








The LUMINA Project: The Demographics of Active Galactic Nuclei from Quasars to Little Red Dots at $z \geq 3$

Xuejian Shen ^{1,*} Oliver Zier ² Aaron Smith ³ Rongrong Liu ² Rahul Kannan ⁴ Teodora-Elena Bulichi ¹ Sonja M. Koehler ² Volker Springel ⁵ Mark Vogelsberger ¹ Lars Hernquist ² Rohan P. Naidu ^{1,†} Anna de Graaff ^{2,6,‡} Elia Pizzati ^{2,§} David M. Alexander ⁷ Luis C. Ho ^{8,9} Vasily Kokorev ^{10,11} Gene Leung ¹ Anna-Christina Eilers ¹ and Ryan C. Hickox ¹²

¹ Department of Physics, Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

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

³ Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA

⁴ Department of Physics and Astronomy, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

⁵ Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany

⁶ Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

⁷ Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK

⁸ Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

⁹ Department of Astronomy, School of Physics, Peking University, Beijing 100871, China

¹⁰ Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

¹¹ Cosmic Frontier Center, The University of Texas at Austin, Austin, TX 78712, USA

¹² Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, NH 03755, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

High-redshift active galactic nuclei (AGN) serve as powerful probes of early black-hole growth, galaxy formation, and the evolving intergalactic medium (IGM). In this work, we use LUMINA, a cosmological radiation-hydrodynamic simulation spanning the epochs of hydrogen and helium reionization, which combines a large (500 cMpc)³ volume with 2×6000^3 resolution elements, to explore high-redshift AGN. The simulation self-consistently follows hundreds of millions of galaxies and supermassive black holes (SMBHs), together with their impact on the ionization and thermal state of the IGM. We exploit this uniquely large dynamic range to predict multi-band AGN luminosity functions (LFs) at $z \geq 3$, from hard X-rays to the mid-infrared. These predictions encompass both moderately luminous quasars and the faint “Little Red Dots” (LRDs) uncovered by *JWST*. We develop an empirical model that maps simulated SMBHs onto observed AGN using bolometric and extinction/absorption corrections for canonical AGN and LRDs, and in which SMBHs with $M_{\text{BH}} \leq 10 M_{\text{seed}} \sim 10^7 M_{\odot}$ stay in the LRD phase with a duty cycle of 30%. This simple framework reproduces the observed LFs and clustering of LRDs. Meanwhile, the pre-*JWST* quasar LF constraints are recovered, although we find that a ~ 0.3 dex log-normal scatter in bolometric luminosity is required to reproduce the bright end. We place the simulated AGN population in the cosmological context by quantifying the redshift evolution of AGN and LRD number densities, and their contributions to the integrated BH mass densities. The same AGN population is the dominant driver for the He II reionization modelled self-consistently in LUMINA. This empirical AGN model paves the way for general population-synthesis models of high-redshift AGN, including LRDs, in a unified cosmological framework.

Key words: quasars: supermassive black holes – galaxies: high-redshift – galaxies: active – methods: numerical

1 INTRODUCTION

Active galactic nuclei (AGN), and luminous quasars in particular, are the observable manifestation of gas accretion onto supermassive black holes (SMBHs) at the centres of most galaxies. The accreted gas forms a hot, dissipative disc (e.g., Shakura & Sunyaev 1973; Rees

1984), whose high radiative efficiency makes such systems among the most luminous persistent sources in the Universe and allows them to be detected out to $z \gtrsim 7$ (e.g., Mortlock et al. 2011; Venemans et al. 2015; Bañados et al. 2018) and now to $z \gtrsim 10$ with *JWST* (Bogdán et al. 2024; Napolitano et al. 2025; Maiolino et al. 2025). The cosmic demography of AGN therefore directly traces the assembly of the SMBH population, as most of the local SMBH mass density was built up during the radiatively efficient accretion phases that we observe as AGN (Soltan 1982). AGN are also among the most important sources of high-energy radiation in the Universe, contributing significantly to the cosmic infrared (IR) and X-ray backgrounds. Their hard extreme

* E-mail: xuejian@mit.edu

† NASA Hubble Fellow

‡ Clay Fellow

§ NASA Einstein Fellow

ultraviolet (EUV) emission is believed to drive the reionization of singly ionized helium (He II; e.g., Miralda-Escudé 1998; Madau et al. 1999; Faucher-Giguère et al. 2008b, 2009; McQuinn et al. 2009; Haardt & Madau 2012; Worseck et al. 2014) and may also contribute to the reionization of hydrogen (e.g., Madau & Haardt 2015; Madau et al. 2024; Asthana et al. 2025), although star-forming galaxies are believed to dominate the hydrogen-ionizing photon budget in most current models (e.g., Faucher-Giguère et al. 2008a; Kuhlen & Faucher-Giguère 2012; Robertson et al. 2015; Haardt & Salvaterra 2015; Giallongo et al. 2015; Onoue et al. 2017; Parsa et al. 2018; Shen et al. 2020a; Kannan et al. 2022; Yeh et al. 2023).

Beyond their role as energy sources, AGN are tightly coupled to the evolution of their host galaxies. The masses of SMBHs are correlated with the masses, luminosities, and velocity dispersions of their host bulges (e.g., Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Gültekin et al. 2009). AGN feedback is widely invoked to regulate or quench star-formation in massive galaxies (e.g., Sanders & Mirabel 1996; Springel et al. 2005; Di Matteo et al. 2005; Bower et al. 2006; Croton et al. 2006; Hopkins et al. 2006; Sijacki et al. 2007; Somerville et al. 2008; Hopkins et al. 2008; Feruglio et al. 2010; Fabian 2012; Cicone et al. 2014) and to resolve the classical “cooling-flow” problem (e.g., Cowie & Binney 1977; Fabian & Nulsen 1977; Fabian et al. 1984; Tabor & Binney 1993; Fabian 1994; Croton et al. 2006). Recent reviews summarize the breadth of these connections (Alexander & Hickox 2012; Fabian 2012; Kormendy & Ho 2013; Heckman & Best 2014; Alexander et al. 2025, and references therein).

