; Myers et al. 2007) have proved that the quasar bias increases with redshift from today to  $z=2\text{--}3$ . However, clustering analyses of quasars at  $z>3$ , there are a few attempts. Shen et al. (2007) utilized 4426 spectroscopically identified luminous quasars with  $2.9 \leq z \leq 5.4$  from the Fifth Data Release (DR5) of the Sloan Digital Sky Survey (SDSS; Schneider et al. 2007) and concluded that quasars typically reside in DMHs with  $(2\text{--}3) \times 10^{12} h^{-1} M_\odot$  at  $2.9 \leq z \leq 3.5$  and  $(4\text{--}6) \times 10^{12} h^{-1} M_\odot$  at  $3.5 \leq z \leq 5.4$ . Eftekharzadeh et al. (2015) measured the clustering signal of quasars from the Baryon Oscillation Spectroscopic Survey (Smee et al. 2013). They estimated the DMH mass for quasars at  $2.2 \leq z \leq 3.4$  as  $0.6\text{--}3 \times 10^{12} h^{-1} M_\odot$ . He et al. (2018) extracted photometrically selected 901 quasars with  $\bar{z}_{\text{phot}} \sim 3.8$  from the early data release of the Subaru Hyper Suprime-Cam Strategic Survey Program (HSC-SSP; Aihara et al. 2018a, 2018b). They added 342 SDSS quasars (Alam et al. 2015) to their sample and evaluated cross-correlation functions (CCF) between the quasars and bright Lyman break galaxies (LBGs) from the HSC-SSP. The typical DMH mass derived from the CCF signal is  $1\text{--}2 \times 10^{12} h^{-1} M_\odot$ . Timlin et al. (2018) measured the clustering signal of photometrically selected quasars with  $2.9 \leq z_{\text{phot}} \leq 5.1$  from SDSS Stripe 82 field (Annis et al. 2014; Jiang et al. 2014) and presumed that characteristic DMH mass is  $1.70\text{--}9.83 \times 10^{12} h^{-1} M_\odot$ . These studies detected significantly large clustering signals, implying that the quasar halo

bias rapidly increases beyond  $z \sim 3$ . In addition, the quasar DMH mass remains approximately constant at  $M_{\text{halo}} \sim 10^{12.5} h^{-1} M_\odot$  from the present day to  $z \sim 4$ , which will be intriguing to see whether these trends continue to higher- $z$ .

Despite intense observational efforts, the clustering measurements have been challenging beyond  $z > 4$ . This is because the clustering analysis requires a quasar sample with sufficient number density, which remarkably decreases toward  $z \sim 6$  (Fan et al. 2022). The sample size and the number density of quasars at  $z \sim 6$  have increased dramatically in the last two decades, but the observable quasar population is limited to high-luminosity obtained by ultra wide-field surveys, hindering the increase in their number density. Increasing the number density of quasars at  $z \sim 6$  has been a major challenge because of the need for wide and deep observations and expensive spectroscopic observations for fainter quasars. HSC (Furusawa et al. 2018; Komiyama et al. 2018; Kawanomoto et al. 2018; Miyazaki et al. 2018) on the Subaru Telescope, which has a large field of view and high sensitivity, has changed the situation. Utilizing the powerful instrument, a wide-field imaging survey program, HSC-SSP, was performed. From the survey data, Subaru High- $z$  Exploration of Low-Luminosity Quasars (SHELLQs; Matsuoka et al. 2016, 2022) has discovered 162 quasars at  $5.66 < z < 7.07$  over  $1200 \text{ deg}^2$ , providing a high number density to allow for clustering measurements.

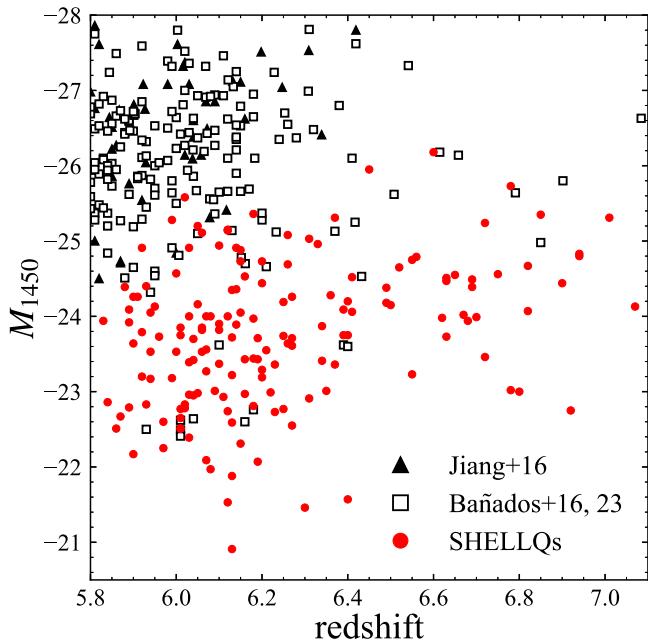
In this paper, we, for the first time, present the clustering analysis of quasars at  $z \sim 6$  by using the SHELLQs sample. We show the samples for our analysis in Section 2. We explain the details of clustering analysis in Section 3. In Section 4, we derive the important physical quantities from the result in the previous section. Finally, we summarize our results in Section 5. We adopt flat  $\Lambda$ CDM cosmology with cosmology parameters  $(h, \Omega_m, \Omega_\Lambda, \sigma_8) = (0.7, 0.3, 0.7, 0.81)$ , namely,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . All magnitudes in this paper are presented in the AB system (Oke & Gunn 1983).

## 2. Data

### 2.1. SHELLQs

Our main quasar sample is from SHELLQs utilizing HSC-SSP data. The HSC data are reduced with HSC pipeline, `hscpipe` (Bosch et al. 2018), which is based on the Large Synoptic Survey Telescope (LSST) pipeline (Jurić et al. 2017; Bosch et al. 2018; Ivezić et al. 2019). The astrometry and photometric calibration are performed based on the data from Panoramic Survey Telescope and Rapid Response System Data Release 1 (Pan-STARRS1; Schlafly et al. 2012; Tonry et al. 2012; Magnier et al. 2013; Chambers et al. 2016). The SHELLQs quasars are a flux-limited ( $m_z < 24.5$  for  $z \sim 6$  and  $m_y < 24$  for  $z \sim 7$ ) sample of quasars at  $z \sim 6\text{--}7$ . These quasars are selected from point sources and by a Bayesian-based probabilistic algorithm, which is applied to the optical HSC-SSP source catalogs. More details of the sample construction are described in Matsuoka et al. (2016). The spectroscopic observation is performed by utilizing the Faint Object Camera and Spectrograph (Kashikawa et al. 2002) on the Subaru Telescope and Optical System for Imaging and low-intermediate-Resolution Integrated Spectroscopy (Cepa et al. 2000) on the Gran Telescopio Canarias.

The advantages of the SHELLQs sample are faintness and high number density thanks to the depth of the HSC-SSP. Figure 1 shows the comparison of the absolute magnitude of



**Figure 1.** Comparison of  $M_{1450}$ , the absolute magnitudes at 1450 Å of SHELLQs quasars (red circles) and those of known quasars. We refer to the  $M_{1450}$  from the quasar sample in Jiang et al. (2016; black triangle) and that of newly identified quasars in Bañados et al. (2016), Bañados et al. (2023; black open square).

quasars detected in SHELLQs, SDSS (Jiang et al. 2016), and Pan-STARRS1 (Bañados et al. 2016, 2023), where it is clear that SHELLQs is exploring a unique regime fainter than those in other surveys. The SHELLQs have a number density ( $0.14 \text{ deg}^{-2}$ ) of quasars  $\sim 30$  times more than the SDSS (Jiang et al. 2016), where 52 quasars are detected in  $11,240 \text{ deg}^2$  at  $5.7 < z \leqslant 6.4$ .

The original SHELLQs sample consists of 162 spectroscopically confirmed quasars. We impose the following four criteria to ensure homogeneity, which yields 93 quasars (see Table 1).

- The first criterion is that the quasars are at  $z \leq 6.5$ . The SHELLQs sample consists of  $z \sim 6$  quasars selected by  $i$ -dropouts and  $z \sim 7$  quasars selected by  $z$ -dropouts. Since the latter has a small sample size and the survey areas of the two do not perfectly match, only the former  $z \sim 6$  quasars are used in this study to ensure uniformity of the sample. The  $z \sim 6$  quasar sample selection criteria (Matsuoka et al. 2022) is

$$m_z < 24.5 \text{ \& } \sigma_z < 0.155 \text{ \& } m_i - m_z > 1.5 \\ \text{\& } 0.7 < \mu/\mu_{\text{PSF}} < 1.2, \quad (3)$$

where  $\mu$  is the adaptive moment of the source averaged over the two image dimensions, and  $\mu_{\text{PSF}}$  is that of the point-spread function (PSF) model.

2. The second criterion is that the quasars are identified in the HSC-SSP S20A region. The SHELLQs sample is still growing. Optical spectroscopic follow-up observations have been completely executed in the S20A survey area. This study uses only SHELLQs quasars spectroscopically confirmed in the S20A region, and the quasars added after S21A are removed to account for the uniformity of the sample.

**Table 1**  
Detail of the Sample Selection

Requirement	Number
All quasars identified by SHELLQs	162
1. $z \leqslant 6.5$	132
2. Identified in HSC-SSP S20A	125
3. Far from bright star masks and edge regions	116
4. Broad line quasars	93

**Note.** We add 14 known quasars into this sample. The additional quasars are listed in Table 2.

- The third criterion is that the quasars are far from bright star masks and edge regions. We remove quasars in areas with poor data quality, such as near the bright star masks and edges by random points covering HSC-SSP S20A region to preserve sample homogeneity.<sup>11</sup>
  - The last criterion is that the quasars have a broad line component in their spectra. According to the unified AGN model (Antonucci 1993), type I AGNs and type II AGNs are the same population, and the difference purely originates from the inclination angle to observers. However, another evolutionary scenario interprets the difference between the two populations in host galaxies. The DMH mass measurements by Hickox et al. (2011) found the differences in the DMH mass between obscured and unobscured quasars through clustering analysis. Onoue et al. (2021) reported that approximately 20% of SHELLQs quasars have narrow Ly $\alpha$  emission lines ( $\text{FWHM}_{\text{Ly}\alpha} < 500 \text{ km s}^{-1}$ ), and one of them can be a type II quasar based on the spectroscopic follow-up. Therefore, as long as it is unclear which interpretation is correct, we decide to be conservative in the study to exclude quasars with narrow Ly $\alpha$  emission lines from the sample.

## 2.2. Other Quasars

We also add to our sample 14 quasars at  $z \sim 6$  that were discovered by other surveys (see Table 2). We select these quasars from the survey, whose area fully covers HSC-SSP S20A field. Most of them are identified by SDSS (e.g., Fan et al. 2004) and Pan-STARRS1 (e.g., Bañados et al. 2014, 2016; Mazzucchelli 2017), which tend to be brighter than the SHELLQs quasars. These quasars also satisfy the requirements imposed on SHELLQs quasars. Ten of the 14 quasars are also detected in the SHELLQs observation but not included in the SHELLQs sample because they had already been found (Matsuoka et al. 2018a). We visually inspect the spectra of all these quasars to confirm that they are actually  $z \sim 6$  quasars. We assume that clustering strength is independent of quasar brightness, which is confirmed at low- $z$  (Croom et al. 2005; Adelberger et al. 2006; Myers et al. 2006; Shen et al. 2009). In fact, when we divide the sample into a bright subsample ( $M_{1450} \leq -24$ ) and faint subsample ( $M_{1450} > -24$ ), the results obtained in Section 3 are consistent with each other within their errors. In summary, our

<sup>11</sup> Specifically, we use the following flags to retrieve random points covering the survey field with a surface number density of 100 arcmin<sup>-2</sup>: `{i,z,y}_pixelflags_edge`, `{i,z,y}_pixelflags_saturatedcenter`, `{i,z,y}_pixelflags_crcenter`, `{i,z,y}_pixelflags_bad`, `{i,z,y}_mask_brightstar_[halo, ghost, blooming]`, and impose `z_inputcount_value ≥ 2`. We exclude quasars that have no random points within 0.12 arcmin from the clustering analysis.

