



No Evidence for a Significant Evolution of M_{\bullet} – M_{*} Relation in Massive Galaxies up to $z \sim 4$

Yang Sun¹, Jianwei Lyu (吕建伟)¹, George H. Rieke¹, Zhiyuan Ji¹, Fengwu Sun², Yongda Zhu¹, Andrew J. Bunker³, Phillip A. Cargile², Chiara Circosta^{4,5}, Francesco D’Eugenio^{6,7,8}, Eiichi Egami¹, Kevin Hainline¹, Jakob M. Helton¹, Pierluigi Rinaldi¹, Brant E. Robertson⁹, Jan Scholtz^{6,7}, Irene Shvaei¹⁰, Meredith A. Stone¹, Sandro Tacchella^{6,7}, Christina C. Williams^{1,11}, Christopher N. A. Willmer¹, and Chris Willott¹²

¹ Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85719, USA; sunyang@arizona.edu

² Center for Astrophysics—Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

³ Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

⁴ European Space Agency (ESA), European Space Astronomy Centre (ESAC), Camino Bajo del Castillo s/n, 28692 Villanueva de la Cañada, Madrid, Spain

⁵ Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

⁶ Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK

⁷ Cavendish Laboratory—Astrophysics Group, University of Cambridge, 19 JJ Thomson Avenue, Cambridge, CB3 0HE, UK

⁸ INAF—Osservatorio Astronomico di Brera, via Brera 28, I-20121 Milano, Italy

⁹ Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

¹⁰ Centro de Astrobiología (CAB), CSIC-INTA, Ctra. de Ajalvir km 4, Torrejón de Ardoz, E-28850, Madrid, Spain

¹¹ NSF’s National Optical-Infrared Astronomy Research Laboratory, 950 North Cherry Avenue, Tucson, AZ 85719, USA

¹² NRC Herzberg, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada

Received 2024 September 09; revised 2024 November 22; accepted 2024 November 23; published 2024 December 27

Abstract

Over the past two decades, tight correlations between black hole masses (M_{\bullet}) and their host galaxy properties have been firmly established for massive galaxies (with stellar mass $\log(M_{*}/M_{\odot}) \gtrsim 10$) at low- z ($z < 1$), indicating coevolution of supermassive black holes and galaxies. However, the situation at high- z , especially beyond cosmic noon ($z \gtrsim 2.5$), is controversial. With a combination of JWST Near Infrared Camera (NIRCam)/wide field slitless spectroscopy (WFSS) from FRESCO, CONGRESS and deep multiband NIRCam/image data from JADES in the GOODS fields, we study the black-hole-to-galaxy mass relation at $z \sim 1$ –4. After identifying 18 broad-line active galactic nuclei (AGNs) at $1 < z < 4$ (with 8 at $z > 2.5$) from the WFSS data, we measure their black hole masses based on broad near-infrared lines (Pa α , Pa β , and He I $\lambda 10833$ Å), and constrain their stellar masses from AGN-galaxy image decomposition or spectral energy distribution decomposition. Taking account of the observational biases, the intrinsic scatter of the M_{\bullet} – M_{*} relation, and the errors in mass measurements, we find no significant difference in the M_{\bullet}/M_{*} ratio for $2.5 < z < 4$ compared to that at lower redshifts ($1 < z < 2.5$), suggesting no evolution of the M_{\bullet} – M_{*} relation at $\log(M_{*}/M_{\odot}) \gtrsim 10$ up to $z \sim 4$.

Unified Astronomy Thesaurus concepts: Active galactic nuclei (16); Supermassive black holes (1663); Active galaxies (17); Galaxy evolution (594)

1. Introduction

With the discovery of quasars (C. Hazard et al. 1963; M. Schmidt 1963), accretion onto supermassive black holes (SMBHs) has been appreciated as the second major source of electromagnetic radiation in the Universe, next to stellar radiation. The argument that the evolution of these two fundamental energy sources might be linked is highly influential in modern astronomy and has fostered numerous investigations on how the correlation between SMBHs and their hosts is established and maintained (see reviews by, e.g., D. M. Alexander & R. C. Hickox 2012; J. Kormendy & L. C. Ho 2013; T. M. Heckman & P. N. Best 2014; C. M. Harrison 2017). The potential scenarios range from a direct causal connection such as feedback due to winds and outflows launched by the active galactic nuclei (AGNs) that regulate the growth of the host galaxy (e.g., V. Springel et al. 2005; P. F. Hopkins et al. 2008a, 2008b; A. C. Fabian 2012), to a simple consequence of the growth of galaxies through

merging without a physical coupling between the galaxy and black hole growth (e.g., C. Y. Peng 2007; K. Jahnke & A. V. Macciò 2011).

Insights into this relationship and the underlying physical mechanisms can be obtained from how it evolves with cosmic time. In the low- z Universe, the masses of SMBHs (M_{\bullet}) have been firmly established to correlate with many properties of their hosts, notably the mass (M_b) of the spheroid component (e.g., J. Magorrian et al. 1998; N. Häring & H.-W. Rix 2004; J. Kormendy & L. C. Ho 2013) and its velocity dispersion (σ_b ; L. Ferrarese & D. Merritt 2000; K. Gebhardt et al. 2000; S. Tremaine et al. 2002). At high- z , due to various observational limitations, the explorations have been largely focused on the ratio between the SMBH masses and total stellar masses (M_{*}) M_{\bullet}/M_{*} of the quasar population up to $z \sim 2$ –2.5, and the results were controversial. At the massive regime ($\log(M_{*}/M_{\odot}) \gtrsim 10$), several observational works claimed that galaxies at $z \sim 2$ tend to have higher M_{\bullet}/M_{*} than the local value (e.g., C. Y. Peng et al. 2006; A. Merloni et al. 2010; Y. Zhang et al. 2023), suggesting a significant evolution of this mass scaling relation over the past ~ 9 Gyr. Meanwhile, many other studies (e.g., M. Schramm & J. D. Silverman 2013; M. Mechtley et al. 2016; X. Ding et al. 2020; H. Suh et al. 2020; J. I. H. Li et al. 2023; G. Mountrichas 2023)

found no significant evolution of M_*/M_* . The lower-mass regime ($\log(M_*/M_\odot) < 10$), is not well explored in the standard scaling relations (A. E. Reines & M. Volonteri 2015; J. E. Greene et al. 2020; but see R. Pucha et al. 2024). A few low-mass galaxies with broad-line AGNs (BL AGNs) at $0.4 < z < 3$ have been reported recently, and they overall have higher M_*/M_* than the local relation (M. Mezcuca et al. 2023, 2024), but this result has not yet been integrated into the broader context of low-mass galaxies overall.

A partial explanation for these discrepancies is that there are multiple measurement biases affecting the M_*/M_* determinations at high redshift. For example, the single-epoch SMBH masses are derived from ultraviolet lines of C IV and Mg II whose widths can be increased by outflows and other effects (e.g., Y. Shen & X. Liu 2012; H. A. N. Le et al. 2020; W. Zuo et al. 2020). The AGNs most readily observed will be the most luminous and hence will tend to be overmassive relative to typical behavior, resulting in a bias toward high values of M_*/M_* (e.g., C. J. Willott et al. 2005; T. R. Lauer et al. 2007), which is the well-known ‘‘Lauer bias.’’ In fact, A. Schulze & L. Wisotzki (2011, 2014) tested this possibility and found that the differences in a number of studies in the slope of the M_*/M_* relation at high and low redshift disappears when corrected for the ‘‘Lauer bias.’’ Taking the selection bias into account, it seems a consensus is emerging that there is little evolution of M_*/M_* up to $z \sim 2.5$ (M. Sun et al. 2015; H. Suh et al. 2020; J. I. H. Li et al. 2023; G. Mountrichas 2023; T. S. Tanaka et al. 2024), although this issue is perhaps not fully settled (e.g., Y. Zhang et al. 2023).

With the successful launch and operation of JWST, the study of the M_*-M_* relation has been pushed to much higher redshifts. Quasars at $z \sim 6$ with direct host stellar emission constraints from Near Infrared Camera (NIRCam; M. J. Rieke et al. 2023) appear to show relatively large values of M_*/M_* ($\log(M_*/M_*) \sim -1$) compared to their lower-redshift counterparts ($\log(M_*/M_*) \sim -2.5$) (X. Ding et al. 2023; M. A. Stone et al. 2024; M. Yue et al. 2024). Significant evolution in this relation has also been suggested for relatively less-massive SMBHs in Seyfert luminosity AGNs at $4 < z < 7$ (Y. Harikane et al. 2023; R. Maiolino et al. 2023; F. Pacucci et al. 2023; H. Übler et al. 2023). These results reveal a sharp contrast with those at $z \lesssim 2.5$, resulting in the need to fill in the redshift gap between $z \sim 2.5$ and $z \sim 4$ to establish a complete picture on the evolution of M_*-M_* relation across the cosmic time.

In this work, we will present a comprehensive study on the M_*-M_* relation in massive galaxies at $z \sim 1-4$ based on a powerful combination of the NIRCam/Wide-Field Slitless Spectroscopy (WFSS) surveys FRESCO (PID: 1895; P. A. Oesch et al. 2023) and CONGRESS (PID: 3577; F. Sun et al. 2024, in preparation) and deep multiband NIRCam images from JADES (D. J. Eisenstein et al. 2023) in GOODS-S and GOODS-N fields (M. Giavalisco et al. 2004). In contrast to previous ground-based $z \sim 1-2$ studies limited to rest-frame UV to optical lines, we are able to accurately constrain the black hole masses from the near-infrared (NIR) broad emission lines (Paschen lines and He I $\lambda 10833$ Å), which are much less affected by dust extinction and galaxy contamination. Moreover, the superior spatial resolution and sensitivity of multiband NIRCam images enable robust AGN-galaxy decomposition with a wide range of wavelength coverage, allowing for accurate measurements of host stellar masses from spectral energy distribution (SED) fittings on the AGN-subtracted galaxy emission. Finally, the survey nature of NIRCam/WFSS

data as well as the deep multiwavelength coverage from X-ray to the radio data in the GOODS fields provide the peerless opportunity to build a complete sample and understand the selection biases.

This paper is organized as follows: We introduce the AGN sample and the relevant JWST data for this project in Section 2. Section 3 describes how the black hole and galaxy masses of our sample are measured. In Section 4, we present the measurements of the black hole to galaxy mass ratios at $z \sim 1-4$ and analyze these results on a basis consistent with the approach used in the lower-redshift studies and evaluate measurement biases with a Monte Carlo (MC) method. We discuss the possible implications of our results for studies at higher redshift in Section 5 and conclude this work in Section 6.

Throughout this paper, we assume a standard cold dark matter (Λ CDM) Universe with cosmological parameters $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.7$, and $\Omega_m = 0.3$.

2. Data and Sample Selection

2.1. Spectroscopic Data

To search for AGN broad-line emission, we used the spectroscopic data obtained with JWST/NIRCam WFSS from the FRESCO survey (P. A. Oesch et al. 2023) in the F444W band ($\lambda \sim 3.9-5.0 \mu\text{m}$) and in both GOODS-S and GOODS-N, and the CONGRESS survey (Sun et al. 2024, in preparation) in the F356W ($\lambda \sim 3.1-4 \mu\text{m}$) band in GOODS-N only. The former survey covers a 7.4×8.4 area in both GOODS fields with the row-direction grisms on both modules of JWST/NIRCam, providing a spectral resolution of $\sim 1590-1680$ from $3.9-5.0 \mu\text{m}$. The CONGRESS program covers the GOODS-N field with a nearly identical footprint to FRESCO and with a spectral resolution of $\sim 1400-1610$ from $3.1-4.0 \mu\text{m}$. The two programs reach similar line sensitivities of $2 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$.

All of these NIRCam/WFSS data were processed by the publicly available reduction routine presented in F. Sun et al. (2023).¹³ We first processed the NIRCam data through the standard JWST stage-1 calibration pipeline¹⁴ v1.11.2. For each individual grism exposure, we assigned world coordinate system (WCS) information, performed flat-fielding, and removed the σ -clipped median sky background. The WCS of the grism exposures is registered to Gaia DR3 (Gaia Collaboration et al. 2023) using the NIRCam short-wavelength imaging data taken at the same time. For each of our targets, we extracted the 2D spectra from individual grism exposures, and coadded them in a common wavelength (1 nm pixel^{-1}) and spatial (0.06 pixel^{-1}) grid. Prior to our scientific spectral extraction, we also extracted spectra of bright point sources ($\lesssim 21 \text{ AB mag}$) to assure the accuracy of spectral tracing function and spectral flux calibration. We also extracted the spectra of galaxies with known ground-based spectroscopic redshifts, measuring the line center of detected Paschen α and β lines to ensure the wavelength calibration error at $< 1 \text{ nm}$. We then optimally extracted the 1D spectra of our targets from coadded 2D spectra using their surface brightness profile in the F444W band (K. Horne 1986).

¹³ doi:10.5281/zenodo.14052875 (F. Sun 2024).

¹⁴ doi:10.5281/zenodo.8140011 (H. Bushouse et al. 2024).

2.2. Imaging Data

Multiband NIRC*am* images with superior spatial resolution and sensitivity are available for both the GOODS-S and GOODS-N fields from JADES (D. J. Eisenstein et al. 2023). We used these images to conduct AGN-galaxy morphology decomposition and stellar mass estimation. The total overlapping area between the JADES and FRESCO footprints is approximately 35 square arcminutes in GOODS-N and about 46 square arcminutes in GOODS-S. JADES also overlaps with CONGRESS in GOODS-N in a similar area since CONGRESS covers almost the same area as FRESCO there. Seven NIRC*am* broadband images were selected to cover a wide range of wavelengths: F090W, F115W, F150W, F200W, F277W, F356W, and F444W. For galaxies at $z \sim 1-3.5$, these NIRC*am* filters nicely cover the rest optical to NIR SEDs of the galaxies, allowing for reasonable constraints on the galaxy stellar properties. The angular resolutions of these images range from $0''.030-0''.145$, corresponding to physical scales of $\sim 0.2-1.2$ kpc for galaxies at $z \sim 1-3.5$.

The four BL AGNs among our sample¹⁵ within the FRESCO area but outside the JADES footprint do not have full NIRC*am* wide-band data, but do have NIRC*am* F182M, F210M, and F444W data from the FRESCO survey. In addition, deep multiband images in the optical to the NIR are also available from HST/ACS and HST/WFC3 (e.g., A. M. Koekemoer et al. 2011). However, given the limited NIRC*am* bands as well as the different spatial resolutions of JWST and HST in the NIR, we decided not to conduct image decompositions for those four targets but to carry out SED decompositions of the integrated photometry (see details in Section 3.2.2).