Studying the cosmic evolution of the AGN population is therefore central to both cosmology and galaxy formation. The measurement of AGN abundance, i.e. luminosity functions (LFs), has a decades-long history in the rest-frame optical/UV (e.g., Schmidt 1968; Boyle et al. 1988; Warren et al. 1994; Fan et al. 2001, 2004; Richards et al. 2006a; Glikman et al. 2011; Ross et al. 2013; McGreer et al. 2013; Kashikawa et al. 2015; Jiang et al. 2016; Niida et al. 2020), soft X-ray (e.g., Maccacaro et al. 1991; Boyle et al. 1993; Jones et al. 1997; Page et al. 1997; Miyaji et al. 2000; Hasinger et al. 2005), hard X-ray (e.g., Ueda et al. 2003; La Franca et al. 2005; Barger et al. 2005; Silverman et al. 2008; Ebrero et al. 2009; Yencho et al. 2009; Aird et al. 2010; Ueda et al. 2014; Aird et al. 2015; Ananna et al. 2019), and IR (e.g., Brown et al. 2006; Matute et al. 2006; Assef et al. 2011; Lacy et al. 2015; Bulichi et al. 2026; Ling et al. 2026) bands. These measurements have revealed strong redshift evolution of AGN LFs (both in normalization and in shape; e.g., Cowie et al. 1996; Barger et al. 2005; Hasinger et al. 2005; Aird et al. 2015; Kulkarni et al. 2019), typical AGN spectral shape (e.g., Wilkes et al. 1994; Green et al. 1995; Vignali et al. 2003; Strateva et al. 2005; Richards et al. 2006b; Steffen et al. 2006; Just et al. 2007; Lusso et al. 2010; Kashikawa et al. 2015; Lusso & Risaliti 2016), and the obscuring column-density distribution (e.g., Hill et al. 1996; Simpson et al. 1999; Willott et al. 2000; Steffen et al. 2003; Ueda et al. 2003; Grimes et al. 2004; Sazonov & Revnivtsev 2004; Barger et al. 2005; Hao et al. 2005; Ueda et al. 2014; Ananna et al. 2019). Syntheses of these multi-band measurements provided a comprehensive pre-*JWST* view of the AGN population from $z \approx 0$ to $z \approx 7$ (Hopkins et al. 2007; Shen et al. 2020a).

Despite this rich, established picture, the rapid decline of the AGN number density beyond $z \approx 6$ made it particularly difficult to constrain AGN LFs at the highest redshifts, which correspond to early SMBH growth phases. Pre-*JWST* wide-field optical and near-IR surveys identified $\gtrsim 500$ luminous quasars at $z \gtrsim 5.3$ (e.g., Willott et al. 2010; Venemans et al. 2013, 2015; Jiang et al. 2016; Reed et al. 2017; Mazzucchelli et al. 2017; Matsuoka et al. 2018; Ross & Cross 2020;

Fan et al. 2023) and pushed the redshift frontier to $z \gtrsim 7$ (e.g., Mortlock et al. 2011; Bañados et al. 2018; Yang et al. 2020; Wang et al. 2021; Fujimoto et al. 2022). These observations provided key probes of early SMBH growth and of the hydrogen reionization history through their absorption spectra (e.g., Miralda-Escudé 1998; Madau & Rees 2000; Fan et al. 2002, 2006). However, these samples were dominated by the rare, hyperluminous tip of the AGN population, and the demographics of moderate- and low-luminosity AGN at high redshift, which likely dominate the integrated SMBH mass-growth budget, remained essentially unconstrained.

JWST has dramatically transformed the landscape of AGN studies at high redshifts. Deep *JWST* NIRSpec and NIRCам Grism surveys have already uncovered hundreds of broad-line AGN (BLAGN) at $4 \lesssim z \lesssim 11$ (e.g., Harikane et al. 2023b; Maiolino et al. 2024; Greene et al. 2024; Matthee et al. 2024; Kokorev et al. 2024; Akins et al. 2025b; Kocevski et al. 2025), with spectroscopically identified candidates extending to $z > 10$ (Bogdán et al. 2024; Napolitano et al. 2025; Maiolino et al. 2025). Meanwhile, deep *JWST* MIRI and X-ray surveys are beginning to probe the obscured high-redshift AGN population (e.g., Bulichi et al. 2026; Hviding et al. 2026; Ling et al. 2026). Faint AGN appear to be far more abundant in the early Universe than predicted by extrapolations of pre-*JWST* quasar LFs, with potentially important implications for the high-redshift ionizing photon budget and build-up of the cosmic SMBH mass density (Madau et al. 2024; Asthana et al. 2025; Inayoshi & Ho 2025).

