

TRINITY IV: Predictions for Supermassive Black Holes at $z \gtrsim 7$

Haowen Zhang (张昊文),^{1*} Peter Behroozi,^{1,2} Marta Volonteri,³ Joseph Silk,^{3,4,5}
Xiaohui Fan,¹ James Aird,^{6,7} Jinyi Yang (杨锦怡),^{1,*} Feige Wang (王飞格),¹
and Philip F. Hopkins⁸

¹University of Arizona, 933 N Cherry Ave., Tucson, AZ 85721, USA

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

³Institut d'Astrophysique de Paris (UMR 7095: CNRS & Sorbonne Universite), 98 bis Bd. Arago, F-75014, Paris, France

⁴Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

⁵BIPAC, Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

⁶Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

⁷Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK

⁸Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

*Strittmatter Fellow

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present predictions for the high-redshift halo–galaxy–supermassive black hole (SMBH) connection from the TRINITY model. Constrained by a comprehensive compilation of galaxy ($0 \leq z \leq 10$) and SMBH datasets ($0 \leq z \leq 6.5$), TRINITY finds: 1) The number of SMBHs with $M_{\bullet} > 10^9 M_{\odot}$ in the observable Universe increases by six orders of magnitude from $z \sim 10$ to $z \sim 2$, and by another factor of ~ 3 from $z \sim 2$ to $z = 0$; 2) The $M_{\bullet} > 10^9/10^{10} M_{\odot}$ SMBHs at $z \sim 6$ live in haloes with $\sim (2-3)/(3-5) \times 10^{12} M_{\odot}$; 3) the new JWST AGNs at $7 \lesssim z \lesssim 11$ are broadly consistent with the median SMBH mass–galaxy mass relation for AGNs from TRINITY; 4) Seeds from runaway mergers in nuclear star clusters are viable progenitors for the SMBHs in GN-z11 ($z = 10.6$) and CEERS_1019 ($z = 8.7$); 5) $z = 6-10$ quasar luminosity functions from wide area surveys by, e.g., *Roman* and *Euclid*, will reduce uncertainties in the $z = 6-10$ SMBH mass–galaxy mass relation by up to ~ 0.5 dex.

Key words: galaxies: haloes – galaxies: evolution – quasars: supermassive black holes

1 INTRODUCTION

Supermassive black holes (SMBHs) are believed to exist in the centres of most, if not all galaxies (Kormendy & Richstone 1995; Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002; Ho 2008; Gültekin et al. 2009; Kormendy & Ho 2013; Heckman & Best 2014). When SMBHs are actively accreting matter, they are called active galactic nuclei (AGN). AGNs are also called quasars when they outshine their host galaxies. With their accretion energy output, it is likely that SMBHs are capable of modulating both their own mass accretion and the star formation of their host galaxies. (Silk & Rees 1998; Bower et al. 2006; Somerville et al. 2008; Sijacki et al. 2015). On the other hand, galaxies can also regulate central black hole growth via the fueling of gas and galaxy mergers. Consequently, both SMBHs and their host galaxies may influence each others' growth, also known as “coevolution.” Hence, to understand galaxy and SMBH assembly histories, it is critical to constrain the interaction between SMBHs and their host galaxies (see, e.g., Hopkins

et al. 2007; Ho 2008; Alexander & Hickox 2012; Kormendy & Ho 2013; Heckman & Best 2014; Brandt & Alexander 2015).

Two lines of observations support the coevolution scenario. First, there are relatively tight scaling relations (~ 0.3 dex scatter) between SMBH masses and host galaxy dynamical properties (e.g., velocity dispersion or bulge mass at $z \sim 0$ (see Häring & Rix 2004; Gültekin et al. 2009; Kormendy & Ho 2013; McConnell & Ma 2013; Savorgnan et al. 2016). Second, the cosmic SMBH accretion rate (CBHAR) density is roughly proportional to cosmic star formation rate (CSFR) density over $0 < z < 4$, and the constant CBHAR/CSFR ratio is $\sim 10^{-4} - 10^{-3}$ (Merloni 2004; Silverman et al. 2008; Shankar et al. 2009; Aird et al. 2010; Delvecchio et al. 2014; Yang et al. 2018). However, it has been difficult, if not impossible, to verify other predictions of the coevolution model (e.g., tight SMBH–galaxy property relationships) at higher redshifts.

At $z = 0$, high spatial resolution spectroscopy and dynamics modeling have been used to measure SMBH–galaxy scaling relations (e.g., Magorrian et al. 1998; Ferrarese & Ford 2005; McConnell & Ma 2013). Beyond $z = 0$, it is difficult to measure individual SMBH masses in the same way, due to lower spatial resolutions. Hence, we have to rely on indirect methods such as re-

* E-mail: hwzhang0595@arizona.edu

verberation mapping (Blandford & McKee 1982; Peterson 1993) or “virial estimates” with single-epoch AGN spectra, which infer SMBH masses from spectral line width and AGN luminosity (Vestergaard & Peterson 2006). All such indirect methods require that SMBHs are actively accreting, and thus: 1) are biased against dormant SMBHs, and 2) make it difficult to measure host galaxy masses at the same time. In addition, overmassive SMBHs are brighter and more likely to be included in the sample, when different SMBHs have similar Eddington ratios. This bias affects the measurement of the galaxy mass–SMBH mass ($M_{\bullet}-M_{*}$) relation with bright quasars (Lauer et al. 2007). Zhang et al. (2023a) quantified this bias as a function of AGN luminosity, and showed that AGNs with bolometric luminosities $L_{\text{bol}} \lesssim 10^{45}$ erg/s are needed to measure the unbiased $M_{\bullet}-M_{*}$ relation at $z = 6$.

Prior to JWST, the highest-redshift AGN ever found was at $z \sim 7.6$ (Wang et al. 2021). AGN luminosity functions at up to $z \sim 7$ have also been measured (Wang et al. 2019; Kulkarni et al. 2019; Shen et al. 2020). Most $z \sim 7$ AGNs were found to be bright quasars with high SMBH masses, accreting at around the Eddington rate (Shen et al. 2019). However, given the capabilities of existing telescopes, it was difficult to ascertain if such mass and Eddington ratio distributions were representative of typical high redshift SMBHs.

JWST has changed the landscape of high-redshift quasar observations. As of the writing of this paper, JWST has found at least two AGN candidates above $z \sim 8$, i.e., CEERS1019 at $z = 8.7$ (Larson et al. 2023) and GN-z11 at $z = 10.6$ (Maiolino et al. 2023a). CEERS1019 has an estimated SMBH mass $M_{\bullet} \sim 10^7 M_{\odot}$, and an Eddington ratio (η) of ~ 1.2 , whereas GN-z11 is estimated to be a $M_{\bullet} \sim 10^6 M_{\odot}$ SMBH accreting at an Eddington ratio of $\eta \sim 5$. JWST’s sensitivity finally allows measuring SMBH masses in AGNs with $L_{\text{bol}} \lesssim 10^{45}$ erg/s at $z \gtrsim 6$ (e.g., CEERS1019 and GN-z11), as well as measuring host galaxy properties for high- z X-ray AGN candidates, like UHZ-1 at $z = 10.1$ (Bogdan et al. 2023; Goulding et al. 2023). With its unique near-infrared (NIR) sensitivity, JWST will continue to provide strong tests/constraints on theoretical models of the halo–galaxy–SMBH connection.

Different theoretical studies have been carried out to ascertain the possible origins and evolution histories of these highest-redshift AGNs. Using a semi-analytical model, the Cosmic Archaeology Tool (CAT, Trinca et al. 2022), Schneider et al. (2023) found a potential evolutionary connection between objects like GN-z11 and CEERS_1019. Scholtz et al. (2023) estimated GN-z11’s host halo mass to be $\sim 3 \times 10^{10} M_{\odot}$, based on the stellar mass–halo mass relation from the empirical UNIVERSE MACHINE model of the halo–galaxy connection (Behroozi et al. 2019). This halo mass scale makes GN-z11 a candidate protocluster core, whose $z = 0$ descendant could reach above $10^{15} M_{\odot}$.

In this paper, we present the SMBH properties at $z \gtrsim 6$ from the empirical TRINITY model. We further put CEERS1019 and GN-z11 in the context of TRINITY, and examine their consistency with our predictions. TRINITY makes statistical connections between dark matter haloes, galaxies, and SMBHs from $z = 0 - 10$, extending the empirical DM halo–galaxy model from Behroozi et al. (2013). Specifically, TRINITY infers SMBH growth histories as a function of halo/galaxy mass in a similar manner as applying the Softan Argument (Softan 1982) to each halo/galaxy population (see also Shankar et al. 2020). Compared to previous empirical models, TRINITY includes joint constraints from a large and diverse compilation of data. Specifically, the following data are used to constrain the halo–galaxy connection: galaxy stellar mass functions (SMFs), galaxy UV luminosity functions (UVLFs), galaxy quenched fractions (QFs), galaxy specific star formation

rates (SSFRs), and cosmic star formation rates (CSFRs). Additionally, the galaxy–SMBH connection is constrained by quasar luminosity functions (QLFs), quasar probability distribution functions (QPDFs), active black hole mass functions (ABHMFs), and the $z = 0$ bulge mass–black hole mass relations. To accommodate the enormous joint constraining power of this dataset, TRINITY has a more flexible parametrization than past approaches (e.g., Marconi et al. 2004; Shankar et al. 2009, 2013; Conroy & White 2013). Last but not least, TRINITY features more realistic modeling of AGN observables by including, e.g., black hole mergers and kinetic AGN luminosities in the model.

Following Behroozi et al. (2013), TRINITY modeling starts with population statistics from a dark matter N-body simulation. The model then generates mock universes, each with different galaxy and SMBH mass–growth rate distributions, as functions of halo mass and redshift. Millions of mock universes are and compared with real observations using a Markov Chain Monte Carlo (MCMC) algorithm to converge on the range of halo–galaxy–SMBH connections consistent with the real Universe.

The paper is organized as follows. In §2, we briefly introduce the TRINITY framework. §3 covers the N-body simulations and galaxy/SMBH connection. For full details about the TRINITY methodology, simulations, and datasets, we refer readers to Zhang et al. (2023c). In §4, we present high redshift predictions from TRINITY. §5 compares CEERS1019 and GN-z11 with TRINITY’s predictions, and discusses how future observations would help constrain the demographics and growth histories of high redshift SMBHs. Finally, our conclusions are given in §6. In this work, we adopt a flat Λ CDM cosmology with parameters ($\Omega_m = 0.307$, $\Omega_{\Lambda} = 0.693$, $h = 0.678$, $\sigma_8 = 0.823$, $n_s = 0.96$) consistent with Planck results (Planck Collaboration et al. 2016). We use datasets that adopt the Chabrier stellar initial mass function (IMF, Chabrier 2003), the Bruzual & Charlot (2003) stellar population synthesis model, and the Calzetti dust attenuation law (Calzetti et al. 2000). Halo masses are calculated following the virial overdensity definition from Bryan & Norman (1998).

2 METHODOLOGY

In §2.1, we explain the key aspects of the TRINITY high-redshift halo–galaxy–SMBH connection that influence the predictions in this work, and how they are constrained by our data compilation. We then give a brief overview on TRINITY’s implementation in §2.2. For full details, we refer readers to §2 of Zhang et al. (2023c).

2.1 Inferring the high-redshift halo–galaxy–SMBH connection with TRINITY

As an empirical model, TRINITY infers the range of halo–galaxy–SMBH connections that match our unique compilation of galaxy+AGN data. Many predictions in this work come from the following two aspects of this connection:

Firstly, the intrinsic $M_{\bullet}-M_{*}$ relation evolves only mildly with redshift (see §4.1). The largest change occurs for low-mass galaxies (i.e., $M_{*} \sim 5 \times 10^9 M_{\odot}$) which have lower SMBH masses at higher redshifts; almost no change in SMBH masses with redshift occurs for larger galaxies with masses $M_{*} \gtrsim 5 \times 10^{10} M_{\odot}$. This redshift evolution is mainly constrained by the combination of QLFs, QPDFs, and the local $M_{\bullet}-M_{\text{bulge}}$ relation. Specifically,

different redshift evolutions of the M_\bullet – M_* relation produce different average SMBH growth rates as functions of host galaxy mass and redshift. Therefore, there is only one unique combination of a redshift-dependent M_\bullet – M_* relation and an AGN energy efficiency that can simultaneously reproduce: 1) the probability distribution of AGN luminosity as a function of galaxy mass and redshift, i.e., QPDFs; 2) QLFs; and 3) the local M_\bullet – M_{bulge} relation. For example, a steeper M_\bullet – M_* relation will not only violate the constraint from the local M_\bullet – M_{bulge} relation, but also induce faster growth of SMBHs at fixed galaxy mass, and thus overproduce AGN activity compared to observed QPDFs.

Secondly, the AGN Eddington ratio distribution becomes narrower towards higher redshifts, along with increasing average Eddington ratios. This is required by the ABHMFs from Kelly & Shen (2013). With the redshift-dependent M_\bullet – M_* relation and AGN duty cycles constrained by QPDFs, QLFs, and the local M_\bullet – M_{bulge} relation, lower Eddington ratios and/or broader Eddington ratio distributions will underpredict the number of SMBHs above certain luminosity limits, i.e., ABHMFs, at higher redshifts.

Of note, the AGN data for constraining TRINITY spans a redshift range of $z = 0 - 6.5$. When making predictions for SMBHs beyond $z \sim 7$, we are effectively examining the observational consequences of a relatively non-evolving SMBH–galaxy relationship at those redshifts. Hence, a failure of TRINITY to predict $z > 6$ observations would indicate different evolution in the SMBH–galaxy mass relationship than that seen at $z < 6$. As we will show in §4.1 and §5.2, the SMBH properties of CEERS1019 and GN-z11 are partly consistent with these extrapolations. The partial inconsistency between TRINITY and these two AGNs does not impose strong constraints on the current TRINITY model—i.e., the inconsistency is not yet very statistically significant. As a result, further observations of high-redshift AGNs are needed to either confirm our prediction, or otherwise provide stronger constraints on the TRINITY model.