2.3. BL AGN Sample Selection

Considering the huge number of galaxies observed by FRESCO and CONGRESS and the frequent galaxy contamination, a blind search for AGN broad-line features from the grism data across the field is not very practical. Therefore, we built an AGN sample with the available multiwavelength data first and then inspected the corresponding NIRC*am*/grism spectra to identify BL AGNs. This selection approach is consistent with previous studies of the mass scaling relation at $z \lesssim 2.5$ that were based on the AGN samples selected by AGN features in multiwavelength ranges and follow-up broad-line detection. In fact, the GOODS-S and GOODS-N fields have been extensively covered by ground- and space-based telescopes with the deepest X-ray, optical, infrared, and radio data in the sky, offering the best resources to build a complete AGN sample. With a combination of Chandra, Hubble, Spitzer, and JVA data, J. Lyu et al. (2022) carried out a comprehensive pre-JWST search of AGN in GOODS-S and reported ~ 900 candidates across the field. With the newly obtained JWST JADES/NIRC*am* and SMILES/MIRI data (S. Alberts et al. 2024; G. Rieke et al. 2024), J. Lyu et al. (2024) further improved the AGN census near the central region of GOODS-S. We combined the AGN catalogs reported in these two papers and extracted the FRESCO NIRC*am* F444W/grism spectra. For GOODS-N, following the same techniques in J. Lyu et al. (2022), a similar panchromatic pre-JWST AGN search with relatively shallower X-ray and radio data has been conducted with ~ 700 AGN candidates revealed

in the field (J. Lyu et al. 2024, in preparation). We started from this GOODS-N AGN sample and extracted the corresponding FRESCO and CONGRESS grism spectra in F444W and F356W.

The final sample analyzed in this work was selected by requiring a broad-line detection in the FRESCO or CONGRESS spectra of the AGN sample described above. We limited the redshift range to be within $1 < z < 4$; thus, Pa α , Pa β , or He I $\lambda 10833$ Å are covered by the FRESCO or CONGRESS wavelength ranges. In other words, our BL AGNs must have at least one of the three broad NIR lines mentioned above in their spectra. For the initial sample selection, we used the redshift values on the SIMBAD website, as reported by previous works. Although most of these galaxies have well-constrained spectroscopic redshifts, some have only photometric redshifts with relatively large uncertainties. Therefore, we also remeasured spectroscopic redshifts for them during the detailed line profile fitting (see Section 3.1).

In the next step, to identify BL AGN candidates, we did a preliminary fit to the line profiles with a Gaussian component, selecting BL AGNs from our initial sample by requiring a broad NIR line ($\text{FWHM} > 1000 \text{ km s}^{-1}$). This line width threshold is consistent with other BL AGN identification works using JWST NIRC*am*/WFSS data (J. Matthee et al. 2023) or NIR BL features (F. Ricci et al. 2022). After we finalized our sample, we had eight AGNs in the GOODS-S field and 11 in the GOODS-N field. After visually inspecting their JWST/NIRC*am* multiband images, we removed one of the BL AGN candidates in the GOODS-N field (GN-1030801) from our sample since it has complicated structures in its central region, which can intrinsically broaden the width of the targeted NIR lines and thus cause an overestimate of the line width of the broad component.

Strong outflows could complicate our determination of AGN properties from line widths. For example, F. D'Eugenio et al. (2024) reported our BL AGN candidate GS-197911 as a post-starburst galaxy hosting an AGN and with strong ionized outflow given its blueshifted and broad [O III] 5008 Å. In this case, the broad-line component of the He I $\lambda 10833$ Å line may be contaminated by the outflow signatures. However, given that its preliminary broad-line FWHM is about 3600 km s^{-1} , which is much larger than the [O III] width ($\sim 1800 \text{ km s}^{-1}$), we conclude that its broad-line component is dominated by AGN broad-line region (BLR) emission, and the outflow contamination would not significantly influence the black hole measurement. Thus, we still keep GS-197911 in our sample. Similarly, we rule out the possibility of outflow-dominated broadening for other BL AGN candidates after conducting detailed multi-component broad-line fitting (see Section 3.1). Therefore, we finalized our sample with 18 AGNs.

The positions on the sky of our sample are shown in Figure 1, and their properties are shown in Table 1. The multiwavelength studies from J. Lyu et al. (2022) reported that all of our BL AGNs have been classified as X-ray AGNs, and the majority of them are also identified in the mid-IR as expected, since they have broad NIR emission lines. Also, their AGN continua are generally obscured in the short-wavelength range. In this case, even though all of our BL AGNs have Chandra X-ray detections (B. Luo et al. 2017), we used the SED-derived bolometric luminosities from J. Lyu et al. (2022) and converted them to get an equivalent X-ray 2–10 keV luminosity using the X-ray bolometric

¹⁵ BL AGN selection will be described in Section 2.3.

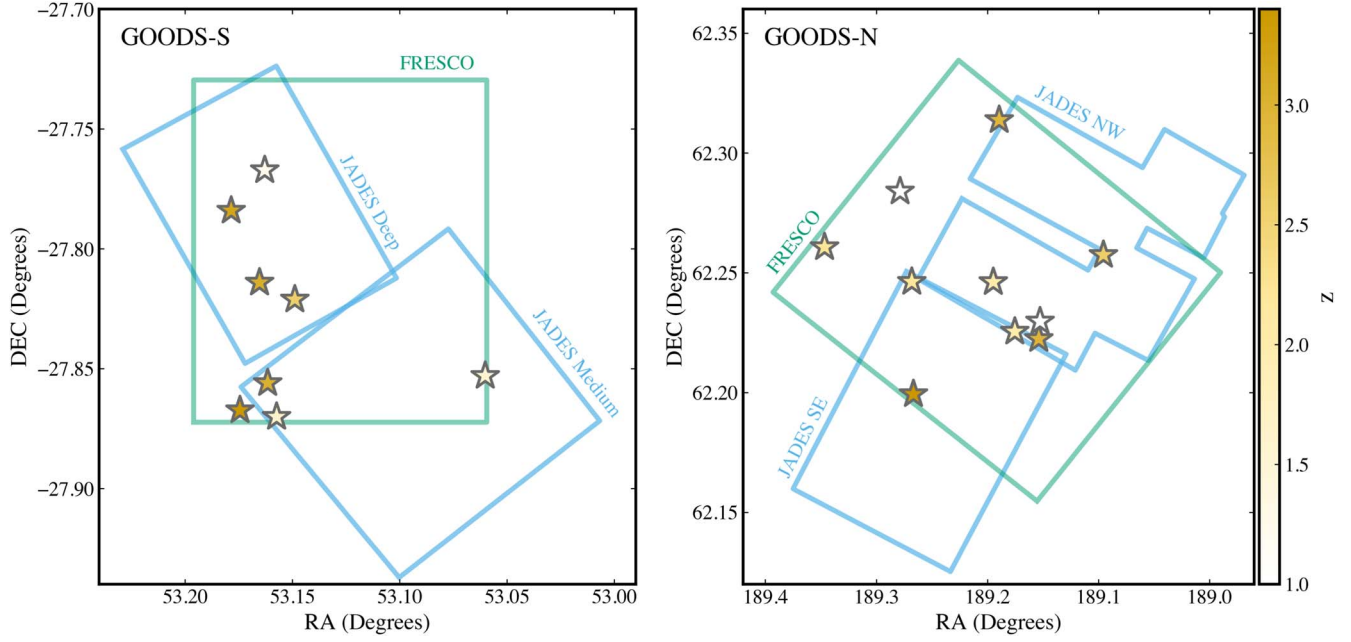


Figure 1. Spatial locations of our 18 $1 < z < 4$ BL AGN sample (stars) in the GOODS-S (left) and GOODS-N (right) fields, color-coded by redshift. The JADES and FRESKO footprints are drawn with blue and green lines, respectively.

Table 1
Properties of the GOODS AGN with NIRC*am*/WFSS NIR Broad-line Detections

ID ^a	R.A. (deg)	Decl. (deg)	z	$\log L_{\text{bol}}^b$ (erg s^{-1})	$\log L_X^c$ (erg s^{-1})	Broad Line	$\text{FWHM}_{\text{int}}^d$ (km s^{-1})	$\log M_*$ (M_\odot)	$\log M_*^e$ (M_\odot)
GS-173091	53.1573639	-27.8701553	1.61	46.17	44.57	Pa α	2391 ± 49	8.07	10.67
GS-243640	53.174408	-27.8674202	3.586	46.18	44.58	He I	3722 ± 671	8.46	11.22*
GS-182930	53.1615181	-27.8560753	3.029	45.78	44.34	He I	1337 ± 68	7.45	10.54
GS-184451	53.0601387	-27.8530674	1.539	44.10	43.06	Pa α	1821 ± 259	7.10	10.90
GS-197911	53.1653061	-27.8141308	3.067	46.25	44.62	He I	3470 ± 1008	8.49	11.19
GS-206907	53.1785088	-27.7841015	3.191	45.20	44.16	He I	5563 ± 419	8.60	9.90
GS-212097	53.1628799	-27.7672272	1.22	45.17	43.93	Pa α	4725 ± 65	8.34	10.83
GS-196290	53.1488495	-27.8211861	2.584	46.38	44.69	He I	1267 ± 69	7.58	11.10
GN-1000721	189.153763	62.2223206	2.948	44.36	43.32	He I	3125 ± 1009	7.68	10.50
GN-1028801	189.095581	62.2574081	2.588	46.32	44.66	Pa β	4941 ± 140	8.75	10.24
GN-1077827	189.266602	62.1992989	3.408	46.15	44.56	He I	4049 ± 1005	8.53	10.26
GN-1095877	189.189438	62.3136978	2.927	45.92	44.43	He I	3352 ± 303	8.30	11.26*
GN-1025078	189.152634	62.22966	0.96	45.14	43.90	Pa α	2058 ± 54	7.61	9.98
GN-1095207	189.278656	62.2839203	1.022	45.63	44.25	Pa α	3460 ± 42	8.23	9.96*
GN-1083261	189.268066	62.2461662	2.217	46.21	44.60	He I	1997 ± 342	7.93	10.18
GN-1094302	189.346619	62.2606583	2.244	45.31	44.03	He I	4884 ± 763	8.42	10.44*
GN-1027287	189.194733	62.2460823	2.004	45.98	44.47	He I	5429 ± 1029	8.73	10.37
GN-1024921	189.175369	62.2253914	2.019	44.59	43.46	Pa β	1453 ± 106	7.10	10.72

Notes.

^a ID is the combination of “GS-” (GOODS-S) or “GN-” (GOODS-N) with JADES ID.

^b AGN bolometric luminosity derived from SED fitting, except for GS-184451, 206907, and GN-1000721 whose L_{bol} is converted from the observed X-ray 2–10 keV luminosity.

^c X-ray 2–10 keV luminosity converted from the L_{bol} using the F. Duras et al. (2020) bolometric correction, except for GS-184451, 206907 and GN-1000721, which is the observed X-ray 2–10 keV luminosity from B. Luo et al. (2017).

^d Intrinsic full-width at half maximum (FWHM) of the broad component corrected for the instrumental and morphological broadening.

^e Stellar mass derived by imaging decomposition and point-spread function (PSF)-subtracted flux SED fitting, while those marked with “*” are derived from SED decomposition due to lack of JWST/NIRC*am* wide-band photometry.

correction by F. Duras et al. (2020) for BH mass estimation (see Section 3.1). This strategy is adopted by default, as the X-ray intrinsic luminosity values reported in, e.g., B. Luo et al. (2017) are based on simple modeling of the source

X-ray flux band ratios, while the SED fitting approach takes care of the obscuration in a more sophisticated way. We note that the SED-derived bolometric luminosities of GS-184451 and 206907 from J. Lyu et al. (2022), and of GN-1000721

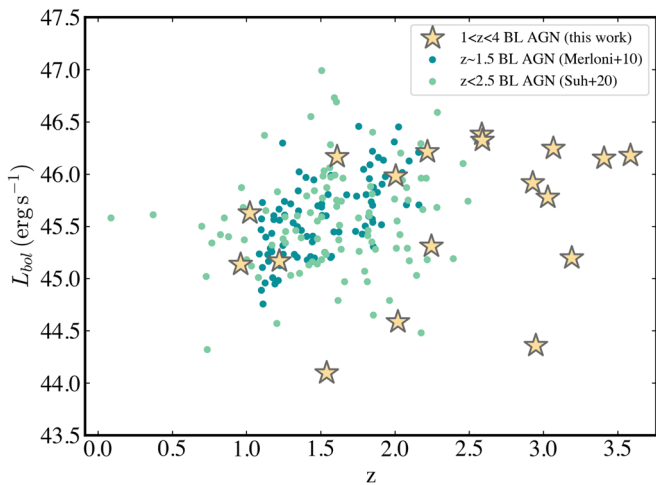


Figure 2. Distribution of the sample of 18 BL AGNs on the $L_{\text{bol}}-z$ plane. The bolometric luminosity of our AGNs ranges from 10^{44} – $10^{46.5}$ erg s^{-1} . Compared to the BL samples at $z \sim 2$ from A. Merloni et al. (2010; dark green) and H. Suh et al. (2020; light green), our $1 < z < 4$ BL AGN sample is on average similar but contains examples with relatively lower L_{bol} ($L_{\text{bol}} < 10^{45}$ erg s^{-1}) and higher redshift ($z > 2.5$).

from Lyu et al. (2024, in preparation), are quite low ($L_{\text{bol}} < 10^{44}$ erg s^{-1}), while they do have strong X-ray emission ($L_{\text{X}} \sim 10^{43}$ – 10^{44} erg s^{-1}) contributed by the AGN (B. Luo et al. 2017). Similar X-ray bright but SED nondetected or unidentified AGNs have been reported in previous work (J. Lyu et al. 2022, 2024). After checking their SED fittings, we concluded that these three sources do not have enough mid-IR photometric data to give meaningful constraints on the AGN component. As a result, we adopted their Chandra X-ray measurements to compute X-ray 2–10 keV luminosities and the bolometric luminosities using the same bolometric correction. The final distribution of our BL AGNs on the AGN bolometric luminosity versus redshift plane is shown in Figure 2. Most of them help bridge the gap from previous measurements between $z \sim 2.5$ and $z = 4$.

3. Black Hole and Galaxy Masses

3.1. Black Hole Masses

Masses of the black holes in our $z \sim 3$ AGNs were estimated by the single-epoch virial mass method, based on the velocity width of the broad-line component and the AGN luminosity in a specific band (M. Vestergaard 2002). F. Ricci et al. (2017) compared the widths of the three lines we use ($\text{Pa}\alpha$, $\text{Pa}\beta$, and $\text{He I } \lambda 10833 \text{ \AA}$) and found that for 90% of their sample, He I widths track closely those in $\text{H}\alpha$; as expected, $\text{Pa}\alpha$ and $\text{Pa}\beta$ also closely track $\text{H}\alpha$. We therefore use the measured line widths interchangeably with no corrections.