**Table 2**  
Additional Quasar Sample

Name	R.A. (J2000)	Decl. (J2000)	Redshift	$M_{1450}$ (mag)	Survey	Reference
SDSS J160254.18+422822.9	16:02:54.18	+42:28:22.9	6.07	-26.82	SDSS	(1)
SDSS J000552.33-000655.6	00:05:52.33	-00:06:55.6	5.855	-26.46	SDSS	(1)
CFHQS J021013-045620 <sup>a</sup>	02:10:13.19	-04:56:20.9	6.44	-24.28	CFHQS <sup>b</sup>	(2)
CFHQS J021627-045534 <sup>a</sup>	02:16:27.81	-04:55:34.1	6.01	-22.21	CFHQS	(3)
CFHQS J022743-060530 <sup>a</sup>	02:27:43.29	-06:05:30.2	6.20	-25.03	CFHQS	(3)
IMS J220417.92+011144.8 <sup>a</sup>	22:04:17.92	+01:11:44.8	5.944	-23.59	IMS <sup>c</sup>	(4)
VIMOS2911001793 <sup>a</sup>	22:19:17.22	+01:02:48.9	6.156	-23.10	Suprime Cam	(5)
SDSS J222843.5+011032.2 <sup>a</sup>	22:28:43.54	+01:10:32.2	5.95	-24.53	SDSS Stripe82	(6)
SDSS J230735.35+003149.4 <sup>a</sup>	23:07:35.35	+00:31:49.4	5.87	-24.93	SDSS	(7)
SDSS J231546.57-002358.1 <sup>a</sup>	23:15:46.57	-00:23:58.1	6.117	-25.38	SDSS	(8)
PSO J183.1124+05.0926	12:12:26.98	+05:05:33.4	6.439	-26.99	Pan-STARRS1	(9)
VIK J121516.88+002324.7 <sup>a</sup>	12:15:16.88	+00:23:24.7	5.93	-24.67	VIKING <sup>e</sup>	(10)
PSO J184.3389+01.5284 <sup>a</sup>	12:17:21.34	+01:31:42.2	6.20	-25.37	Pan-STARRS1	(11)
PSO J187.3050+04.3243	12:29:13.21	+04:19:27.7	5.89	-25.4	Pan-STARRS1	(12)

**Notes.** (1) Fan et al. (2004), (2) Willott et al. (2010), (3) Willott et al. (2009), (4) Kim et al. (2015), (5) Kashikawa et al. (2015), (6) Zeimann et al. (2011), (7) Jiang et al. (2009), (8) Jiang et al. (2008), (9) Mazzucchelli (2017), (10) Venemans et al. (2015), (11) Bañados et al. (2016), and (12) Bañados et al. (2014).

<sup>a</sup> The quasar is recovered by SHELLQs project (Matsuoka et al. 2018a).

<sup>b</sup> Canada-France High- $z$  Quasar Survey.

<sup>c</sup> Infrared Medium-deep Survey.

<sup>d</sup> VISTA Kilo-degree Infrared Galaxy Public Survey.

final sample consists of 107 quasars. The distributions of absolute magnitude and redshift are shown in Figure 2.

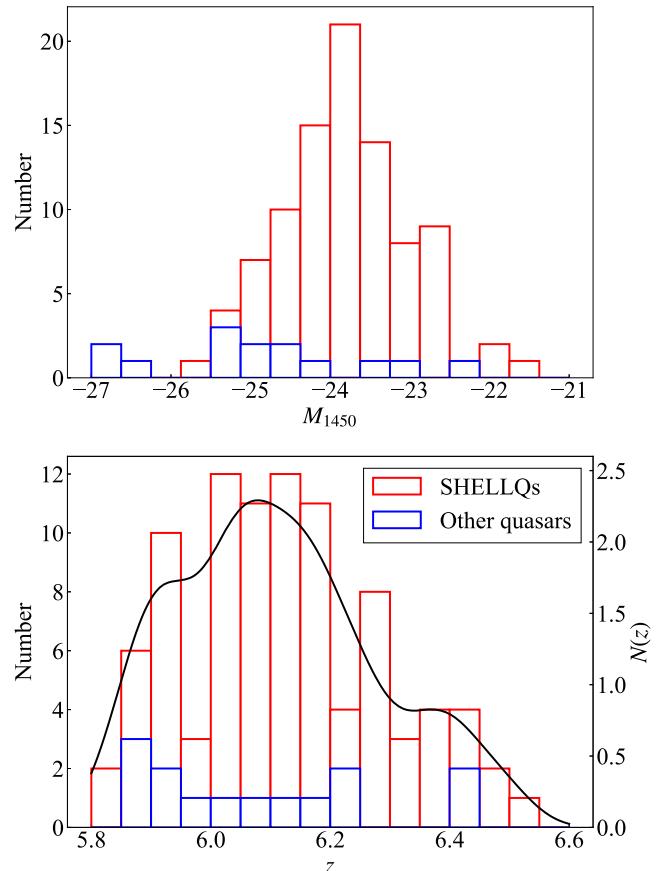
### 2.3. Homogeneity of the Sample

Our sample is distributed over 891 deg<sup>2</sup> of three fields: HECTOMAP, Autumn, and Spring. The sky distribution of our sample quasars is shown in Figure 3. Sample homogeneity is of utmost importance for clustering analysis. Since the spectroscopy is completely executed for the candidates in S20A, the spatial homogeneity of the photometric data in selecting candidate objects should be examined. We verify the homogeneity by calculating detection completeness over the survey region. The detection completeness is defined as the ratio of the number of quasars recovered by `hscpipe` ver. 8.4<sup>12</sup> to the number of mock quasars scattered at random points on the HSC image in the same manner in Matsuoka et al. (2018b). The PSF of the input mock quasars is generated to be the same as that measured at each image position. The PSF is modeled by PSFEx (Bertin 2011), which can extract precise PSF model from images processed by SExtractor (Bertin & Arnouts 1996). A small region (patch) of 12' × 12' is randomly selected per tract, which consists of 81 patches, over the HSC-SSP region. We embed more than 3000 mock quasars in the survey field per patch with  $m_z = 21\text{--}28$ , which are randomly spread over the HSC-SSP coadded  $z$  band images using Balrog (Suchyta et al. 2016). We perform photometry on the coadded  $z$ -band images embedded mock quasars utilizing `hscpipe` and detect the mock quasars. The detection completeness is estimated on 662 patches in total.

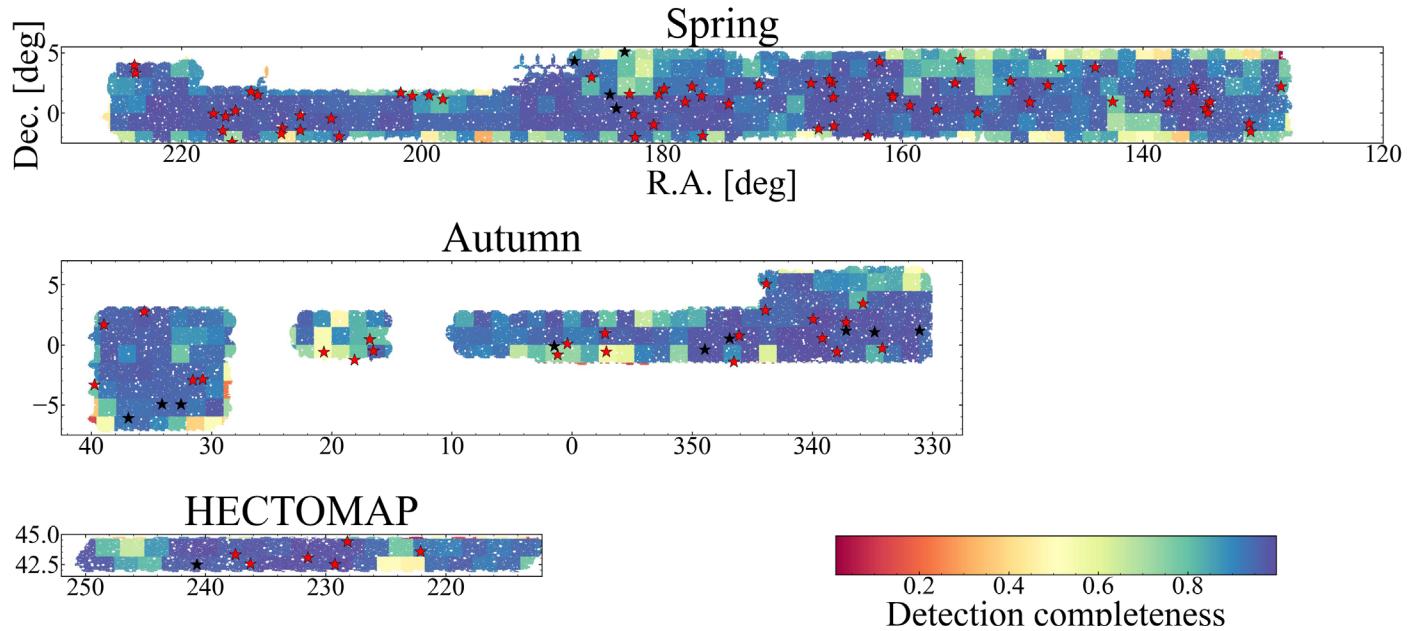
The detection completeness is fitted with a function in Sejeant et al. (2000);

$$f_{\text{det}}(m_z) = \frac{f_{\text{max}} - f_{\text{min}}}{2} \{\tanh[\alpha(m_z^{50} - m_z)] + 1\} + f_{\text{min}}, \quad (4)$$

<sup>12</sup> [https://hsc.mtk.nao.ac.jp/pipedoc/pipedoc\\_8\\_e/index.html](https://hsc.mtk.nao.ac.jp/pipedoc/pipedoc_8_e/index.html)



**Figure 2.** Top: red and blue histograms show 1450 Å absolute magnitude distribution for SHELLQs quasars and other quasars respectively. Bottom: red and blue histograms show the redshift distribution for SHELLQs quasars and other quasars respectively. The black line represents  $N(z)$ , the redshift distribution of SHELLQs quasars estimated by kernel density estimation, which is used in Section 3.1.



**Figure 3.** Detection completeness map of the HSC-SSP S20A region. The color represents the detection completeness at  $m_z = 24.5$  of each tract. The red and black stars represent the SHELLQs and other quasars, respectively.

where  $f_{\max}$ ,  $f_{\min}$ ,  $\alpha$ , and  $m_z^{50}$  represent the detection completeness at the brightest magnitude and the faintest magnitude, the sharpness of the function, and the magnitude at which the detection completeness is 50%, respectively. Our measurement of each tract is presented in Figure 4, and the best-fit parameters with  $1\sigma$  error for the median completeness are  $f_{\max} = 0.978 \pm 0.015$ ,  $f_{\min} = 0.016 \pm 0.008$ ,  $\alpha = 2.7 \pm 0.8$ , and  $m_z^{50} = 25.08 \pm 0.36$ , which are also denoted in the figure. Almost all the functions have similar parameters with a  $m_z^{50}$  scatter as small as  $\sigma(m_z^{50}) = 0.36$ . Figure 3 shows the completeness map at  $m_z = 24.5$  of the survey region overplotted with the sample quasars. The completeness holds more than 70% (80%) over 85% (77%) of the entire survey region, and >50% over almost all the area at the  $z$ -band limiting magnitude, and there are few areas of singularly lowered completeness. It is noted that the following results hardly change when the area with  $f_{\det} < 0.5$  is excluded. Therefore, we conclude that the whole survey area is homogeneous enough to conduct the clustering analysis.

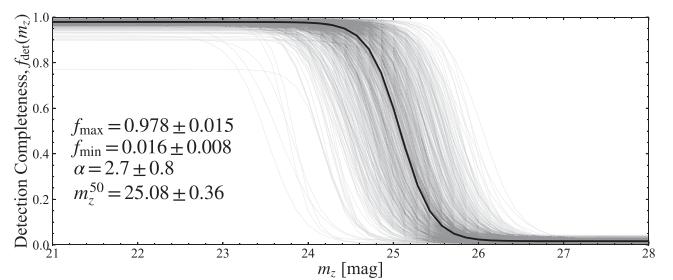
### 3. Clustering Analysis

#### 3.1. Autocorrelation Function of the Quasars

We first measure a projected correlation function  $\omega_p(r_p)$  in Section 3.1.1 that can be directly related to real-space clustering. At the same time, to check the robustness of the result, we also measure an angular correlation function  $\omega(\theta)$  without redshift information in Section 3.1.2 and a redshift-space correlation function  $\xi(s)$ , which includes redshift–space distortion, in Section 3.1.3.