2.2 Implementation overview

TRINITY makes self-consistent and statistical halo–galaxy–SMBH connections from $z = 0 - 10$. Instead of following individual haloes across cosmic time like the UNIVERSE MACHINE (Behroozi et al. 2019), TRINITY is built upon halo statistics from N-body simulations. The halo–galaxy connection is made via the SFR – M_{peak} relation, where M_{peak} is the peak halo mass of haloes throughout their assembly histories. With a given redshift-dependent SFR – M_{peak} relation, TRINITY calculates the average star formation histories for different halo populations, and integrates them over time to obtain average galaxy stellar masses (M_*) as functions of redshift and halo mass. Average galaxy total stellar masses are then converted into average bulge masses (M_{bulge}) with a redshift-dependent but fixed M_{bulge} – M_* based on SDSS (Mendel et al. 2014) and CANDELS (Lang et al. 2014) observations. We parametrize the M_\bullet – M_{bulge} relation as a redshift-dependent power-law for *all* galaxies, i.e., SMBH hosts and non-hosts. The log-normal scatter around the M_\bullet – M_{bulge} relation, σ_{BH} , is a free parameter to be constrained in MCMC. We assume that σ_{BH} is redshift-independent, due to the lack of data to constrain its redshift dependence. As shown in Zhang et al. (2023c), our results do not change significantly if we instead parameterize the M_\bullet – M_* relation as a power-law. SMBH occupation fractions are parameterized as a sigmoid function of halo mass, which evolves with redshift. Together, these choices fully parameterize the SMBH mass distribution as a function of galaxy mass and redshift ($P(M_\bullet|M_*,z)$). Av-

erage black hole accretion rates (BHARs) are calculated by comparing the average M_\bullet of the same halo population between two consecutive snapshots. To forward model SMBH observables, we convert average BHARs into AGN Eddington ratio distributions using parametrized AGN efficiency, duty cycles, and Eddington ratio distribution shapes (i.e., $P(L_{\text{bol}}|M_\bullet, M_*, z)$). Before comparing predicted observables with real data, we also correct them for systematic effects like errors in stellar masses and SFRs from SED fitting, and Compton-thick AGN obscuration. To get the posterior distribution of model parameters, we compare predicted observables with the compiled data for ~ 2 million points in parameter space with a custom made Metropolis Markov Chain Monte Carlo (MCMC) algorithm (based on Haario et al. 2001). From this posterior distribution, we obtain both the best-fitting model parameter set given the observational constraints, as well as the uncertainties in the halo–galaxy–SMBH connection and AGN properties, e.g., efficiency and duty cycles. The whole parametrization of TRINITY contains 54 free parameters.

2.3 Calculating SMBH statistics

In this work, we present various statistics of SMBHs at $z \gtrsim 6$ from the best-fitting TRINITY model. Here, we briefly introduce how we calculate them.

To calculate the median M_\bullet – M_* relation for bright quasars above a certain bolometric luminosity threshold ($L_{\text{bol,lim}}$), we firstly calculate the SMBH mass function at each given stellar mass and redshift:

$$\phi_\bullet(M_\bullet|M_*, L_{\text{bol}} > L_{\text{bol,lim}}, z) = \int_0^\infty \phi_h(M_{\text{peak}}, z) \int_{L_{\text{bol,lim}}}^\infty P(M_\bullet|M_*, z) P(M_*|M_{\text{peak}}, z) P(L_{\text{bol}}|M_{\text{peak}}, z) dL_{\text{bol}} dM_{\text{peak}}, \quad (1)$$

where $P(L_{\text{bol}}|M_h, z)$ is obtained by integrating the number density of black holes emitting at the corresponding Eddington ratio:

$$P(L_{\text{bol}}|M_{\text{peak}}, z) = \int_0^\infty P(\eta_{\text{rad}}(L_{\text{bol}}, M_\bullet)|M_\bullet, M_{\text{peak}}, z) \times P(M_\bullet|M_{\text{peak}}, z) dM_\bullet. \quad (2)$$

We then calculate the median M_\bullet and the 16th–84th percentile range from $\phi_\bullet(M_\bullet|M_*, L_{\text{bol}} > L_{\text{bol,lim}}, z)$. The M_\bullet – M_* relation for bright quasars is then obtained by repeating this process for a series of M_* at fixed z .

Similarly, SMBH mass functions are convolutions of halo mass functions and the conditional probability distribution of M_\bullet given M_{peak} :

$$\phi(M_\bullet, z) = \int_0^\infty \phi_h(M_{\text{peak}}, z) P(M_\bullet|M_{\text{peak}}, z) dM_{\text{peak}}, \quad (3)$$

where $P(M_\bullet|M_{\text{peak}}, z)$ is the convolution of: 1) the probability distribution of M_\bullet given M_* , $P(M_\bullet|M_*, z)$; and 2) the probability distribution of M_\bullet given M_{peak} , $P(M_*|M_{\text{peak}}, z)$:

$$P(M_\bullet|M_{\text{peak}}, z) = \int_0^\infty P(M_\bullet|M_*, z) P(M_*|M_{\text{peak}}, z) dM_*. \quad (4)$$

The cosmic number density of SMBHs above a mass threshold

$M_{\bullet, \text{lim}}$, $n_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, z)$, is a convolution of $P(M_{\bullet}|M_{\text{peak}}, z)$ and the halo mass function $\phi_h(M_{\text{peak}})$:

$$n_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, z) = \int_{M_{\bullet, \text{lim}}}^{\infty} P(M_{\bullet}|M_{\text{peak}}, z) \phi_h(M_{\text{peak}}, z) dM_{\text{peak}}. \quad (5)$$

The differential number of SMBHs above $M_{\bullet, \text{lim}}$ per unit redshift, $dN_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, z)/dz$, is then:

$$\frac{dN_{\bullet}}{dz}(M_{\bullet} > M_{\bullet, \text{lim}}, z) = n_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, z) \times \frac{dV}{dz}, \quad (6)$$

where dV/dz is the differential comoving volume per unit redshift, calculated based on the adopted cosmology (see §1). The cumulative number of SMBHs above $M_{\bullet, \text{lim}}$ at above redshift z , $N_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, > z)$ is thus:

$$N_{\bullet}(M_{\bullet} > M_{\bullet, \text{lim}}, > z) = \int_z^{\infty} \frac{dN_{\bullet}}{dz}(M_{\bullet} > M_{\bullet, \text{lim}}, z') dz'. \quad (7)$$

3 SIMULATIONS AND DATA CONSTRAINTS

3.1 Dark Matter Halo Statistics

As noted in §2.2, TRINITY requires only halo population statistics from dark matter simulations, as opposed to individual halo merger trees. We use the peak historical mass (M_{peak}) halo mass functions for central and satellite haloes from Behroozi et al. (2013), for the cosmology specified in the introduction. These mass functions are based on central halo mass functions from Tinker et al. (2008), with adjustments to include satellite halo number densities as well as to use M_{peak} instead of the present day halo mass. These adjustments were based on the Bolshoi & Consuelo simulations (Klypin et al. 2011). We refer readers to Appendix G of Behroozi et al. (2013) for full details. With these calibrations, the halo statistics used in this work are suitable for studying the evolution of halos from $10^{10} M_{\odot}$ to $10^{15} M_{\odot}$.

Haloes grow by smooth accretion and halo mergers. For average halo mass accretion histories, we use the fitting formulae in Appendix H of Behroozi et al. (2013). For halo mergers, we fit merger rates from the UNIVERSEMACHINE (Behroozi et al. 2019), with full details and formulae in Appendix B of Zhang et al. (2023c).

In addition to statistical halo assembly histories, we also use individual haloes from the MultiDark Planck 2 simulation (MDPL2, Klypin et al. 2016) to study the evolution of $z = 6$ haloes hosting SMBHs with $M_{\bullet} > 10^9 M_{\odot}$. The MDPL2 simulation assumes Planck cosmology (Planck Collaboration et al. 2016), and has a simulation box size of 1 Gpc/h. The mass of each dark matter particle is $1.51 \times 10^9 M_{\odot}/h$, which enables the simulation to resolve haloes down to $M_{\text{peak}} \sim 10^{11} M_{\odot}$. As we will show in this work, the combination of the mass resolution and simulation box volume makes MDPL2 suitable for studying the growth histories of haloes hosting SMBHs with $M_{\bullet} > 10^9 M_{\odot}$ at $z \sim 6$, as well as the descendants of typical haloes hosting GNz-11 and CEERS_1019.

3.2 Observational Data Constraints

To constrain the 54 free parameters in TRINITY, we have originally compiled a diverse dataset including galaxy and SMBH observables. The galaxy data include: galaxy stellar mass functions (SMFs), quenched fractions (QFs), cosmic star formation rates (CSFRs), average specific star formation rates (SSFRs), and UV

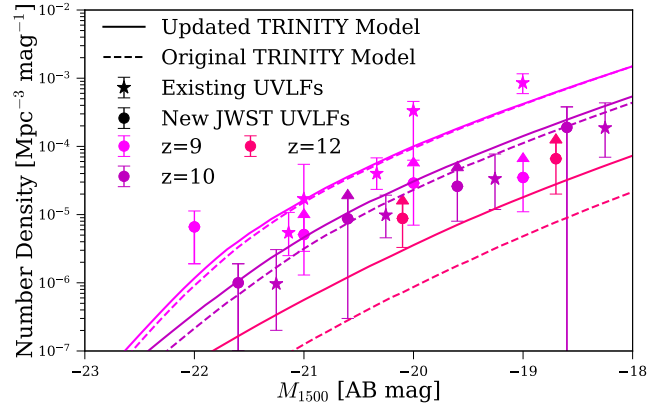


Figure 1. The comparison between the original TRINITY model (dashed lines, Zhang et al. 2023c), the updated TRINITY model (solid lines), the existing galaxy UVLF data (stars, Ishigaki et al. 2018; Oesch et al. 2018; Bouwens et al. 2019, also see Table 8 of Zhang et al. 2023c), and the new galaxy UVLFs from JWST (Harikane et al. 2023a). The difference between the original and updated TRINITY models increases with redshift. See §3.2.

luminosity functions (UVLFs). Collectively, these data cover a redshift range of $0 \leq z \leq 10$. The SMBH observables are: quasar luminosity functions (QLFs), quasar probability distribution functions (QPDFs), active black hole mass functions (ABHMFs), the $z = 0$ M_{\bullet} – M_{bulge} relation, and the observed M_{\bullet} distribution of high redshift bright quasars. These SMBH observables collectively cover a redshift range of $0 \leq z \leq 6.5$. These datasets are summarized in Tables 3–10 of Zhang et al. (2023c), and we refer readers to Appendix C of Behroozi et al. (2019) and Appendix D of Zhang et al. (2023c) for the details about the homogenization of galaxy and AGN data, respectively.

Since its launch, JWST has found more high-redshift UV-bright galaxies than predicted by theoretical models such as the UNIVERSEMACHINE. A higher number density of high-redshift galaxies likely implies more high-redshift SMBHs, which would systematically change our prediction for the early Universe. We thus added the $9 \lesssim z \lesssim 13$ galaxy UVLFs from Harikane et al. (2023b) to our previous data compilation, which includes only galaxies confirmed with spectroscopy. In TRINITY, we calculate galaxy UV magnitudes with a scaling relation between UV magnitudes and SFR, to get statistically equivalent results to the Flexible Stellar Population Synthesis code (FSPS, Conroy & White 2013). For full details of calculating galaxy UV magnitudes, we refer the readers to Appendix B of Zhang et al. (2023c).

The addition of $9 \lesssim z \lesssim 13$ galaxy UVLFs does not induce any qualitative changes in the results presented by Zhang et al. (2023c), Zhang et al. (2023a), or Zhang et al. (2023b). Quantitatively, it does increase the number of UV-bright galaxies at $z \gtrsim 9$, and ultimately, the galaxy masses at fixed halo masses. This increase in galaxy number density and typical galaxy mass also leads to more SMBHs at fixed SMBH masses, because the redshift-dependent M_{\bullet} – M_{\star} relation is unchanged with the addition of new galaxy UVLFs. As shown in Fig. 1, this change in galaxy number density increases towards higher redshifts, especially above $z = 10$. This is because the new JWST UVLFs (solid circles) are largely consistent with existing UVLFs in our data compilation (solid stars), but we lacked constraints at $z \sim 12$ in the original TRINITY model. Without $z \sim 12$ constraints, the original TRINITY underproduced UV-bright galax-

ies, leading to the larger difference at this redshift. The results presented in this work are based on the *updated* TRINITY model after adding these $9 \lesssim z \lesssim 13$ galaxy UVLFs to our data constraints. We have also updated the plots in previous TRINITY papers [here](#).

4 RESULTS

We present the offset between the $z = 6 - 10$ $M_\bullet - M_*$ relations for high redshift bright quasars vs. all SMBHs in §4.1, and the cosmic abundance of very massive black holes ($M_\bullet > 10^9 M_\odot$ and $10^{10} M_\odot$) in §4.2. In §4.3, we show the $z = 0$ descendant mass distribution of the $z = 6$ haloes hosting $M_\bullet > 10^9 M_\odot$ and $10^{10} M_\odot$ SMBHs, and discuss the prospect of detecting these descendants in the local Universe.

4.1 The biased $M_\bullet - M_*$ relation for bright quasars between $6 \lesssim z \lesssim 11$

In [Zhang et al. \(2023a\)](#), we showed the median $M_\bullet - M_*$ relation for bright quasars at $z \sim 6$, and concluded that overmassive SMBHs compared to the observed local $M_\bullet - M_*$ relation are consistent with arising purely from selection bias against less massive and thus fainter quasars. In Fig. 2, we extend our predictions to even higher redshifts, i.e., $6 \lesssim z \lesssim 11$. The solid lines denote the intrinsic (i.e., including active and inactive SMBHs) median $M_\bullet - M_*$ relation at different redshifts. The non-solid curves are the $M_\bullet - M_*$ relations for quasars above various bolometric luminosity limits, which are labeled next to each line. Given their weak redshift dependence, we only show these quasar $M_\bullet - M_*$ relations at $z = 11$ for visual clarity. The shaded regions correspond to $1 - \sigma$ log-normal scatter of ~ 0.6 dex around these quasar $M_\bullet - M_*$ relations, which includes: 1) the intrinsic scatter around the $M_\bullet - M_*$ relation (~ 0.3 dex, according to the best-fit TRINITY model); and 2) the random scatter around M_\bullet from virial estimates (0.5 dex, see, e.g., [McLure & Dunlop 2004](#); [Vestergaard & Osmer 2009](#)). The scatter is nearly luminosity-independent, so we only show the scatter for the brightest quasars with $\log L_{\text{bol}} > 48$ and the deepest samples with $\log L_{\text{bol}} > 45$, also for clarity.