We fitted the spectra to determine the central wavelength and the line widths. We began by using the preexisting spectroscopic or photometric redshift to select the wavelength range containing the line (± 5000 km s^{-1}). Then, we set the peak wavelength as an initial guess for the central wavelength. We used one or two Gaussian components to model the profile of the broad NIR lines ($\text{Pa}\alpha$, $\text{Pa}\beta$, and $\text{He I } \lambda 10833 \text{ \AA}$). Since the typical FWHM of the broad component is around 2000 km s^{-1} (e.g., H. Landt et al. 2008), we selected $0.05 \mu\text{m}$ wide regions to either side to define the continuum, lying beyond two to three times the FWHM of the broad component

(± 4000 – 6000 km s^{-1} , varied case by case to capture a clean continuum). We applied linear fits to these regions to determine the continuum level.¹⁶

After subtracting the continuum, we used the Python package `lmfit` (M. Newville et al. 2014) to fit the line profile. We first tried two Gaussian components. The center of the narrow component was allowed to shift within ± 200 km s^{-1} , while the center of the broad component was more flexible (± 1000 km s^{-1}), given that previous works have found the broad component can be blue- or redshifted relative to the narrow component (e.g., T. E. Zastrocky et al. 2024). We also required the FWHM of the narrow (broad) component to be no larger (smaller) than 1000 km s^{-1} (except for GN-1000721 and 1095207, whose narrow component slightly exceeds this threshold). After visual inspection, we found the above procedure works for most of our AGNs, except for the following:

1. For GS-182930, 196290, and GN-1024921, we found their narrow component is either too faint compared to the broad component or too narrow (< 50 km s^{-1}). Therefore, we think the signal of the narrow component of the line is washed out by the noise. Thus, we only used one Gaussian component to fit the line and treated it as the broad component.
2. For GN-1027287, 1077827, 1083261, 1094302, and 1095877, the Paschen γ line (10941 \AA) has nonnegligible emission and is blended with the targeted $\text{He I } \lambda 10833 \text{ \AA}$ line. Therefore, an additional Gaussian component was assigned for modeling Paschen γ , where the components have the same offset and width as the narrow component of the $\text{He I } \lambda 10833 \text{ \AA}$ line.
3. For GN-1028801, an intermediate component needed to be assigned for the Paschen β line to improve the fitting. Multiple broad components have been commonly seen in BL AGNs (e.g., F. Ricci et al. 2022; L. Kuhn et al. 2024), which might trace the complicated geometry and dynamics of the BLR of AGNs (B. M. Peterson 2006; L. Č. Popović et al. 2019).

Examples of the fitting results are shown in Figure 3 (the rest of them are shown in Appendix A and Figure A1). The observed FWHMs of the broad components of our AGNs are between 1300 and 6000 km s^{-1} . The spectroscopic redshift was also calculated based on the wavelength offset of the center of the narrow component (for two-/three-component cases) or the broad component (for one-component cases) relative to its rest-frame wavelength.

Then, to quantify the measurement errors of the line profile parameters and the spectroscopic redshift, we applied an MC simulation, i.e., generating 100 mock spectra based on the observed flux errors and re-fitting them to get the distributions of best-fit parameters.

Next, we derived the intrinsic FWHM of the broad component by correcting the instrumental and morphological broadening effects. Given the NIRCam/WFSS instrumental resolution at the wavelength range of 3 – $5 \mu\text{m}$ is about 1500 , the typical instrumental broadening of FWHM ($\text{FWHM}_{\text{inst}}$) is 200 km s^{-1} . To evaluate the morphological broadening effect, we collapsed the imaging data of each source at the observed wavelength of the broad line along the dispersion direction, and

¹⁶ For GS-196290, we only fit the red continuum region with a width of $0.1 \mu\text{m}$, since the targeted line lies at the blue end of the spectrum.

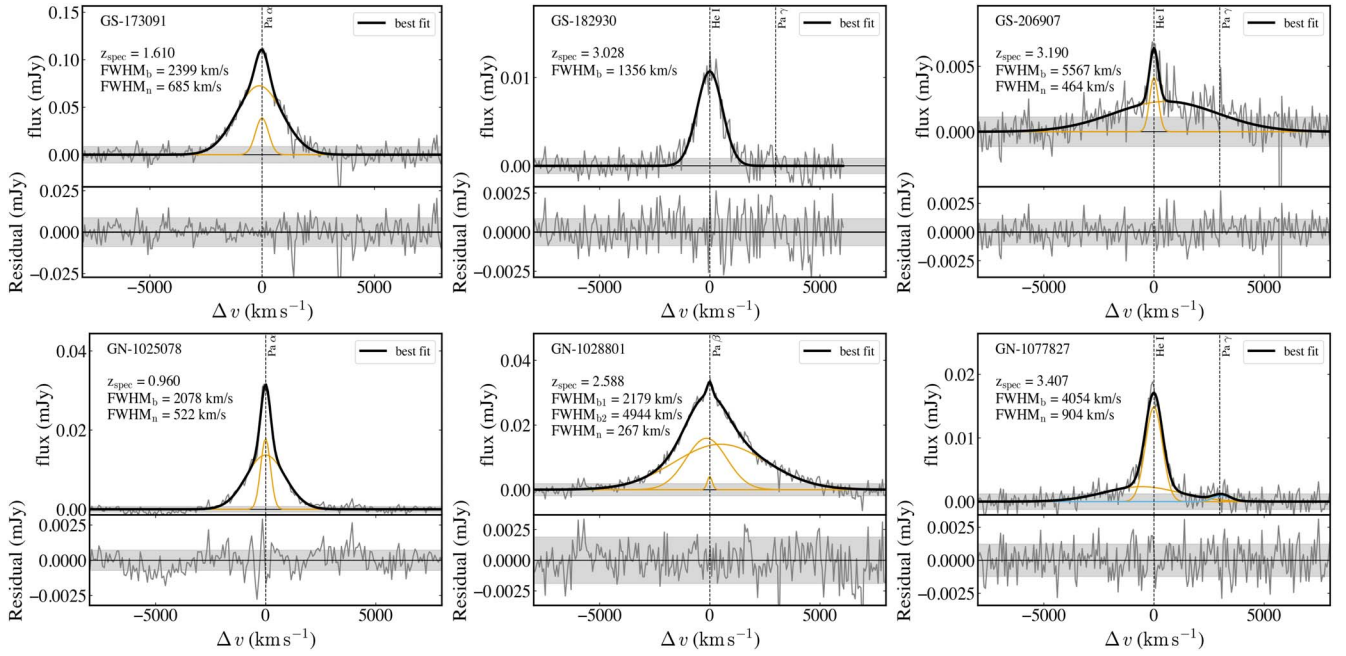


Figure 3. Examples of line profile fitting for the broad Pa α , Pa β , and He I λ 10833 \AA lines. In each panel, we show the extracted 1D spectrum (continuum subtracted) and 1σ error level (gray line and shading), the best-fit line profile (black), and each component (orange lines are for the targeted lines (Pa α /Pa β /He I). The blue line in the bottom-right panel is an additional component for close Pa γ), and the residual of the best fit.

derived a Gaussian convolution kernel such that the convolved point-spread function (PSF) profile would match the observed brightness profile of the source. The angular FWHM of this convolution kernel then corresponds to an FWHM of the grism spectrum, which is the morphological broadening width of the observed broad spectral line ($\text{FWHM}_{\text{morph}}$). The FWHM of the convolution kernel for our samples ranges from $0''.09$ – $0''.3$, corresponding to the $\text{FWHM}_{\text{morph}}$ range of 90 – 350 km s^{-1} . We then derived the intrinsic FWHM of the broad component by $\text{FWHM}_{\text{int}} = \sqrt{\text{FWHM}_{\text{obs}}^2 - \text{FWHM}_{\text{inst}}^2 - \text{FWHM}_{\text{morph}}^2}$. We notice that such correction for our sample is small ($<50 \text{ km s}^{-1}$).

Then, using the errors derived from this MC simulation and then requiring the intrinsic FWHM of the broad component to be at least 1σ higher than 1000 km s^{-1} and the peak flux of the broad component to be 1σ higher than the 1σ error level of the spectrum, we confirmed all 18 of our BL AGNs have a detected broad NIR line.

We finally used the virial mass estimation relation reported by F. Ricci et al. (2017):

$$\log\left(\frac{M_{\bullet}}{M_{\odot}}\right) = 8.03 + 2 \log\left(\frac{\text{FWHM}_{\text{NIR,int}}}{10^4 \text{ km s}^{-1}}\right) + 0.5 \log\left(\frac{L_X}{10^{42} \text{ erg s}^{-1}}\right), \quad (1)$$

where L_X is the AGN X-ray luminosity at 2 – 10 keV ($L_{2-10 \text{ keV}}$). This relation was measured with a scatter of 0.4 dex and a virial factor of $f = 4.31$. Also, this virial BH mass relation can be applied to all three NIR lines (Pa α , Pa β , and He I) given that F. Ricci et al. (2017) found a good agreement between the FWHMs of H α and of those three NIR emission lines. L_X of our AGNs was determined as described at the end

of Section 2.3. The final M_{\bullet} of our AGN sample ranges from 10^7 – $10^9 M_{\odot}$, and the individual measurements are shown in Table 1. The typical uncertainty of the BH mass propagated from the uncertainty of the FWHM measurements is about 0.1 dex. Combined with the systematic uncertainty of F. Ricci et al.’s (2017) relation (0.4 dex), the overall uncertainty of our BH mass measurements is ~ 0.4 dex.

In addition, for cases with more than one component, we can also measure the velocity offset of the broad component relative to the narrow one. We found, other than the known outflow host GS-197911, that there are six more BL AGNs (GS-184451, 173091, GN-1027285, 1095207, 1077827, and 1000721) with a broad component blueshifted more than 100 km s^{-1} from the systemic velocity, which could be alternatively explained by galactic outflows rather than AGN BLRs. However, we rule out the possibility of outflow-dominated broadening for all of them:

1. For GS-184451, GN-1024921, 1027285, 1077827, and 1000721, we detected a nearby narrow forbidden line (e.g., [S III] λ 5931 \AA or [Fe II] λ 1.257, $1.644 \mu\text{m}$) in the same grism spectrum, for which the width is much narrower than the broad component of the targeted BLR-indicator line. This indicates that their broad NIR lines are dominated by AGN BLR emission rather than outflows.
2. For GS-173091 and GN-1095207, the grism spectra do not show any other significant emission lines except for the targeted NIR broad line. However, given that their broad components are only blueshifted by ~ 100 and 400 km s^{-1} , respectively, which are common in BL AGN observations (Y. Shen et al. 2016; T. E. Zastrocky et al. 2024), and are dramatically broad (2399 and 3470 km s^{-1} , respectively), we conclude that their broad lines are dominated by AGN BLRs as well.

3.2. Host Galaxies

In this Section, we introduce the two approaches applied to measuring the stellar mass of the BL AGN hosts. For the AGNs that have JADES JWST/NIRCam broadband images, we conduct AGN-galaxy imaging decomposition to measure the fluxes of the hosts, then use SED fitting with a custom setup for the AGN component (J. Lyu et al. 2022) in *Prospector* (B. D. Johnson et al. 2021) to derive the stellar masses. For the four BL AGNs outside of the JADES footprint for which we cannot do the same imaging decomposition as the other 14 AGNs, we estimate the stellar masses by decomposing their UV to mid-IR SEDs with a modified *Prospector* SED fitting using the semiempirical AGN component.

3.2.1. AGN-galaxy Decomposition

We used the Python package *galight* (X. Ding et al. 2020) to decompose the light from AGNs and galaxies on JADES NIRCam broadband images (from F090W to F444W), by fitting them with a PSF component and a Sérsic component (with a free Sérsic index (n) convolved with a PSF. We used a single Sérsic profile; although this approach is not able to determine the substructure, if any, our primary goal is to obtain the overall mass of the galaxy, which is not sensitive to the detailed morphology.

First, the target images were cut out from all of the JADES JWST/NIRCam broadband drizzled images in the GOODS-S and GOODS-N field (with a nominal resolution of ~ 0.03), with a field of view of $6'' \times 6''$ to make sure all light coming from our sources is included. The PSFs applied during the fitting are from the JADES collaboration (see Appendix A in Z. Ji et al. 2023 for details). All nearby objects (detected in the segmentation map at a signal-to-noise ratio (S/N) > 3 in the JADES photometric catalog) were masked before fitting. There is a bright companion to the southeast of GS-206907 shown in all NIRCam images, even though it can only be deblended by the segmentation maps from F090W to F277W. Therefore, we used an additional Sérsic model to fit it simultaneously in these five short-wavelength bands, instead of masking.

The full fitting routine supported by *galight* starts with the Particle Swarm Optimizer (J. Kennedy & R. Eberhart 1995) to find the best-fit model. Then, the inferred minimized parameters are passed to the Markov Chain Monte Carlo (MCMC) routine to estimate the posterior parameter distributions. The values at the peak of each posterior parameter distribution are the final best-fit parameters. Finally, the fluxes of the AGNs and host galaxies are measured by summing up the model light of the best-fit PSF and Sérsic components, respectively.

Figure 4 shows an example of the *galight* fits in the seven NIRCam broadband images. The “data – point source” images reveal the light from the host galaxies, and the 1D surface brightness profiles tell us whether the AGN or the galaxy dominates the total brightnesses of the sources. By visually inspecting the host galaxy images (“data – point source”) and then requiring the host to be well detected at least in the rest-frame $1 \mu\text{m}$ band, we confirmed that the hosts of the 14 BL AGNs that have images covering all NIRCam broad bands are well detected. In addition, the “residual” images in Figure 4 illustrate the difference between the *galight* best-fit model (PSF + Sérsic) and the observed fluxes.

We also tested the robustness of our galaxy flux measurements by subtracting the *galight* best-fit point-source flux from the PSF-convolved KRON fluxes for the galaxies from the JADES DR2 catalog.¹⁷ We confirmed that this measure of galaxy flux (JADES total flux–PSF flux) is similar to the Sérsic galaxy flux.

The MCMC results provide the uncertainties of the best-fit parameters, but there are some concerns that they might be underestimated (X. Ding et al. 2023; T. S. Tanaka et al. 2024). Therefore, we used the MC method to derive the errors to see if they agree with the MCMC errors reported by *galight*. We generated 100 mock images for each target based on the observed flux errors and re-fitted them to get the distributions of best-fit parameters. Overall, the MC error of the galaxy magnitude is comparable to the MCMC one, which is typically ~ 0.01 mag. Also, the error in stellar mass propagated from the galaxy flux error is < 0.1 dex, which is much smaller than the error associated with the SED fitting to convert the flux to mass (~ 0.3 dex; see Section 3.2.2). Therefore, we still use the MCMC errors as the galaxy flux errors.

Even though studying host galaxy properties other than stellar mass, e.g., morphology and size, is beyond the scope of this work, we briefly point out here that the distribution of the Sérsic indices, especially those in the long-wavelength filters, peaks at $n \sim 1$ – 2 . However, as J. Krywult et al. (2017) have pointed out, the Sérsic index for bulge and disk galaxies decreases with redshift. Namely, $n \sim 1$ at $1 < z < 4$ does not necessarily mean the galaxies are disk-like. Therefore, we are conservative about whether our AGNs are mostly hosted by late-type (disk-dominated) galaxies, although the host galaxy type (i.e., evolutionary stage) of the AGN sample at different redshifts is critical for studying the time evolution of M_* – M_* relation (see Section 5). Meanwhile, the best-fit galaxy effective radius of our 18 BL AGNs is typically $0.2''$, corresponding to physical scales of ~ 1.5 kpc for galaxies at $z \sim 1$ – 3.5 . In this case, it is very difficult to tell the existence of a bulge or a disk, and further to separate its light contribution to the total galaxy flux to derive the bulge or disk mass. Therefore, we are only able to study the M_* – M_* relation for our BL AGN sample at $1 < z < 4$, rather than the M_* – M_b relation.