Among the most striking *JWST* discoveries is the population of “Little Red Dots” (LRDs), compact sources with characteristic “V-shaped” spectral energy distributions (SEDs) that feature a blue UV continuum, a steep red optical continuum, and a turnover near the Balmer limit (e.g., Matthee et al. 2024; Kokorev et al. 2024; Akins et al. 2025b; Kocevski et al. 2025; Setton et al. 2025b). Spectroscopic follow-up has revealed broad Balmer lines in the majority of the observed LRDs (e.g., Greene et al. 2024; Matthee et al. 2024; Maiolino et al. 2024; Hviding et al. 2025), supporting an AGN origin for at least the dominant red continuum component. Yet LRDs differ in essentially every other respect from canonical Type-I/Type-II AGN: they are systematically X-ray weak or undetected (e.g., Yue et al. 2024; Ananna et al. 2024; Maiolino et al. 2025), faint or absent in the mid- and far-IR (e.g., Setton et al. 2025a; Casey et al. 2025), and exhibit little or no continuum variability (at least in the majority; e.g., Hayes et al. 2024; Kokubo & Harikane 2025; Tee et al. 2025; Zhang et al. 2025b; Liu et al. 2026, but see also Zhang et al. 2025a). Many show prominent Balmer breaks and Balmer absorption features (e.g., Matthee et al. 2024; Naidu et al. 2025; de Graaff et al. 2025c; Setton et al. 2025b; Ji et al. 2025) that are difficult to explain by simple dust-reddened Type-I AGN. The same sources are often inferred to host BHs that are over-massive relative to their host stellar masses (e.g., Pacucci et al. 2023; Maiolino et al. 2024; Inayoshi & Ho 2025; Durodola et al. 2025), although the interpretation depends sensitively on the assumed line-broadening mechanism (e.g., Naidu et al. 2025; Brazzini et al. 2025; Chang et al. 2026; Ji et al. 2026a; Rusakov et al. 2026) and selection effects (e.g., Li et al. 2025). Whether LRDs represent a transient, early-growth phase of more typical SMBHs, a fundamentally distinct population, or a heterogeneous mixture of both remains an open question.

Interpreting these new observational findings, however, is complicated by the fact that observations in any single band are subject to selection effects, host-galaxy contamination, and obscuration in a wavelength-dependent manner. Although AGN are intrinsically very luminous in the optical/UV, dust extinction along certain viewing angles (e.g., Antonucci 1993; Urry & Padovani 1995) can render them difficult to detect, and the optical/UV emission of heavily ob-

scured AGN is easily contaminated by stellar UV light from their host galaxies (e.g., see review of Hickox & Alexander 2018). In the X-ray, which is much less affected by dust, the search for Compton-thick (CTK) AGN, which are heavily attenuated and constitute 20%–50% of the AGN population (e.g., Burlon et al. 2011; Ricci et al. 2015; Carroll et al. 2021) and remain highly incomplete in current surveys. Mid-IR observations from the ground are limited by atmospheric absorption and contamination by warm dust in host galaxies. In the far-IR to millimetre regime ($30\mu\text{m} - 10\text{mm}$), emission from cold dust in host galaxies can dominate over the AGN, blunting AGN identification. LRDs clearly highlight this problem: if we had relied solely on X-ray or IR data, nearly all of these BHs would have gone undetected. Furthermore, LF measurements based on a single survey are limited in their luminosity coverage and volume probed and are subject to non-negligible biases and uncertainties in completeness corrections. It is therefore essential to develop a framework that links observations across multiple bands, providing a more comprehensive picture of AGN demographics over cosmic time.

Specifically for LRDs, the physical interpretation is evolving rapidly. One emerging class of models invokes accreting BHs embedded in extremely dense, optically thick gas envelopes that reprocess the emergent emission from the central engine into a “star-like” (but not identical to stars) optical spectrum (e.g., Inayoshi & Maiolino 2025; Naidu et al. 2025; de Graaff et al. 2025c; Kido et al. 2025; Rusakov et al. 2026), possibly forming a quasi-star configuration (e.g., Begelman et al. 2006, 2008; Volonteri & Begelman 2010; Begelman & Dexter 2026). A complementary scenario invokes super-Eddington accretion onto modest-mass seed BHs as the explanation for X-ray weakness, and the lack of variability in LRDs (e.g., Lambides et al. 2024; Pacucci & Narayan 2024; Madau & Haardt 2024; Lupi et al. 2024; Inayoshi et al. 2025). Whether LRDs trace heavy seeds (e.g., direct-collapse BHs), light seeds undergoing efficient early growth, or some combination of both remains a major theoretical question that connects directly to early SMBH-seeding scenarios (e.g., Bromm & Loeb 2003; Begelman et al. 2006; Lodato & Natarajan 2006; Devecchi & Volonteri 2009; Xiao et al. 2021; Shen et al. 2025a; Jiang et al. 2026). A unified cosmological framework that can simultaneously address the abundance and large-scale environments of *JWST*-selected faint AGN together with luminous quasars is therefore needed.

Theoretical efforts to interpret the high-redshift AGN population have progressed in parallel with *JWST* observations. Large-volume cosmological hydrodynamic simulations, such as Illustris (Vogelsberger et al. 2013, 2014a,b), EAGLE (Schaye et al. 2015; Crain et al. 2015), IllustrisTNG (Pillepich et al. 2018b,a; Springel et al. 2018), SIMBA (Davé et al. 2019), Horizon-AGN (Dubois et al. 2016; Volonteri et al. 2016), BlueTides (Feng et al. 2016), ASTRID (Bird et al. 2022; Ni et al. 2022), MillenniumTNG (Pakmor et al. 2023), and FLAMINGO (Schaye et al. 2023), have established a coherent framework for modelling the growth and feedback of SMBHs alongside their co-evolution with host galaxies in a cosmological context. Cross-comparison studies (e.g., Habouzit et al. 2021, 2022) highlight that the predicted high-redshift AGN LFs and BH-host scaling relations are sensitive to the choice of seeding prescription, sub-resolution recipes for BH accretion and AGN feedback, motivating more theoretical experiments. Smaller-volume zoom and more dedicated suites, such as BRAHMA (Bhowmick et al. 2024a,b), NewHorizon (Dubois et al. 2021), and OBELISK (Trebitsch et al. 2021), systematically explored seed channels, accretion regimes, and feedback prescriptions, while semi-analytic and empirical frameworks, such as CAT (Trinca et al. 2022, 2024), DELPHI (Dayal et al. 2014, 2019), TRINITY (Zhang et al. 2023), L-Galaxies (Bonoli et al.

2025), and the Santa Cruz SAM (Yung et al. 2021), efficiently explored the high-redshift SMBH model parameter space.