##### 3.1.1. Projected Correlation Function

We evaluate the projected correlation function  $\omega_p$  of the sample. In this analysis, the comoving distance is calculated from the spectroscopic redshift. We separate  $s$ , the 3D distance between two objects, into  $r_p$ , perpendicular to the line of sight,



**Figure 4.** Detection completeness as a function of  $z$ -band magnitude. The black solid line denotes the median completeness of all patches. The thin gray solid lines show the detection completeness of each patch. Each median parameter and  $1\sigma$  error are denoted in the left of the figure.

and  $\pi$ , parallel with it ( $s = \sqrt{r_p^2 + \pi^2}$ ). We estimate the 2D correlation function  $\xi(r_p, \pi)$  from Landy & Szalay (1993);

$$\xi(r_p, \pi) = \frac{\text{DD}(r_p, \pi) - 2\text{DR}(r_p, \pi) + \text{RR}(r_p, \pi)}{\text{RR}(r_p, \pi)} \quad (5)$$

where  $\text{DD}(r_p, \pi)$ ,  $\text{DR}(r_p, \pi)$ ,  $\text{RR}(r_p, \pi)$  represent data–data, data–random, and random–random pair counts within perpendicular distance separation  $r_p$  and parallel distance separation  $\pi$ , respectively. The survey area is divided into three independent fields; therefore, we count the pairs in each field to sum them up before being normalized by all pairs. The random points are retrieved from the random catalog in HSC-SSP DR3, which has random points scattered over the entire effective survey area, excluding mask areas, at a surface density of  $1 \text{ arcmin}^{-2}$ . The total number of random points is 3,209,416. The redshift of random points is assigned to follow the  $N(z)$ , which is the redshift distribution of SHELLQs estimated by kernel density estimation (see Figure 2). To count pairs, we use `Corrfunc`,<sup>13</sup> which is a Python package containing routines for clustering

<sup>13</sup> <https://github.com/manodeep/Corrfunc>

analysis. We also use the package in Sections 3.1.2, 3.1.3, and 3.2. The projected correlation function  $\omega_p(r_p)$  is derived by integrating  $\xi(r_p, \pi)$  with  $\pi$  direction.

$$\omega_p(r_p) = 2 \int_0^{\pi_{\text{cutoff}}} \xi(r_p, \pi) d\pi; \quad (6)$$

where  $\pi_{\text{cutoff}}$ , which is the optimum limit above which the signal is almost negligible, is fixed to  $\pi_{\text{cutoff}} = 80 h^{-1}\text{Mpc}$  after sufficient trial and error. The redshift distortion is eliminated through the integration (Eftekharzadeh et al. 2015); though, the angular scale of the redshift distortion is much smaller ( $< 20 h^{-1}\text{Mpc}$ ) than the scales of our measurements.

The uncertainty of  $\omega_p(r_p)$  is evaluated by Jackknife resampling (Zehavi et al. 2005). In the  $k$ th resampling, we exclude  $k$ th subregion, and calculate the correlation function,  $\omega_{p,k}(r_p)$ . We divide the survey area into  $N=21$  subregions using the  $k$ -means method. In this case, the covariance matrix is defined as

$$C_{ij} = \frac{N-1}{N} \sum_{k=1}^N (\omega_{p,k}(r_{p,i}) - \bar{\omega}_p(r_{p,i})) \times (\omega_{p,k}(r_{p,j}) - \bar{\omega}_p(r_{p,j})) \quad (7)$$

where  $\omega_{p,k}(r_{p,i})$  and  $\bar{\omega}_p$  represent the value of  $k$ th projected correlation function for  $i$ th  $r_p$  bin and the mean of the projected correlation function, respectively. The uncertainty of  $\omega_p(r_{p,i})$ ,  $\sigma_i$ , is evaluated as  $\sigma_i = \sqrt{C_{ii}}$ , which is used only for plotting process.

The projected function is related to the real-space correlation function  $\xi(r)$  as (Davis & Peebles 1983)

$$\omega_p(r_p) = 2 \int_{r_p}^{\infty} \frac{r \xi(r)}{\sqrt{r^2 - r_p^2}} dr. \quad (8)$$

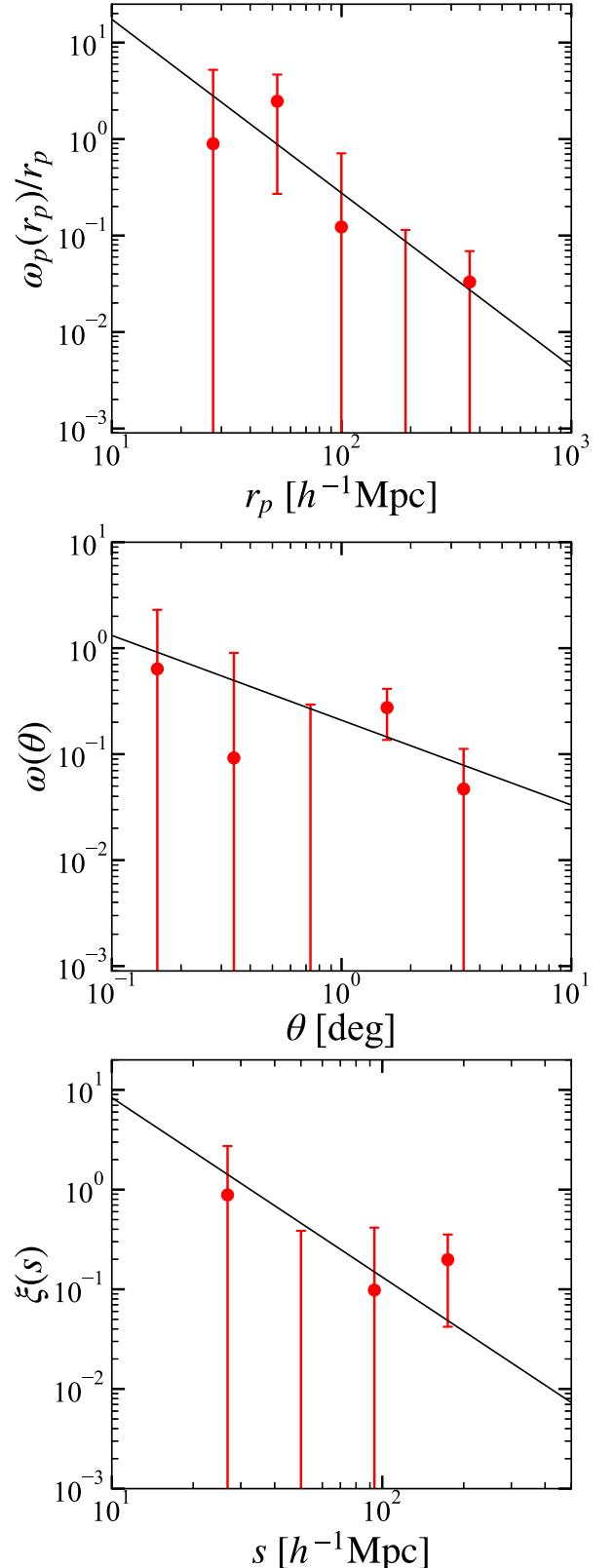
Assuming that the real space correlation function is regarded as a power-law function,  $\xi(r) = (r/r_0)^{-\gamma}$ , the fitted function  $\omega_{p,\text{fit}}(r_p)$  is represented as

$$\frac{\omega_{p,\text{fit}}(r_p)}{r_p} = B \left( \frac{\gamma-1}{2}, \frac{1}{2} \right) \left( \frac{r_p}{r_0} \right)^{-\gamma}, \quad (9)$$

where  $\gamma$  is a power-law index of the DM correlation function,  $B$  represents the beta function, and  $r_0$  is the correlation length, which represents the scale of clustering. In this study, we fit  $\gamma$  to a fiducial value ( $\gamma = 1.8$ ; Peebles 1980). The black solid line in the top panel of Figure 5 represents the power-law function fitted to the projected correlation function based on the  $\chi^2$  fit. Then, we obtain  $r_0 = 23.7 \pm 11 h^{-1}\text{Mpc}$  as listed in Table 3. The goodness-of-fit is evaluated by

$$\chi^2 = \sum_{i,j} [\omega_p(r_{p,i}) - \omega_{p,\text{fit}}(r_{p,i})] C_{ij}^{-1} [\omega_p(r_{p,j}) - \omega_{p,\text{fit}}(r_{p,j})]. \quad (10)$$

To see the robustness of the clustering signal, we integrate the real-space correlation function  $\xi(r)$  within  $r_{\min} \leq r \leq r_{\max}$



**Figure 5.** The top, middle, and bottom panels show the projected correlation function, the angular correlation function, and the redshift-space correlation function, respectively. The error bars represent the  $1\sigma$  error from the Jackknife resampling. The black solid lines represent the power-law functions, which are fitted by  $\chi^2$  fit.

**Table 3**  
DMH Mass from the Clustering Analysis.

Estimator	Correlation Length ( $h^{-1}\text{Mpc}$ )	Bias	DMH Mass ( $10^{12} h^{-1}M_{\odot}$ )	Reduced- $\chi^2$
Projected, $\omega_p$	$23.7 \pm 11$	$20.8 \pm 8.7$	$5.0^{+7.4}_{-4.0}$	0.87
Angular, $\omega$	$27.0 \pm 8.4$	$23.4 \pm 6.6$	$6.9^{+6.1}_{-4.1}$	0.92
Redshift space, $\xi$	$32.5 \pm 19$	$27.7 \pm 15$	$10.6^{+17.5}_{-9.3}$	0.91
CCF, $\omega_{\text{QG}}$		$19.5 \pm 16$	$4.0^{+14.8}_{-4.0}$	0.52

(e.g., Shen et al. 2007);

$$\xi_{100} = \frac{3}{r_{\max}^3} \int_{r_{\min}}^{r_{\max}} \xi(r) r^2 dr, \quad (11)$$

where  $r_{\min} = 10 h^{-1}\text{Mpc}$ , and  $r_{\max} = 100 h^{-1}\text{Mpc}$ , over which the observed signal is detected in this study. Since we assume  $\xi(r) = (r/r_0)^{-\gamma}$ , Equation (11) reduces to

$$\xi_{100} = \frac{3r_0^\gamma}{(3-\gamma)r_{\max}^3} (r_{\max}^{3-\gamma} - r_{\min}^{3-\gamma}). \quad (12)$$

Adopting  $r_0 = 23.7 \pm 11 h^{-1}\text{Mpc}$ , we obtain  $\xi_{100} = 0.175 \pm 0.147$ . Although the uncertainty of each individual data point is large, the overall clustering signal is found to be positive with a significance of more than  $1\sigma$ .

We also test the robustness in terms of whether the signal can be obtained by chance from a random sample. We extract the same number of random points as our quasar sample, treat them as data points, and evaluate the projected correlation function. Based on 10,000 iteration, we evaluate the probability of obtaining a clustering signal as shown in Figure 5. Counting the number of the projected correlation function that has positive signals in the same bins in Figure 5, we find that there is only a 4% probability of obtaining the clustering signal observed in this study. Hence, we conclude that the signal is not artificial.

### 3.1.2. Angular Correlation Function

We also evaluate the angular correlation function. Then, we use the estimator from Landy & Szalay (1993);

$$\omega(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \quad (13)$$

where  $DD(\theta)$ ,  $DR(\theta)$ ,  $RR(\theta)$  represent the normalized data–data, data–random, and random–random pair counts normalized by whole pair counts within an angular separation  $\theta$ , respectively. The random points are retrieved from the random catalog in HSC-SSP DR3.

The uncertainty of  $\omega(\theta)$  is evaluated by Jackknife resampling in the same manner as that in the previous section. We evaluate the uncertainty from the diagonal elements of the covariance matrix derived from Equation (7) replacing  $\omega_p$  for  $\omega$ . The uncertainty of  $\omega(\theta_i)$ ,  $\sigma_i$ , is evaluated based on the diagonal element of the covariance matrix.

The middle panel of Figure 5 represents the result of the angular correlation function. The black solid line represents the best fit of a single power-law model,  $\omega_{\text{true}}(\theta) = A_\omega \theta^{-\beta}$ , considering the effect of the limited survey area to the

correlation function. Then, we assume the following function:

$$\omega(\theta) = A_\omega \theta^{-\beta} - IC, \quad (14)$$

where  $A_\omega$  is the amplitude,  $\beta$  is the power-law index, and  $IC$  is the integral constraint, which is a negative offset as the survey region is limited (Groth & Peebles 1977). We fix  $\beta$  to 0.8 ( $= \gamma - 1$ ) for consistency with the projected correlation function. We evaluate the integral constraint based on Woods & Fahlman (1997);

$$IC = \frac{1}{\Omega^2} \iint \omega_{\text{true}}(\theta) d\Omega_1 d\Omega_2, \quad (15)$$

where  $\Omega$  represents the solid angle of the survey field. The integral constraint becomes considerably smaller than the clustering signal in all fields. Therefore, the integral constraint is ignored in this study. We assume the error follows the Gaussian function, and evaluate the goodness-of-fit of the fitted function through Equation (10) replacing  $\omega_p$  for  $\omega$ .