3.2.2. Galaxy Stellar Mass Measured by *Prospector* SED Fitting

The stellar masses of the AGN hosts were measured by fitting the PSF-subtracted fluxes with the software *Prospector*. A delayed-tau star formation history with Kroupa initial mass function and the M. Kriek & C. Conroy (2013) extinction law for the stellar continuum were assumed. Since the point-source flux is removed by PSF subtraction, we did not add any AGN component, but only incorporated stellar components and associated nebular emission-line components and then derived the stellar mass from the best-fit models. One example of the galaxy SED fitting is shown in Figure 5. The typical statistical uncertainty of the *Prospector* stellar mass is < 0.1 – 0.15 dex, which is smaller than the systematic errors introduced by different SED model assumptions, such as star formation history (0.2 – 0.3 dex). As a result, we only consider systematic errors during the following analysis.

For GS-243640, GN-1095877, 1095207, and 1094302, which do not have sufficient multiband NIRCam images for

¹⁷ <https://archive.stsci.edu/hlsp/jades>

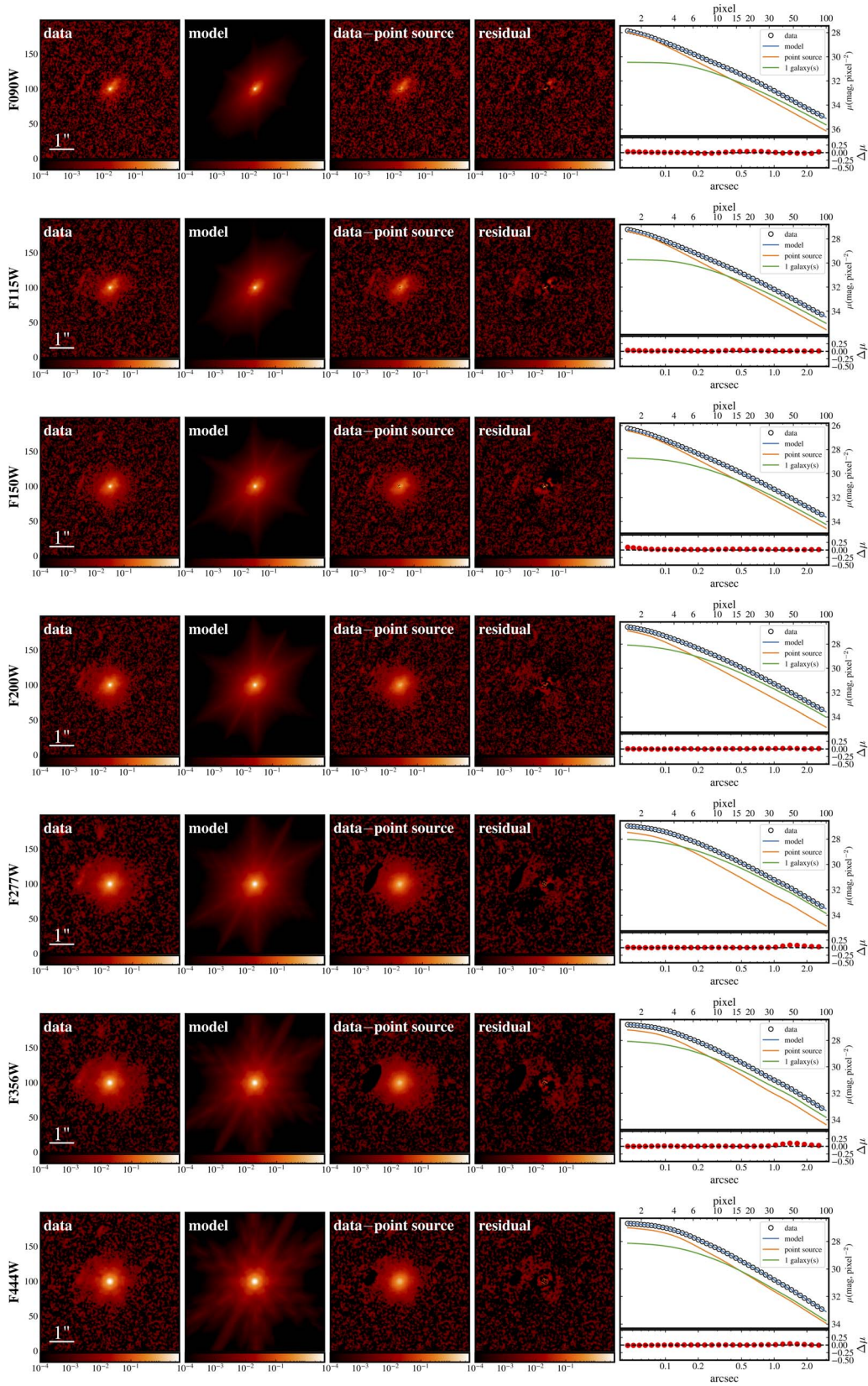


Figure 4. Example *galight* fits of all seven JWST/NIRCam broadband images (F090W to F444W from top to bottom) for GN-1083261. The columns from left to right show the data, model (PSF+Sérsic), data – point source (host galaxy), residual (data – model) images, and 1D surface brightness profiles. From the data – point source image, we confirmed the detection of the host galaxy.

accurate imaging decomposition, we derived their stellar masses using SED decomposition, i.e., fitting the UV-to-mid-IR integrated galaxy emission separated from the total SED.

We used the measurements based on CANDELS/SHARDS photometry from G. Barro et al. (2019) and fitted with the modified Prospector code described in J. Lyu et al. (2024),

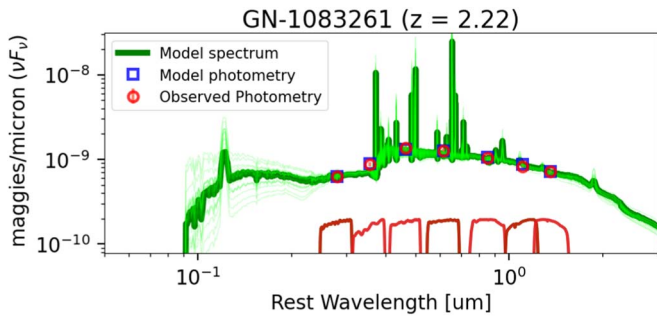


Figure 5. Example of the `Prospector` galaxy SED fitting for GN-1083261, using the PSF-subtracted fluxes derived from `galight`. The PSF-subtracted fluxes (red circles) are fitted with an SED model (green lines) that contains the stellar component and the nebular emissions. The blue squares represent the best-fit model photometry. The bottom red enclosed regions show the transmission curves of F090W, F115W, F150W, F200W, F277W, F356W, and F444W from left to right.

where a semiempirical model for AGN UV to mid-IR continua with nebular emission lines and dust attenuation is introduced to replace the default AGN torus model in `Prospector`. We compared the stellar mass derived from the SED decomposition method with the imaging-decomposition stellar mass for the BL AGNs with imaging-decomposition stellar mass as well, confirming that the difference between them is < 0.3 dex (the typical error of the stellar mass by SED fitting), and thus using SED-derived stellar mass for these four BL AGNs does not have any impacts on our results of the $M_{\bullet}-M_{*}$ relation. The stellar masses of our AGNs are listed in Table 1; overall, they are hosted by massive galaxies ($\log(M_{*}/M_{\odot}) \gtrsim 10$).

4. Results

4.1. Approach

To study the evolution of the $M_{\bullet}-M_{*}$ relation up to $z \sim 4$, we need to put our analysis on the same basic foundation as the studies for $z = 1$ to 2.5. The M_{\bullet}/M_{*} ratio is best behaved for the galaxy bulges (J. Kormendy & L. C. Ho 2013), but for high-redshift samples, it is not always possible to isolate galaxy bulge, especially beyond cosmic noon, as we pointed out in Section 3.2.1. In this case, relations to the integrated stellar output are used, particularly those that A. E. Reines & M. Volonteri (2015) and J. E. Greene et al. (2020) developed for local galaxies. Since local AGNs are typically in late-type (disk-dominated, relatively lower-mass) galaxies, their relation is usually assumed. However, high-redshift studies identify AGNs primarily in relatively massive, early-type (bulge-dominated) host galaxies. For example, most of the AGN hosts (26 out of 38) in J. I. H. Li et al. (2023) at $z \sim 0.2-0.8$ are early-type systems (light-blue points in Figure 6). Therefore, the studies of AGN hosts at $z \gtrsim 0.2$ find that M_{\bullet}/M_{*} for them lies well above the local late-type/AGN relation from A. E. Reines & M. Volonteri (2015), approaching the relation for early-type galaxies (see Figure 6). To maintain consistency, studies of the BH–galaxy mass relation should be based on a reasonable reference M_{\bullet}/M_{*} ratio appropriate for the inspected stellar mass range. In Section 5, we will discuss the issue of the M_{\bullet}/M_{*} ratio comparison for samples at different evolutionary stages.

A second consideration is how to quantify the observational biases (e.g., T. R. Lauer et al. 2007). We do this with an MC program similar to that introduced by J. Li et al. (2021), in

which the observations are simulated according to (1) a distribution of galaxy masses (i.e., stellar mass function, SMF); (2) an assumed intrinsic behavior of M_{\bullet}/M_{*} with a scatter; (3) a BL AGN fraction (fraction of galaxies containing a BL AGN); (4) an input Eddington ratio distribution function (ERDF); and (5) measurement uncertainties for stellar mass and black hole mass. Each MC trial is for a mock galaxy having a true stellar mass from the SMF but with observables according to an intrinsic M_{\bullet}/M_{*} ratio sampled from the assumed intrinsic relation, BL AGN fraction, ERDF, and errors. The trial then yields a mock galaxy with observed stellar mass and black hole mass, as well as AGN luminosity and broad line width if it hosts an active nucleus. By applying the observational limit, we will obtain the ensemble of results distributed according to these uncertainties and subject to the “Lauer bias.”

A feature of our approach is that we simulate the actual observations we could take in an area-limited survey like FRESCO, i.e., with the appropriate number of BL AGNs for a single set of observations. Using this actual observation design, we compare the observed distribution of $\log(M_{\bullet}/M_{*})$ with the simulated ones. We use the Kolmogorov–Smirnov (K-S) test to determine the probability that the observed distribution could be drawn by chance from the distribution defined by the many simulated observations.

4.2. Determination of the $M_{\bullet}-M_{*}$ Relation

The offset in the $M_{\bullet}-M_{*}$ relation in the local Universe relative to that *observed* in AGNs at moderately high redshift (Figure 6) is likely due to a selection bias, but must be taken into account when evaluating evidence for evolution. The different choices of the reference benchmark can cause disagreements in the existence of redshift evolution in $M_{\bullet}-M_{*}$ relation for $z \leq 2$, even though the derived intrinsic M_{\bullet}/M_{*} ratios at $1 < z < 2$ generally agree with each other reasonably well. For example, for $1 < z < 2.5$, H. Suh et al. (2020) found that the average ratio of $\log(M_{\bullet}/M_{*})$ is -2.64 for galaxies with $8.5 < \log(M_{\bullet}/M_{*}) < 9.5$ and -3.00 for galaxies with $7 < \log(M_{\bullet}/M_{*}) < 8.5$, and -2.50 when their sample is combined with the A. Merloni et al. (2010) sample. Similarly, without differentiating for black hole mass, G. Mountichias (2023) found an intrinsic average value of $\log(M_{\bullet}/M_{*}) \sim -2.50$, and K. Setoguchi et al. (2021) obtained -2.22 for this redshift range. These values are consistent with each other, while not comparable with the ratio often used from A. E. Reines & M. Volonteri (2015), for which $\log(M_{\bullet}/M_{*}) \sim -3.83$. We notice that a dozen rare low-mass AGN-host galaxies ($\log(M_{*}/M_{\odot}) \lesssim 9.5$) recently reported by M. Mezcuca et al. (2023, 2024) have significantly higher M_{\bullet}/M_{*} ratios ($\log(M_{\bullet}/M_{*}) \sim -1.5$) compared to the values for massive AGN hosts. Given that this work focuses on the mass scaling relation at the massive end, we do not include those BL AGNs in low-mass galaxies in the comparison. We will briefly discuss their different M_{\bullet}/M_{*} behaviors in Section 5.

Therefore, to capture the range of observed values for massive galaxies for $1 < z < 2.5$, we will carry out analyses for $\log(M_{\bullet}/M_{*}) = -2.50$ with an intrinsic scatter of 0.5 dex as the baseline (hereafter we refer to this as the “ $z \sim 1$ relation”), and also with $\log(M_{\bullet}/M_{*}) = -3.00$. These two values are appropriate for comparison with the other high-redshift studies at a similar stellar mass range, and using both gives assurance that our conclusions are not strongly dependent on the choice of fiducial mass ratio.

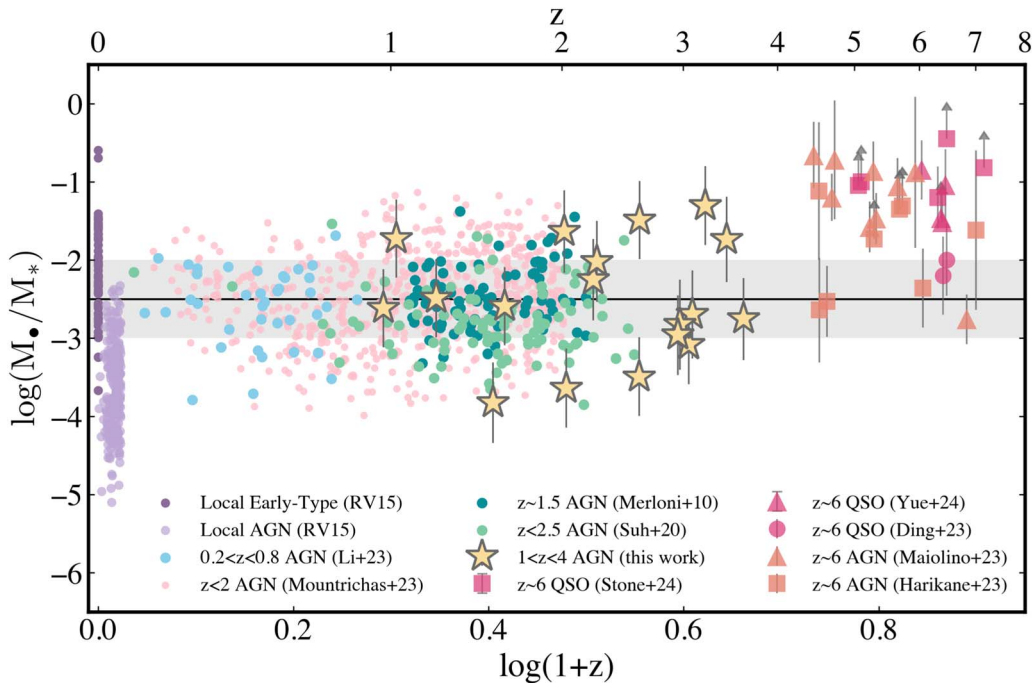


Figure 6. Summary of the M_{\bullet}/M_{*} behavior from the local Universe to $z = 7$. We include the A. Merloni et al. (2010) and H. Suh et al. (2020) samples, along with local AGNs and early-type galaxies from A. E. Reines & M. Volonteri (2015; light and dark purple), $0.2 < z < 0.8$ AGNs (light blue) from J. I. H. Li et al. (2023), and $z < 2$ AGNs (pink) from G. Mountrichas (2023). The $z > 4$ faint AGNs (magenta) from R. Maiolino et al. (2023) and Y. Harikane et al. (2023) and quasars (orange) from X. Ding et al. (2023), M. A. Stone et al. (2024), and M. Yue et al. (2024) are also shown as a reference. Our new BL AGN sample (yellow stars) at $1 < z < 4$ fills the redshift gap in mass scaling relation studies. The $z \sim 1$ relation ($\log(M_{\bullet}/M_{*}) = -2.5$ with a scatter of 0.5 dex) is represented by the black solid line and the gray shaded region. For $z \gtrsim 0.1$, the average $\log(M_{\bullet}/M_{*})$ values of massive AGNs fall well above the typical behavior for local AGNs (usually hosted by late-type galaxies) and are slightly below the behavior of local early-type galaxies (A. E. Reines & M. Volonteri 2015).