In this paper, we use LUMINA — a large-volume (500cMpc^3) cosmological radiation-hydrodynamic simulation with an extreme dynamical range of 2×6000^3 adaptive resolution elements that follows galaxies, AGN, and the intergalactic medium (IGM) through hydrogen and helium reionization down to $z = 3$ — to make predictions for multi-band AGN LFs at $z \geq 3$ across the electromagnetic spectrum, from hard X-rays to the mid-IR. LUMINA is designed to play a complementary role to the theoretical efforts described above. The large volume samples bright quasars together with a statistically representative, well-resolved population of galaxies and faint AGN, enabling joint predictions for AGN, host galaxies, and their distributions on large scales coherently. In addition, LUMINA is the first simulation in this regime to include on-the-fly multi-frequency radiative transfer through both hydrogen and He II reionization, allowing self-consistent modelling of the ionization and thermal state of the IGM impacted by galaxies and AGN. This combination enables the unified, multi-band view of high-redshift AGN in the cosmological context that we develop in this paper. In Figure 1, we show a multi-scale view of LUMINA at $z \approx 4$ from the large-scale dark matter (DM) distribution to a He III bubble driven by a bright AGN to the smaller-scale gas distribution and BHs, eventually to the host galaxy of the AGN. In Figure 2, we give a more complete overview of the topology of He II reionization at $z \approx 4$, featuring large $\sim 10 - 100\text{cMpc}$ size He III bubbles created by AGN along with islands of He II in voids.

Based on LUMINA, we develop an empirical model to bridge BH particles in the simulation with observed AGN, including the puzzling LRD population, and then use it to test whether the simulated SMBH population can simultaneously reproduce several important observables of high-redshift AGN, including LFs, large-scale clustering measurements, and integrated cosmological constraints. The paper is organized as follows. Section 2 summarizes the LUMINA simulation, including its galaxy formation and SMBH model, the on-the-fly radiative transfer, and the structure-finding pipeline. Section 3 describes the empirical model that maps simulated SMBHs to observed AGN. Section 4 presents the predicted multi-band AGN LFs at $z \approx 3 - 6$, the large-scale clustering of LRDs, and the role of AGN in the cosmological context. Section 5 discusses the underlying assumptions and limitations of our framework, and compares it to other theoretical models. Conclusions are presented in Section 6.

2 SIMULATION

LUMINA is a cosmological radiation-hydrodynamic simulation that evolves 2×6000^3 resolution elements in an $L_{\text{box}} = 500\text{cMpc}$ box from $z_{\text{init}} = 49$ down to $z = 3$, spanning the epochs of hydrogen and He II reionization. This combination of volume and resolution provides the dynamic range required to follow simultaneously the galaxy population that drives hydrogen reionization and the rare, bright quasars that drive He II reionization. The simulation is run with the moving-mesh code AREPO (Springel 2010; Pakmor et al. 2016; Weinberger et al. 2020), that solves the cosmological Euler equations on an unstructured moving Voronoi mesh using a finite volume Godunov method (Godunov & Bohachevsky 1959). Gravity is solved with an adaptive Tree-PM scheme (Bagla 2002; Bagla & Ray 2003). It is coupled to the GPU-accelerated radiative-transfer module AREPO-RT (Kannan et al. 2019; Zier et al. 2024). Similar to MTNG (Pakmor et al. 2023), magnetic fields are not evolved. We adopt a flat Λ CDM cosmology consistent with *Planck* 2018 (Planck Collaboration et al. 2020), which we summarize in Table 1. Initial