We convert the amplitude  $A_\omega$  into the correlation length  $r_0$  based on Limber (1953), which formulated

$$r_0 = \left\{ A_\omega \frac{c}{H_0 H_\gamma} \frac{\left[ \int N(z) dz \right]^2}{\int N^2(z) \chi(z)^{1-\gamma} E(z) dz} \right\}^{1/\gamma}, \quad (16)$$

where

$$H_\gamma = B\left(\frac{\gamma-1}{2}, \frac{1}{2}\right), \quad (17)$$

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}, \quad (18)$$

$$\text{and } \chi(z) = \frac{1}{H_0} \int_0^z \frac{1}{E(z')} dz'. \quad (19)$$

Finally, we obtain  $r_0 = 27.0 \pm 8.4 h^{-1}\text{Mpc}$ . The result is listed in Table 3.<sup>14</sup> Adopting  $r_0 = 27.0 \pm 8.4 h^{-1}\text{Mpc}$ , Equation (12) gives  $\xi_{100} = 0.222 \pm 0.124$ , which suggests that the clustering signal is actually detected.

### 3.1.3. Redshift-Space Correlation Function

We also evaluate the redshift–space correlation function of the quasars. All redshifts for SHELLQs quasars are measured in Ly  $\alpha$  emission lines, which has an uncertainty up to  $\Delta z \sim 0.1$ , in particular for those without clear Ly $\alpha$  emission (Matsuoka et al. 2022). This uncertainty and the redshift distortion due to peculiar velocity induce systematic bias in the redshift–space correlation. We derive the redshift–space correlation function  $\xi(s, \mu)$ , where  $s$  is the 3D distance, and  $\mu$  represents the cosine of the angle to the line of sight, utilizing the estimator from Landy & Szalay (1993);

$$\xi(s, \mu) = \frac{DD(s, \mu) - 2DR(s, \mu) + RR(s, \mu)}{RR(s, \mu)}, \quad (20)$$

where  $DD(s, \mu)$ ,  $DR(s, \mu)$ ,  $RR(s, \mu)$  represent data–data, data–random, and random–random pair counts within a separation  $s$  and an angular separation  $\arccos\mu$ , respectively. The correlation function of the entire survey is evaluated by summing the

<sup>14</sup> We note that the obtained amplitude is consistent with that in T. Shinohara et al. (2023, in preparation), which also evaluate the angular correlation function from 92 quasars at  $5.88 < z < 6.49$  including 81 SHELLQs quasars, although samples are not an exact match.

whole pair counts. As mentioned in Section 3.1.1, the redshift of random points is assigned to follow  $N(z)$ , the distribution function in Figure 2. The redshift-space correlation function decomposed into multipoles  $\xi_l(s)$  is derived by integrating  $\xi(s, \mu)$  by  $\mu$  (Marín et al. 2016);

$$\xi_l(s) = \frac{2l+1}{2} \int_{-1}^1 \xi(s, \mu) L_l(\mu) d\mu, \quad (21)$$

where  $L_l$  is the Legendre polynomial of order  $l$ . We evaluate the monopole ( $l=0$ ) of the redshift-space correlation function.

The bottom panel of Figure 5 represents the result of the redshift-space correlation function. Taking the redshift distortion into account, the redshift-space correlation function  $\xi_0(s)$  is related to the real-space correlation function  $\xi(r)$  suggested by Kaiser (1987);

$$\xi_0(s) = \left( b^2 + \frac{2}{3} bf + \frac{1}{5} f^2 \right) \xi(r), \quad (22)$$

where  $b$  is the bias parameter, which is defined in Equation (2), and  $f$  is the gravitational growth factor. However, the effect of the redshift distortion is negligible because our clustering signal is measured on a large scale, beyond the small scale where redshift distortion can be observed. We fit the power-law function  $\xi(s) = (s/s_0)^{-\gamma}$ , black solid line in the bottom panel of Figure 5, to the redshift-space correlation function in place of the Kaiser's function by  $\chi^2$  fit. Based on Equation (10) replacing  $\omega$  for  $\xi$ , we obtain  $s_0 = 32.5 \pm 19 h^{-1}\text{Mpc}$  as the correlation length in the redshift space, which is almost consistent with that in the real space. Adopting  $s_0 = 32.5 \pm 19 h^{-1}\text{Mpc}$  as the correlation length in the real space, we obtain  $\xi_{100} = 0.310 \pm 0.326$ , which suggests that the significance of the clustering signal of the redshift-space correlation function is marginal.

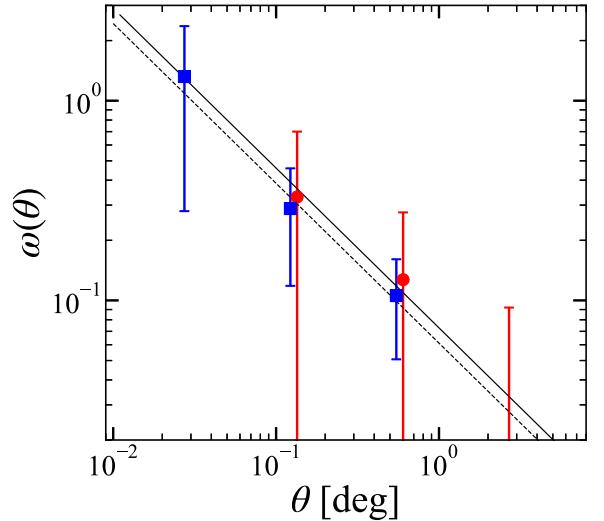
As shown in Table 3, consistent correlation lengths are obtained using three different correlation functions. The  $\xi_{100}$  shows that the clustering signal is barely detected. However, the errors for each correlation length are relatively large. This is probably due to the sample size not being large enough yet.

### 3.2. Cross-correlation Function with Galaxies

We evaluate CCF with our quasars and their neighboring LBGs at the similar redshift. The LBGs at  $z \sim 6$  are retrieved from Great Optically Luminous Dropout Research Using Subaru HSC (Harikane et al. 2022). The LBGs in the wide layer are not suitable for clustering analysis due to their low number density; therefore, we use the LBG sample in the Deep and Ultra Deep layers (COSMOS and SXDS) over  $8.7 \text{ deg}^2$  of HSC-SSP S18A (Aihara et al. 2019) where the SHELLQs quasars reside. As a result, the number of quasars and LBGs to calculate CCF is limited to 3 and 200, respectively. The limiting magnitude of LBGs is  $m_{\text{UV}} = 25.15$ . They are not spectroscopically confirmed; therefore, only angular correlation function can be evaluated. We use the estimator of CCF from the following equation (Landy & Szalay 1993; Cooke et al. 2006):

$$\omega_{\text{QG}}(\theta) = \frac{\text{QG}(\theta) - \text{QR}(\theta) - \text{GR}(\theta) + \text{RR}(\theta)}{\text{RR}(\theta)}, \quad (23)$$

where  $\text{QG}(\theta)$ ,  $\text{QR}(\theta)$ ,  $\text{GR}(\theta)$ ,  $\text{RR}(\theta)$  represent quasar-galaxy, quasar-random, galaxy-random, and random-random pairs of



**Figure 6.** The CCF between quasars and LBGs at  $z \sim 6$  (red) and ACF of LBGs (blue). The solid line and the dashed line show the power-law function fitted to the CCF and the ACF by  $\chi^2$  fit, respectively. For visibility, CCF plots are offset toward  $x$ -axis direction.

the given separation normalized by total pairs, respectively. The random points are retrieved from the random catalog in HSC-SSP DR2 utilizing the same flags of Table 2 in Harikane et al. (2022) at the surface number density of  $1 \text{ arcmin}^{-2}$ .

Figure 6 represents the result of CCF (red circles) and autocorrelation function (ACF) of LBGs at  $z \sim 6$  (blue squares;  $\omega_{\text{GG}}$ ), which is found to be consistent with Harikane et al. (2022). We fit the power-law functions to the CCF and the ACF by  $\chi^2$  fit, and the results are shown as the solid line and the dashed line, respectively. We confirmed that the angular scale at which we see the CCF signal is large enough to exceed the small scale ( $\lesssim 20''$ ), where the one-halo term is dominant. The errors are also evaluated by Jackknife resampling mentioned in Section 3.1.1 with  $N=5$ . The goodness-of-fit is calculated based on Equation (10). Our quasar sample has a cross-correlation strength similar to that of the autocorrelation of LBGs at the same redshift. Although the UV luminosity of the host galaxy, on which the clustering strength of galaxies strongly depends, in our quasar sample is not known, the result seems natural, given that quasars are stochastic processes that all galaxies experience at some period.

The correlation length is derived from the amplitude  $A_\omega$  of the power-law function fitted to the CCF. In CCFs, the Limber's equation is formalized as (Croom & Shanks 1999)

$$r_0 = \left\{ A_\omega \frac{c}{H_0 H_\gamma} \frac{\int N_Q(z) dz \int N_G(z) dz}{\int N_Q(z) N_G(z) \chi(z)^{1-\gamma} E(z) dz} \right\}^{1/\gamma}. \quad (24)$$

The suffixes of Q and G in Equation (24) denote quasars and LBGs, respectively. The redshift distribution of LBGs is assumed to be the same as that in Harikane et al. (2022). Finally, we obtain  $r_0 = 17.7 \pm 8.0 h^{-1}\text{Mpc}$  as the correlation length of quasars and galaxies. It should be noted that our LBG sample is photometrically selected, and the contamination of low- $z$  interlopers, the fraction of which is unknown, reduces the amplitude of cross-correlation. Ono et al. (2018) concluded that the contamination rate in the  $i$ -dropout galaxies may be small based on the fact that all 31 spectroscopic  $i$ -dropout galaxies

have  $z > 5.5$ , but it is difficult to know the exact contamination rate in the sample in this study down to the limiting magnitude.

### 3.3. The Bias Parameter

We assume that quasars reside in the peak of DM distribution and trace the distribution of underlying DM (Sheth & Tormen 1999). The bias parameter is derived from the ratio of clustering strength between quasars and underlying DM at a scale of  $r = 8 h^{-1} \text{Mpc}$ ;

$$b = \sqrt{\frac{\xi(8, z)}{\xi_{\text{DM}}(8, z)}}, \quad (25)$$

assuming that the real space correlation function  $\xi(r)$  is approximated by the power-law function. The correlation function of DM is generated by halomod<sup>15</sup> (Murray et al. 2013, 2021), assuming the bias model of Tinker et al. (2010), the transfer function of CAMB, and the growth model of Carroll et al. (1992). We evaluate the bias parameter as  $b = 20.8 \pm 8.7$ ,  $23.4 \pm 6.6$ , and  $27.7 \pm 15$  from the projected correlation function, the angular correlation function, and the redshift-space correlation function, respectively. We also evaluate the bias parameters  $b_{\text{QG}}$  and  $b_{\text{GG}}$  from the CCF between quasars and LBGs and the ACF of LBGs, respectively. We derive the bias parameter of quasars  $b_{\text{QQ}}$  from Mountrichas et al. (2009),

$$b_{\text{QG}} \sim b_{\text{QQ}} b_{\text{GG}}. \quad (26)$$

We obtain  $b_{\text{QG}} = 16.1 \pm 6.6$ , and  $b_{\text{GG}} = 13.3 \pm 2.3$  from the same analysis, yielding  $b_{\text{QQ}} = 19.5 \pm 16$ , and they are summarized in Table 3. The bias parameters derived by four independent methods are consistent with each other within their errors.

### 3.4. DMH Mass of $z \sim 6$ Quasars

We derive typical DMH mass from bias parameters of correlation functions. Under the assumption that quasars are the tracer for the underlying DM distribution, we adopt the bias model in Tinker et al. (2010), which is formalized as

$$b(\nu) = 1 - A \frac{\nu^a}{\nu^a + \delta_c^a} + B\nu^b + C\nu^c, \quad (27)$$

where  $\nu$  is the peak height, which is defined as  $\nu = \delta_c/\sigma(M)$ ,  $\delta_c$  is the critical density for the collapse of DMHs ( $\delta_c = 1.686$ ), and  $\sigma(M)$  is the linear matter variance at the radius of each DMH. We use the other parameters as they are in Table 2 of Tinker et al. (2010) for  $\Delta = 200$ , which represents the ratio between mean density and background density. The linear variance is defined as

$$\sigma^2(M) = \frac{1}{2\pi^2} \int P(k, z) \hat{W}^2(k, R) k^2 dk, \quad (28)$$

where  $P(k, z)$  is the matter power spectrum generated by CAMB<sup>16</sup> with our cosmology parameters, and  $\hat{W}$  is the spherical

top-hat function defined as

$$\hat{W}(k, R) = \frac{3}{(kR)^3} [\sin(kR) - kR \cos(kR)]. \quad (29)$$

This model is based on the clustering of DMHs in cosmological simulations of the flat  $\Lambda$ CDM cosmology. We obtain the radius of DMH  $R_{\text{halo}}$  by solving Equation (28). Finally, we evaluate the DMH mass  $M_{\text{halo}}$  assuming the spherical DMH;

$$M_{\text{halo}} = \frac{4}{3} \pi R_{\text{halo}}^3 \bar{\rho}_m. \quad (30)$$

We adopt  $\bar{\rho}_m = 2.78 \times 10^{11} \Omega_m h^2 M_\odot$ . Our DMH mass from each estimator is summarized in Table 3. The bias and halo mass of the CCF are slightly smaller than the other three, but this may be due to the contamination of the low- $z$  interlopers to the  $z \sim 6$  LBG sample (see Section 3.2).