The average BH-to-galaxy mass ratio ($\log(M_{\bullet}/M_{*})$) for our BL AGN sample, ($\log(M_{\bullet}/M_{*}) = -2.61^{+0.90}_{-0.59}$), is consistent with the ratio for the $z \sim 1$ relation even before correcting for the observational biases (see the bias test in Section 4.3).

4.3. Modeling Observational Biases

In this Section, we apply an MC simulation to explore how observational biases affect the apparent $M_{\bullet}-M_{*}$ relation from our observations. Basically, we generate a mock AGN sample that represents the underlying AGN population at a specific redshift epoch and that follows the $z \sim 1$ relation; we then apply the observation effects to this population to make a mock “observable” AGN sample, and then we test how likely it is that the observed distribution on the $M_{\bullet}-M_{*}$ diagram is consistent with the mock distribution. In Section 4.3.1, we will first start with an ideal scenario in which the survey area is infinitely large to compare the intrinsic distribution and the observations on the $M_{\bullet}-M_{*}$ plane. In Section 4.3.2, we will simulate mock observations by restricting the simulated survey area to the size of our actual observations and applying a BL AGN fraction related to the whole galaxy population. The third and final step is to properly test how likely the $z \sim 1$ relation can reproduce the observed distribution. We discuss the procedures for generating mock AGN populations in Appendix B.

4.3.1. Infinitely Large Survey

We use 10^7 trials for randomly creating 10^7 mock BL AGNs with true stellar mass $M_{*,\text{true}}$ and redshift z_{true} from the galaxy SMF (J. R. Weaver et al. 2023). This arbitrary number is much

larger than the BL AGN population size in the FRESCO fields ($\sim 10^2$; see Section 4.3.2), so it can represent an infinitely large survey scenario. Assuming that the BL AGNs follow the $z \sim 1$ relation, we determine the true BH mass $M_{\bullet,\text{true}}$ and then the bolometric luminosity (L_{bol}) by adopting the intrinsic ERDF at $z \sim 2.15$ by B. C. Kelly & Y. Shen (2013).

Next, assuming the virial BH mass and stellar mass measurements are not biased from the true mass but are subject only to measurement errors, we randomly add a Gaussian error with a dispersion of 0.4 dex to the $M_{\bullet,\text{true}}$, and 0.3 dex to the $M_{*,\text{true}}$ to obtain $M_{\bullet,\text{obs}}$ and $M_{*,\text{obs}}$. These errors were determined from our mass measurement errors mentioned in Sections 3.1 and 3.2.

Finally, we apply the observational limits to select the “observable” mock AGNs. The first observational bias comes from the BL AGN selection. In this work, only the AGNs that have a measurable broad line ($\text{FWHM} > 1000 \text{ km s}^{-1}$) can be identified as BL AGNs and be used for single-epoch M_{\bullet} estimation. To estimate the FWHM of the broad line for mock AGNs, we use $M_{\bullet,\text{true}}$ and L_{bol} based on the F. Ricci et al. (2017) single-epoch BH mass relation. Another limit arises because of the spectral sensitivity of FRESCO/CONGRESS. Typically, they can reach to a line flux sensitivity of $\sim 5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ for observing a broad line of 1000 km s^{-1} (but shallower for detecting a broader line; see Appendix). We convert the line flux sensitivity to the spectroscopic AGN bolometric luminosity limit $L_{\text{lim,spec}}$ based on z_{true} (see Appendix B for details). Therefore, the mock AGNs with $L_{\text{bol}} < L_{\text{lim,spec}}$ are undetectable. Considering the two observational limits mentioned above, we simulate an

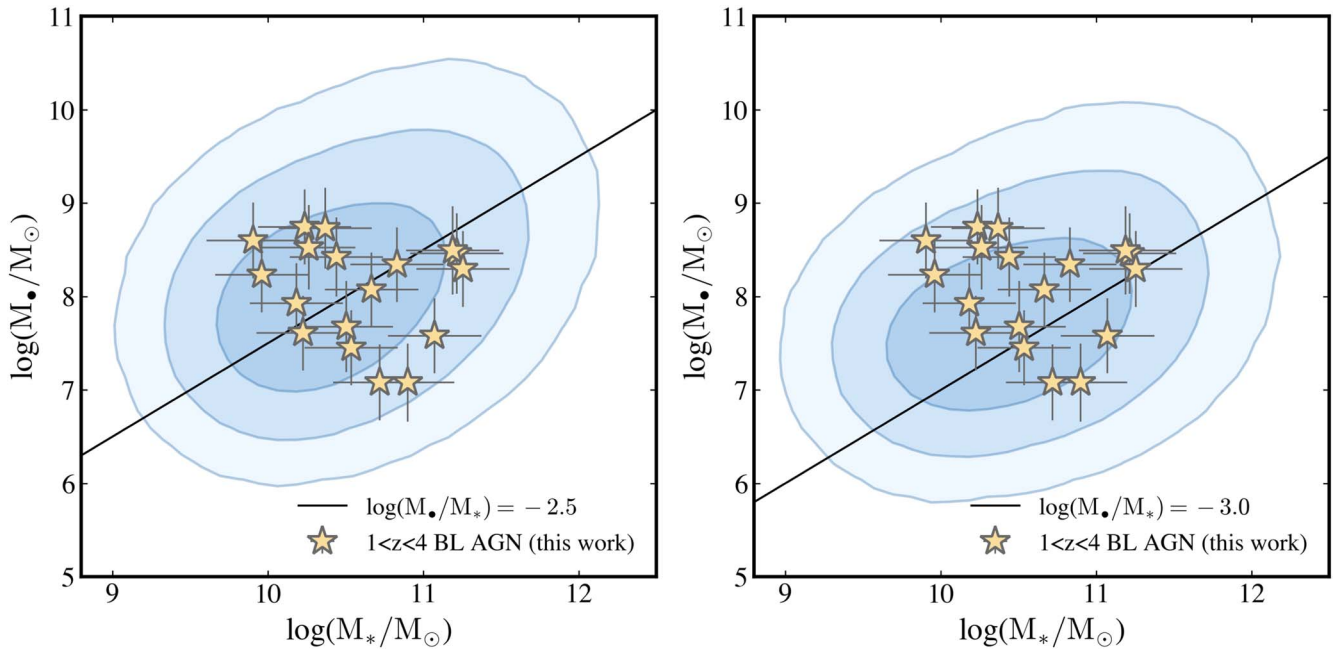


Figure 7. The distributions on the M_{\bullet} - M_* diagram at $1 < z < 4$ based on the assumption of $\log(M_{\bullet}/M_*) = -2.5$ (left) and -3.0 (right). The contours show the predicted distribution from our Monte Carlo simulation that includes both errors and biases. Specifically, the contours fall slightly above the input $\log(M_{\bullet}/M_*)$ primarily because of the “Lauer bias.”

“observable” BL AGN sample from the whole mock BL AGN sample.

The left panel of Figure 7 shows that the “observable” $M_{\bullet,\text{obs}}-M_{*,\text{obs}}$ distribution at $1 < z < 4$ is only slightly biased upward from the intrinsic $z \sim 1$ relation ($\log(M_{\bullet}/M_*) = -2.5$), indicating that our BL AGN sample at this redshift range is not much affected by observational biases. Also, all of our BL AGNs are enclosed within the 2σ contour of the “observable” distribution; namely, our mock AGN simulation can successfully reproduce the observed AGNs without invoking any evolution in M_{\bullet}/M_* .

To test the robustness of the no-evolution statement, we also check the result if we adopt $\log(M_{\bullet}/M_*) = -3.0$ as the $z \sim 1$ relation. We find, even for this lower ratio, our $1 < z < 4$ BL AGNs are still located in the central region of the “observable” distribution. This simulation and the results are very similar to the recent paper by J. Li et al. (2024), which showed that with many trials (i.e., our “infinite” sample), the characteristics of $4 < z < 7$ AGNs recently claimed to be above the local M_{\bullet}/M_* relation could be reproduced without requiring evolution in this ratio.

4.3.2. Area-limited Survey

The fact that the data points fall within a distribution does not necessarily indicate that these data points are consistent with the distribution. With a large enough sample, the distribution can successfully include the outliers, which may just reproduce some observed targets by chance. Moreover, in practice, our $1 < z < 4$ BL AGNs are selected from area-limited surveys rather than an infinitely large survey, such as we modeled in Section 4.3.1: the $1 < z < 4$ BL AGNs are blindly selected from FRESCO and CONGRESS.¹⁸

Therefore, to further quantify the probability that the BL AGNs observed in the FRESCO and CONGRESS areas follow the mock distribution on the M_{\bullet} - M_* plane assuming the $z \sim 1$ relation, we ran another MC simulation to model the observations of the BL AGN population within an FRESCO-like survey area. In this case, we are doing an apples-to-apples comparison in which the mock AGN observations will have a comparable sample size and parameter ranges (M_* and z) as the actual observations.

First, using the SMF, we estimate the expected number of $1 < z < 4$ (i.e., the redshift range of our BL AGN sample) galaxies appearing in the footprint of FRESCO’s survey area (122 arcmin^2) to be $\sim 1.3 \times 10^3$ over the stellar mass range of $\log M_*/M_{\odot} > 9.9$. Next, we run this number of trials to simulate the mock galaxy population following the $z \sim 1$ relation ($\log(M_{\bullet}/M_*) = -2.5$). We then apply a BL AGN fraction (the number ratio of BL AGNs to galaxies) from A. Schulze et al. (2015; $3\% \pm 1.2\%$, where the error is from their Figure 22) to this population to estimate the number of galaxies (or trials) to have a BL AGN ($\sim 40 \pm 15$ BL AGNs). After assigning an Eddington ratio, mass measurement uncertainties, and observational limits, as we did in Section 4.3.1, we generate a mock BL AGN observation. Finally, to exactly match the range of stellar mass between the mock observation and the true observation, we exclude the mock AGNs whose $M_{*,\text{obs}}$ is outside of the range of $9.9 < \log(M_*/M_{\odot}) < 11.3$.

By repeating the procedure of mock observation generation 10,000 times, the distribution of the “observable” AGN number counts in each mock observation with the intrinsic ratio of $\log(M_{\bullet}/M_*) = -2.5$ is plotted in the top-left panel of Figure 8, which demonstrates that the median number count of the simulated “observable” BL AGNs at $z \sim 2$ within an FRESCO-like survey area is about 24 ± 3 ; the observed number count ($N_{\text{obs}} = 18$) is within the expected 3σ error although slightly less than the predicted median. The generally successful

¹⁸ The FRESCO survey mapped the 61 arcmin^2 areas in both the GOODS-S field and GOODS-N field, and the CONGRESS survey covers a very similar footprint as the FRESCO GOODS-N field.

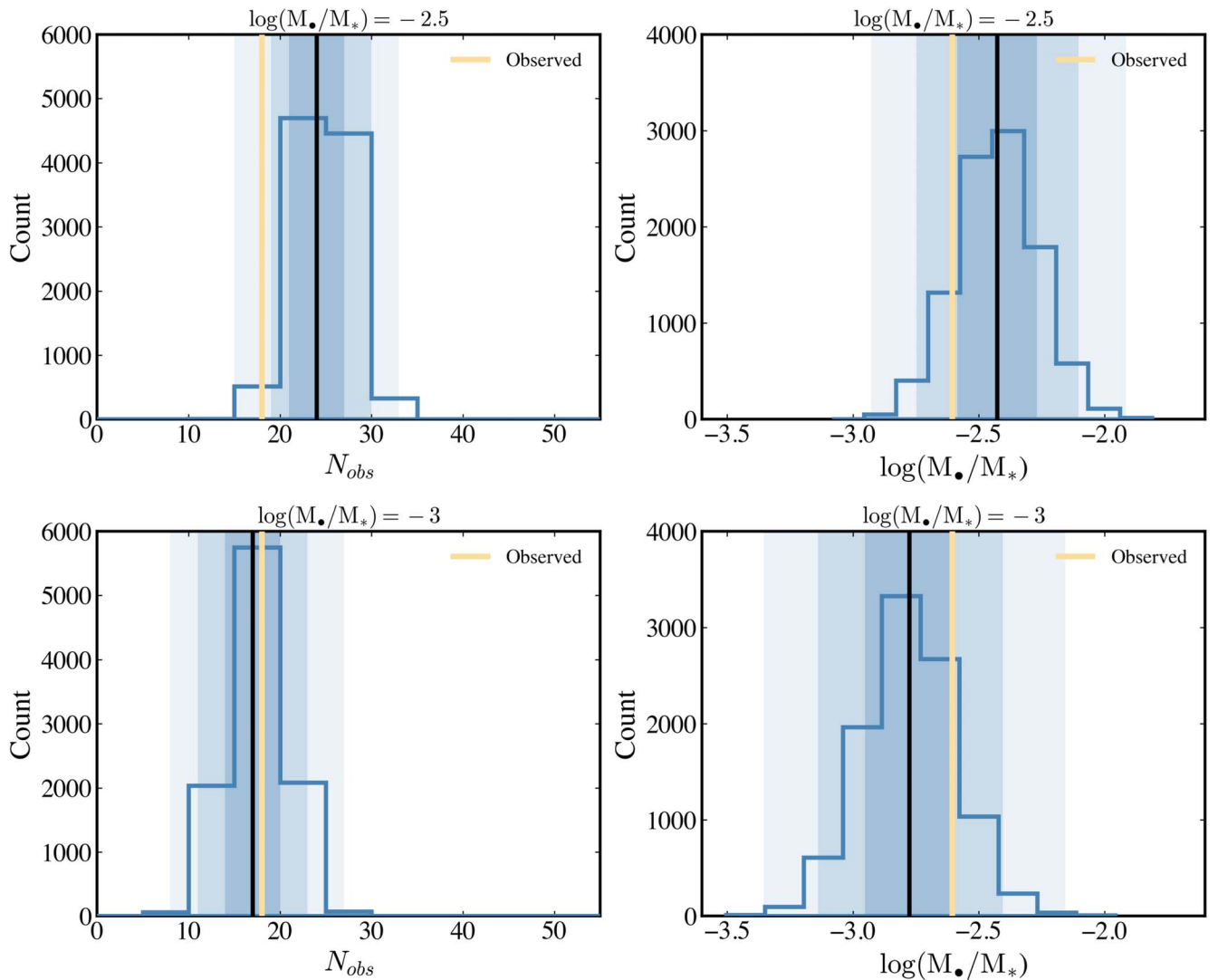


Figure 8. MC simulation results for the mock AGN sample assuming the underlying intrinsic relation is either $\log(M_*/M_*) = -2.5$ (top) or -3.0 (bottom). The blue histograms represent the mock distributions, the black lines represent the median values of the mock distributions, and the yellow lines represent the observed value of our AGN sample. Left panels: the distribution of the simulated “observable” mock AGN number count in a FRESCO-like area-limited survey. The observed number count ($N_{\text{obs}} = 18$) is slightly lower than the predicted number count if $\log(M_*/M_*) = -2.5$ ($N_{\text{obs}} = 24$, the difference is $<3\sigma$), while consistent with the predicted number count if $\log(M_*/M_*) = -3.0$ ($N_{\text{obs}} = 17$, the difference is $<1\sigma$). Right panels: the average $\log(M_*/M_*)$ distribution (blue open histograms) for 10,000 mock observations in the FRESCO fields, for which the ranges of M_* and z are matched with our observed BL AGN sample. The black lines represent the median $\log(M_*/M_*)$ of the observed AGNs. The differences between the peak values of the distribution for the simulated “observable” mock AGNs and the observed $\log(M_*/M_*)$ ratio of our $1 < z < 4$ AGN sample have significances of less than 2σ and 1σ , respectively, for the two simulations. Overall, the results show that our observed BL AGNs at $1 < z < 4$ are consistent with the simulated “observable” sample with either $\log(M_*/M_*) = -2.5$ or -3.0 .

matching of observable BL AGN number count between the simulation and observation confirms that our simulation accurately incorporates the correct assumptions of properties of the underlying galaxy and AGN populations (i.e., SMF, ERDF, BL AGN fraction), as well as the observational biases and measurement uncertainties. A marginally lower observed BL AGN number count compared to the prediction is not surprising given that we did not inspect every spectrum in FRESCO and CONGRESS, and we might also have rejected a few AGNs in our line width test.