Similar to those at $z \sim 6$, the $M_\bullet - M_*$ relations for bright quasars continue to be higher than the intrinsic relation for all SMBHs. This is because at high redshifts, different SMBHs share very similar Eddington ratio distributions (see §2.1). When selected by bolometric luminosity, more massive SMBHs are brighter and naturally over-represented in the sample. At fixed bolometric luminosity limits, the $M_\bullet - M_*$ relations for quasars barely evolve with redshift, despite the evolution in the *intrinsic* $M_\bullet - M_*$ relation for *all* SMBHs. This is because at fixed SMBH mass, the Eddington ratio distribution evolves very mildly beyond $z \sim 6$. Thus, selecting quasars by their luminosities yields similar SMBH masses across different redshifts, which are more and more overmassive SMBHs compared to the lower intrinsic $M_\bullet - M_*$ relations towards higher redshifts.

We also note that the typical M_\bullet for quasars depends only weakly on host galaxy mass below $\log M_* \sim 9.5$. This is because, on average (i.e., including their active and dormant phases), high-redshift SMBHs grow at slightly below the Eddington rate, and hyper-Eddington AGNs are very rare. This upper limit in Eddington ratios translates into a lower limit in the typical M_\bullet , regardless of host galaxy mass. For brighter quasars (e.g., $\log L_{\text{bol}} \geq 48$), the typical SMBH mass remains above $3 \times 10^8 M_\odot$ even for galaxies below the same mass. This implies that if we detected any $\log L_{\text{bol}} \geq 48$

quasars in a $M_* < 3 \times 10^8 M_\odot$ galaxy, they either: 1) have Eddington ratios much higher than predicted by TRINITY; and/or 2) they are in the outsized black hole galaxy (OBG) phase ([Agarwal et al. 2013](#); [Natarajan et al. 2017](#)).

Of note, when quantifying this luminosity dependence of the $M_\bullet - M_*$ relation (i.e., the Lauer bias), we assume that *every* AGN above those luminosity limits is included in the samples, with SMBH mass measurements via virial estimates. In other words, we do not account for the possibility that these AGNs may be hidden in the light from host galaxies. This assumption is more likely to hold for brighter quasars. For fainter AGNs, e.g., $\log L_{\text{bol}} \sim 45$, stronger contamination from the host galaxies makes it more difficult to measure SMBH masses, even if the AGN luminosities are technically above the limits of current instruments. As a result, overmassive SMBHs would still be preferred among fainter AGNs (see, e.g., [Volonteri et al. 2023](#)). It takes detailed modeling of high-redshift galaxy and SMBH spectral energy distributions (SEDs) to properly account for this bias, which we defer to a future work. Therefore, our predicted $M_\bullet - M_*$ relations for fainter AGNs (e.g., with $\log L_{\text{bol}} > 45$) should be regarded as *lower limits*.

We also show the four new highest-redshift AGNs detected by JWST, CEERS_00717 ([Harikane et al. 2023a](#), the triangle), CEERS1019 ([Larson et al. 2023](#), the pentagon), UHZ1 ([Bogdan et al. 2023](#); [Goulding et al. 2023](#), the diamond) and GN-z11 ([Maiolino et al. 2023a](#), the star) in Fig. 2. Since the SMBH masses of CEERS_00717, CEERS_1019, and GN-z11 are estimated with virial estimates, we adopt an M_\bullet uncertainty of 0.5 dex for these three AGNs. The M_\bullet of UHZ1 was estimated from X-ray spectral information, modeling jointly luminosity and obscuring column density, and further assuming Eddington-limited accretion. Given that SMBHs accreting at five times the Eddington rate have been detected at similar redshifts (i.e., GNz-11), we adopt a higher M_\bullet uncertainty of 0.7 dex (i.e., a factor of five) to reflect the uncertainty in Eddington ratio. Galaxy mass uncertainties are directly taken from these source papers. We note that the stellar masses have in several cases been determined without including an AGN component in the fit, and that the uncertainties in the star formation histories can give large systematic uncertainties (see, e.g., the summary in Table 1 of [Volonteri et al. 2023](#)).

Compared to the brighter quasars with median $\log L_{\text{bol}} \sim 47$ compiled by [Izumi et al. \(2021\)](#), all these JWST AGNs have lower SMBH masses. This is consistent with TRINITY’s prediction for the luminosity-dependent bias (i.e., the Lauer bias) of the $M_\bullet - M_*$ relation. At face value, the SMBH in CEERS_00717 seems much more massive than TRINITY’s prediction for faint AGNs. But given the substantial uncertainty in galaxy mass, we decide to compare our predictions with CEERS_00717 after more precise measurements of galaxy mass are available. We do note that both CEERS1019 and GN-z11 lie above our $M_\bullet - M_*$ relation for AGNs with $\log L_{\text{bol}} > 45$, at $\sim 1.5\sigma$ and $\sim 0.8\sigma$, respectively. This could be due to: 1) small number statistics; 2) the aforementioned selection bias due to smaller SMBHs being harder to detect against the host galaxy’s light; and/or 3) TRINITY underpredicting SMBH masses compared to the real Universe. Further observations and better characterizations of all selection biases are needed to ascertain the nature of this difference. In §5.2, we compare CEERS1019 and GN-z11 with other $6 \lesssim z \lesssim 11$ predictions from TRINITY.

Finally, we make a few remarks about UHZ1. Assuming Eddington accretion, the M_\bullet of UHZ1 is more than 1 dex higher than expected for AGNs above similar luminosities ($\log L_{\text{bol}} \geq 45.7$, the dash-dotted line). But, at $z \gtrsim 10$, another super-Eddington AGN has been found, i.e., GN-z11. If UHZ1 is in fact also accreting at a

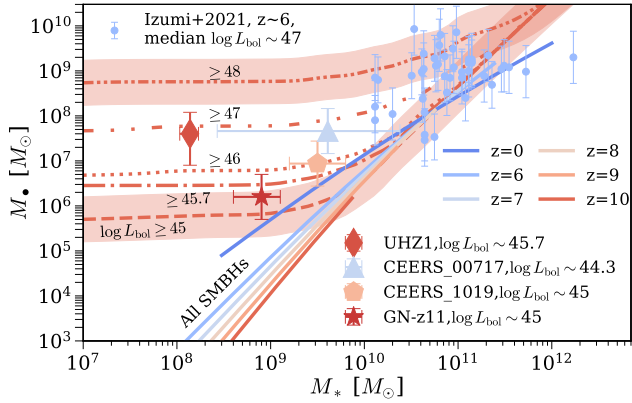


Figure 2. The observed median M_{\bullet} – M_* relations for bright quasars selected by different bolometric luminosity thresholds at $z = 6$ – 11 . The dashed, dash-dotted, dotted, and dash-double-dotted curves are the M_{\bullet} – M_* relations for bright quasars above different luminosity limits, which are annotated next to each line. These M_{\bullet} – M_* relations are nearly redshift-independent, so we only show the ones at $z = 10$ for visual clarity. The shaded regions are the $1 - \sigma$ log-normal scatter (~ 0.6 dex) around the M_{\bullet} – M_* relations for bright quasars, which includes the intrinsic scatter around the M_{\bullet} – M_* relation (~ 0.3 dex), as well as the typical random scatter in M_{\bullet} from virial estimates (~ 0.5 dex). Since the magnitude of this scatter is nearly luminosity-independent, we only show it for the brightest and faintest quasars, for clarity. The solid lines of different colours represent the intrinsic median M_{\bullet} – M_* relations for *all* SMBHs at each redshift. UHZ1 (Bogdan et al. 2023; Goulding et al. 2023), CEERS_00717 (Harikane et al. 2023a), CEERS_1019 (Larson et al. 2023) and GN-z11 (Maiolino et al. 2023a) are shown in the diamond, the triangle, the pentagon, and the star, respectively. M_{\bullet} uncertainties of 0.5 dex are assumed for CEERS_00717, CEERS_1019, and GN-z11. For UHZ1, the upper uncertainty of M_{\bullet} is assumed to be 0.5 dex, but the lower uncertainty corresponds to changing the assumed Eddington ratio from 1 to 5, i.e., 0.7 dex. The galaxy mass uncertainties for these four high-redshift AGNs are taken from the source papers. We also show the $z \sim 6$ bright quasars with median $\log L_{\text{bol}} \sim 47$ compiled by Izumi et al. (2021) in solid circles with error bars (~ 0.5 dex). See §4.1.

similar Eddington ratio, the resulting M_{\bullet} would be marginally consistent with TRINITY predictions. A caveat of this argument is that the current estimate of UHZ1’s luminosity (and thus M_{\bullet} as well) is only a 1σ lower limit from X-ray spectral fitting. Such values were reported instead of the actual best-fitting values due to the strong degeneracy between the intrinsic X-ray luminosity and gas column density used in the fitting process. Thus, choosing higher intrinsic luminosity values will push UHZ1 further away from TRINITY’s prediction. This would imply that UHZ1 may be experiencing an OBG phase, where the SMBH is not evolving in sync with the host galaxy and could become more massive than the galaxy. If confirmed, such AGNs in OBGs will provide critical information for future versions of TRINITY so that it can include these extreme objects.

4.2 Very massive black holes at $z > 6$

The top panel of Fig. 3 shows the redshift density of very massive black holes ($M_{\bullet} > 10^9 M_{\odot}$ and $10^{10} M_{\odot}$), i.e., dN/dz , in the whole observable Universe. The solid lines represent all (i.e., both active and dormant) SMBHs with *intrinsic* masses above these thresholds. From $z = 10$ to $z \sim 1 - 2$, dN/dz increases rapidly with time due to high black hole accretion rates. Below $z \sim 1 - 2$, dN/dz decreases due to the combination of: 1) the decreasing observable volume; 2) the decline in SMBH growth. The middle panel of Fig.

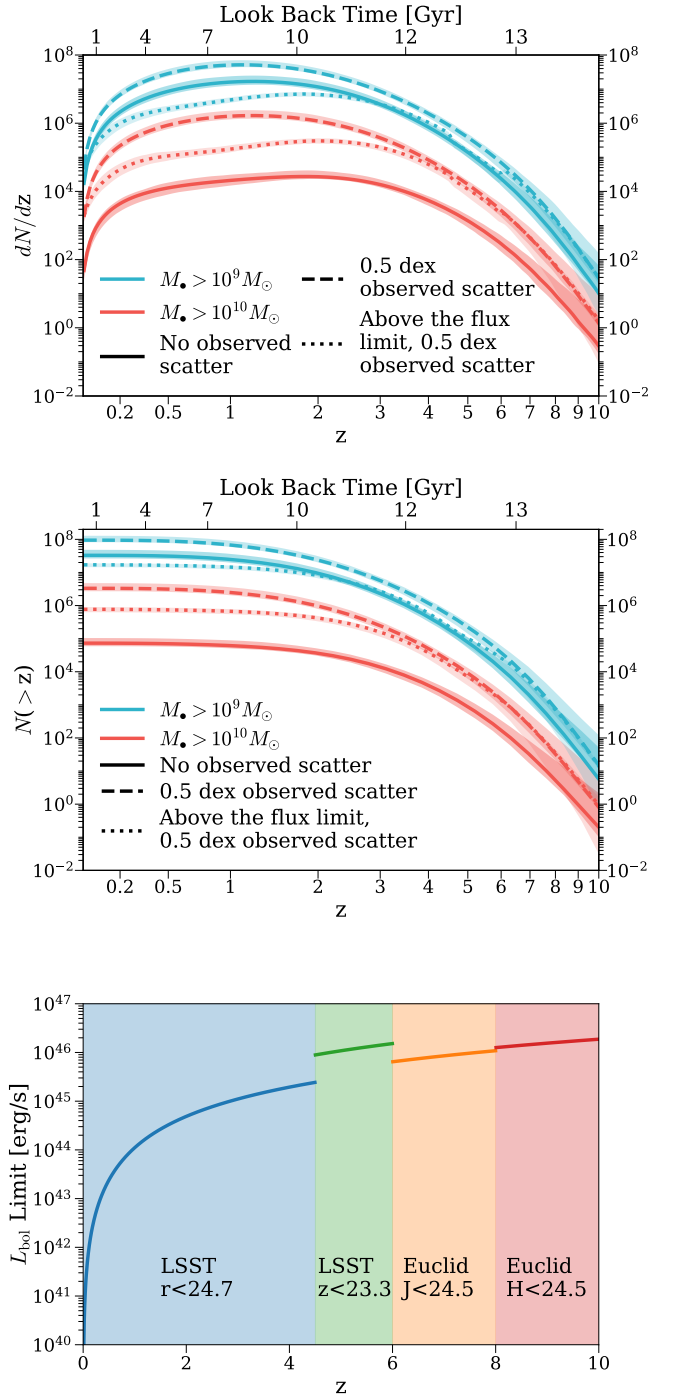


Figure 3. **Top Panel:** The number of very massive black holes *per unit redshift* in the whole observable universe (i.e., dN/dz). The solid lines show results assuming no observational scatter in the measured black hole masses, and the dashed lines show results assuming an observational scatter of $\sigma_{\text{vir}} = 0.5$ dex, typical for virial estimates (see §4.2 for details). The dotted lines show the numbers of SMBHs above the bolometric luminosity limit shown in the bottom panel. **Middle Panel:** The *cumulative* number of very massive black holes as a function of redshift. The shaded regions show the 68% confidence intervals inferred from the model posterior distribution. **Bottom Panel:** The bolometric luminosity limit to produce the dotted lines in the top and middle panels, as a function of redshift. See §4.2.