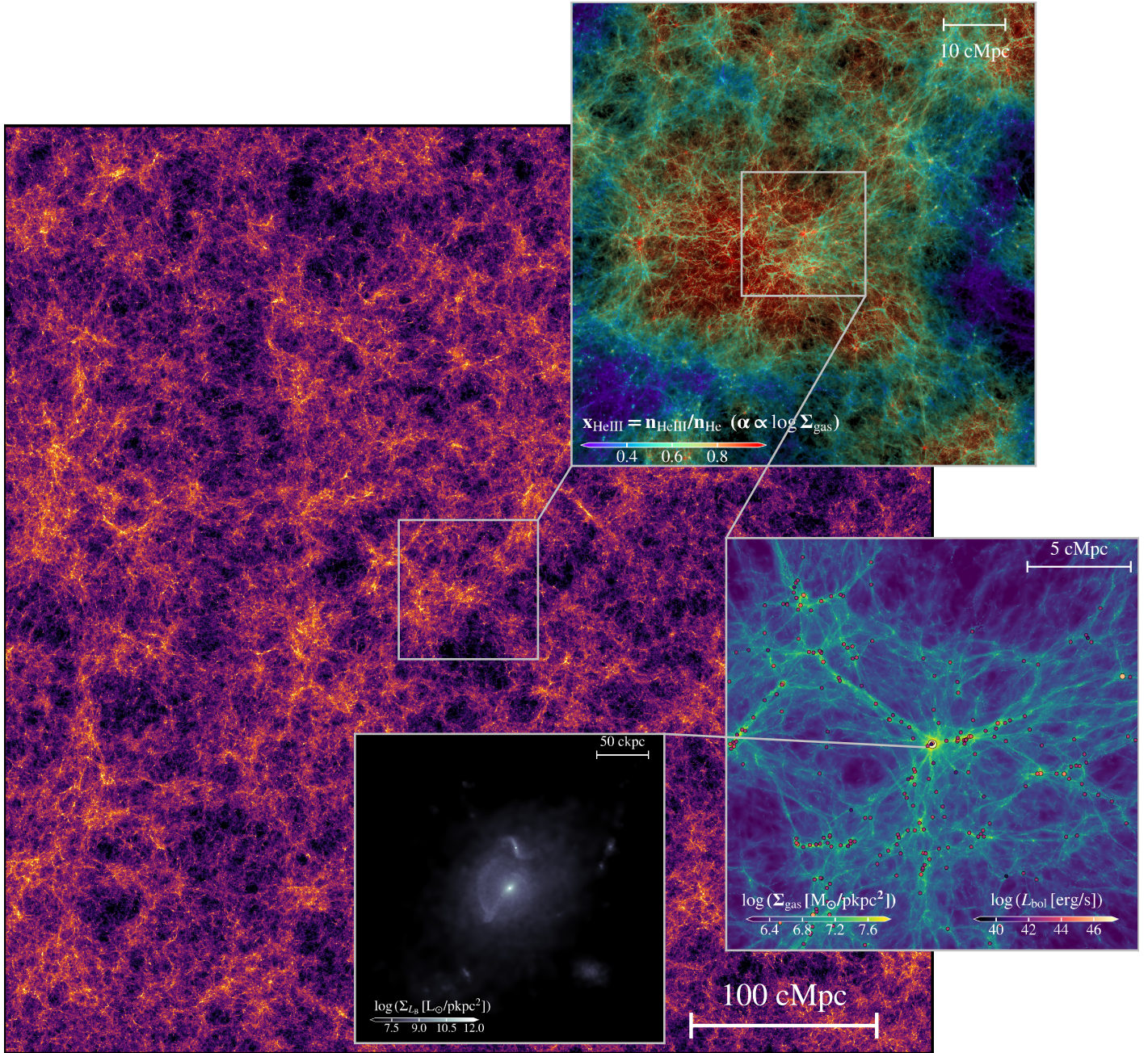


Figure 1. DM surface density field of a slice of the simulation volume centred on the most luminous AGN at $z \approx 4$. The thickness of the slice is 150 cMpc. We first zoom into a 75×75 cMpc² subfield and present the blending of the mass-weighted He III fraction and the gas surface density field. This illustrates the He III bubble driven by the hard ionizing radiation from the central AGN. We further zoom in on the 25×25 cMpc² neighbourhood around the central AGN and show the gas surface density map with individual circles showing locations of nearby AGN, colour-coded by their bolometric luminosities. The brightest central AGN has $L_{\text{bol}} \sim 10^{47}$ erg s⁻¹. Finally, we zoom into the host galaxy of this AGN and show a surface brightness map of the galaxy and AGN (as saturated point sources) in the B band. One can spot the two AGN in the centres of a merging pair of galaxies. This overview highlights the exceptionally broad dynamical range and wealth of physical information captured by LUMINA.

conditions are generated from separate baryon and CDM transfer functions computed with CAMB (Lewis et al. 2000; Howlett et al. 2012), which captures the delayed growth of baryonic perturbations relative to CDM at high redshift as well as the associated large-scale streaming velocity (Eisenstein & Hu 1998; Tseliakhovich & Hirata 2010; Fialkov et al. 2012). Below we summarize the physics ingredients most relevant to the AGN modelling presented in this paper,

and refer the reader to Zier et al. (2026) for a complete description of the simulation campaign.

2.1 Galaxy Formation Model

LUMINA adopts a slightly modified version of the well-tested IllustrisTNG galaxy formation model (Weinberger et al. 2017; Pillepich