Next, we test how likely it is that our observed $1 < z < 4$ BL AGNs are statistically consistent with the mock observations following the $z \sim 1$ relation ($\log(M_*/M_*) = -2.5$) by comparing their M_*/M_* distributions. The top-right panel of Figure 8 illustrates the distribution of the median $\log(M_*/M_*)$ of mock observations; the observed median $\log(M_*/M_*)$ falls within the

2σ range of the distribution of the mock observations. The bottom panels show the same information for $\log(M_*/M_*) = -3.0$; again there is no evidence for evolution.

4.4. Kolmogorov–Smirnov Tests

To further compare the full M_*/M_* distribution rather than just the median, we apply a K-S test between the mock observations and the actual one to investigate their consistency. We use two different approaches of the K-S test to make this comparison:

1. *One-sample Test.* For this test, we average all of the mock observations to obtain a noise-free reference distribution and do a K-S test between this reference and the actual observations. If the p -value of the K-S test is higher than 0.05, we can reject the hypothesis that the observed

AGNs are not consistent with the mock reference distribution. To make the reference distribution, we randomly select 18 AGNs from each mock observation if the predicted number count is no less than 18, which also, more or less, accounts for the potential incompleteness of our BL AGN sample. We rank the 18 $\log(M_*/M_*)$ values of all mock observations and then average all of the mock values at each rank to get the final averaged distribution.¹⁹ This approach reduces the uncertainty of the reference distribution introduced by the noise of each random simulated observation, providing a virtually noiseless distribution of observations from 18 targets.

2. *Two-sample Test.* In this case, we do a K-S test between each mock observation and the actual observation; namely, we make 10,000 K-S tests and calculate the fraction of test results for which we can reject the hypothesis that the observed AGN sample does not come from the same distribution as the mock AGNs (the p -value of the K-S test is higher than 0.05). This approach fully accounts for the fluctuations that exist in our small sample of observations and provides a statistical probability of rejecting the null hypothesis.

The one-sample test returns a p -value of 0.97 for $\log(M_*/M_*) = -2.5$ (see Figure 9). In $>99\%$ (9916/10,000) of the realizations of the two-sample test for $\log(M_*/M_*) = -2.5$, the p -values are >0.05 , indicating we can reject the hypothesis that the observed $\log(M_*/M_*)$ distribution is different from the mock distribution. All tests strongly indicate that there is no statistically significant difference of the observed $1 < z < 4$ sample from the expected no-evolution behavior at $z \sim 1-2.5$.

Also, we tested the case of $\log(M_*/M_*) = -3.0$ to see if the large dispersion of the M_*/M_* ratio at $z \sim 1-2.5$ could result in a different conclusion on the M_*/M_* evolution. The MC distributions of mock BL AGN number count and median $\log(M_*/M_*)$ are shown in the bottom panels of Figure 8. The K-S test between the reference mock $\log(M_*/M_*)$ distribution and the observed one is plotted in the bottom panel of Figure 9, with a p -value of 0.78. Additionally, $>96\%$ (9676/10,000) of the realizations of the two-sample test have a p -value higher than 0.05. We confirmed that, even for $\log(M_*/M_*) = -3.0$, there is no statistically significant deviation of the observed sample from the expected no-evolution behavior at $z \sim 1-2.5$.

Therefore, the observed $1 < z < 4$ BL AGNs are consistent with the mock distribution that is based on the $z \sim 1$ relation after taking into account the selection biases, suggesting no evolution of the M_*/M_* ratio at the massive end ($\log(M_*/M_\odot) \gtrsim 10$) at least up to $z \sim 4$. We have already shown this result in Figure 6. Figure 10 is another presentation that highlights the lack of change in the relation between M_* and M_* at $\log(M_*/M_\odot) \gtrsim 10$ between the previously studied lower-redshift AGNs and our sample.

5. Discussion

Determining the masses of high-redshift AGN-host galaxies directly from their stellar populations or stellar dynamics, in a similar manner to the basis for the determination of the local M_*/M_* relation, is paramount for characterizing the possible

¹⁹ We discard the mock observations with a predicted number count of <18 (1% for the $\log(M_*/M_*) = -2.5$ case and 50% for the $\log(M_*/M_*) = -3.0$ case) because it cannot be ranked with other mock observations including 18 AGNs.

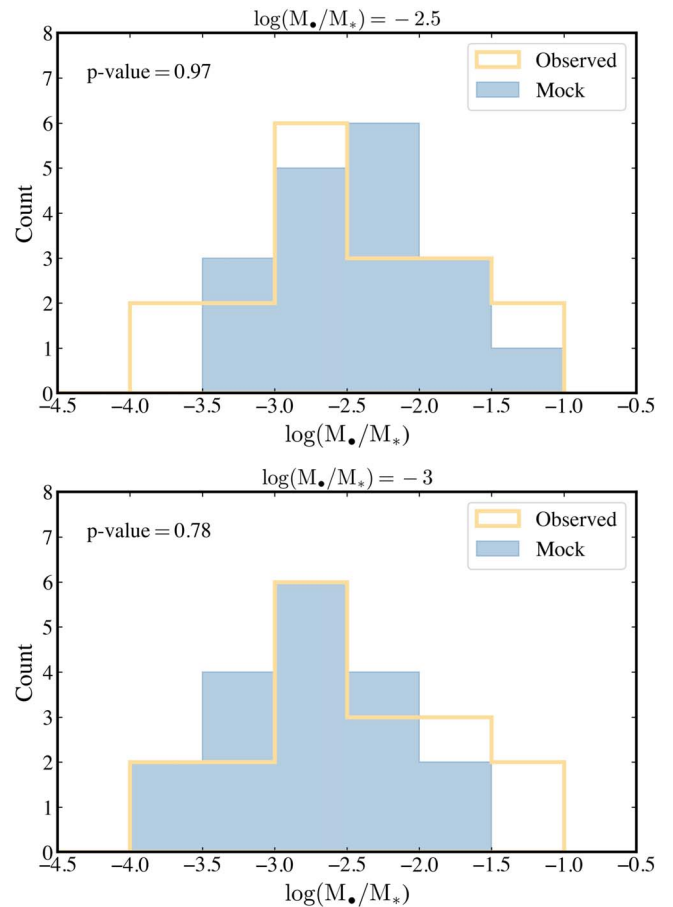


Figure 9. Comparison between the reference distribution of $\log(M_*/M_*)$ derived by averaging all of the mock observations (blue-filled histogram) and the observed distribution (yellow-open histogram), assuming the underlying intrinsic relation is either $\log(M_*/M_*) = -2.5$ (top) or -3.0 (bottom). The p -value of the K-S test suggests the two distributions have no significant difference, supporting the idea of no-evolution of the M_*/M_* ratio from $z = 1$ to 4 .

evolution of the SMBH–galaxy correlations. As a step toward this goal, we have shown that the observed M_*/M_* ratio for massive galaxies has no significant change from $z \sim 1$ to 4 . Along with previous findings of no-evolution up to $z \sim 1$ (e.g., H. Suh et al. 2020; J. I. H. Li et al. 2023; G. Mountrichas 2023), we conclude that the $M_- - M_*$ relation at $\log(M_*/M_\odot) \gtrsim 10$ up to $z \sim 4$ is consistent with the local one.

Our study indicates some cautions on studying the M_*/M_* evolution using the local A. E. Reines & M. Volonteri (2015) relation as the baseline. It should be noted that the local AGN sample in A. E. Reines & M. Volonteri (2015) is based on lower-mass late-type systems, while AGNs at $1 < z < 4$ are usually detected in more massive galaxies. The higher-redshift AGN hosts may also be bulge-dominated galaxies (e.g., J. I. H. Li et al. 2023); although, the morphology study of AGN hosts is still limited above $z \sim 1-2$ even in the JWST era. Therefore, the shift in the mixture of galaxy types from the AGN sample in high- z to that in the local Universe might be the reason for the significant change of the M_*/M_* ratio when comparing the value at high redshift with the local A. E. Reines & M. Volonteri (2015) value. Future detailed discussion on this issue will be provided in a forthcoming paper.

In addition, some previous studies at $z < 2.5$ only used an M_*/M_* value to quantify the $M_- - M_*$ relation, namely, by

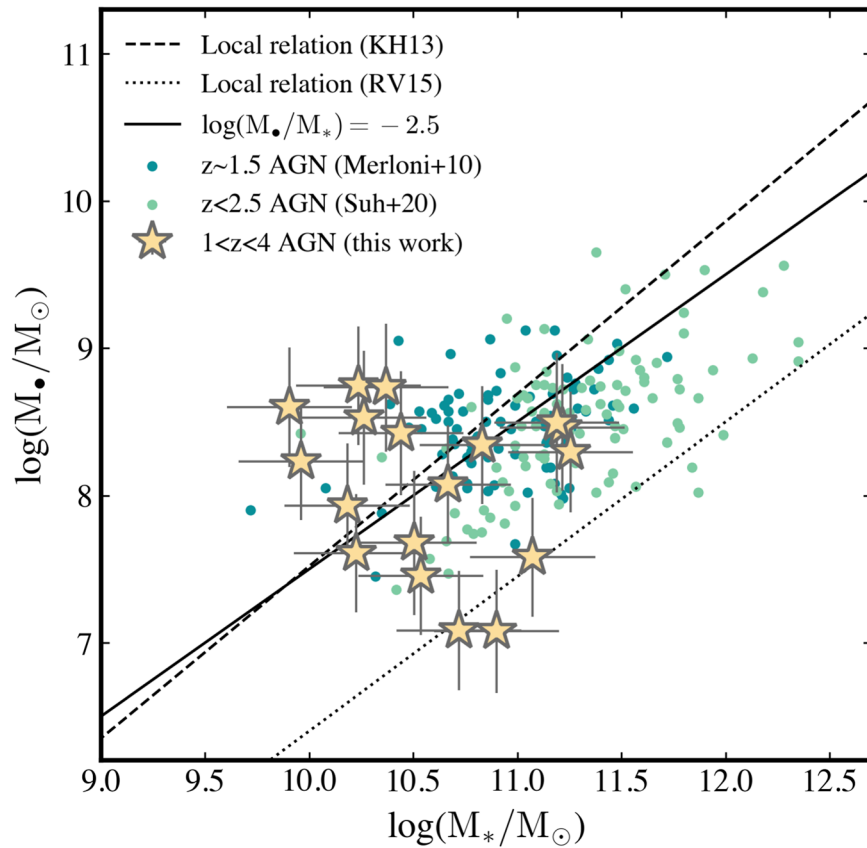


Figure 10. M_{\bullet} vs. M_* relation for the GOODS-S and GOODS-N AGN sample at $1 < z < 4$ (yellow stars). As a comparison, the AGN samples at cosmic noon ($z \sim 2$) from A. Merloni et al. (2010; dark-green dots) and from H. Suh et al. (2020; light-green dots) are plotted. The local M_{\bullet} – M_* relation for the bulge-dominated galaxies from J. Kormendy & L. C. Ho (2013) and that for the AGN hosts from A. E. Reines & M. Volonteri (2015) are illustrated by the black dashed line and the black dotted line, respectively. Also, we show the intrinsic BH-to-stellar mass ratio at $1 < z < 2$ ($\log(M_{\bullet}/M_*) = -2.5$) determined by previous studies (e.g., H. Suh et al. 2020; G. Mountrichas 2023) using the solid black line.

simply assuming the M_{\bullet}/M_* is independent of M_{\bullet} or M_* (the slope is always 1). Indeed, the relation from A. E. Reines & M. Volonteri (2015),

$$\log(M_{\bullet}/M_{\odot}) = 7.45 + 1.05 \times \log(M_*/10^{11}M_{\odot}), \quad (2)$$

suggests that M_{\bullet} is almost linearly proportional to the stellar mass M_* , so a simple re-normalization might seem to be adequate. However, other work has found the M_{\bullet}/M_* ratio can be different across a wide range of black hole masses or stellar masses. For example, H. Suh et al. (2020) found $\log(M_{\bullet}/M_*) \sim -2.64$ for $8.5 < \log(M_{\bullet}) < 9.5$ and $\log(M_{\bullet}/M_*) \sim -3.00$ for $7.0 < \log(M_{\bullet}) < 8.5$ at $z \sim 1.5$. These observations indicate that not only can the general host galaxy properties in the high- z samples be different from the low- z samples (e.g., those studied in A. E. Reines & M. Volonteri 2015), but also that the range of BH masses or stellar masses being sampled at high- z are also different than the range at low- z , resulting in an intrinsic difference in the M_{\bullet}/M_* ratios. Such behavior does not influence our results here since the stellar masses of our sample are relatively high and span a narrow range (~ 1 dex). However, it could complicate studies of BL AGNs in low-mass galaxies beyond the local Universe, especially the new JWST discoveries of faint AGNs in low-mass galaxies at $z > 4$, where the host galaxy properties are even more challenging to determine.

In this work, we have also evaluated the net errors and biases in our measurements based on the MC approach originally suggested by J. Li et al. (2021) with some additional improvements. We have shown that some considerably larger M_{\bullet}/M_* ratios similar to those observed at high- z can be produced by chance, sometimes after tens of millions of trials. However, the critical question is whether the observed ratios are likely in observational programs with a limited number of targets. The targets constitute the parent sample from which the observations are drawn, and a similar philosophy must be applied in the simulations. This distinction is important at high redshift, where the available samples are small. In a forthcoming paper, we will extend the test of the redshift evolution of M_{\bullet} – M_* relation up to $z \sim 7$ by using a similar analysis as developed in this paper.