The DMH masses derived by four independent methods are consistent with each other within their errors. However, we note that the DMH mass estimation is sensitive to  $\sigma_8$ . For simplicity, the following discussions will use the bias and halo mass obtained from the projected correlation function, but note that there is variation in these evaluations as shown in Table 3.

## 4. Discussion

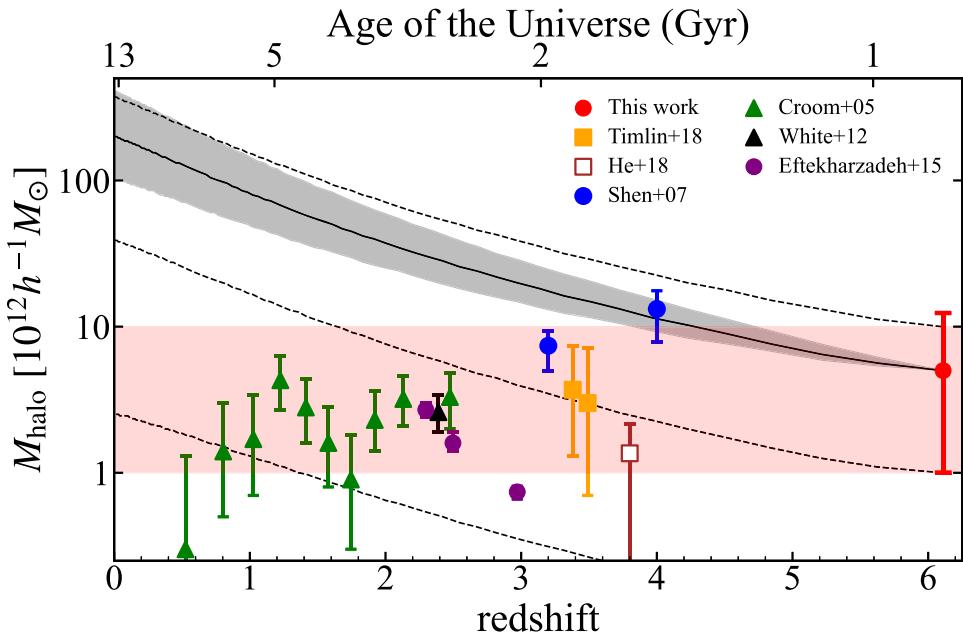
### 4.1. Comparison of DMH Mass with Other Studies

This study is the first to obtain the typical DMH mass of quasars at  $z \sim 6$  from clustering analysis, and not many previous studies have obtained DMH mass at  $z \sim 6$  using other methods. Shimasaku & Izumi (2019) estimated DMH mass of 49  $z \sim 6$  quasars, assuming that the FWHM of [C II] corresponds to the circular velocity of the DMH. They estimated that the median DMH mass of the whole samples is  $1.2_{-0.6}^{+2.2} \times 10^{12} M_\odot$ , which is slightly lower than our measurement; though, it is consistent within the errors. Chen et al. (2022) estimated the typical DMH mass of a bright quasar ( $M_{1450} < -26.5$ ) at  $z \sim 6$  is  $2.2_{-1.8}^{+3.4} \times 10^{12} h^{-1} M_\odot$  by measuring the intergalactic medium density around these luminous quasars, which is also consistent with our result within the error. Furthermore, a cosmological  $N$ -body simulation (Springel et al. 2005) predicts that the virial mass of DMH of quasars at  $z = 6.2$  is  $3.9 \times 10^{12} h^{-1} M_\odot$ , which is consistent with our result.

Figure 7 shows a compilation of the previous DMH mass measurements based on clustering analysis at lower- $z$ . In the figure, we convert the bias parameter in each previous research into DMH mass adopting our cosmology to reduce the effect of different  $\sigma_8$  among this work and the previous research. Some previous studies use different fitting formulae to infer a DMH mass from clustering, but we confirmed that the difference produces a few percent discrepancy in the DMH mass estimate. We conclude that the definition does not have a large impact on the DMH mass measurement. We also plot the mass evolution of DMH with  $M_{\text{halo}} = 10^{13}, 10^{12}, 10^{11} h^{-1} M_\odot$  (dotted lines) from the sample mean redshift,  $z = 6.1$ , to  $z = 0$  based on the extended Press–Schechter theory (Bond et al. 1991; Bower 1991; Lacey & Cole 1993). Our quasar sample with  $M_{\text{halo}} = 5.0 \times 10^{12} h^{-1} M_\odot$ , at  $z = 6.1$  grows to  $2.0_{-1.0}^{+2.2} \times 10^{14} h^{-1} M_\odot$  (black solid line), at  $z = 0$ , which is comparable to a rich galaxy cluster at present (Bhattacharya et al. 2013), implying that quasars reside in the most massive DMHs in the early universe.

<sup>15</sup> <https://github.com/halomod/halomod>

<sup>16</sup> <https://github.com/cmbant/CAMB>



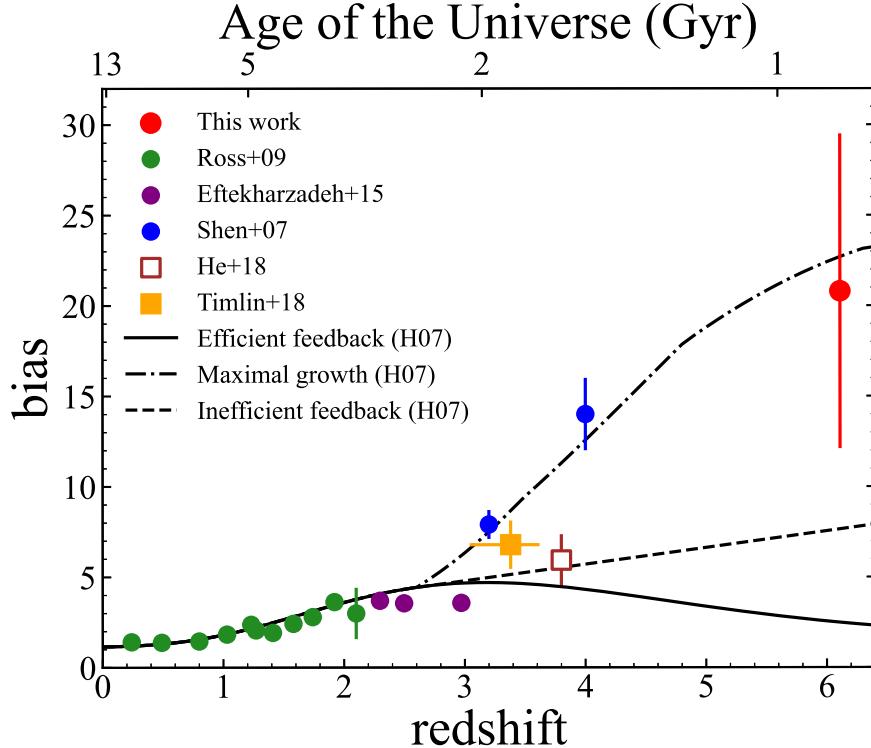
**Figure 7.** The DMH mass from clustering analysis as a function of redshift. Our result from the projected correlation function is represented as a red circle. Other symbols represent the DMH masses from previous studies. The circle, square, and triangle show studies that derive the projected, angular, and redshift-space correlation function, respectively. The filled symbols denote estimates from ACF signal, while others from CCF signal. The references are shown in the upper right of the figure. These masses are recalculated based on each bias parameter by using the cosmology in the study and the bias model of Tinker et al. (2010). The solid line with gray shade denotes the DMH mass evolution with its error based on the extended Press-Schechter theory. The three dashed lines represent the mass evolution of DMHs with  $M_{\text{halo}} = 10^{13}, 10^{12}, 10^{11} h^{-1} M_{\odot}$  at  $z \sim 6$  from the top. The red-shaded region represents the DMH mass range where most of quasars are expected to reside.

Interestingly, the DMH mass of quasars has remained almost constant  $\sim 10^{12.5} h^{-1} M_{\odot}$  across the cosmic time. Although the errors of each data point and variations even at the same epoch are large, and the DMH mass tends to decrease slightly from  $z=1$  to 0, it appears to remain roughly  $M_{\text{halo}} \sim 10^{12} - 10^{13} h^{-1} M_{\odot}$ . A quite constant halo mass of quasars as a function of redshift has been suggested up to  $z \sim 4$  by the previous studies (Trainor & Steidel 2012; Shen et al. 2013; Timlin et al. 2018), and this study confirms that the trend continues up to  $z \sim 6$  for the first time. This is in clear contrast to the standard growth of DMHs (the dashed lines in Figure 7). Greiner et al. (2021) also concluded from the quasar pair statistics that there is no strong evolution in clustering strength from  $z \sim 6$  to  $z \sim 4$ . McGreer et al. (2016) also used pair statistics to constrain the correlation length at  $z \sim 5$  as  $r_0 \gtrsim 20 h^{-1} \text{Mpc}$ , which is consistent with the trend. The observed trend is also consistent with that from the model (e.g., Lidz et al. 2006) that the characteristic mass of quasar host halos should evolve only weakly with redshift to reproduce the quasar luminosity function; though, their constraints are predicted only at  $0 < z < 3$ . Even though quasars at  $z=0-6$  reside in similar host halos of  $10^{12.5} h^{-1} M_{\odot}$ , this means that, as seen in the next section, higher- $z$  quasars are hosted in DMHs, which are more massive (higher bias) for the mass at that time. In other words, quasars appear in the most massive halos at  $z \sim 6$ , but they appear in less extreme halos at a later time.

Our result that quasars at  $z \sim 6$  reside in a fairly massive-end halo implies that they could be in overdense regions. However, observational evidence is far from conclusive, with some studies (e.g., Stiavelli et al. 2005; Morselli et al. 2014; Mignoli et al. 2020) finding quasars in the overdense region and others (e.g., Willott et al. 2005; Bañados et al. 2013; Mazzucchelli et al. 2017) finding no sign of it. This may be due to differences

in the depth and survey area of the overdense regions explored, or different selection criteria for surrounding galaxies, which may have led to a lack of consensus. A recent James Webb Space Telescope (JWST) observation (Kashino et al. 2023), which assessed the galaxy distribution around quasars at  $z \sim 6$  on the scale of up to  $\sim 10 \text{ Mpc}$  in the comoving coordinate, showed a clear overdensity of [O III] emitters around an ultra-luminous quasar at  $z = 6.327$ . Another JWST observation by Wang et al. (2023), which performed an imaging and spectroscopic survey of quasars utilizing NIRCam/WFSS, discovered ten [O III] emitters around a quasar at  $z = 6.6$ , and the galaxy overdensity corresponds to  $\delta_{\text{gal}} = 12.6^{+5.9}_{-5.0}$  over a  $637 \text{ Mpc}^3$  volume in the comoving space. A large number of such deep and wide observations will provide clearer insights into the large-scale environments of  $z \sim 6$  quasars.

In the low-mass regime below  $M_{\text{halo}} < 10^{12} h^{-1} M_{\odot}$ , the quasars with small black hole mass, small stellar mass and extremely low luminosity may not be detected observationally. In this case, the apparent lower limit of the observed halo mass may be due to observational bias. In contrast, there may be an upper limit of halo mass rather than a typical halo mass at which quasar activity appears. In Figure 7, there appears to be an upper limit where the quasar DMH mass never exceeds  $10^{13} h^{-1} M_{\odot}$ , i.e., most quasars reside in DMHs with  $M_{\text{halo}} < 10^{13} h^{-1} M_{\odot}$  across most of the cosmic history. Fanidakis et al. (2013) used GALFORM, a semianalytic model, and concluded that quasars live in average mass halos and do not reside in the most massive DMHs at any redshift. In their model, the quasar activity, which is maintained by the cold gas accretion onto a central SMBH, will be suppressed by the radio-mode AGN feedback in a massive halo larger than  $10^{13} h^{-1} M_{\odot}$ . If the halo mass of quasars does not exceed  $10^{13} h^{-1} M_{\odot}$  at any cosmic time, then such physics may ubiquitously operate. This is supported by the observation by



**Figure 8.** Evolution of the quasar bias parameter. This figure is the extension of Figure 12 in Timlin et al. (2018). The red circle represents our result of the projected correlation function, and the other symbols show the result of previous studies. Three black lines represent the bias parameters as a function of redshift based on the theoretical models: *efficient feedback* (solid line), *inefficient feedback* (dashed line), and *maximal growth* (dashed-dotted line). The *efficient feedback* model assumes that the growth of quasars only occurs in their active phase. The *inefficient feedback* model presumes that quasars continue growing periodically even after their active phase until  $z \sim 2$ . The *maximal growth* model premises that quasars keep growing with their host DMH until  $z \sim 2$ . In these models, as the feedback becomes inefficient, the DMH mass gets more massive.