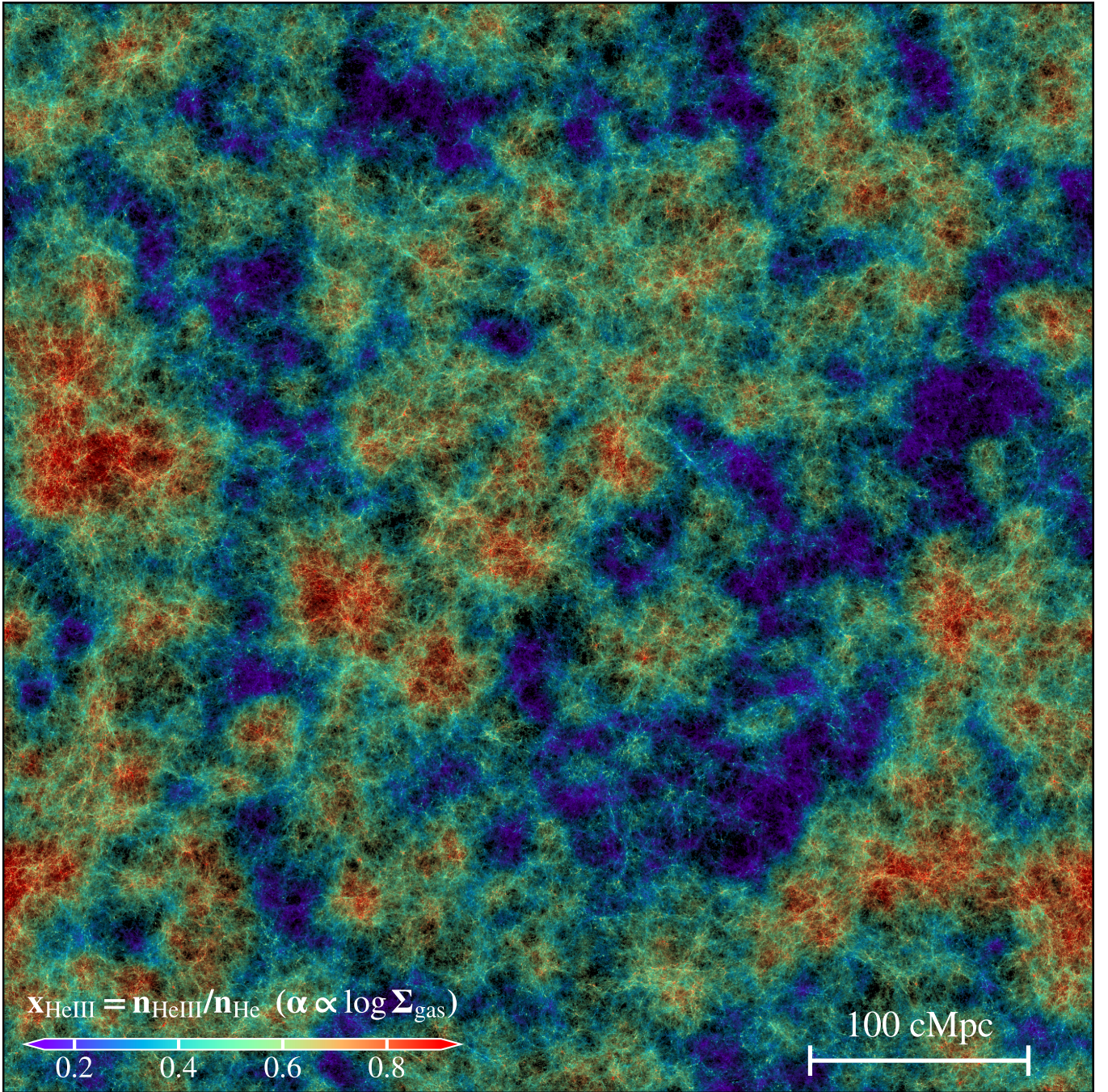


Figure 2. A blending of the gas surface density map and the mass-weighted fraction of He III for a slice of the simulation volume centred on the most luminous AGN at $z \approx 4$. The He III fraction is colour-mapped while the gas surface density dictates the transparency. The thickness of the slice is 150 cMpc. This image illustrates the large-scale topology of He II reionization driven mostly by AGN in LUMINA.

et al. 2018a), which is itself an updated version of the original Illustris framework (Vogelsberger et al. 2013, 2014a,b). In the following, we briefly summarize the relevant components of this model and highlight the modifications adopted for LUMINA. The IllustrisTNG model includes:

(i) A two-phase effective equation-of-state (EOS) model for the interstellar medium (ISM), incorporating sub-resolution pressurization from stellar feedback (Springel & Hernquist 2003). The EOS acts as a temperature floor for gas above the density threshold

$n_{\text{H}} \geq 0.106 \text{ cm}^{-3} \equiv n_{\text{H,SF}}$, while gas cells heated above the EOS are modelled as a single hot phase.

(ii) Primordial and metal-line cooling of gas (Smith et al. 2008; Wiersma et al. 2009) below $n_{\text{H,SF}}$ in the presence of a redshift-dependent, spatially uniform UV radiation background (Faucher-Giguère et al. 2009) with corrections for self-shielding (Katz et al. 1992, 1996; Rahmati et al. 2013). In LUMINA, these equilibrium cooling routines for low-density gas are replaced by a non-equilibrium