Lastly, as we pointed out in Section 4.2, this work only focuses on the M_{\bullet} – M_* relation evolution between $1 < z < 4$ at the high-mass end because of the stellar mass range of our new BL AGN sample. M. Mezcua et al. (2024) recently discovered a dozen BL AGNs hosted by low-mass star-forming galaxies at $1 < z < 3$. They found those low-mass AGN-host galaxies are significantly above any of the local relations, which seems to indicate the presence of the M_{\bullet} – M_* relation evolution at the low-mass regime. Also, when taking into account the future evolutionary pathway of those AGNs in low-mass galaxies, the authors speculated that about 30% of their AGNs could merge into the normal relation with an evolved and increased stellar mass at $z \sim 0$. However, we do not include those low-mass

samples in our mass scaling relation evolution analysis nor draw any conclusion at the low-mass regime in this paper, because (1) the local $M_{\bullet}-M_{\star}$ baseline relation at the lower-mass regime ($\log(M_{\star}/M_{\odot}) < 10$) is still highly uncertain due to the limited number of BL AGNs in low-mass galaxies observed with uniform BH mass measurements (A. E. Reines & M. Volonteri 2015; J. E. Greene et al. 2020; but see R. Pucha et al. 2024); (2) the sample size of such a kind of AGN in higher redshift is also small and sparsely distributed across redshift (see Figure 6); and (3) most importantly, those AGNs in low-mass galaxies suffer from stronger observational biases compared to the massive ones (see Appendix B), which were not evaluated in M. Mezcuca et al. (2023, 2024). To better constrain the $M_{\bullet}-M_{\star}$ relation at the low-mass end and examine its evolution with redshift, future observations of a larger sample of BL AGNs in low-mass galaxies with consistent measurements of both BH mass and stellar mass, and a comprehensive understanding of their observational biases, are essential.

6. Conclusion

In this work, we built a new sample of 18 BL AGNs with broad rest-frame NIR spectral lines (Paschen lines and He I $\lambda 10833 \text{ \AA}$) at $1 < z < 4$ in the GOODS-S and GOODS-N fields using the FRESCO and CONGRESS JWST/NIRCam grism spectra. We measure the BH masses of our BL AGN sample ($7.1 < \log(M_{\bullet}/M_{\odot}) < 8.8$) using the single-epoch virial BH mass method based on the three NIR lines. We derived their stellar masses ($9.9 < \log(M_{\star}/M_{\odot}) < 11.3$) by either applying AGN-galaxy imaging decomposition on the JADES JWST/NIRCam broadband images and fitting the host-only SED, or doing the SED decomposition if targets are outside of the JADES footprints.

We calculated the average M_{\bullet}/M_{\star} ratio for our new $1 < z < 4$ BL AGNs ($\log(M_{\bullet}/M_{\star}) = -2.61^{+0.90}_{-0.59}$) and found it is consistent with the values at $z \sim 1-2.5$ from previous studies (Figures 6 and 10). In other words, using the $z \sim 1$ relation ($\log(M_{\bullet}/M_{\star}) = -2.5$) as the baseline, we do not observe an evolution of M_{\bullet}/M_{\star} ratio up to $z \sim 4$ (Figure 6). We also tested the impact of observational biases on the observed $M_{\bullet}-M_{\star}$ distribution by running an MC simulation to model both the mock $1 < z < 4$ AGN population in an infinitely large survey and the mock observations in an FRESCO-like (area-limited) survey, following the $z \sim 1$ relation. The observed $M_{\bullet}-M_{\star}$ distribution is in good agreement with the intrinsic distribution derived with the infinite sample (Figure 7). Also, by applying the K-S test to compare the $\log(M_{\bullet}/M_{\star})$ distribution between the mock observation and the true observation, we confirmed there is no significant difference between them (Figure 9), which indicates that there is no evolution of the M_{\bullet}/M_{\star} ratio at $\log(M_{\star}/M_{\odot}) \gtrsim 10$ up to $z \sim 4$, even accounting for observational biases.

This work sets up an appropriate baseline for comparison with the other high-redshift studies and an improved MC tool to simulate mock observations with biases and errors for a fair

comparison with true observations. In an accompanying paper, we will apply this developed tool to test the redshift evolution of the $M_{\bullet}-M_{\star}$ relation up to $z \sim 7$ by using a similar analysis as developed in this paper on higher-redshift ($4 < z < 7$) samples.

Acknowledgments

We thank Roberto Maiolino for helpful discussions and comments. This work was supported by the JWST Mid-Infrared Instrument (MIRI) Science Team Lead grant, 80NSSC18K0555, from NASA Goddard Space Flight Center to the University of Arizona. Y.S., Z.J., F.S., Y.Z., P.A.C., E. E., K.H., J.M.H., P.R., B.E.R., M.A.S., G.H.R., and C.N.A.W. also acknowledge support from the NIRCam Science Team contract to the University of Arizona, NAS 5-02015. I.S. acknowledges funding support from the Atracción de Talento program, grant No. 2022-T1/TIC-20472, of the Comunidad de Madrid, Spain. S.T. acknowledges support by the Royal Society Research grant G125142. The work of C.C.W. is supported by NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. A.J.B. has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Advanced grant 789056 “First Galaxies.” F.D.E. and J.S. acknowledge support by the Science and Technology Facilities Council (STFC) and by the ERC through Advanced grant 695671 “QUENCH,” and the UKRI Frontier Research grant RISEandFALL.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The data were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. These observations are associated with programs 1180, 1181, 1895, and 3577, for which the data can be accessed via [10.17909/8tdj-8n28](https://doi.org/10.17909/8tdj-8n28) (M. Rieke et al. 2023), and [10.17909/gdyc-7g80](https://doi.org/10.17909/gdyc-7g80) (P. Oesch & D. Magee 2023). The authors sincerely thank the FRESCO team (PI: Pascal Oesch) for developing and executing their observing program.

Facility: JWST (NIRCam).

Software: AstroPy (Astropy Collaboration et al. 2013, 2018, 2022), galight (X. Ding et al. 2020), lmfit (M. Newville et al. 2014), Prospector (B. D. Johnson et al. 2021), SciPy (P. Virtanen et al. 2020).

Appendix A Broad-line Fitting of the BL AGNs

We use the Python package lmfit (M. Newville et al. 2014) to model the broad-line ($\text{Pa}\alpha$, $\text{Pa}\beta$, and He I $\lambda 10833 \text{ \AA}$ lines) profile of our BL AGNs. In addition to the six examples shown in Figure 3, the rest (12) of the 18 BL AGNs are shown in Figure A1.

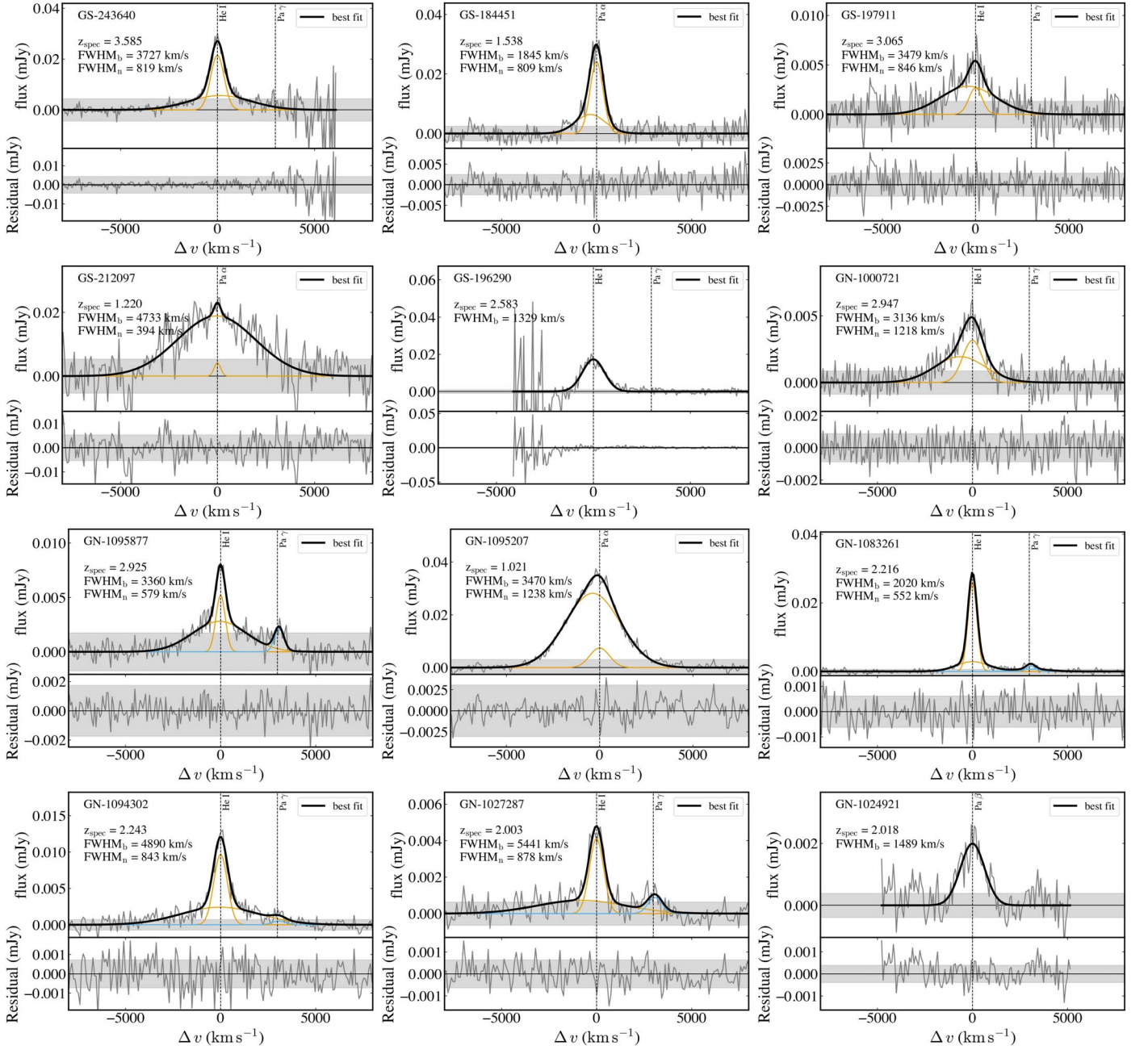


Figure A1. Line profile fitting on the broad Pa α , Pa β , and He I λ 10833 Å lines, for the remaining (12/18) of our BL AGNs identified in this work, the same as Figure 3.

Appendix B Details of Building a Mock AGN Population

For the infinitely large survey scenario (Section 4.3.1), to build a mock AGN population, we first randomly generate 10^7 mock galaxies with true stellar mass $M_{*,\text{true}}$ and redshift z_{true} from the COSMOS2020 SMF at $z \sim 2$ (J. R. Weaver et al. 2023). The mock stellar mass is sampled from $10^{9.9}$ – $10^{12} M_{\odot}$ with the grid size of 0.005 dex in logarithmic scale, and the redshift z_{true} is sampled in steps of $\Delta z = 0.01$ within $1 < z < 4$ (corresponding to the redshift range of our AGN sample). We limit the lower boundary of simulated stellar mass to $10^{9.9} M_{\odot}$ to match with the lowest stellar mass of our AGN sample, because there should be an observational limit determining the lowest stellar mass of galaxies that we can detect and can be

well-measured through imaging-decomposition using JWST/NIRCam data. However, since it is hard to quantify and thus is not included in the observational limits that we carefully model in this work, we simply limit the simulated stellar mass above a threshold to model this observational limit.

In the next step, we assign a true BH mass $M_{\bullet,\text{true}}$ for each $M_{*,\text{true}}$ by assuming the $z \sim 1$ relation with the intrinsic scatter σ_{int} :

$$P(\log M_{\bullet} | \log M_{*}) = \frac{1}{\sqrt{2\pi} \sigma_{\text{int}}} \times \exp\left(-\frac{(\log M_{\bullet} - (\log M_{*} + a))^2}{2\sigma_{\text{int}}^2}\right), \quad (\text{B1})$$

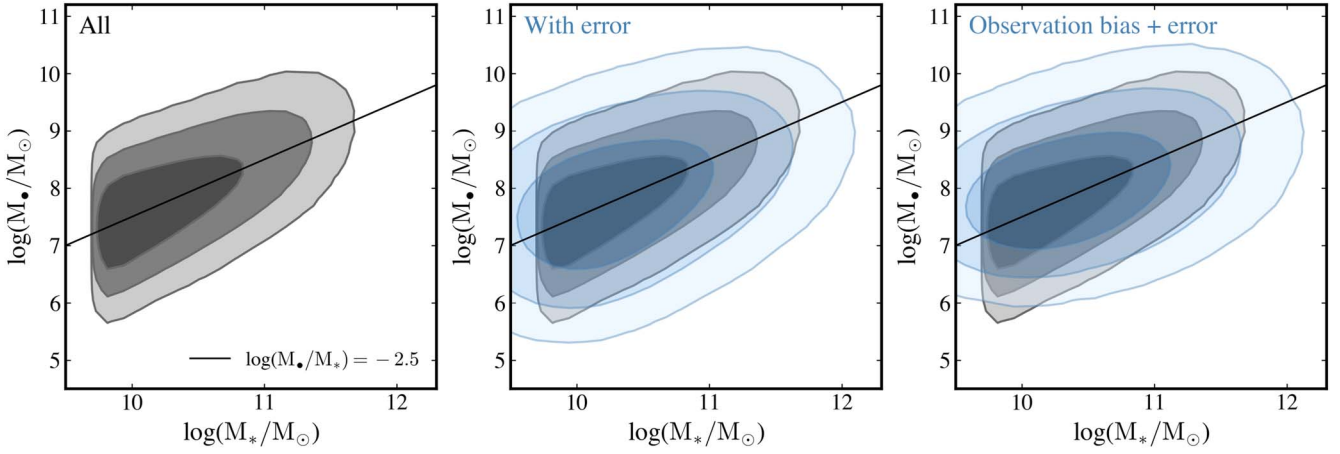


Figure B1. The distributions of mock $1 < z < 4$ AGNs that follow the $z \sim 1$ relation on the $M_{\text{bol}}-M_*$ diagram. The gray contour in all panels represents the $M_{\text{bol,true}}-M_{*,\text{true}}$ distribution of the whole mock AGN population, without applying mass measurement errors. In the middle column, the blue contours show the $M_{\text{bol,obs}}-M_{*,\text{obs}}$ distributions of the whole mock AGN populations, namely, the observed distributions with the errors of M_{bol} (0.4 dex) and M_* (0.3 dex). In the right panel, the blue contours show the $M_{\text{bol,obs}}-M_{*,\text{obs}}$ distributions of the “observable” mock AGN populations, i.e., the errors of M_{bol} and M_* , and the observational limits are added. All of the contours indicate the 1σ to 3σ levels.

where a is the intrinsic M_{bol}/M_* ratio ($\log(M_{\text{bol}}/M_*)$). Here we use the $z \sim 1$ relation for AGNs, which have $a = -2.5$ or -3.0 and $\sigma_{\text{int}} = 0.5$ dex. Then, we simply assume all of the generated mock galaxies have an AGN, given that we are modeling an infinitely large survey. In most cases, the AGNs are at very low luminosity and are undetectable, as reflected in the ERDF. The true distribution of our mock AGN sample in the M_{bol} versus M_* plane is shown in the left panel of Figure B1. For each mock AGN, we then calculate its bolometric luminosity $L_{\text{bol,mock}}$ by assigning an Eddington ratio ($\lambda_{\text{Edd}} = L_{\text{Edd}}/L_{\text{bol}}$) that is randomly sampled from the $z \sim 2.15$ ERDF by B. C. Kelly & Y. Shen (2013) and also the X-ray luminosity at 2–10 keV ($L_{2-10 \text{ keV}}$) using the X-ray bolometric correction reported by F. Duras et al. (2020).