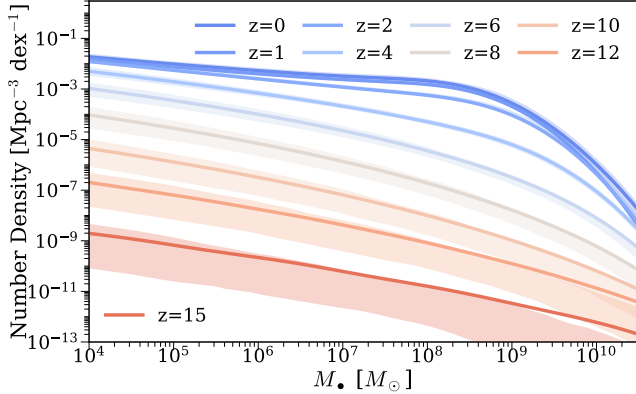


Figure 4. Total SMBH mass functions from $z = 0 - 10$. The shaded regions show the 68% confidence intervals inferred from the model posterior distribution. See §4.2. This figure is reproduced from Zhang et al. (2023c) to help readers understand the trends in Fig. 3.

3 shows cumulative numbers of very massive black holes at different redshifts in the observable Universe. Beyond the local Universe, virial estimates remain the only way to measure black hole masses for large samples. These methods result in random scatter of 0.4–0.6 dex in observed black hole masses around the intrinsic values (McLure & Jarvis 2002; Vestergaard & Peterson 2006). This scatter causes an Eddington bias (Eddington 1913), because there are more low-mass black holes that can be upscattered than high-mass black holes that can be downscattered. This elevates the observed number of massive black holes above the true number, as shown by the dashed curves in the top and middle panels of Fig. 3. Overall, an observational scatter of 0.5 dex elevates observed massive black hole number counts by a factor of $\sim 3 - 40$. This boost in black hole numbers is stronger for more massive objects and/or lower redshifts. This is because the massive end of the SMBH mass function becomes steeper towards lower redshifts, as is shown in Fig. 4. The steeper mass function leads to a larger population of upscattered M_{\bullet} measurements, and thus a stronger boost in the numbers of large SMBHs. With this Eddington bias, the cumulative numbers of $M_{\bullet} > 10^9, 10^{10} M_{\odot}$ SMBHs increase by six orders of magnitude from $z \sim 10$ to $z \sim 2$.

According to TRINITY, a $10^9 M_{\odot}$ SMBH would be $\sim 3 - 4$ dex above the median $M_{\bullet} - M_{*}$ relation at $z = 10$ (the bottom right panel of Fig. 2). The rarity of such SMBHs is also shown in the bottom panel of Fig. 3: we expect to see only $\lesssim 10$ such SMBHs at $z \gtrsim 10$ over the whole sky, whether they are active or not. Such rarity results from the extreme difficulty of growing a $10^9 M_{\odot}$ SMBH by $z \sim 10$, which requires an appropriate combination of SMBH seed mass, Eddington ratios, duty cycle, and AGN efficiency (Pacucci & Loeb 2022). The detection of massive BHs at $z \gtrsim 10$ will therefore provide stringent constraints on SMBH seeding mechanisms, as well as their early growth histories. We will further compare the possible progenitor masses and growth speeds of $10^9 M_{\odot}$ SMBHs from Pacucci & Loeb (2022) and TRINITY in §5.3.

To better understand the SMBH detection capability of future wide field surveys, we also made predictions of expected SMBH numbers from the *Legacy Survey of Space and Time* (LSST) and the *Euclid* wide survey. To calculate the bolometric luminosity limits as a function of redshift, we adopted the magnitude limits from Ivezić et al. (2019) and Euclid Collaboration et al. (2022). We fur-

ther convert observed magnitude limits in different wavebands (labeled in the bottom panel of Fig. 3) into bolometric luminosities, by assuming: 1) that the bolometric correction to the rest-frame i -band luminosity is 12 (Richards et al. 2006); and 2) the UV-to-optical quasar SED is a power-law with a power-law index of $\alpha_{\nu} = -0.44$ (Vanden Berk et al. 2001). The luminosity limit as a function of redshift is shown in the bottom panel of Fig. 3, and the differential (cumulative) number of massive black holes is shown in dotted lines in the top (middle) panel of Fig. 3.

At $z \gtrsim 4$, the vast majority of SMBHs with $M_{\bullet, \text{obs}} > 10^9 M_{\odot}$ are active, so the bulk of these SMBHs are above the flux limits of LSST and Euclid. We note that the number of SMBHs in Fig. 3 include both obscured and unobscured SMBHs. Given the uncertainties in the obscured AGN fraction towards higher redshifts, we opt not to show results for unobscured AGNs in Fig. 3, but only give a rough estimate: at these high luminosity limits ($\gtrsim 10^{46}$ erg/s), we expect $\sim 40\%$ (80%) of such SMBHs to be unobscured, based on the obscured fractions as functions of AGN X-ray luminosity given by Merloni et al. (2014) and Ueda et al. (2014). Below $z \sim 4$, massive black holes experience a significant decline in activity level. Consequently, smaller and smaller fractions of massive black holes at lower redshifts will lie above the flux limit of LSST.

4.3 Host halos of massive black holes at $z = 6$ and their descendants at $z = 0$

We show the distribution of host halo masses for SMBHs with $M_{\bullet} > 10^9 M_{\odot}$ (solid curves) and $M_{\bullet} > 10^{10} M_{\odot}$ (dashed curves) at $z = 6$, in the top panel of Fig. 5. When using virial estimates to measure M_{\bullet} , there is random scatter in the measured M_{\bullet} around the intrinsic values. So, we also show the host halo mass functions for observed $M_{\bullet} > 10^9 M_{\odot}$ in blue and red curves, respectively. In this comparison, we assume a random scatter of 0.5 dex in measured M_{\bullet} . This random scatter induces Eddington bias, i.e., more lower-mass haloes hosting intrinsically smaller SMBHs have their observed M_{\bullet} upscattered than the opposite. Thus, the host halo mass function (HMF) is both enhanced in normalization and shifted towards lower masses. The median host halo mass for SMBHs with intrinsic $M_{\bullet} > 10^9 M_{\odot}$ is $\sim 3 \times 10^{12} M_{\odot}$, compared to $\sim 2 \times 10^{12} M_{\odot}$ for SMBHs with observed $M_{\bullet} > 10^9 M_{\odot}$. For $M_{\bullet} > 10^{10} M_{\odot}$ SMBHs, their host haloes are typically $\sim 5 \times 10^{12} M_{\odot}$ and $\sim 3 \times 10^{12} M_{\odot}$, for intrinsic and observed M_{\bullet} , respectively. But overall, host haloes of $M_{\bullet} > 10^9 (10^{10}) M_{\odot}$ SMBHs only make up $\sim 10\%$ (1%) of all the haloes (orange solid curve, calculated from the MDPL2 simulation) at their respective median host halo masses, which correspond to $\sim 4.5 - 5.5\sigma$ peaks in Gaussian random fields (Eisenstein & Hu 1998; Diemer 2018). In such haloes, gas reservoirs available for rapid SMBH growth could include cold flows (e.g., Feng et al. 2014 and Latif et al. 2022), accretion triggered by protogalaxy mergers (e.g., Li et al. 2007), or a combination of the two (e.g., Dubois et al. 2012). Observationally, we expect to see overdensities of galaxies around the SMBHs in these massive haloes (Wang et al. 2023). In future work, we will create realistic mock galaxy+AGN catalogs to quantify the typical magnitude and stochastic variance of galaxy overdensities around $z \gtrsim 6$ quasars, and compare them with results from, e.g., the ASPIRE (Wang et al. 2023) and EIGER (Kashino et al. 2023) surveys.

Ever since the discovery of SMBHs with $M_{\bullet} > 10^9 M_{\odot}$, there have been studies investigating where their descendant SMBHs can be found in the local Universe (e.g., Angulo et al. 2012; Fanidakis et al. 2013; Tenneti et al. 2018). To address this question with TRINITY, we follow each $z = 6$ halo in the MDPL2 simulation

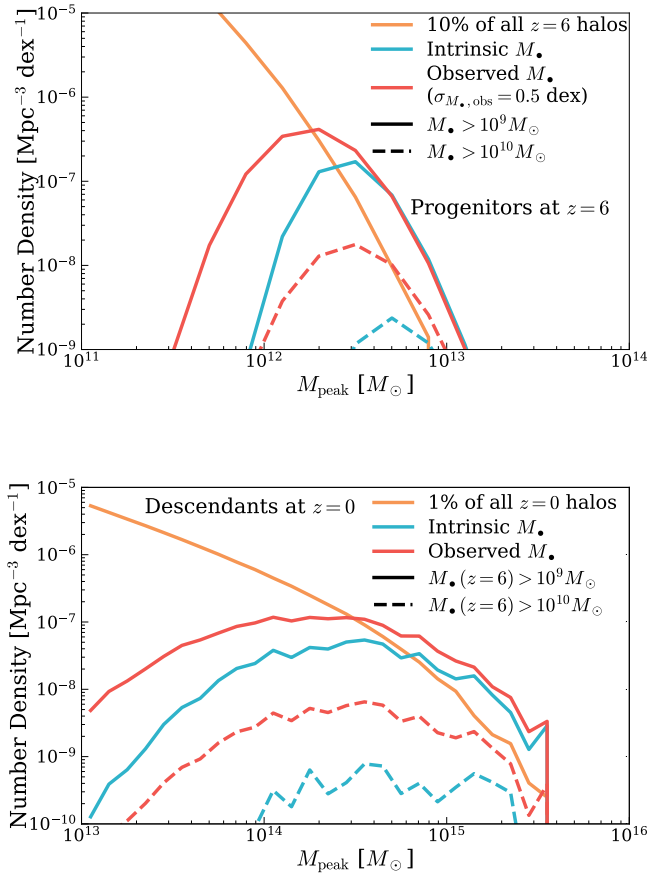


Figure 5. Top Panel: The host halo mass functions of SMBHs with intrinsic (blue lines) and observed (red lines) $M_{\bullet} > 10^9 M_{\odot}$ (solid lines) and $M_{\bullet} > 10^{10} M_{\odot}$ (dashed lines) at $z = 6$, calculated by TRINITY. The orange solid line is the $z = 6$ halo mass function from the MDPL2 simulation, scaled down to 10% of the original value for ease of comparison. **Bottom Panel:** The $z = 0$ descendant halo mass functions of the $z = 6$ haloes hosting SMBHs with intrinsic (blue lines) and red (red lines) $M_{\bullet} > 10^9 M_{\odot}$ (solid lines) and $M_{\bullet} > 10^{10} M_{\odot}$ (dashed lines). They are calculated by following individual haloes in the MDPL2 simulation across time, and weighting each of them by the ratios between the host halo mass functions and the total halo mass functions in the top panel. The orange solid line is the $z = 0$ halo mass function in the MDPL2 simulation, scaled down by a factor of 100. See §4.3.

down to $z = 0$, and examine their descendant halo masses. We weight each progenitor–descendant pair by the fraction of halos with the progenitor’s mass at $z = 6$ that would host a $> 10^9/10^{10} M_{\odot}$ SMBH. As shown in the bottom panel of Fig. 5, the descendant HMFs are much broader than the progenitor HMFs. This is due to the diversity in the assembly histories of individual haloes. At $z = 0$, descendants with $M_{\bullet}(z = 6) > 10^9 M_{\odot}$ typically have $M_{\text{peak}} \sim 3 - 7 \times 10^{14} M_{\odot}$. However, these haloes only make up $\lesssim 1\%$ of the haloes at similar mass scales (the orange solid line). Combined with the small comoving volume in the local Universe, we would expect only ~ 1 such relic SMBH closer than $z = 0.05$. For $M_{\bullet}(z = 6) > 10^{10} M_{\odot}$, their host halo masses are similar to those with $M_{\bullet}(z = 6) > 10^9 M_{\odot}$, but they are much rarer (i.e., a factor of $\sim 30 - 100$). This makes it even harder to find the descendant haloes of these more massive high-redshift SMBHs in the local Universe.

5 DISCUSSION

In §5.1, we discuss the physical implications of the steeper intrinsic $M_{\bullet}-M_{*}$ relation between $7 \lesssim z \lesssim 10$ from TRINITY; In §5.2, we compare CEERS_1019 and GN-z11 with more TRINITY predictions at $6 \lesssim z \lesssim 11$; In §5.3, we compare the progenitor masses and growth rates of $\gtrsim 10^9 M_{\odot}$ SMBHs at $z \sim 9$ from Pacucci & Loeb (2022) with those from TRINITY; §5.4 discusses the constraining power of $z = 6 - 10$ quasar luminosity functions on the halo–galaxy–SMBH connection.

5.1 Physical Implications of the $M_{\bullet}-M_{*}$ relation at $z \gtrsim 6$ from TRINITY

According to TRINITY, the intrinsic $M_{\bullet}-M_{*}$ relation between $7 \lesssim z \lesssim 10$ is relatively static, with the largest difference being that the relation is mildly steeper at higher redshifts than at lower redshifts (see also Volonteri & Stark 2011). This means that SMBHs are less massive than in local galaxies at fixed stellar masses below $5 \times 10^{10} M_{\odot}$ (§4.1). Physically, this redshift evolution implies that SMBH growth in $M_{*} < 10^{11} M_{\odot}$ galaxies is modestly slower than the host galaxy’s growth. Several physical mechanisms could contribute to such limited growth: 1) strong supernova feedback in low-mass galaxies preventing efficient SMBH accretion (as in, e.g., IllustrisTNG, Dubois et al. 2015; Bower et al. 2017; Pillepich et al. 2018); 2) erratic motions of SMBHs due to shallow potential wells (as in, e.g., Habouzit et al. 2017, Bellovary et al. 2019, and Dekel et al. 2019, though c.f. Choksi et al. 2017); and 3) Feedback from SMBHs themselves (e.g., as in ROMULUS, Tremmel et al. 2017). Limited SMBH growth at $z \gtrsim 7$ also implies that AGN are unlikely to dominate the cosmic reionization of intergalactic medium, as also suggested by, e.g., Finkelstein et al. (2019).