Uchiyama et al. (2018), which concluded that few of the most massive protocluster candidates were found around quasars at  $z \sim 4$ . In other words, at  $z \sim 4$ , quasars do not exist in overdense regions exceeding  $10^{13} h^{-1} M_{\odot}$ , but in medium-weight overdense regions below  $10^{13} h^{-1} M_{\odot}$ .

However, it only appears that the halo mass does not exceed  $10^{13} h^{-1} M_{\odot}$  in Figure 7, and what is measured from the clustering is the average DMH mass of quasars in each period, and it is therefore strictly inconclusive whether there are no quasars in the halo with a mass exceeding  $10^{13} h^{-1} M_{\odot}$ .

#### 4.2. Implication to AGN Feedback

We compare our bias parameter with theoretical models in Hopkins et al. (2007), which predicted a bias parameter evolution at  $z \gtrsim 3$  for three models with simple assumptions: “efficient feedback,” “inefficient feedback,” and “maximal growth,” as shown in Figure 8. In the *efficient feedback* model, quasars only grow during their active phase, and the growth thoroughly terminates after the phase. The bias parameter is predicted to become smaller at higher- $z$  if feedback is efficient. In “inefficient feedback” model, quasars and their central SMBHs continue growing periodically even after their active phase until  $z \sim 2$ . Since their feedback is inefficient, the quasars do not stop growing and shine episodically. In contrast to the previous model, the quasars tend to reside in more massive DMHs, which makes the bias parameter larger at  $z \gtrsim 3$ . In the last model, *maximal growth*, quasars keep growing at the same rate with their host DMHs simultaneously until  $z \sim 2$ . The central SMBHs retain Eddington accretion all the time, and

their growth is rapid. The feedback of quasars is less efficient than that from the second model. Therefore, the DMHs are the most massive among these models, which is apparent in Figure 8. It should be noted that the model only predicts the evolution of the bias parameter, and no prediction of other observables (e.g., luminosity function,  $M-\sigma$  relation) is given for each assumption.

Our result is most consistent with a large bias parameter, favoring the *maximal model*, which assumes Eddington accretion and the feedback is highly inefficient at  $z \sim 6$ . This result is consistent with the fact that the Eddington ratio of quasars at  $z \sim 6$  tends to be higher than that in local (Yang et al. 2021). However, it is noted that the Eddington ratio of quasars at  $z < 4$  is usually smaller than unity (Shen et al. 2008), being inconsistent with “maximal growth” model. The measurement of bias parameters at  $z \sim 4$  has not yet been settled, as the results are largely divided into large (Shen et al. 2007) and small (He et al. 2018; Timlin et al. 2018) values. Therefore, since Hopkins’ models simply attempt to explain the evolution from  $z = 2$ , to  $z = 6$ , with a single physical mechanism, it is not necessary only to support this *maximal growth* model at  $4 < z < 6$  as it is. For example, by assuming that the feedback is inefficient at  $z \sim 6$  while it becomes more efficient until  $z \sim 4$ , an evolution model that the bias keeps low until  $z \sim 4$  and increases rapidly to  $z \sim 6$  does not conflict with our observational result. Alternatively, it could be explained by intermittent black hole growth (Inayoshi et al. 2022; Li et al. 2023). To further restrict the models, the measurement of the bias parameter at  $z \sim 5$  is a key.

### 4.3. Duty Cycle

We also evaluate the duty cycle of quasars, which represents the fraction of DMHs that host active quasars. At first, following the traditional approach (Haiman & Hui 2001; Martini & Weinberg 2001), we assume that a DMH with more than the threshold  $M_{\min}$  can host a quasar, which activates randomly for a certain period. Under this assumption, the duty cycle  $f_{\text{duty}}$  is defined as the ratio of the number of observed quasars to the number of the whole host halos above  $M_{\min}$ . Therefore,  $f_{\text{duty}}$  is evaluated as

$$f_{\text{duty}} = \frac{\int_{L_{\min}}^{\infty} \Phi(L) dL}{\int_{M_{\min}}^{\infty} n(M) dM}, \quad (31)$$

where  $\Phi(L)$  is the quasar luminosity function at  $z \sim 6$  derived by Matsuoka et al. (2018b),  $L_{\min}$  is the minimum luminosity of the quasar sample,  $n(M)$  represents the DMH mass function at  $z \sim 6$  derived by Sheth & Tormen (1999), and  $M_{\min}$  represents the DMH minimum mass to host a quasar. The quasar luminosity function is evaluated based on the sample almost equivalent to that in this study, by excluding type II quasars. We adopt the DMH mass function from Sheth & Tormen (1999);

$$n(M) = -A \sqrt{\frac{2a}{\pi}} \frac{\rho_0}{M} \frac{\delta_c(z)}{\sigma^2(M)} \frac{d\sigma(M)}{dM} \times \left\{ 1 + \left[ \frac{\sigma^2(M)}{a\delta_c(z)} \right]^p \right\} \exp \left[ -\frac{a\delta_c(z)}{2\sigma^2(M)} \right], \quad (32)$$

where  $A = 0.3222$ ,  $a = 0.707$ ,  $p = 0.3$ , and  $\delta_c(z) = \delta_c/D(z)$ . The  $D(z)$  represents the growth factor from Carroll et al. (1992). The minimum mass is estimated from the effective bias, which is expressed as

$$b_{\text{eff}} = \frac{\int_{M_{\min}}^{\infty} b(M, z) n(M) dM}{\int_{M_{\min}}^{\infty} n(M) dM}, \quad (33)$$

where  $b(M, z)$  is the bias parameter of the given DMH mass at a given redshift from the model (Tinker et al. 2010). Based on the effective bias determined from the clustering analysis,  $M_{\min}$  is evaluated to be  $4.5 \times 10^{12} h^{-1} M_{\odot}$ . In this case, we obtain  $f_{\text{duty}} = 6.3 \pm 2.7$ , exceeding unity, which is unreasonable given its definition.

We consider that it is too simple to assume that all halos above a certain minimum mass can host a quasar, as expressed in Equation (31). Equation (31) is correct when luminosity and mass are proportional, and  $L_{\min}$  corresponds to  $M_{\min}$ , but this is not the case of quasars. In fact, as seen in Figure 7, the halo mass of quasars is within a certain narrow range over cosmic time, and there seems to exist an upper limit to the halo mass of quasars. Although it is difficult to determine the exact mass range, we here simply assume that the DMHs with  $12 \leq \log(M_{\text{halo}}/h^{-1} M_{\odot}) \leq 13$  can host quasars based on Figure 7. In this case, Equation (31) can be expressed as

$$f_{\text{duty}} = \frac{\int_{L_{\min}}^{\infty} \Phi(L) dL}{\int_{M_1}^{M_2} n(M) dM}, \quad (34)$$

where  $M_1 = 10^{12} h^{-1} M_{\odot}$ , and  $M_2 = 10^{13} h^{-1} M_{\odot}$ . This equation gives  $f_{\text{duty}} = 0.019 \pm 0.008$ . The derived  $f_{\text{duty}}$  corresponds to 1.9% of the age of the universe, namely,  $\sim 1.7 \times 10^7$  yr, as the lifetime of quasars at  $z \sim 6$ . While this is consistent with the lifetime obtained from the clustering analysis at low- $z$  (e.g., White et al. 2012), it is about equal to the upper limit obtained from the proximity zone size measurements at  $z \sim 6$  (Eilers et al. 2021).

We derive the duty cycle based on the new definition, which cannot be simply compared with previous results at low- $z$ . Based on Equation (34), we recalculate  $f_{\text{duty}}$  at  $z \sim 4$  from the luminosity function (Akiyama et al. 2018) and obtain  $f_{\text{duty}} = 0.012 \pm 0.001$ , which is consistent with the conventional estimate,  $f = 0.001\text{--}0.06$  (He et al. 2018), and  $f_{\text{duty}}$  at  $z \sim 6$  by this study. On the other hand, in the case of Eftekharzadeh et al. (2015) at  $z \sim 3$ , we obtain  $f_{\text{duty}} = 0.0060 \pm 0.0008$ . Based on Croom et al. (2005), we obtain  $f_{\text{duty}} = 0.0039 \pm 0.0005$ ,  $0.0043 \pm 0.0005$ ,  $0.0042 \pm 0.0006$ ; at  $z = 0.804$ , 1.579, 2.475, respectively. They are slightly smaller than those at  $z > 4$ .

However, note that there is no justification for the mass range used for integration here. Unless we know the exact mass distribution of halos that can host quasars, we cannot precisely obtain the denominators in either Equation (31) or Equation (34). Also, the numerator in these equations are the number of quasars observed, and  $f_{\text{duty}}$  will inevitably increase as the limiting magnitude deepens in the future, that is, as  $L_{\min}$  decreases. Because of this physical discrepancy,  $f_{\text{duty}}$  should be considered to give only a very rough estimate.

### 4.4. Stellar Mass and Dynamical Mass

We evaluate the stellar mass of host galaxies based on the empirical stellar ( $M_*$ )-to-halo ( $M_{\text{halo}}$ ) mass ratio (SHMR) from Behroozi et al. (2019). The SHMR at  $z \sim 6$  has only evaluated up to  $M_{\text{halo}}^{\max} = 10^{12} h^{-1} M_{\odot}$  and needs to be extended beyond this point to reach the observed halo mass,  $M_{\text{halo}} = 5 \times 10^{12} h^{-1} M_{\odot}$ . However, the pivot mass ( $M_{\text{halo}}^{\max}$ ) is just where the slope of this relationship changes, and the slope at the high-mass regime tends to become shallower toward higher- $z$  (Behroozi et al. 2019); therefore, this extension involves a large uncertainty. Assuming conservatively here that the ratio above  $M_{\text{halo}}^{\max}$  does not change from SHMR  $\sim 0.013$  at  $M_{\text{halo}}^{\max}$ , the stellar mass is evaluated as  $M_* = 6.5_{-5.2}^{+9.6} \times 10^{10} h^{-1} M_{\odot}$ , where the error is estimated from the uncertainty of  $M_{\text{halo}}$  only and does not take into account the uncertainty of the SHMR extrapolation.

On the other hand, the dynamical mass evaluated by the [C II] 158  $\mu\text{m}$  observation is often used as the surrogate of the stellar mass (e.g., Willott et al. 2015a; Venemans et al. 2016; Izumi et al. 2018) although the dynamical mass is essentially different from the stellar mass. They estimated the dynamical mass with an assumption of a thin rotation disk with a diameter  $D = 1\text{--}2 h^{-1} \text{kpc}$ . (Wang et al. 2013; Willott et al. 2015b, 2017). The stellar mass was evaluated by the [C II] observations of seven SHELLQs quasars as  $M_* = (0.91 - 20) \times 10^{10} h^{-1} M_{\odot}$  (Izumi et al. 2018, 2019). Neeleman et al. (2021) evaluated the mean stellar mass of 27 brighter quasars as  $M_{\text{dyn}} = (3.5 \pm 2.5) \times 10^{10} h^{-1} M_{\odot}$ , which is consistent with that from Izumi et al. (2018, 2019).

The stellar mass based on the clustering analysis with the SHMR is consistent with those independently measured from [C II] observation. It is a bit surprising that they both agree, albeit with large uncertainties: in addition to the SHMR being

uncertain at the massive end, there is no guarantee that the quasar hosts will have the same SHMR as the normal galaxy. On the other hand, there is an implicit assumption that [C II] dynamical mass is a good proxy for the stellar mass of the bulge.