We further consider the uncertainty of the observed mass measurements ($M_{\text{bol,obs}}$ and $M_{*,\text{obs}}$) related to the true masses ($M_{\text{bol,true}}$ and $M_{*,\text{true}}$). In our case, the uncertainty of the virial BH mass method built by F. Ricci et al. (2017) is about 0.4 dex, and the SED-based stellar mass measurement error is about 0.3 dex. By adding random Gaussian errors to both $M_{\text{bol,true}}$ and $M_{*,\text{true}}$, we obtain $M_{\text{bol,obs}}$ and $M_{*,\text{obs}}$. The $M_{\text{bol,obs}}-M_{*,\text{obs}}$ distribution of the mock AGN sample is shown in the middle panel of Figure B1, which is significantly broader than the initial distribution.

Finally, we apply the observational limits to select the “observable” mock AGNs. The first observational bias comes from the BL AGN selection. In this work, only the AGNs that have a measurable broad line ($\text{FWHM} > 1000 \text{ km s}^{-1}$) can be identified as BL AGNs and be used for single-epoch M_{bol} estimation. Therefore, we use $M_{\text{bol,mock}}$ and $L_{2-10 \text{ keV}}$ to estimate the FWHM of the targeted NIR line based on the F. Ricci et al. (2017) single-epoch BH mass relation for each mock AGN. The mock AGNs with a broad-line width $< 1000 \text{ km s}^{-1}$ will be left out of the “observable” sample. Another limit arises because of the spectral sensitivity of FRESKO. It is about $5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ for detecting a broad component with a width of 1000 km s^{-1} . For detecting a broader line, the flux line sensitivity decreases as $(\text{FWHM}/200)^{0.5} \times 2 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$, where 200 km s^{-1}

is the instrumental broadening with spectral resolution $R \sim 1500$. We convert the line flux sensitivity to the line luminosity limit based on z_{true} . Given that most of our BL AGN samples are identified using the FRESKO F444W grism spectra with a wavelength coverage of 3.1–4.0 μm , we assume the broad lines that we could detect in the FRESKO spectra for mock AGNs at $1 < z < 2$, $2 < z < 2.5$, and $2.5 < z < 4$ are $\text{Pa}\alpha$, $\text{Pa}\beta$, and $\text{He I } \lambda 10833 \text{ \AA}$, respectively. By fitting the luminosity relation between $\text{H}\alpha$ and those three NIR lines using the H. Landt et al. (2008) local BL AGN sample, we can convert the targeted line luminosity limit to the $\text{H}\alpha$ line luminosity limit. We further convert the $\text{H}\alpha$ line luminosity to the spectroscopic AGN bolometric luminosity limit $L_{\text{lim,spec}}$ using the relation from J. E. Greene & L. C. Ho (2005). We classify the mock AGNs with $L_{\text{bol,obs}} < L_{\text{lim,spec}}$ as undetectable. Considering the two observational limits mentioned above, we made an “observable” mock AGN subsample from the whole AGN sample. The final “observable” $M_{\text{bol,obs}}-M_{*,\text{obs}}$ distribution is shown in the right panel of Figure B1. Comparing the “observable” $M_{\text{bol,obs}}-M_{*,\text{obs}}$ distribution (blue) with the whole AGN $M_{\text{bol,true}}-M_{*,\text{true}}$ distribution (black), it is straightforward to illustrate the significant effects of observational limits on biasing the real mass scaling relation to the observed one, which is known as the “Lauer bias” (T. R. Lauer et al. 2007).

For simulating the actual observation in an FRESKO-like survey (Section 4.3.2), we apply a similar routine to that for the infinitely large survey scenario. The only two differences are: (1) we now simulate the exact galaxy population within an FRESKO-like survey area instead of an arbitrarily large population. We first estimate the galaxy population size within an FRESKO-like survey field (122 arcmin^2) using the COSMOS2020 SMF, which results in 4170 galaxies with stellar mass ranging from 10^9 – $10^{12} M_{\odot}$ and redshift in the range of $1 < z < 4$. (2) To estimate the BL AGN population underlying the whole galaxy population, we assume the BL AGN fraction at $1 < z < 4$ as a constant $3\% \pm 1.2\%$ (A. Schulze et al. 2015) of the whole galaxy sample, given that A. Schulze et al. (2015) found the BL AGN fraction at $z \sim 2$ is almost constant when $\log(M_{\text{bol}}/M_{\odot}) < 9$.

ORCID iDs

Yang Sun  <https://orcid.org/0000-0001-6561-9443>
 Jianwei Lyu (吕建伟)  <https://orcid.org/0000-0002-6221-1829>
 George H. Rieke  <https://orcid.org/0000-0003-2303-6519>
 Zhiyuan Ji  <https://orcid.org/0000-0001-7673-2257>
 Fengwu Sun  <https://orcid.org/0000-0002-4622-6617>
 Yongda Zhu  <https://orcid.org/0000-0003-3307-7525>
 Andrew J. Bunker  <https://orcid.org/0000-0002-8651-9879>
 Phillip A. Cargile  <https://orcid.org/0000-0002-1617-8917>
 Chiara Circosta  <https://orcid.org/0000-0001-8522-9434>
 Francesco D'Eugenio  <https://orcid.org/0000-0003-2388-8172>
 Eiichi Egami  <https://orcid.org/0000-0003-1344-9475>
 Kevin Hainline  <https://orcid.org/0000-0003-4565-8239>
 Jakob M. Helton  <https://orcid.org/0000-0003-4337-6211>
 Pierluigi Rinaldi  <https://orcid.org/0000-0002-5104-8245>
 Brant E. Robertson  <https://orcid.org/0000-0002-4271-0364>
 Irene Shivaiei  <https://orcid.org/0000-0003-4702-7561>
 Meredith A. Stone  <https://orcid.org/0000-0002-9720-3255>
 Sandro Tacchella  <https://orcid.org/0000-0002-8224-4505>
 Christina C. Williams  <https://orcid.org/0000-0003-2919-7495>
 Christopher N. A. Willmer  <https://orcid.org/0000-0001-9262-9997>
 Chris Willott  <https://orcid.org/0000-0002-4201-7367>

References

- Alberts, S., Lyu, J., Shivaiei, I., et al. 2024, *ApJ*, 976, 224
 Alexander, D. M., & Hickox, R. C. 2012, *NewAR*, 56, 93
 Astropy Collaboration, Price-Whelan, A. M., Lim, P. M., et al. 2022, *ApJ*, 935, 167
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
 Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, *ApJS*, 243, 22
 Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2024, JWST Calibration Pipeline, v1.11.2, Zenodo, doi:10.5281/ZENODO.6984365
 D'Eugenio, F., Perez-Gonzalez, P., Maiolino, R., et al. 2024, *NatAs*, 8, 1443
 Ding, X., Onoue, M., Silverman, J. D., et al. 2023, *Natur*, 621, 51
 Ding, X., Silverman, J., Treu, T., et al. 2020, *ApJ*, 888, 37
 Duras, F., Bongiorno, A., Ricci, F., et al. 2020, *A&A*, 636, A73
 Eisenstein, D. J., Willott, C., Alberts, S., et al. 2023, arXiv:2306.02465
 Fabian, A. C. 2012, *ARA&A*, 50, 455
 Ferrarese, L., & Merritt, D. 2000, *ApJL*, 539, L9
 Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023, *A&A*, 674, A1
 Gebhardt, K., Bender, R., Bower, G., et al. 2000, *ApJL*, 539, L13
 Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJL*, 600, L93
 Greene, J. E., & Ho, L. C. 2005, *ApJ*, 630, 122
 Greene, J. E., Strader, J., & Ho, L. C. 2020, *ARA&A*, 58, 257
 Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, *ApJ*, 959, 39
 Häring, N., & Rix, H.-W. 2004, *ApJL*, 604, L89
 Harrison, C. M. 2017, *NatAs*, 1, 0165
 Hazard, C., Mackey, M. B., & Shimmings, A. J. 1963, *Natur*, 197, 1037
 Heckman, T. M., & Best, P. N. 2014, *ARA&A*, 52, 589
 Hopkins, P. F., Cox, T. J., Kereš, D., & Hernquist, L. 2008a, *ApJS*, 175, 390
 Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008b, *ApJS*, 175, 356
 Horne, K. 1986, *PASP*, 98, 609
 Jahnke, K., & Macciò, A. V. 2011, *ApJ*, 734, 92
 Ji, Z., Williams, C. C., Tacchella, S., et al. 2023, *ApJ*, 974, 135
 Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, *ApJS*, 254, 22
 Kelly, B. C., & Shen, Y. 2013, *ApJ*, 764, 45
 Kennedy, J., & Eberhart, R. 1995, in IEEE International Conference on Neural Networks, 4 (Pisataway, NJ: IEEE), 1942
 Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36
 Kormendy, J., & Ho, L. C. 2013, *ARA&A*, 51, 511
 Kriek, M., & Conroy, C. 2013, *ApJL*, 775, L16
 Krywult, J., Tasca, L. A. M., Pollo, A., et al. 2017, *A&A*, 598, A120
 Kuhn, L., Shangguan, J., Davies, R., et al. 2024, *A&A*, 684, A52
 Landt, H., Bentz, M. C., Ward, M. J., et al. 2008, *ApJS*, 174, 282
 Lauer, T. R., Tremaine, S., Richstone, D., & Faber, S. M. 2007, *ApJ*, 670, 249
 Le, H. A. N., Woo, J.-H., & Xue, Y. 2020, *ApJ*, 901, 35
 Li, J., Silverman, J. D., Ding, X., et al. 2021, *ApJ*, 922, 142
 Li, J., Silverman, J. D., Shen, Y., et al. 2024, arXiv:2403.00074
 Li, J. I. H., Shen, Y., Ho, L. C., et al. 2023, *ApJ*, 954, 173
 Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, *ApJS*, 228, 2
 Lyu, J., Alberts, S., Rieke, G. H., & Rujopakarn, W. 2022, *ApJ*, 941, 191
 Lyu, J., Alberts, S., Rieke, G. H., et al. 2024, *ApJ*, 966, 229
 Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285
 Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2023, *A&A*, 691, A145
 Matthee, J., Naidu, R. P., Brammer, G., et al. 2023, *ApJ*, 963, 129
 Mechtley, M., Jahnke, K., Windhorst, R. A., et al. 2016, *ApJ*, 830, 156
 Merloni, A., Bongiorno, A., Bolzonella, M., et al. 2010, *ApJ*, 708, 137
 Mezcuca, M., Pacucci, F., Suh, H., Siudek, M., & Natarajan, P. 2024, *ApJL*, 966, L30
 Mezcuca, M., Siudek, M., Suh, H., et al. 2023, *ApJL*, 943, L5
 Mountrichas, G. 2023, *A&A*, 672, A98
 Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. 2014, LMFIT: Non-linear Least-square Minimization and Curve-Fitting for Python, v0.8.0, Zenodo, doi:10.5281/zenodo.11813
 Oesch, P., & Magee, D. 2023, The JWST FRESCO Survey, STScI/MAST, doi:10.17909/GDYC-7G80
 Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, *MNRAS*, 525, 2864
 Pacucci, F., Nguyen, B., Carniani, S., Maiolino, R., & Fan, X. 2023, *ApJL*, 957, L3
 Peng, C. Y. 2007, *ApJ*, 671, 1098
 Peng, C. Y., Impey, C. D., Rix, H.-W., et al. 2006, *ApJ*, 649, 616
 Peterson, B. M. 2006, in Physics of Active Galactic Nuclei at all Scales, ed. D. Alloin, Vol. 693 (Berlin: Springer), 77
 Popović, L. Č., Kovačević-Dojčinović, J., & Marčeta-Mandić, S. 2019, *MNRAS*, 484, 3180
 Pucha, R., Juneau, S., Dey, A., et al. 2024, arXiv:2411.00091
 Reines, A. E., & Volonteri, M. 2015, *ApJ*, 813, 82
 Ricci, F., La Franca, F., Onori, F., & Bianchi, S. 2017, *A&A*, 598, A51
 Ricci, F., Treister, E., Bauer, F. E., et al. 2022, *ApJS*, 261, 8
 Rieke, G., Alberts, S., Shivaiei, I., et al. 2024, arXiv:2406.03518
 Rieke, M., Robertson, B., Tacchella, S., et al. 2023, Data from the JWST Advanced Deep Extragalactic Survey (JADES), STScI/MAST, doi:10.17909/8TDJ-8N28
 Rieke, M. J., Kelly, D. M., Misselt, K., et al. 2023, *PASP*, 135, 028001
 Schmidt, M. 1963, *Natur*, 197, 1040
 Schramm, M., & Silverman, J. D. 2013, *ApJ*, 767, 13
 Schulze, A., Bongiorno, A., Gavignaud, I., et al. 2015, *MNRAS*, 447, 2085
 Schulze, A., & Wisotzki, L. 2011, *A&A*, 535, A87
 Schulze, A., & Wisotzki, L. 2014, *MNRAS*, 438, 3422
 Setoguchi, K., Ueda, Y., Toba, Y., & Akiyama, M. 2021, *ApJ*, 909, 188
 Shen, Y., Brandt, W. N., Richards, G. T., et al. 2016, *ApJ*, 831, 7
 Shen, Y., & Liu, X. 2012, *ApJ*, 753, 125
 Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, 435, 629
 Stone, M. A., Lyu, J., Rieke, G. H., Alberts, S., & Hainline, K. N. 2024, *ApJ*, 964, 90
 Suh, H., Civano, F., Trakhtenbrot, B., et al. 2020, *ApJ*, 889, 32
 Sun, F. 2024, nircam_grism v3.0, Zenodo, doi:10.5281/zenodo.14052875
 Sun, F., Egami, E., Pirzkal, N., et al. 2023, *ApJ*, 953, 53
 Sun, M., Trump, J. R., Brandt, W. N., et al. 2015, *ApJ*, 802, 14
 Tanaka, T. S., Silverman, J. D., Ding, X., et al. 2024, arXiv:2401.13742
 Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, *ApJ*, 574, 740
 Übler, H., Maiolino, R., Curtis-Lake, E., et al. 2023, *A&A*, 677, A145
 Vestergaard, M. 2002, *ApJ*, 571, 733
 Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, 17, 261
 Weaver, J. R., Davidzon, I., Toft, S., et al. 2023, *A&A*, 677, A184
 Willott, C. J., Percival, W. J., McLure, R. J., et al. 2005, *ApJ*, 626, 657
 Yue, M., Eilers, A.-C., Simcoe, R. A., et al. 2024, *ApJ*, 966, 176
 Zastrocky, T. E., Brotherton, M. S., Du, P., et al. 2024, *ApJS*, 272, 29
 Zhang, Y., Ouchi, M., Gebhardt, K., et al. 2023, *ApJ*, 948, 103
 Zuo, W., Wu, X.-B., Fan, X., et al. 2020, *ApJ*, 896, 40