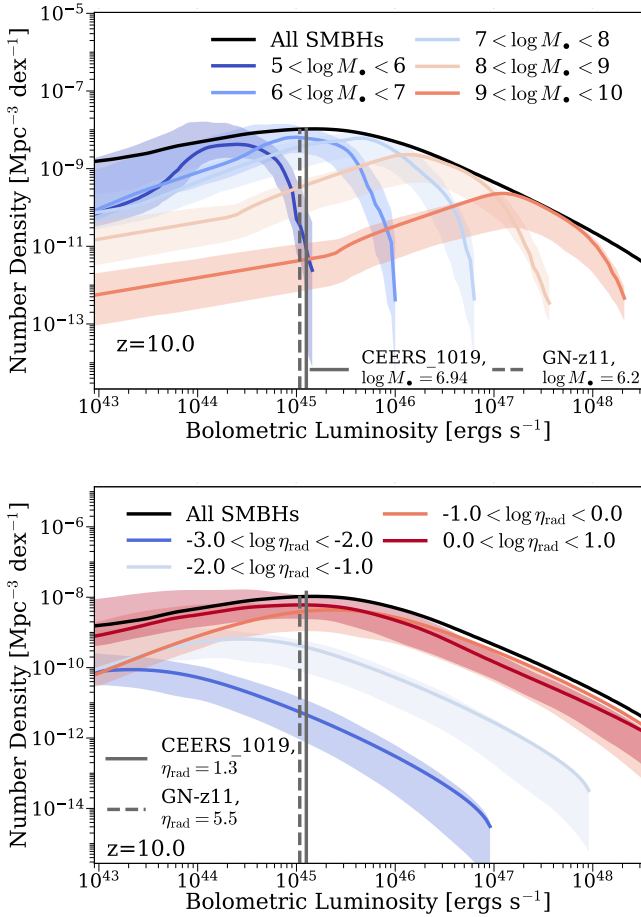
5.2 CEERS_1019 and GN-z11 in the context of TRINITY

5.2.1 SMBH mass and Eddington ratios

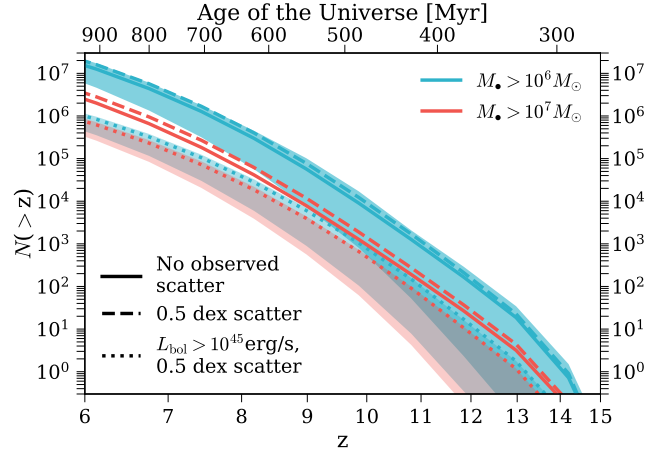
Fig. 6 shows the $z = 10$ bolometric quasar luminosity function (QLF) decomposed into bins of SMBH mass (M_{\bullet} , top panel) and Eddington ratio (η_{rad}), predicted by TRINITY, which are relevant to help better understand the measured properties of CEERS_1019 and GN-z11. At $z \sim 10$, the QLF peaks at around 10^{45} erg/s. In addition, SMBHs with $6 < \log M_{\bullet} < 7$ (top panel) and $0 < \log \eta_{\text{rad}} < 1$ dominate the QLF at around 10^{45} erg/s. Given that both GN-z11 and CEERS1019 have bolometric luminosities of $\sim 10^{45}$ erg/s, their masses ($\log M_{\bullet} \sim 6.2$ and 6.94) and Eddington ratios ($\eta_{\text{rad}} \sim 5.5$ and 1.3) are consistent with TRINITY’s predictions (shown in the solid and dashed vertical lines). We also note that in TRINITY, $z \sim 10$ SMBHs grow at **slightly below** the Eddington rate, when averaged over both their active and dormant phases. Consequently, although GN-z11 is estimated to be accreting at a super-Eddington rate, such fast accretion phases are likely intermittent. Two objects are far from a statistically representative sample, so more detections and measurements of $z \gtrsim 9$ SMBHs are needed to further confirm/disprove our predictions on the dominant SMBH mass and Eddington ratios for these $z \gtrsim 8$ AGNs.

5.2.2 Abundance of $M_{\bullet} > 10^6/10^7 M_{\odot}$ SMBHs between $6 < z < 11$

Fig. 7 shows the cumulative number of SMBHs above $10^6 M_{\odot}$ (blue curves) and $10^7 M_{\odot}$ (red curves) from $z = 6 - 11$. The solid



curves represent the number of all (i.e., active+dormant) SMBHs with *intrinsic* M_\bullet above these mass limits, whereas the dashed curves denote the number of all SMBHs with *observed* M_\bullet above the limits, assuming a random scatter of 0.5 dex in observed M_\bullet . The dotted curves are similar to the dashed curves, but for SMBHs above the bolometric luminosity limit of 10^{45} erg/s. This limit has been chosen to match the luminosities of CEERS1019 and GN-z11. Similar to Fig. 3, these curves represent the numbers of SMBHs we expect to find *over the full sky*. Compared to more massive SMBHs in Fig. 3, the scatter in observed M_\bullet causes a smaller boost in the number of SMBHs, i.e., a factor of ~ 1.5 . This is because the total BHMF is flatter at the less massive end, which reduces the magnitude of Eddington bias. According to TRINITY, we expect to find ~ 6000 and ~ 1500 SMBHs with observed $M_\bullet > 10^6 M_\odot$ and $10^7 M_\odot$ at $z > 10$, respectively. However, only part of the underlying SMBH populations are bright enough to be detected. Given the nearly mass-independent Eddington ratio distributions from TRINITY, fewer less massive SMBHs will survive the same luminosity limit. Consequently, when the luminosity limit is applied, the numbers of SMBH with $M_\bullet > 10^6 M_\odot$ and $10^7 M_\odot$ at $z > 10$ decrease to ~ 600 and ~ 400 over the whole sky, respectively.



5.2.3 Host halos of CEERS1019 and GN-z11

In the top panel of Fig. 8, we show the host halo mass functions of SMBHs with intrinsic (blue curves) and observed (red curves) $M_\bullet > 10^6 M_\odot$ at $z = 10.6$ (solid curves) and $M_\bullet > 10^7 M_\odot$ at $z = 8.7$ (dashed curves). These redshifts and mass limits are chosen to roughly match the observed properties of CEERS_1019 and GN-z11. Similar to the host halo mass functions of $M_\bullet > 10^9 M_\odot$ at $z = 6$ (§4.3, Fig. 5), the difference between the red and blue curves comes from the Eddington bias that upscatters undermassive halos hosting intrinsically smaller SMBHs. GN-z11 is likely to be hosted in a $(2-3) \times 10^{11} M_\odot$ halo, whereas CEERS1019's typical host halo mass is $(3-6) \times 10^{11} M_\odot$. These host halo masses correspond to $\gtrsim 7\sigma$ peaks in Gaussian random fields at their respective redshifts, making them candidates for protocluster cores. This means that the typical host haloes for CEERS_1019 and GN-z11 are rarer than the luminous quasars at $z \sim 6$ at their respective redshifts. We note that our estimated host halo mass for GN-z11 is ~ 1 dex higher than the one by Scholtz et al. (2023). Scholtz et al. made this estimate based on the stellar mass measurement by Tacchella et al. (2023) and the stellar mass–halo mass relation from Behroozi et al. (2019). However, TRINITY predicts that the AGN duty cycle in low mass haloes are much lower. In other words, low mass haloes are much less likely to host AGNs like GN-z11. These predictions are driven by extrapolating the QPDFs from Aird et al. (2018) to high redshifts. Therefore, TRINITY finds more massive host haloes for GN-z11, despite the very similar galaxy mass–halo mass relations at $z \gtrsim 10$.

The bottom panel of Fig. 8 shows the $z = 0$ descendant halo mass functions of haloes hosting $M_\bullet > 10^6 M_\odot$ at $z = 10.6$ and $M_\bullet > 10^7 M_\odot$ at $z = 8.7$. These descendant halo mass functions are calculated using the method introduced in §4.3. The typical descendant halo masses for GN-z11- and CEERS_1019-like objects are $\sim (1-3) \times 10^{14} M_\odot$, although the distribution is broad, due to the diversity in individual halo assembly histories. Similar to the

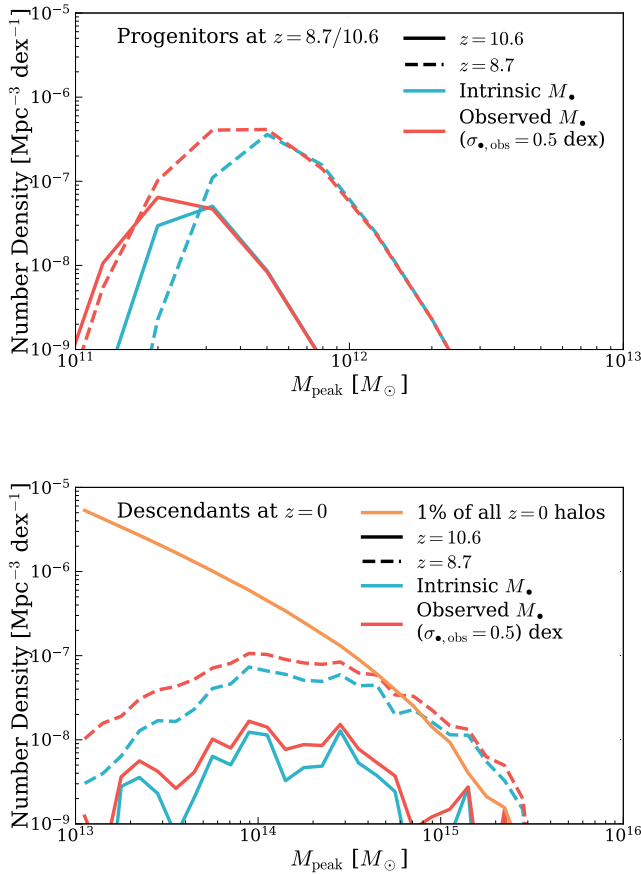


Figure 8. Top Panel: The host halo mass functions of SMBHs with intrinsic (blue curves) and observed (red curves) $M_{\bullet} > 10^6 M_{\odot}$ at $z = 10.6$ (solid curves) and $M_{\bullet} > 10^7 M_{\odot}$ at $z = 8.7$ (dashed curves). **Bottom Panel:** The $z = 0$ descendant halo mass functions of halos hosting intrinsic (blue curves) and observed (red curves) $M_{\bullet} > 10^6 M_{\odot}$ at $z = 10.6$ (solid curves) and $M_{\bullet} > 10^7 M_{\odot}$ at $z = 8.7$ (dashed curves). The orange solid curve is the total host halo mass functions $z = 0$, scaled down by a factor of 100. See §5.2.3.

$z = 0$ descendants of $z \sim 6$ massive black holes, the descendants of GN-z11- and CEERS_1019-like objects make up only $\lesssim 1\%$ of all the haloes at similar mass scales. This makes them too rare to find in the local Universe. Compared to the descendant halo mass of GN-z11 obtained by Scholtz et al. (2023) based on the Millennium Simulation (Springel et al. 2005; Chiang et al. 2013), our descendant mass scale is lower by ~ 1 dex. This is because Scholtz et al. (2023) chose to look at the average $z = 10.6$ progenitor halo masses of $z = 0$ haloes above $10^{15} M_{\odot}$, and found that the average progenitor halo mass coincides with the halo mass estimate of GNz-11. Due to the scatter in individual halo assembly histories, this is different from looking at the $z = 0$ descendant halo masses of halo populations with certain masses at higher redshifts, and will lead to a larger $z = 0$ descendant halo mass. In this work, we chose the opposite method because it is more representative of the expected descendant masses; the Scholtz et al. (2023) method is more appropriate for finding $z = 0$ objects whose typical progenitors look like GNz-11.

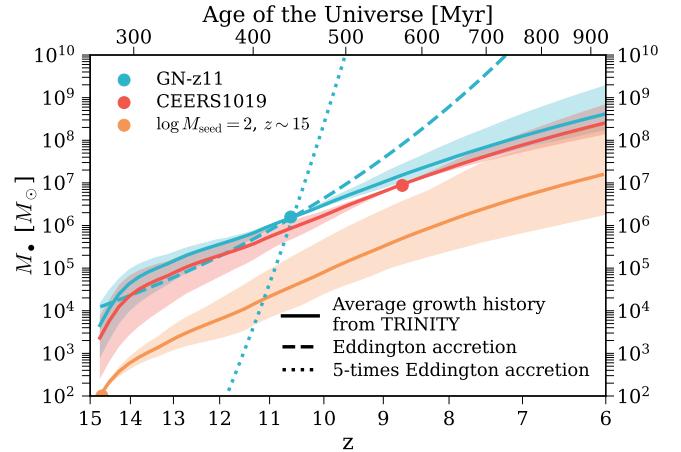


Figure 9. The average mass growth curves of GN-z11 (the blue solid curve and circle), CEERS1019 (the red curve and circle), and a seed with $M_{seed} = 100 M_{\odot}$ at $z \sim 15$ (the orange curve and circle), predicted by TRINITY. The blue dashed curve shows the mass growth history of GN-z11 assuming Eddington-limited accretion. The shaded regions reflect the statistical uncertainties in the MCMC process in TRINITY, given the existing data constraints. See §5.2.4.

5.2.4 Potential descendants and progenitors of CEERS1019 and GN-z11

To further understand the potential progenitors and descendants of GN-z11 and CEERS_1019, we calculated the average¹ Eddington ratio as a function of M_{\bullet} and time, and calculated the average mass growth curves (shown in Fig. 9 for: 1) GN-z11 (blue solid curve); 2) CEERS_1019 (red solid curve); and 3) a seed of $100 M_{\odot}$ at $z \sim 15$ (orange solid curve). The shaded regions represent the statistical uncertainties from the MCMC, which reflects the uncertainty in the input observational constraints (see §2). In other words, they quantify how uncertain the average mass growth curves are given all the existing data constraints, but does not show the scatter in the growth histories among individual SMBHs. The blue dashed(dotted) curve shows the SMBH mass growth with a constant Eddington ratio of 1(5).