We compare the dynamical mass between the clustering analysis and the [C II] observation. We estimate the dynamical mass at  $D = 1\text{--}2 h^{-1}\text{kpc}$ , where [C II] dynamical mass is estimated, from our  $M_{\text{halo}}$  measurement, as follows. The virial radius  $r_{\text{vir}}$  can be estimated by using the spherical collapse model (Barkana & Loeb 2001);

$$r_{\text{vir}} = 0.756 \left( \frac{M_{\text{halo}}}{10^8 h^{-1} M_{\odot}} \right)^{1/3} \left[ \frac{\Omega_m}{\Omega_m(z) 200} \Delta_c \right]^{-1/3} \left( \frac{1+z}{10} \right)^{-1} h^{-1}\text{kpc}, \quad (35)$$

where

$$\Omega_m(z) = \frac{\Omega_m(1+z)^3}{E(z)^2}, \quad (36)$$

and

$$\Delta_c = 18\pi^2 + 82[\Omega_m(z) - 1] - 39[\Omega_m(z) - 1]^2 \quad (37)$$

is the overdensity at the halo collapse, and we obtain  $r_{\text{vir}} = 64 h^{-1}\text{kpc}$ , which is much larger than the scale on which the [C II] dynamical mass is estimated. When assuming a rotation-dominated disk with the flat rotation of DMHs, i.e., rotation velocity does not depend on the radius, the dynamical mass is estimated as  $M_{\text{dyn}} = (0.83\text{--}1.7) \times 10^{11} h^{-1} M_{\odot}$  at the scale of  $1\text{--}2 h^{-1}\text{kpc}$ , which is larger than the [C II] dynamical mass, and in other words, the [C II] rotation velocity is much slower than the halo circular velocity. This suggests either that the area where [C II] is detected is substantially central to the halo, where the rotation velocity has not yet reached the maximum halo circular velocity, or the rotation of [C II] is independent of the rotation of the halo. These considerations make it difficult to regard the dynamical mass obtained from [C II] as that of the entire system. Nevertheless, the stellar masses of both estimates agree, which could be a coincidence due to the large uncertainties in both.

It should be noted that recent direct observation by JWST/NIRCam for host galaxies of a couple of SHELLQs quasars (Ding et al. 2022) applies the spectral energy distribution fitting to derive the stellar mass, which is comparable to that inferred from the halo mass measurement. Marshall et al. (2023) also used JWST/NIRSpec to detect [O III]  $\lambda 5008$  emitting regions, which are more extended than the [C II], of the host galaxy, giving a slightly higher dynamical mass. More observations should be made in the future to increase the number of direct measurements of the stellar mass of quasar host galaxies. Also, we should keep in mind that the halo mass obtained in this study is still accompanied by a large error.

## 5. Summary

We conduct a clustering analysis of 107 quasars at  $z \sim 6$ , mainly composed of SHELLQs, which has increased the number density of quasars at  $z \sim 6$  by more than 30 times that from SDSS. This study is the first attempt to measure the DMH

mass of quasars at  $z \sim 6$ . The main results are summarized below.

1. The quasars are spectroscopically identified in the HSC-SSP wide layer over  $891 \text{ deg}^2$ . The completeness holds 70% (80%) over 85% (77%) of the entire survey regions. We evaluate the three types of ACF for our sample: projected correlation function  $\omega_p(r_p)$ , angular correlation function  $\omega(\theta)$ , and redshift-space correlation function  $\xi(s)$ . We also evaluate the angular CCF between our quasar sample and LBG sample at  $z \sim 6$  in the HSC-SSP Deep layer. The DMH mass at  $z \sim 6$  is evaluated as  $5.0^{+7.4}_{-4.0} \times 10^{12} h^{-1} M_{\odot}$  with the bias parameter,  $b = 20.8 \pm 8.7$  by the projected correlation function. The other three estimators agree with these values; though, the uncertainties are large due to the small sample size. Using extended Press-Schechter theory, we find that the DMH with  $5.0 \times 10^{12} h^{-1} M_{\odot}$  at  $z \sim 6$  will grow into  $2.0^{+2.2}_{-1.0} \times 10^{14} h^{-1} M_{\odot}$  at  $z = 0$ , which is comparable to the rich clusters of galaxies today.
2. The DMH mass of quasars is found to be nearly constant  $\sim 10^{12.5} h^{-1} M_{\odot}$  throughout the cosmic epoch. While there is broad agreement in previous studies that the quasar halo mass remains approximately constant up to  $z \sim 4$ , this study confirms, for the first time, that this trend continues up to  $z \sim 6$ . This means that there is a characteristic mass of DMH where quasars are always activated. As a result, quasars appear in the most massive halos at  $z \sim 6$ , but in less extreme halos thereafter. The mass of the quasar DMH is unlikely to exceed its upper limit of  $10^{13} h^{-1} M_{\odot}$ . This suggests that most quasars reside in  $M_{\text{halo}} < 10^{13} h^{-1} M_{\odot}$  DMHs across most of the cosmic history. This is consistent with the model by Fanidakis et al. (2013). In the model, the quasar activity, which is maintained by cold gas accretion into the central SMBH, is suppressed by radio-mode AGN feedback for massive halos larger than  $10^{13} h^{-1} M_{\odot}$ . If the quasar halo mass does not exceed  $10^{13} h^{-1} M_{\odot}$  at any time, then such physics may be ubiquitously at work.
3. Our result that the bias parameter  $b$  is as large as  $b = 20.8 \pm 8.7$  at  $z \sim 6$  supports the “maximum model” proposed by Hopkins et al. (2007), which assumes that feedback is highly inefficient during  $z \sim 4\text{--}6$ . Without the observational constraint at  $z \sim 5$ , our result along with the previous observations can also be explained by a bias evolution model in which feedback is inefficient at  $z \sim 6$  but becomes progressively more efficient at  $z \sim 4$ .
4. We estimate the quasar duty cycle  $f_{\text{duty}}$  at  $z \sim 6$ . We find that the conventional definition of  $f_{\text{duty}}$  yields an unphysical result as the  $f_{\text{duty}}$  becomes greater than unity. We propose a new method to estimate the duty cycle in line with the observational result that DMH mass is nearly constant. We assume that DMHs with  $12 \leq \log(M_{\text{halo}}/h^{-1} M_{\odot}) \leq 13$  can host quasars. Using the number density of DMHs in the mass interval and the quasar luminosity function at  $z \sim 6$ , we achieve  $f_{\text{duty}} = 0.019 \pm 0.008$ , which is consistent with that at  $z \sim 4$ .
5. Assuming that the empirical SHMR at  $z \sim 6$  is constant at  $M_{\text{halo}} > 10^{12} h^{-1} M_{\odot}$ , the average stellar mass of quasar host galaxies at  $z \sim 6$  is evaluated from the observed DMH mass to be  $M_* = 6.5^{+9.6}_{-5.2} \times 10^{10} h^{-1} M_{\odot}$ , which is

found to be consistent with those derived from [C II] observations.

The clustering signal measurement utilizing the quasar candidates at  $z \sim 5$  identified by HSC will soon be made, which can constrain the feedback models more rigidly. More stellar mass measurements from [C II] observations by Atacama Large Millimeter/submillimeter Array will lead to a rigorous comparison with halo masses derived in the study and a constraint on the SHMR at the massive-end at  $z \sim 6$ . The high sensitivity of JWST will allow us to directly measure the host stellar mass and the dynamical mass of quasars and investigate the environment around quasars (e.g., overdensity). In the future, more powerful surveys (e.g., Legacy Survey of Space and Time; LSST Science Collaboration et al. 2009) will contribute to the larger quasar sample at high- $z$ , which will lead to the detection of clearer clustering signals. In addition, promising instruments, such as Nancy Roman Space Telescope and Euclid Satellite, will be expected to identify quasars at  $z > 7$ . These next-generation instruments will make the sample size deeper and larger, which will have a huge impact on our understanding of the coevolution in the early universe.

### Acknowledgments

We appreciate the anonymous referee for constructive comments and suggestions. We thank Tomo Takahashi, Takumi Shinohara, and Teruaki Suyama for fruitful discussions and Kazuhiro Shimasaku, and Yuichi Harikane for useful suggestions.

J.A. is supported by the International Graduate Program for Excellence in Earth-Space Science (IGPEES). N.K. was supported by the Japan Society for the Promotion of Science through Grant-in-Aid for Scientific Research 21H04490. Y.M. was supported by the Japan Society for the Promotion of Science KAKENHI grant Nos. JP17H04830 and 21H04494. K. I. acknowledges support by grant PID2019-105510GB-C33 funded by MCIN/AEI/10.13039/501100011033 and “Unit of excellence María de Maeztu 2020-2023” awarded to ICCUB (CEX2019-000918-M). M.O. is supported by the National Natural Science Foundation of China (12150410307).

The Hyper Suprime-Cam (HSC) collaboration includes the astronomical communities of Japan and Taiwan, and Princeton University. The HSC instrumentation and software were developed by the National Astronomical Observatory of Japan (NAOJ), the Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), the University of Tokyo, the High Energy Accelerator Research Organization (KEK), the Academia Sinica Institute for Astronomy and Astrophysics in Taiwan (ASIAA), and Princeton University. Funding was contributed by the FIRST program from the Japanese Cabinet Office, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Japan Society for the Promotion of Science (JSPS), Japan Science and Technology Agency (JST), the Toray Science Foundation, NAOJ, Kavli IPMU, KEK, ASIAA, and Princeton University. This paper is based [in part] on data collected at the Subaru Telescope and retrieved from the HSC data archive system, which is operated by Subaru Telescope and Astronomy Data Center (ADC) at NAOJ. Data analysis was in part carried out with the cooperation of Center for Computational Astrophysics (CfCA) at NAOJ. We are honored and grateful for the opportunity of observing the Universe from

Maunakea, which has the cultural, historical, and natural significance in Hawaii.

This paper makes use of software developed for Vera C. Rubin Observatory. We thank the Rubin Observatory for making their code available as free software at <http://pipelines.lsst.io/>.

The Pan-STARRS1 Surveys (PS1) and the PS1 public science archive have been made possible through contributions by the Institute for Astronomy, the University of Hawaii, the Pan-STARRS Project Office, the Max Planck Society and its participating institutes, the Max Planck Institute for Astronomy, Heidelberg, and the Max Planck Institute for Extraterrestrial Physics, Garching, The Johns Hopkins University, Durham University, the University of Edinburgh, the Queens University Belfast, the Harvard-Smithsonian Center for Astrophysics, the Las Cumbres Observatory Global Telescope Network Incorporated, the National Central University of Taiwan, the Space Telescope Science Institute, the National Aeronautics and Space Administration under grant No. NNX08AR22G issued through the Planetary Science Division of the NASA Science Mission Directorate, the National Science Foundation grant No. AST-1238877, the University of Maryland, Eotvos Lorand University (ELTE), the Los Alamos National Laboratory, and the Gordon and Betty Moore Foundation.

*Facility:* Subaru.

*Software:* astropy (Astropy Collaboration et al. 2013, 2018), CAMB (Lewis & Challinor 2011), Corrfunc (Sinha & Garrison 2019, 2020), halomod (Murray et al. 2013, 2021), hscpipe (Bosch et al. 2018) Numpy (Harris et al. 2020), Matplotlib (Hunter 2007), Pandas (McKinney et al. 2010), Scipy (Virtanen et al. 2020).