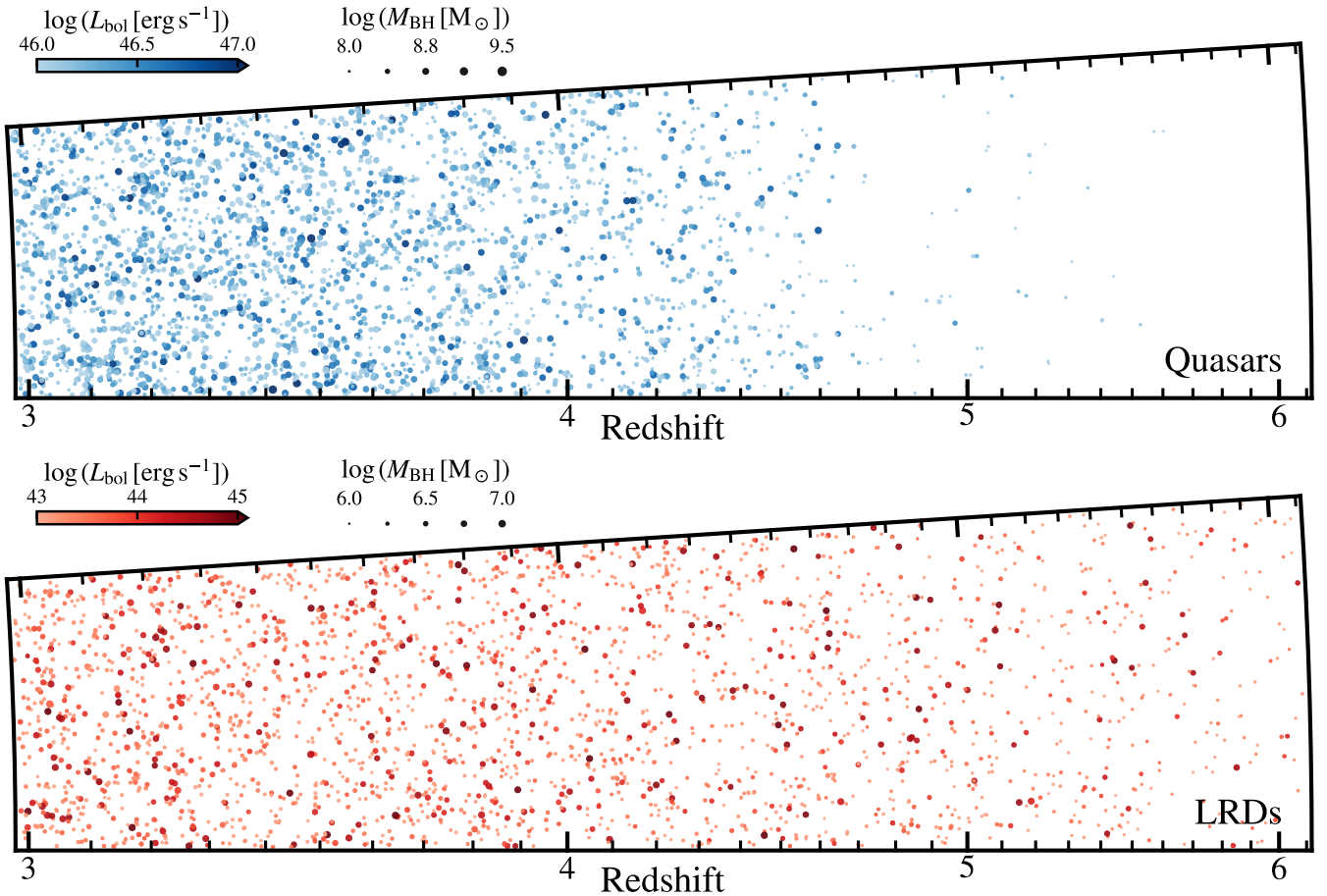


Figure 3. BH lightcone wedges from LUMINA. The wedge shows a $3.6^\circ \times 3.6^\circ$ pencil-beam (full opening angle 0.063 rad) cut from the lightcone (corrected for redshift space distortions) from $z \approx 3$ to $z \approx 6$. Each filled circle marks a BH with sightline comoving distance plotted along the long axis and the transverse position along the short axis. Marker colour encodes the bolometric luminosity, while marker size encodes the BH mass. In the top panel, we show the distribution of massive bright AGN in LUMINA (considered as moderately luminous quasars) with $L_{\text{bol}} \geq 10^{46} \text{ erg s}^{-1}$ and $M_{\text{BH}} \geq 10^8 M_\odot$. In the bottom panel, we show the LRDs selected according to our empirical model at $M_{\text{BH}} \lesssim 10^7 M_\odot$ and with $L_{\text{bol}} \geq 10^{43} \text{ erg s}^{-1}$ (downsampled by a factor of 100 to avoid overcrowding the figure). It demonstrates the ability of LUMINA in connecting SMBHs with their evolving large-scale environments. The quasars are notably rarer at higher redshifts, where LRDs start to dominate.

chemical network coupled to the radiative-transfer solver (see Section 2.3 below).

(iii) Star-formation in dense ($n_{\text{H}} \geq n_{\text{H,SF}}$) gas, following the empirical volumetric Kennicutt–Schmidt relation, with stellar particles spawned stochastically as single stellar populations following the Chabrier (2003) initial mass function (IMF). Metal enrichment from asymptotic-giant-branch (AGB) stars and supernovae (SNe) follows Pillepich et al. (2018a), except that we track only the total gas metallicity rather than individual elemental species.

(iv) Galactic winds driven by SNe, implemented with a non-local thermal-plus-kinetic scheme. Randomly selected star-forming gas cells are converted into wind particles that decouple from the hydrodynamics, leave the dense ISM, and re-couple to the Voronoi mesh at low density. The mass and energy loading factors, described in detail by Pillepich et al. (2018a), depend on the SNe rate, the local DM velocity dispersion, redshift, and gas metallicity.

2.2 Seeding, accretion, and feedback of SMBHs

The BH model largely follows Weinberger et al. (2017). BHs are seeded in haloes identified by a friends-of-friends (FoF) algorithm once the halo mass first exceeds $5 \times 10^{10} h^{-1} M_\odot$ and no BH is yet present. The gas cell with the lowest gravitational potential is converted into a BH particle with an initial mass of $M_{\text{seed}} = 8 \times 10^5 h^{-1} M_\odot$.

BHs grow by accreting from nearby gas cells at the Bondi–Hoyle rate (Bondi & Hoyle 1944; Bondi 1952)

$$\dot{M}_{\text{Bondi}} = \frac{4 \pi G^2 M_{\text{BH}}^2 \rho_{\text{gas}}}{c_s^3}, \quad (1)$$

where G is the gravitational constant, M_{BH} is the BH mass, ρ_{gas} is the ambient gas density, and c_s is the local sound speed. Ambient quantities are estimated with a smoothed-particle-hydrodynamics (SPH)-like kernel over the 128 nearest gas cells. To prevent unphysically high Bondi rates in low-density environments, which can arise from excessively large smoothing lengths, we suppress the accretion rate in the high-accretion (quasar) mode discussed below by a factor $(P_{\text{ext}}/P_{\text{ref}})^2$ whenever the kernel-weighted external pressure P_{ext}

Parameter	Value
Ω_m	0.3096
$100 \times \Omega_b$	4.897
Ω_Λ	0.6903
H_0 [km s ⁻¹ Mpc ⁻¹]	67.66
σ_8	0.8102
n_s	0.9665
z_{init}	49
T_{init} [K]	54.17
$T_{\text{CMB},0}$ [K]	2.7255
$T_{\nu,0}$ [K]	1.9524
N_{neutrino}	3

Table 1. Key cosmological parameters from the TT,TE,EE+lowE + lensing + BAO data set of Planck Collaboration et al. (2020) used in LUMINA. We list the matter density Ω_m , baryon density Ω_b , dark-energy density Ω_Λ , the Hubble constant H_0 , the amplitude of fluctuation σ_8 and spectral index n_s of the primordial power spectrum, the starting redshift of our simulations z_{init} , the initial baryon temperature T_{init} (Fixsen 2009), the present-day CMB temperature $T_{\text{CMB},0}$, the present-day temperature of the cosmic neutrino background $T_{\nu,0}$, and the number of massless neutrino species used N_{neutrino} . We absorb the slightly larger effective number of relativistic species, $N_{\text{eff}} = 3.046$, into the effective neutrino background temperature.

falls below the reference pressure P_{ref} defined in Vogelsberger et al. (2013, Equation 23). Because LUMINA does not evolve magnetic fields, P_{ext} here does not include the magnetic-pressure contribution adopted in the original IllustrisTNG simulations, which can otherwise enhance the accretion rate in low-density, magnetically dominated regions.