The high Eddington ratios at $z \gtrsim 6$ mean that SMBHs grow exponentially, although the e -folding time scale increases due to the decline in Eddington ratio with time. Consequently, the uncertainty in M_{\bullet} increases as we move away from the starting redshift. This effect is the most significant for the $100 M_{\odot}$ seed, whose growth was traced for more than 600 Myrs in one single direction. As a comparison, the average growth histories of GN-z11 and CEERS_1019 are only traced for at most ~ 450 Myrs in each direction. At $z \sim 10.5$ and $z \sim 8.7$, the uncertainty in the $100 M_{\odot}$ seed's descendant mass does not cover either GN-z11 or CEERS_1019. By $z = 6$, the uncertainty in its descendant mass is ~ 2 dex. This means that with current observational constraints on TRINITY, the potential progenitors of GN-z11, CEERS_1019 are unlikely to be as small as $100 M_{\odot}$ at $z \sim 15$.

On average, SMBHs grow at above(below) the Eddington rate above(below) $z \sim 13$. Thus, even though GN-z11 is accreting at five times the Eddington rate at the time of observation, such fast accretion episodes are likely intermittent (also see §5.2.1). This differ-

¹ I.e., averaged over all active and dormant SMBHs of the same mass at the same time

ence between the average and the instantaneous Eddington ratio is vital in estimating SMBH growth histories, given the exponential nature of early SMBH growth. As shown in Fig. 9, a constant Eddington ratio of 5 enables the fast mass buildup of $\sim 10^{10} M_\odot$ within ~ 120 Myrs. Conversely, it would take ~ 500 Myrs to achieve the same amount of mass growth with the Eddington rate. The drastically different SMBH growth histories also point to different SMBH progenitor masses, e.g., the constant super-Eddington growth implies the possibility of creating a GN-z11-like SMBH from with $100 M_\odot$ progenitor at $z \sim 12$, which is strongly disfavored by the constant Eddington-limited scenario. In light of this, it is critical to use more accurate estimates of average Eddington ratios, rather than instantaneous values, when inferring potential SMBH progenitor masses. Because it is not possible to measure the average historical Eddington ratios for individual SMBH, it is essential to have models like TRINITY that infer average Eddington ratios for high-redshift SMBHs, based on as many data constraints as possible.

We then discuss the potential seeding mechanisms for early SMBHs like GN-z11 and CEERS_1019. As can be seen from Fig. 9, the typical $z \sim 15$ progenitor masses of GN-z11 and CEERS_1019 are $\sim 10^3 - 10^4 M_\odot$, although CEERS_1019 could be descendants of seeds as small as $\sim 300 M_\odot$ at $z \sim 15$. Such a progenitor mass range is consistent with the scenario where the seeds are created via runaway mergers in nuclear star clusters (see, e.g., Devecchi et al. 2012; Lupi et al. 2014; Schleicher et al. 2022). In addition, when SMBH seeds form in star clusters, their dynamics are determined by the cluster mass rather than the seed mass, which makes them less susceptible to erratic motions than PopIII remnants. Compared to PopIII remnant seeds, they are thus also more likely to stay in high gas density regions and experience consistent growth as predicted by TRINITY. Consequently, we believe that if the SMBHs are seeded at $z \sim 15$, the star cluster scenario is more plausible for objects like GN-z11 and CEERS_1019, compared to the PopIII remnant scenario. We also note that since the 68th percentile range of GN-z11’s and CEERS_1019’s progenitor masses cover $\sim 10^4 M_\odot$ at $z \sim 11 - 14$, we *cannot* rule out the supermassive star scenario that would produce such heavy seed SMBHs (see, e.g., Inayoshi et al. 2020 and the references therein).

5.2.5 AGN fraction among $6 < z < 11$ galaxies

Fig. 10 shows the fraction of galaxies hosting AGNs with bolometric luminosities above 10^{45} erg/s, f_{gal} , as a function of galaxy mass and redshift. These fractions include both obscured (including Compton-thick obscured) and non-obscured AGNs. The stochastic uncertainties in f_{gal} at $z = 6$ ($z = 11$) are shown in blue (orange) shaded regions. We also show the galaxy masses of GN-z11 and CEERS_1019 in red solid and dashed lines, respectively.

At $z \gtrsim 6$, f_{gal} is a strong function of galaxy mass, increasing from $\lesssim 10^{-6}$ at $M_* \sim 10^9 M_\odot$ to $\sim 100\%$ at $M_* \sim 3 \times 10^{10} M_\odot$. This is an extrapolation based on quasar probability distribution functions (QPDFs) from Aird et al. (2018), which requires AGN duty cycle to be low in less massive galaxies. In addition, the AGNs in Fig. 10 are selected with a fixed luminosity. The fact that lower-mass galaxies host smaller SMBHs also naturally decreases the incidence of such AGNs at the low-mass end. The high incidence of AGN in massive galaxies suggests that it is essential to properly account for any AGN contribution when trying to characterize high-redshift massive galaxies (Volonteri et al. 2017). On the flip

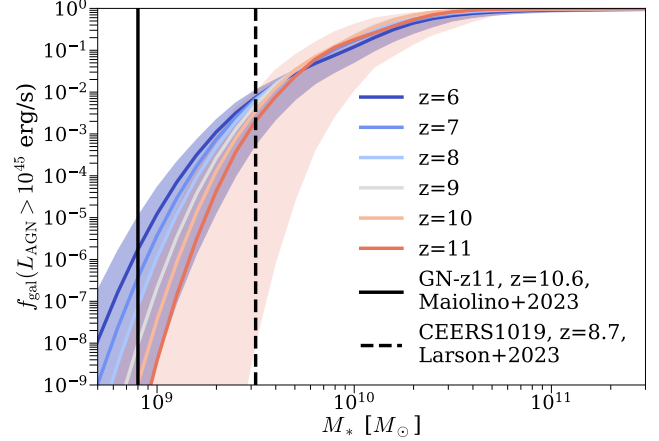


Figure 10. The fraction of galaxies that host AGNs above the bolometric luminosity of 10^{45} erg/s as a function of galaxy mass (M_*) and redshift. The solid and dashed vertical lines denote the estimated galaxy masses for Maiolino et al. (2023a) and Larson et al. (2023), respectively. For visual clarity, we only show the 68th percentile ranges (from MCMC) for the $z = 6$ blue shaded region) and the $z = 11$ (orange shaded region) cases. See §5.2.5.

side, it also means that detections of high-redshift massive galaxies that have not included AGN contributions likely need to be revisited (e.g., Labbé et al. 2023). In the low-mass regime, it is *statistically* more acceptable to ignore potential AGN contributions when characterizing galaxies. However, it is still essential to account for AGN contributions when characterizing host galaxies of *confirmed* AGNs like GN-z11 and CEERS_1019. It also implies that GN-z11 and CEERS_1019 are rare cases where low-mass galaxies do host bright AGNs, which might be in tension with observations given the survey areas of JADES and CEERS (see also Maiolino et al. 2023b for an estimate of AGN fraction). As mentioned in §5.2.2, in addition to a better understanding of galaxy mass measurements and various selection effects, non-detection over a much bigger survey area or a larger sample of low-mass galaxies with similar AGNs is needed to confirm/disprove this prediction from TRINITY, as well as provide valuable constraints on f_{gal} at high redshifts.

5.2.6 Using CEERS_1019 and GN-z11 as constraints on TRINITY

We conclude §5.2 by quantifying the constraint that CEERS_1019 and GN-z11 can impose on the current TRINITY model when added to the current data compilation. Given the uncertainties in their host galaxy masses, we chose not to compare them with the number density of AGNs hosted by similar-mass galaxies, as in Fig. 10. Instead, we converted the detection of CEERS_1019 and GN-z11 into the cumulative number density of AGNs above $L_{\text{bol}} = 10^{45}$ erg/s between $z = 8 - 12$. To calculate the comoving volume, we combined the survey areas of CEERS, GOODS-S, and GOODS-N fields, which amount to 260 arcmin². The fiducial TRINITY model predicts ~ 0.15 AGNs above $L_{\text{bol}} = 10^{45}$ erg/s between $z = 8 - 12$ in these fields. Using the two detected AGNs as a lower limit and assuming Poisson statistics, this corresponds to a $\chi^2_{z>9} \sim 9.2$. We added this additional number density constraint to TRINITY, and reran the MCMC. We found that $\chi^2_{z>9}$ is too small to significantly change the posterior distribution of model parameters, compared to the total $\chi^2_{\text{tot}} \sim 760$ from other data. In other words,

this means that the constraint imposed by the AGN number density from CEERS_1019 and GN-z11 is not strong enough to change the empirical galaxy–SMBH connection that is already constrained by all the other existing galaxy and SMBH data. This further confirms that we need either non-detections of such high-redshift AGNs over a larger volume and better host galaxy measurements to confirm TRINITY’s predictions, or more detections per unit volume to provide strong enough constraints on TRINITY.

If future observations (with selection effects corrected) confirmed that the AGN fraction is indeed much higher for $M_* \lesssim 10^9 M_\odot$ galaxies than predicted by TRINITY (Fig. 10), then TRINITY would likely have underestimated the following SMBH properties: 1) the typical SMBH masses; 2) typical AGN Eddington ratios; and/or 3) AGN duty cycles in these low-mass galaxies. Physically, this implies that high-redshift SMBHs in low-mass galaxies likely grow faster and more consistently than their low-redshift counterparts, which is not captured by extrapolating low-redshift data with the current parameterization. One possible solution to this tension is to introduce new model parameters to describe the high-redshift behavior of these SMBH properties that are not captured by the current extrapolation. We will defer this exploration until we have a sufficient number of $z \gtrsim 8$ AGNs with robust measurements of host galaxy properties.

5.3 Constraints on the progenitor masses and growth rates of $\gtrsim 10^9 M_\odot$ SMBHs at $z \sim 9$

In this section, we compare the potential progenitor masses and growth rates of $\gtrsim 10^9 M_\odot$ SMBHs from Pacucci & Loeb (2022) and TRINITY. Since the methodologies of TRINITY and the Monte Carlo simulation by Pacucci & Loeb are very different, we will only make qualitative statements on any differences, to avoid over-interpretation. Pacucci & Loeb (2022) argued that the detection of $\gtrsim 10^9 M_\odot$ SMBHs at $z \gtrsim 9$ would strongly disfavour the PopIII remnant seeding mechanism, due to the insufficient time to grow substantially within the Eddington limit. Based on their Monte Carlo simulation, they conclude that the mean seed mass, M_{seed} , for $\gtrsim 10^9 M_\odot$ SMBHs at $z \gtrsim 9$ is $\log M_{\text{seed}} = 5.34$ at $z \sim 25$. Since the TRINITY model starts at $z \sim 15$, it is impossible to extrapolate towards higher redshifts with TRINITY. We thus opt to redo the Monte Carlo simulation as described in Pacucci & Loeb (2022), and compare the $z \sim 15$ progenitor masses of $\gtrsim 10^9 M_\odot$ SMBHs at $z \gtrsim 9$ with TRINITY.

We find after this exercise that the $z \sim 15$ progenitor masses peak around $\log M_{*, z \sim 15} \sim 7$. To calculate the $z \sim 15$ progenitor masses from TRINITY, we make use of the predicted mass-independence of the Eddington ratio at high redshifts. Specifically, the $z \sim 9$ descendant of GN-z11 is predicted to have $\log M_* \sim 7.5$, according to Fig. 9. Its $z \sim 15$ progenitor has a mass of $\log M_* \sim 3$. Given that Eddington ratio is nearly mass-independent, a $z \sim 15$ SMBH must be at least $\log M_* \sim 3 + (9 - 7.5) = 4.5$ to reach $10^9 M_\odot$ by $z \sim 9$. This is roughly 2.5 dex lower than the typical $z \sim 15$ progenitor masses from Pacucci & Loeb (2022). Several factors could lead to this difference: 1) Pacucci & Loeb (2022) adopted a *flat* prior in $\log M_{\text{seed}}$, which is very different from the high-redshift BHMFs inferred by TRINITY (see Fig. 4). Consequently, the massive seeds will be given much larger weights in Pacucci & Loeb (2022) than in TRINITY; 2) In Pacucci & Loeb (2022), SMBHs are not allowed to accrete at super-Eddington rates at any time. On the contrary, SMBHs are predicted to accrete at super-Eddington rates at $z \gtrsim 10.5$ in TRINITY, even when averaged over active and dor-

mant phases. This allows smaller SMBHs to grow into $\gtrsim 10^9 M_\odot$ SMBHs by $z \gtrsim 9$ in TRINITY.

With constant or slowly changing Eddington ratios, SMBH experience exponential growth:

$$M_*(t) = M_{\text{seed}} \times \exp\left(\mathcal{D} \eta \frac{1 - \epsilon}{\epsilon} \frac{\Delta t}{t_{\text{Edd}}}\right), \quad (8)$$

where \mathcal{D} is the AGN duty cycle, i.e., the fraction of time during which SMBHs are accreting at an Eddington ratio η . ϵ is the AGN radiative efficiency, Δt is the time interval between SMBH seeding and the time of observation t , and $t_{\text{Edd}} \equiv 450$ Myr. Since it is the product of \mathcal{D} , η , and $(1 - \epsilon)/\epsilon$ that determines the growth rate, we define:

$$\lambda \equiv \mathcal{D} \eta \frac{1 - \epsilon}{\epsilon}, \quad (9)$$

so that

$$M_*(t) = M_{\text{seed}} \times \exp\left(\lambda \frac{\Delta t}{t_{\text{Edd}}}\right), \quad (10)$$

and compare the average λ values from TRINITY with the peak value of the λ distribution from the Monte Carlo simulation by Pacucci & Loeb (2022). Among $\gtrsim 10^9 M_\odot$ SMBHs by $z \sim 9$, the λ distribution from Pacucci & Loeb (2022) peaks at $\lambda \sim 8.5$. In TRINITY, the average λ decreases from ~ 60 at $z \sim 15$ to ~ 8.8 by $z \sim 9$. In other words, SMBHs in TRINITY grow significantly faster than simulated by Pacucci & Loeb (2022). Given the similar AGN efficiency and duty cycle values, this difference is mostly caused by allowing/disallowing super-Eddington accretion in TRINITY/Pacucci & Loeb (2022). As mentioned in the last paragraph, this difference in growth rate directly affects the plausible progenitor masses for $\gtrsim 10^9 M_\odot$ SMBHs by $z \sim 9$ and could give different implications on SMBH seeding mechanisms. Our comparison here demonstrates again the importance of characterizing early SMBH growth rates in determining their seeding scenarios.