### ORCID iDs

Junya Arita  <https://orcid.org/0009-0007-0864-7094>  
Nobunari Kashikawa  <https://orcid.org/0000-0003-3954-4219>

Yoshiki Matsuoka  <https://orcid.org/0000-0001-5063-0340>

Wanqiu He  <https://orcid.org/0000-0001-7759-6410>

Kei Ito  <https://orcid.org/0000-0002-9453-0381>

Yongming Liang  <https://orcid.org/0000-0002-2725-302X>

Rikako Ishimoto  <https://orcid.org/0000-0002-2134-2902>

Takehiro Yoshioka  <https://orcid.org/0000-0002-3800-0554>

Yoshihiro Takeda  <https://orcid.org/0000-0001-7154-3756>

Kazushi Iwasawa  <https://orcid.org/0000-0002-4923-3281>

Masafusa Onoue  <https://orcid.org/0000-0003-2984-6803>

Yoshiki Toba  <https://orcid.org/0000-0002-3531-7863>

Masatoshi Imanishi  <https://orcid.org/0000-0001-6186-8792>

### References

- Adelberger, K. L., Steidel, C. C., Kollmeier, J. A., & Reddy, N. A. 2006, *ApJ*, **637**, 74
- Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, *PASJ*, **71**, 114
- Aihara, H., Arimoto, N., Armstrong, R., et al. 2018a, *PASJ*, **70**, S4
- Aihara, H., Armstrong, R., Bickerton, S., et al. 2018b, *PASJ*, **70**, S8
- Akiyama, M., He, W., Ikeda, H., et al. 2018, *PASJ*, **70**, S34
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
- Annis, J., Soares-Santos, M., Strauss, M. A., et al. 2014, *ApJ*, **794**, 120
- Antonucci, R. 1993, *ARA&A*, **31**, 473
- 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., Schindler, J.-T., Venemans, B. P., et al. 2023, *ApJS*, **265**, 29

- Bañados, E., Venemans, B., Walter, F., et al. 2013, *ApJ*, **773**, 178
- Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, *ApJS*, **227**, 11
- Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, *Natur*, **553**, 473
- Bañados, E., Venemans, B. P., Morganson, E., et al. 2014, *AJ*, **148**, 14
- Barkana, R., & Loeb, A. 2001, *PhR*, **349**, 125
- Behroozi, P., Wechsler, R. H., Hearn, A. P., & Conroy, C. 2019, *MNRAS*, **488**, 3143
- Bertin, E. 2011, in ASP Conf. Ser. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans et al. (San Francisco, CA: ASP), 435
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Bhattacharya, S., Habib, S., Heitmann, K., & Vikhlinin, A. 2013, *ApJ*, **766**, 32
- Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, *ApJ*, **379**, 440
- Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, *PASJ*, **70**, S5
- Bower, R. G. 1991, *MNRAS*, **248**, 332
- Carroll, S. M., Press, W. H., & Turner, E. L. 1992, *ARA&A*, **30**, 499
- Cepa, J., Aguiar, M., Escalera, V. G., et al. 2000, *Proc. SPIE*, **4008**, 623
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
- Chen, H., Eilers, A.-C., Bosman, S. E. I., et al. 2022, *ApJ*, **931**, 29
- Cole, S., & Kaiser, N. 1989, *MNRAS*, **237**, 1127
- Cooke, J., Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2006, *ApJ*, **652**, 994
- Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, *MNRAS*, **356**, 415
- Croom, S. M., & Shanks, T. 1999, *MNRAS*, **303**, 411
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, **267**, 465
- Dekel, A., & Lahav, O. 1999, *ApJ*, **520**, 24
- Di Matteo, T., Khandai, N., DeGraf, C., et al. 2012, *ApJL*, **745**, L29
- Ding, X., Onoue, M., Silverman, J. D., et al. 2022, arXiv:2211.14329
- Eftekharzadeh, S., Myers, A. D., White, M., et al. 2015, *MNRAS*, **453**, 2779
- Eilers, A.-C., Hennawi, J. F., Davies, F. B., & Simcoe, R. A. 2021, *ApJ*, **917**, 38
- Fan, X., Banados, E., & Simcoe, R. A. 2022, arXiv:2212.06907
- Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, *AJ*, **128**, 515
- Fanidakis, N., Macciò, A. V., Baugh, C. M., Lacey, C. G., & Frenk, C. S. 2013, *MNRAS*, **436**, 315
- Ferrarese, L. 2002, *ApJ*, **578**, 90
- Furusawa, H., Koike, M., Takata, T., et al. 2018, *PASJ*, **70**, S3
- Granato, G. L., De Zotti, G., Silva, L., Bressan, A., & Danese, L. 2004, *ApJ*, **600**, 580
- Greiner, J., Bolmer, J., Yates, R. M., et al. 2021, *A&A*, **654**, A79
- Groth, E. J., & Peebles, P. J. E. 1977, *ApJ*, **217**, 385
- Haiman, Z., & Hui, L. 2001, *ApJ*, **547**, 27
- Harikane, Y., Ono, Y., Ouchi, M., et al. 2022, *ApJS*, **259**, 20
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- He, W., Akiyama, M., Bosch, J., et al. 2018, *PASJ*, **70**, S33
- Heckman, T. M., & Best, P. N. 2014, *ARA&A*, **52**, 589
- Hickox, R. C., Myers, A. D., Brodwin, M., et al. 2011, *ApJ*, **731**, 117
- Hopkins, P. F., Lidz, A., Hernquist, L., et al. 2007, *ApJ*, **662**, 110
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Inayoshi, K., Nakatani, R., Toyouchi, D., et al. 2022, *ApJ*, **927**, 237
- Izumi, T., Onoue, M., Matsuoka, Y., et al. 2019, *PASJ*, **71**, 111
- Izumi, T., Onoue, M., Shirakata, H., et al. 2018, *PASJ*, **70**, 36
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *ApJ*, **873**, 111
- Jiang, L., Fan, X., Annis, J., et al. 2008, *AJ*, **135**, 1057
- Jiang, L., Fan, X., Bian, F., et al. 2009, *AJ*, **138**, 305
- Jiang, L., Fan, X., Bian, F., et al. 2014, *ApJS*, **213**, 12
- Jiang, L., McGreer, I. D., Fan, X., et al. 2016, *ApJ*, **833**, 222
- Jing, Y. P. 1998, *ApJL*, **503**, L9
- Jurić, M., Kantor, J., Lim, K.-T., et al. 2017, in ASP Conf. Ser. 512, Astronomical Data Analysis Software and Systems XXV, ed. N. P. F. Lorento, K. Shortridge, & R. Wayth (San Francisco, CA: ASP), 279
- Kaiser, N. 1987, *MNRAS*, **227**, 1
- Kashikawa, N., Aoki, K., Asai, R., et al. 2002, *PASJ*, **54**, 819
- Kashikawa, N., Ishizaki, Y., Willott, C. J., et al. 2015, *ApJ*, **798**, 28
- Kashino, D., Lilly, S. J., Matthee, J., et al. 2023, *ApJ*, **950**, 66
- Kawanomoto, S., Uraguchi, F., Komiyama, Y., et al. 2018, *PASJ*, **70**, 66
- Kim, Y., Im, M., Jeon, Y., et al. 2015, *ApJL*, **813**, L35
- Komiyama, Y., Obuchi, Y., Nakaya, H., et al. 2018, *PASJ*, **70**, S2
- Koptelova, E., & Hwang, C.-Y. 2022, arXiv:2212.05862
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Kormendy, J., & Richstone, D. 1995, *ARA&A*, **33**, 581
- Lacey, C., & Cole, S. 1993, *MNRAS*, **262**, 627
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, **412**, 64
- Lewis, A., & Challinor, A. 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library, ascl:1102.026
- Li, W., Inayoshi, K., Onoue, M., & Toyouchi, D. 2023, *ApJ*, **950**, 85
- Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006, *ApJ*, **641**, 41
- Limber, D. N. 1953, *ApJ*, **117**, 134
- Lynden-Bell, D. 1969, *Natur*, **223**, 690
- LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
- Magnier, E. A., Schlafly, E., Finkbeiner, D., et al. 2013, *ApJS*, **205**, 20
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, **115**, 2285
- Mandelbaum, R., Tasitsiomi, A., Seljak, U., Kravtsov, A. V., & Wechsler, R. H. 2005, *MNRAS*, **362**, 1451
- Marin, F. A., Beutler, F., Blake, C., et al. 2016, *MNRAS*, **455**, 4046
- Marshall, M. A., Perna, M., Willott, C. J., et al. 2023, arXiv:2302.04795
- Martini, P., & Weinberg, D. H. 2001, *ApJ*, **547**, 12
- Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2018a, *ApJS*, **237**, 5
- Matsuoka, Y., Strauss, M. A., Kashikawa, N., et al. 2018b, *ApJ*, **869**, 150
- Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2022, *ApJS*, **259**, 18
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, *ApJ*, **828**, 26
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2019, *ApJL*, **872**, L2
- Mazzucchelli, C., Bañados, E., Decarli, R., et al. 2017, *ApJ*, **834**, 83
- Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017, *ApJ*, **849**, 91
- McGreer, I. D., Eftekharzadeh, S., Myers, A. D., & Fan, X. 2016, *AJ*, **151**, 61
- McKinney, W., et al. 2010, in Proc. of the 9th Python in Sci. Conf., ed. S. van der Walt & J. Millman (Austin, TX: Scipy), 56
- Mignoli, M., Gilli, R., Decarli, R., et al. 2020, *A&A*, **642**, L1
- Miyazaki, S., Komiyama, Y., Kawanomoto, S., et al. 2018, *PASJ*, **70**, S1
- Morselli, L., Mignoli, M., Gilli, R., et al. 2014, *A&A*, **568**, A1
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Natur*, **474**, 616
- Mountreichas, G., Sawangwit, U., Shanks, T., et al. 2009, *MNRAS*, **394**, 2050
- Murray, S., 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., Richards, G. T., et al. 2006, *ApJ*, **638**, 622
- Myers, A. D., Brunner, R. J., Richards, G. T., et al. 2007, *ApJ*, **658**, 99
- Neeleman, M., Novak, M., Venemans, B. P., et al. 2021, *ApJ*, **911**, 141
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, **266**, 713
- Ono, Y., Ouchi, M., Harikane, Y., et al. 2018, *PASJ*, **70**, S10
- Onoue, M., Matsuoka, Y., Kashikawa, N., et al. 2021, *ApJ*, **919**, 61
- Peebles, P. J. E. 1980, Large-Scale Structure of the Universe (Princeton, NJ: Princeton Univ. Press)
- Porciani, C., Magliocchetti, M., & Norberg, P. 2004, *MNRAS*, **355**, 1010
- Salpeter, E. E. 1964, *ApJ*, **140**, 796
- Schlafly, E. F., Finkbeiner, D. P., Jurić, M., et al. 2012, *ApJ*, **756**, 158
- Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, *AJ*, **134**, 102
- Seljak, U., & Warren, M. S. 2004, *MNRAS*, **355**, 129
- Serjeant, S., Oliver, S., Rowan-Robinson, M., et al. 2000, *MNRAS*, **316**, 768
- Shankar, F., Lapi, A., Salucci, P., De Zotti, G., & Danese, L. 2006, *ApJ*, **643**, 14
- Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, *ApJ*, **680**, 169
- Shen, Y., McBride, C. K., White, M., et al. 2013, *ApJ*, **778**, 98
- Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, *AJ*, **133**, 2222
- Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, *ApJ*, **697**, 1656
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, **323**, 1
- Sheth, R. K., & Tormen, G. 1999, *MNRAS*, **308**, 119
- Shimasaku, K., & Izumi, T. 2019, *ApJL*, **872**, L29
- Sinha, M., & Garrison, L. H. 2019, in Software Challenges to Exascale Computing, ed. A. Majumdar & R. Arora (Singapore: Springer), 3
- Sinha, M., & Garrison, L. H. 2020, *MNRAS*, **491**, 3022
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, *AJ*, **146**, 32
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Stiavelli, M., Djorgovski, S. G., Pavlovsky, C., et al. 2005, *ApJL*, **622**, L1
- Suchyna, E., Huff, E. M., Aleksić, J., et al. 2016, *MNRAS*, **457**, 786
- Timlin, J. D., Ross, N. P., Richards, G. T., et al. 2018, *ApJ*, **859**, 20
- Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, *ApJ*, **724**, 878
- Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, *ApJ*, **750**, 99
- Totsuki, H., & Kihara, T. 1969, *PASJ*, **21**, 221
- Trainor, R. F., & Steidel, C. C. 2012, *ApJ*, **752**, 39
- Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, *ApJ*, **574**, 740
- Uchiyama, H., Toshikawa, J., Kashikawa, N., et al. 2018, *PASJ*, **70**, S32
- Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, *MNRAS*, **453**, 2259
- Venemans, B. P., Walter, F., Zschaechner, L., et al. 2016, *ApJ*, **816**, 37
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Wang, F., Yang, J., Hennawi, J. F., et al. 2023, *ApJL*, **951**, L4
- Wang, R., Wagg, J., Carilli, C. L., et al. 2013, *ApJ*, **773**, 44
- White, M., Myers, A. D., Ross, N. P., et al. 2012, *MNRAS*, **424**, 933
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010, *AJ*, **140**, 546

- Willott, C. J., Bergeron, J., & Omont, A. 2015a, *ApJ*, **801**, 123  
Willott, C. J., Bergeron, J., & Omont, A. 2017, *ApJ*, **850**, 108  
Willott, C. J., Carilli, C. L., Wagg, J., & Wang, R. 2015b, *ApJ*, **807**, 180  
Willott, C. J., Delorme, P., Reylé, C., et al. 2009, *AJ*, **137**, 3541  
Willott, C. J., Percival, W. J., McLure, R. J., et al. 2005, *ApJ*, **626**, 657  
Woods, D., & Fahlman, G. G. 1997, *ApJ*, **490**, 11  
Yang, J., Wang, F., Fan, X., et al. 2020, *ApJL*, **897**, L14  
Yang, J., Wang, F., Fan, X., et al. 2021, *ApJ*, **923**, 262  
Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, *ApJ*, **630**, 1  
Zeimann, G. R., White, R. L., Becker, R. H., et al. 2011, *ApJ*, **736**, 57