5.4 Constraining power of $z \gtrsim 6$ quasar luminosity functions

As mentioned in §2.1, there is currently no $z > 6.5$ SMBH data in the observational constraints for TRINITY. Therefore, our predictions at $z \gtrsim 7$ are effectively extrapolations based on SMBH data at $z \leq 6.5$, in addition to galaxy data between $0 < z < 10$. Consequently, we see increasing statistical uncertainties towards higher redshifts in many predictions (e.g., Figs. 3 and 4). However, new instruments like JWST, *Roman Space Telescope*, and *Euclid* will provide an unprecedented amount of $z > 6.5$ SMBH observations. Even before these new instruments, large ground-based telescopes already give constraints on QLFs up to $z \sim 7$ (e.g., Matsuoka et al. 2023). These high-redshift data will be vital for understanding SMBH formation and evolution in the early Universe. In this section, we demonstrate this by quantifying the constraining power of $z > 6.5$ quasar luminosity functions (QLFs) on the halo–galaxy–SMBH connection in TRINITY.

In the top panel of Fig. 11, we show the uncertainty in M_* as a function of M_{peak} and z from the fiducial TRINITY model, where no SMBH data at $z \gtrsim 6$ were included to constrain our model. The M_* uncertainty is a quadratic sum of: 1) the statistical uncertainty from the MCMC algorithm, which reflects the constraining power of all the data points when their values and uncertainties are perfectly accurate; and 2) the typical systematic uncertainty in M_* mass measurement from virial estimates, i.e., 0.5 dex. Due to the

lack of high-redshift SMBH data, the M_\bullet uncertainty increases significantly towards higher redshifts at $z \gtrsim 6$, reaching ~ 1 dex at $z \sim 10$ and $\gtrsim 2$ dex at $z \gtrsim 12$.

We then added mock $z = 6 - 10$ QLFs to see whether/how much these data can help constrain the halo–galaxy–SMBH connection at $z \gtrsim 6$. These QLFs are extrapolated based on the modeling (global fit B) of $z \lesssim 6$ QLFs by Shen et al. (2020). These extrapolated QLFs are based on obscuration-corrected lower-redshift QLFs, so in principle include both obscured and unobscured AGNs. We determine the upper and lower luminosity limits of the QLFs assuming the characteristics of Roman High Latitude Wide Area Survey (HLS, $\sim 2000 \text{ deg}^2$); and 2) the Euclid Wide Survey (EWS, $\sim 14500 \text{ deg}^2$). To calculate the minimum luminosities of the mock QLFs, we convolved the quasar SED from Temple et al. (2021) with the filter transmission curves, and compare the fluxes with the near infrared magnitude limits of HLS (26.5 mag, Wang et al. 2022) and EWS (24.5 mag, Euclid Collaboration et al. 2022). This yields minimum bolometric quasar luminosities of $10^{44.7}$ and $10^{45.5}$ erg/s for HLS and EWS, respectively. We truncate the bright end of mock QLFs at the luminosities where only one quasar will be detected in the redshift–luminosity bins. This maximum luminosity value is further capped at 10^{48} erg/s, which is the Eddington luminosity for a $10^{10} M_\odot$ SMBH. For HLS (EWS), the maximum luminosity is 10^{48} (10^{48}) erg/s at $z = 6$, and $10^{46.5}$ ($10^{47.5}$) erg/s at $z = 10$.

As shown by the bottom panel of Fig. 11, the $z = 6 - 10$ QLFs from HLS will greatly reduce the M_\bullet uncertainty in the early Universe. In particular, the total M_\bullet uncertainties are now dominated by the virial estimate uncertainties in M_\bullet for haloes with $M_{\text{peak}} \gtrsim 10^{11} M_\odot$ at $z \lesssim 10$. This means that $z = 6 - 10$ QLFs strongly constrain SMBH growth history between $z = 6 - 10$, when they are combined with (the high-redshift extrapolation of) all the other existing data included by TRINITY. $z = 6 - 10$ SMBH growth histories are so strongly constrained that we can precisely calculate $z = 6 - 10$ SMBH masses by rewinding the growth histories of their $z \lesssim 6$ descendants. As such, our knowledge about $z = 6 - 10$ SMBH masses only depends on the accuracy of their $z \lesssim 6$ descendant SMBH masses. We do not show the results with EWS, because there is only very slight improvement from HLS to EWS. This improvement comes from the better sky coverage of Euclid, which provides better statistics for bright quasars.

The inclusion of high-redshift QLFs also does not significantly improve the constraints on early SMBH growth histories. In Fig. 12, we show the average SMBH mass growth curves for the same objects as in Fig. 9, but now with simulated $z = 6 - 10$ QLFs as additional constraints. The uncertainties in both progenitor and descendant masses of these objects only decrease slightly. This is because the redshift dependence of AGN Eddington ratio distributions is already well constrained (see §2.1), and the addition of QLFs ends up mostly constraining SMBH mass. Nonetheless, we do caution that the high-redshift AGN Eddington ratios in TRINITY are effectively extrapolated from low-redshift data. The shaded regions shown in Figs. 9 and 12 thus only represent the statistical uncertainties from the MCMC process, but do not include the systematic uncertainties from our input assumptions for such extrapolations, e.g., the exact parameterization of Eddington ratio distribution shape as a function of redshift.

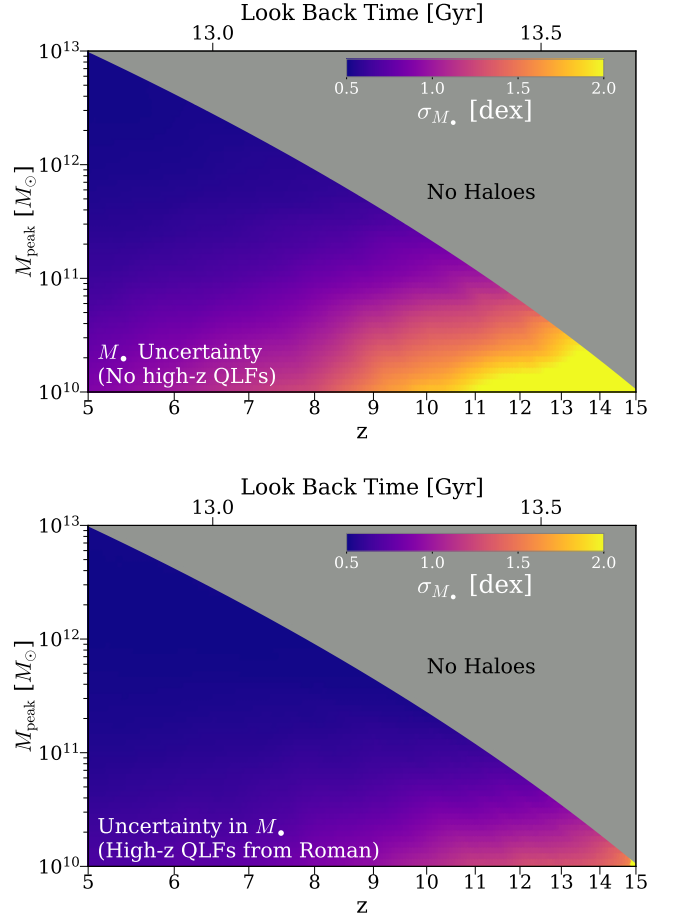


Figure 11. Top Panel: The uncertainty in black hole mass (M_\bullet), as a function of M_{peak} and z , with no quasar luminosity function (QLF) constraints between $z = 6 - 10$. **Bottom Panel:** same as the top panel, but with $z = 6 - 10$ mock QLF constraints from Roman High Latitude Wide Area Survey ($\sim 2000 \text{ deg}^2$). The uncertainties are the quadratic sums of the those from MCMC process and the typical M_\bullet uncertainty from virial estimates, which is 0.5 dex.

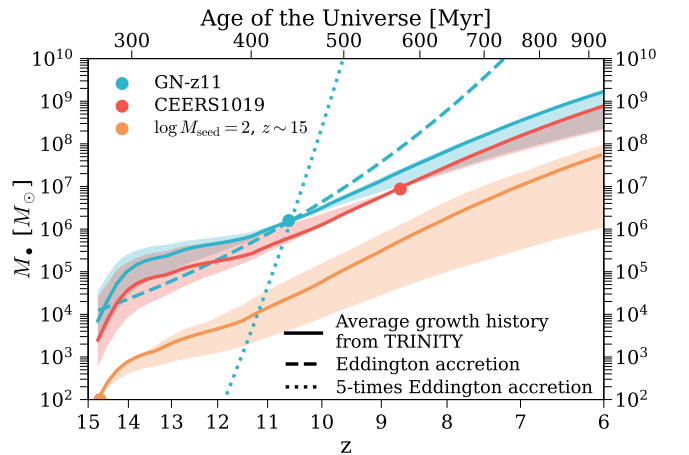


Figure 12. Similar average SMBH mass growth curves to Fig. 9, but with extra constraints from simulated $z = 6 - 10$ QLFs from Roman High Latitude Wide Area Survey. See §5.4.

6 CONCLUSIONS

In this work, we updated the TRINITY model of the halo–galaxy–SMBH connection, by adding the latest $9 \lesssim z \lesssim 13$ galaxy UV luminosity functions from JWST to our galaxy+SMBH data compilation. This additional constraint does not qualitatively change results presented in previous TRINITY papers (Zhang et al. 2023c,a,b), but does increase the number of galaxies and SMBHs at fixed masses beyond $z \sim 9$. With the updated TRINITY model, we predict the $z > 6$ M_\bullet – M_* relation for AGNs, the number density of massive ($M_\bullet > 10^9 M_\odot$) SMBHs, and the host halo mass of massive SMBHs at $z \sim 6$. **Key results are as follows:**

- From $7 \lesssim z \lesssim 10$, the median M_\bullet – M_* relation for AGNs ($\log L_{\text{bol}} \geq 45$) is biased higher than the intrinsic relation for *all* SMBHs, due to Lauer bias. The M_\bullet – M_* relation for quasars is also roughly redshift-independent. $z \gtrsim 7$ AGNs detected by JWST are broadly consistent with our predicted M_\bullet – M_* relation given their AGN luminosities and mass measurement uncertainties. (§4.1, Fig. 2).

- The cumulative numbers of very massive ($M_\bullet > 10^9, 10^{10} M_\odot$) black holes above a certain redshift over the whole sky increase by six orders of magnitude from $z \sim 10$ to $z \gtrsim 2$, due to fast SMBH growth. Below $z \sim 2$, the SMBHs numbers increase slowly, both because SMBH growth slows down and the comoving observable volume is less. Including observational scatter in measured M_\bullet inflates the numbers of very massive black holes by a factor of ~ 3 – 40 due to Eddington bias. This overestimation is stronger for more massive black holes at lower redshifts, where the SMBH mass functions are steeper. (§4.2, Figs. 3–4).

- $z \sim 6$ bright quasars with $M_\bullet > 10^9/10^{10} M_\odot$ are likely living in haloes with $M_{\text{peak}} \sim (2-3)/(3-5) \times 10^{12} M_\odot$, which are ~ 4.5 – 5.5σ peaks in Gaussian random fields. The $z = 0$ relics of these quasars are typically hosted by haloes with $M_{\text{peak}} \gtrsim 3 \times 10^{14} M_\odot$. However, these host haloes only make up $\lesssim 1\%$ of all the haloes at this mass scale, making it very difficult to find the relic SMBHs at $z \sim 0$. We expect to find only $\lesssim 1$ such relic SMBH within $z \lesssim 0.05$ (§4.3, Fig. 5).

We also compared our predictions at $z \gtrsim 9$ with the two highest-redshift AGNs detected by JWST, GN-z11 ($z = 10.6$) and CEERS_1019 ($z = 8.7$). Our main findings are:

- The SMBH masses and Eddington ratios of GN-z11 and CEERS_1019 are consistent with TRINITY’s prediction that Eddington to super-Eddington AGNs with $6 \lesssim \log M_\bullet \lesssim 7$ dominate the quasar luminosity function at $\log L_{\text{bol}} \sim 45$ and $z \sim 9$ – 10 (§5.2.1, Fig. 6).

- The host halo masses are estimated to be $\sim (2-3) \times 10^{11} M_\odot$ and $\sim (3-6) \times 10^{11} M_\odot$ for GN-z11 and CEERS_1019, respectively. These halo masses correspond to $\gtrsim 7\sigma$ peaks in the density field at those redshifts, which are extremely biased regions of the Universe. The $z = 0$ descendants of GN-z11 and CEERS_1019 are likely living in $(1-3) \times 10^{14} M_\odot$ haloes, but they are also very rare, similar to the local relics of $z = 6$ quasars (§5.2.3, Fig. 8).

- On average, SMBHs grow at above(below) the Eddington rate above(below) $z \sim 13$. The predicted average growth histories make GN-z11 a potential progenitor of $z \sim 6$ quasars with $\log M_\bullet \gtrsim 9$, which is less likely to be the case for CEERS_1019. Runaway mergers in nuclear star clusters represent a viable SMBH seeding mechanism for GN-z11 and CEERS_1019 (§5.2.4, Fig. 9).

- The fraction of galaxies hosting AGNs with bolometric luminosities above 10^{45} erg/s, f_{gal} is a strong function of galaxy mass.

At $z \gtrsim 6$, TRINITY predicts that f_{gal} increases from $\lesssim 10^{-6}$ at $M_* \sim 10^9 M_\odot$ to $\sim 100\%$ at $M_* \sim 3 \times 10^{10} M_\odot$ (§5.2.5, Fig. 10).

- The cumulative $z = 8$ – 12 quasar luminosity function calculated from CEERS_1019 and GN-z11 is not a strong enough constraint on the current TRINITY model to cause significant shifts in the galaxy–SMBH connection. To impose a strong enough constraint on TRINITY, more detections of AGNs above $L_{\text{bol}} = 10^{45}$ erg/s and/or better measurements of host galaxy mass would be needed. Alternatively, non-detection of such AGNs over a larger area would be necessary to confirm the prediction of AGN number densities by TRINITY.

Finally, we showed that future $6 < z < 10$ AGN data will provide significant constraints on the early halo–galaxy–SMBH mass connection. Specifically, $z = 6$ – 10 quasar luminosity functions (QLFs) measured by future wide area surveys (e.g., the Roman High Latitude Wide Area Survey and the Euclid Wide Survey) will reduce SMBH mass uncertainty from ~ 1.0 dex to ~ 0.5 dex for haloes with $M_{\text{peak}} \gtrsim 10^{11} M_\odot$ at $z \lesssim 10$. (§5.4, Fig. 11).

DATA AVAILABILITY

The parallel implementation of TRINITY, the compiled datasets, and the posterior distribution of model parameters are available [online](#).

ACKNOWLEDGEMENTS

We thank Gurtina Besla, Frank van den Bosch, Haley Bowden, Alison Coil, Ryan Endsley, Sandy Faber, Nickolay Gnedin, Richard Green, Jenny Greene, Kate Grier, Melanie Habouzit, Kevin Hainline, Andrew Hearin, Yun-Hsin Huang, Raphael Hviding, Stephanie Juneau, David Koo, Andrey Kravtsov, Daniel Lawther, Jianwei Lyu, Chung-Pei Ma, Vasileios Paschalidis, Joel Primack, Eliot Quataert, George Rieke, Marcia Rieke, Jan-Torge Schindler, Xuejian Shen, Yue Shen, Dongdong Shi, Rachel Somerville, Fengwu Sun, Wei-Leong Tee, Yoshihiro Ueda, Marianne Vestergaard, Ben Weiner, Christina Williams, Charity Woodrum, and Minghao Yue for very valuable discussions.

Support for this research came partially via program number HST-AR-15631.001-A, provided through a grant from the Space Telescope Science Institute under NASA contract NAS5-26555. PB was partially funded by a Packard Fellowship, Grant #2019-69646. PB was also partially supported by a Giacconi Fellowship from the Space Telescope Science Institute. Finally, PB was also partially supported through program number HST-HF2-51353.001-A, provided by NASA through a Hubble Fellowship grant from the Space Telescope Science Institute, under NASA contract NAS5-26555.

Data compilations from many studies used in this paper were made much more accurate and efficient by the online WEBPLOT-DIGITIZER code.² This research has made extensive use of the arXiv and NASA’s Astrophysics Data System.

This research used the Ocelote supercomputer of the University of Arizona. The allocation of computer time from the UA Research Computing High Performance Computing (HPC) at the University of Arizona is gratefully acknowledged. The Bolshoi-Planck simulation was performed by Anatoly Klypin within the Bolshoi

² <https://apps.automeris.io/wpd/>

project of the University of California High-Performance Astro-Computing Center (UC-HiPACC; PI Joel Primack).

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE, www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de).

This paper has been typeset from a \LaTeX file prepared by the author.

REFERENCES

- Agarwal B., Davis A. J., Khochfar S., Natarajan P., Dunlop J. S., 2013, *MNRAS*, **432**, 3438
- Aird J., et al., 2010, *MNRAS*, **401**, 2531
- Aird J., Coil A. L., Georgakakis A., 2018, *MNRAS*, **474**, 1225
- Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, **56**, 93
- Angulo R. E., Springel V., White S. D. M., Cole S., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, **425**, 2722
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, **770**, 57
- Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, *MNRAS*, **488**, 3143
- Bellovary J. M., Cleary C. E., Munshi F., Tremmel M., Christensen C. R., Brooks A., Quinn T. R., 2019, *MNRAS*, **482**, 2913
- Blandford R. D., McKee C. F., 1982, *ApJ*, **255**, 419
- Bogdan A., et al., 2023, *arXiv e-prints*, p. [arXiv:2305.15458](https://arxiv.org/abs/2305.15458)
- Bouwens R. J., Stefanon M., Oesch P. A., Illingworth G. D., Nanayakkara T., Roberts-Borsani G., Labbé I., Smit R., 2019, *ApJ*, **880**, 25
- Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, *MNRAS*, **370**, 645
- Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, **465**, 32
- Brandt W. N., Alexander D. M., 2015, *A&ARv*, **23**, 1
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Bryan G. L., Norman M. L., 1998, *ApJ*, **495**, 80
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, **533**, 682
- Chabrier G., 2003, *PASP*, **115**, 763
- Chiang Y.-K., Overzier R., Gebhardt K., 2013, *ApJ*, **779**, 127
- Choksi N., Behroozi P., Volonteri M., Schneider R., Ma C.-P., Silk J., Moster B., 2017, *MNRAS*, **472**, 1526
- Conroy C., White M., 2013, *ApJ*, **762**, 70
- Dekel A., Lapiner S., Dubois Y., 2019, *arXiv e-prints*, p. [arXiv:1904.08431](https://arxiv.org/abs/1904.08431)
- Delvecchio I., et al., 2014, *MNRAS*, **439**, 2736
- Devecchi B., Volonteri M., Rossi E. M., Colpi M., Portegies Zwart S., 2012, *MNRAS*, **421**, 1465
- Diemer B., 2018, *ApJS*, **239**, 35
- Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, *MNRAS*, **420**, 2662
- Dubois Y., Volonteri M., Silk J., Devriendt J., Slyz A., Teyssier R., 2015, *MNRAS*, **452**, 1502
- Eddington A. S., 1913, *MNRAS*, **73**, 359
- Eisenstein D. J., Hu W., 1998, *ApJ*, **496**, 605
- Euclid Collaboration et al., 2022, *A&A*, **662**, A112
- Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, *MNRAS*, **436**, 315
- Feng Y., Di Matteo T., Croft R., Khandai N., 2014, *MNRAS*, **440**, 1865
- Ferrarese L., Ford H., 2005, *Space Sci. Rev.*, **116**, 523
- Ferrarese L., Merritt D., 2000, *ApJL*, **539**, L9
- Finkelstein S. L., et al., 2019, *ApJ*, **879**, 36
- Gebhardt K., et al., 2000, *ApJL*, **539**, L13
- Goulding A. D., et al., 2023, *arXiv e-prints*, p. [arXiv:2308.02750](https://arxiv.org/abs/2308.02750)
- Gültekin K., et al., 2009, *ApJ*, **698**, 198
- Haario H., Saksman E., Tamminen J., 2001, *Bernoulli*, **7**, 223
- Habouzit M., Volonteri M., Dubois Y., 2017, *MNRAS*, **468**, 3935
- Harikane Y., et al., 2023a, *arXiv e-prints*, p. [arXiv:2303.11946](https://arxiv.org/abs/2303.11946)
- Harikane Y., Nakajima K., Ouchi M., Umeda H., Isobe Y., Ono Y., Xu Y., Zhang Y., 2023b, *arXiv e-prints*, p. [arXiv:2304.06658](https://arxiv.org/abs/2304.06658)
- Häring N., Rix H.-W., 2004, *ApJL*, **604**, L89
- Heckman T. M., Best P. N., 2014, *ARAA*, **52**, 589
- Ho L. C., 2008, *ARAA*, **46**, 475
- Hopkins P. F., Bundy K., Hernquist L., Ellis R. S., 2007, *ApJ*, **659**, 976
- Inayoshi K., Visbal E., Haiman Z., 2020, *ARAA*, **58**, 27
- Ishigaki M., Kawamata R., Ouchi M., Oguri M., Shimasaku K., Ono Y., 2018, *ApJ*, **854**, 73
- Ivezić Ž., et al., 2019, *ApJ*, **873**, 111
- Izumi T., et al., 2021, *ApJ*, **914**, 36
- Kashino D., Lilly S. J., Matthee J., Eilers A.-C., Mackenzie R., Bordoloi R., Simcoe R. A., 2023, *ApJ*, **950**, 66
- Kelly B. C., Shen Y., 2013, *ApJ*, **764**, 45
- Klypin A. A., Trujillo-Gomez S., Primack J., 2011, *ApJ*, **740**, 102
- Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, *MNRAS*, **457**, 4340
- Kormendy J., Ho L. C., 2013, *ARAA*, **51**, 511
- Kormendy J., Richstone D., 1995, *ARAA*, **33**, 581
- Kulkarni G., Worbeck G., Hennawi J. F., 2019, *MNRAS*, **488**, 1035
- Labbé I., et al., 2023, *Nature*, **616**, 266
- Lang P., et al., 2014, *ApJ*, **788**, 11
- Larson R. L., et al., 2023, *arXiv e-prints*, p. [arXiv:2303.08918](https://arxiv.org/abs/2303.08918)
- Latif M. A., Whalen D. J., Khochfar S., Herrington N. P., Woods T. E., 2022, *Nature*, **607**, 48
- Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, *ApJ*, **670**, 249
- Li Y., et al., 2007, *ApJ*, **665**, 187
- Lupi A., Colpi M., Devecchi B., Galanti G., Volonteri M., 2014, *MNRAS*, **442**, 3616
- Magorrian J., et al., 1998, *AJ*, **115**, 2285
- Maiolino R., et al., 2023a, *arXiv e-prints*, p. [arXiv:2305.12492](https://arxiv.org/abs/2305.12492)
- Maiolino R., et al., 2023b, *arXiv e-prints*, p. [arXiv:2308.01230](https://arxiv.org/abs/2308.01230)
- Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, *MNRAS*, **351**, 169
- Matsuoka Y., et al., 2023, *ApJL*, **949**, L42
- McConnell N. J., Ma C.-P., 2013, *ApJ*, **764**, 184
- McLure R. J., Dunlop J. S., 2004, *MNRAS*, **352**, 1390
- McLure R. J., Jarvis M. J., 2002, *MNRAS*, **337**, 109
- Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, *ApJS*, **210**, 3
- Merloni A., 2004, *MNRAS*, **353**, 1035
- Merloni A., et al., 2014, *MNRAS*, **437**, 3550
- Natarajan P., Pacucci F., Ferrara A., Agarwal B., Ricarte A., Zackrisson E., Cappelluti N., 2017, *ApJ*, **838**, 117
- Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, *ApJ*, **855**, 105
- Pacucci F., Loeb A., 2022, *MNRAS*, **509**, 1885
- Peterson B. M., 1993, *PASP*, **105**, 247
- Pillepich A., et al., 2018, *MNRAS*, **475**, 648
- Planck Collaboration et al., 2016, *A&A*, **594**, A13
- Richards G. T., et al., 2006, *ApJS*, **166**, 470
- Savorgnan G. A. D., Graham A. W., Marconi A. r., Sani E., 2016, *ApJ*, **817**, 21
- Schleicher D. R. G., et al., 2022, *MNRAS*, **512**, 6192
- Schneider R., Valiante R., Trinca A., Graziani L., Volonteri M., 2023, *arXiv e-prints*, p. [arXiv:2305.12504](https://arxiv.org/abs/2305.12504)
- Scholtz J., et al., 2023, *arXiv e-prints*, p. [arXiv:2306.09142](https://arxiv.org/abs/2306.09142)
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, *ApJ*, **690**, 20
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2013, *MNRAS*, **428**, 421
- Shankar F., et al., 2020, *MNRAS*, **493**, 1500
- Shen Y., et al., 2019, *ApJ*, **873**, 35
- Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020, *MNRAS*, **495**, 3252
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, **452**, 575
- Silk J., Rees M. J., 1998, *A&A*, **331**, L1
- Silverman J. D., et al., 2008, *ApJ*, **679**, 118
- Sothman A., 1982, *MNRAS*, **200**, 115

- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, **391**, 481
- Springel V., et al., 2005, *Nature*, **435**, 629
- Tacchella S., et al., 2023, *arXiv e-prints*, p. [arXiv:2302.07234](#)
- Temple M. J., Hewett P. C., Banerji M., 2021, *MNRAS*, **508**, 737
- Tenneti A., Di Matteo T., Croft R., Garcia T., Feng Y., 2018, *MNRAS*, **474**, 597
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, **688**, 709
- Tremaine S., et al., 2002, *ApJ*, **574**, 740
- Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, *MNRAS*, **470**, 1121
- Trinca A., Schneider R., Valiante R., Graziani L., Zappacosta L., Shankar F., 2022, *MNRAS*, **511**, 616
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *ApJ*, **786**, 104
- Vanden Berk D. E., et al., 2001, *AJ*, **122**, 549
- Vestergaard M., Osmer P. S., 2009, *ApJ*, **699**, 800
- Vestergaard M., Peterson B. M., 2006, *ApJ*, **641**, 689
- Volonteri M., Stark D. P., 2011, *MNRAS*, **417**, 2085
- Volonteri M., Reines A. E., Atek H., Stark D. P., Trebitsch M., 2017, *ApJ*, **849**, 155
- Volonteri M., Habouzit M., Colpi M., 2023, *MNRAS*, **521**, 241
- Wang F., et al., 2019, *ApJ*, **884**, 30
- Wang F., et al., 2021, *ApJL*, **907**, L1
- Wang Y., et al., 2022, *ApJ*, **928**, 1
- Wang F., et al., 2023, *ApJL*, **951**, L4
- Yang G., et al., 2018, *MNRAS*, **475**, 1887
- Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Hopkins P. F., Yang J., Aird J., 2023a, *arXiv e-prints*, p. [arXiv:2303.08150](#)
- Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Aird J., Yang J., Hopkins P. F., 2023b, *arXiv e-prints*, p. [arXiv:2305.19315](#)
- Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Hopkins P. F., Yang J., Aird J., 2023c, *MNRAS*, **518**, 2123