Accretion rate is capped at the Eddington limit

$$\dot{M}_{\text{Edd}} \equiv \frac{4\pi G M_{\text{BH}} m_{\text{p}}}{\epsilon_{\text{rad}} \sigma_{\text{T}} c}, \quad (2)$$

where m_{p} is the proton mass, c is the speed of light, σ_{T} is the Thomson scattering cross-section, and $\epsilon_{\text{rad}} = 0.2$ is the radiative efficiency, so that the effective accretion rate is $\dot{M}_{\text{BH}} = \min(\dot{M}_{\text{Bondi}}, \dot{M}_{\text{Edd}})$. BHs further grow through mergers with neighbouring BHs and, to mimic the effect of unresolved dynamical friction, are repositioned at each globally synchronized time step to the local minimum of the gravitational potential.

AGN feedback is implemented in a dual-mode prescription. At high accretion rates (the ‘‘quasar’’ mode), the feedback is a pure thermal energy injection

$$\dot{E}_{\text{therm}} = \epsilon_{\text{f,high}} \epsilon_{\text{rad}} \dot{M}_{\text{BH}} c^2, \quad (3)$$

with efficiency $\epsilon_{\text{f,high}} = 0.1$. At low accretion rates (the ‘‘radio’’ mode), the feedback is injected in a pulsed, kinetic form,

$$\dot{E}_{\text{kin}} = \epsilon_{\text{f,kin}} \dot{M}_{\text{BH}} c^2, \quad (4)$$

with the efficiency

$$\epsilon_{\text{f,kin}} = \min\left(\frac{n_{\text{H}}}{0.05 n_{\text{H,SF}}}, 0.2\right). \quad (5)$$

The transition occurs at the mass-dependent critical Eddington ratio

$$\lambda_{\text{Edd}} = \frac{\dot{M}_{\text{BH}}}{\dot{M}_{\text{Edd}}} = \min\left[0.002 \left(\frac{M_{\text{BH}}}{10^8 M_{\odot}}\right)^2, 0.1\right], \quad (6)$$

so that more massive BHs can more easily enter the kinetic mode at fixed Eddington ratios. Throughout this paper, the bolometric luminosity of an accreting BH is taken to be

$$L_{\text{bol}} = \epsilon_{\text{rad}} (1 - \epsilon_{\text{f,high}}) \dot{M}_{\text{BH}} c^2, \quad (7)$$

i.e. the radiated power after the energy allocated to AGN feedback in the quasar mode has been subtracted, consistent with the AGN source treatment in the radiative transfer described in the following section. Following the Illustris/IllustrisTNG prescription, we suppress radiation from BHs accreting at Eddington ratios $\lambda_{\text{Edd}} < 2 \times 10^{-3}$, where we assume an advection-dominated accretion flow (ADAF; e.g., Narayan & Yi 1994).

2.3 Radiative transfer

We solve the radiation-transport equations on the moving Voronoi mesh using a moment-based finite-volume method with the M1 closure (Levermore 1984; Dubroca & Feugeas 1999; Kannan et al. 2019; Zier et al. 2024). For $z > 4.75$, the radiation field is discretized into six frequency bins: three photoionizing bands associated with H I, He I, and He II, plus three higher-energy X-ray bins. At $z < 4.75$, after hydrogen reionization is complete, the radiation field below 54.42 eV is replaced with the spatially uniform UV background of Faucher-Giguère et al. (2009), and only a single high-energy bin relevant to He II reionization is evolved on the fly. As shown in Zier et al. (2026), AGN dominate the photon production in this remaining bin, so radiation from stellar particles and from other X-ray sources described below is deactivated at $z < 4.75$. Absorption is coupled to a non-equilibrium thermo-chemical network for primordial gas (Kannan et al. 2019), while metal-line cooling follows the tables of Vogelsberger et al. (2013) and is not coupled to the local radiation field. We employ the reduced-speed-of-light approximation (Gnedin & Abel 2001) with $\tilde{c} = 0.2c$, which has been shown to yield a converged reionization history for our setup (Kannan et al. 2022; Garaldi et al. 2022; Smith et al. 2022). Radiation transport and chemistry are subcycled relative to the hydrodynamics by a factor of 64 during hydrogen reionization and by a factor of 256 for $z < 4.75$. During these RT subcycles, the Voronoi mesh is held fixed to enable efficient GPU acceleration. We include three classes of radiation sources.

(i) *Stellar populations.* The ionizing output of stellar particles is computed with the BPASS v2.2.1 binary population-synthesis models (Eldridge et al. 2017; Stanway & Eldridge 2018), assuming a Chabrier IMF with an upper-mass cutoff of 100 M_{\odot} . Unresolved ISM absorption is captured by a constant stellar escape fraction, $f_{\text{esc},\star} = 0.18$, calibrated against the observed timing of hydrogen reionization (see Zier et al. 2026).

(ii) *Accreting SMBHs.* We adopt the canonical Type-I AGN spectrum of Shen et al. (2020a), which combines the optical/UV/IR template of Krawczyk et al. (2013), the extreme-UV power law of Lusso et al. (2015), and an X-ray component normalized through X-ray-optical luminosity scaling relations (Steffen et al. 2006; Lusso et al. 2010). A more detailed description is given in Section 3.2 below. The hard EUV tail of this template provides the photons that drive the He II reionization signal analysed in Section 4.7. A luminosity-dependent correction following Vogelsberger et al. (2013) is applied at the source

$$f_{\text{esc}} \equiv \frac{L_{\text{bol}}^{\text{esc}}}{L_{\text{bol}}} = 0.3 \left(\frac{L_{\text{bol}}}{10^{46} \text{ erg s}^{-1}}\right)^{0.07}, \quad (8)$$

which effectively encapsulates the escape fraction of ionizing photons through unresolved circumnuclear gas and dust and the uncertainties in the EUV spectra of AGN.

(iii) *X-ray emission from star-forming processes.* We include X-ray emission from shock-heated ISM gas and high-mass X-ray binaries (HMXBs). For the ISM component, we adopt a thermal bremsstrahlung SED with $k_{\text{B}} T_{\text{ISM}} = 240 \text{ eV}$ (Mineo et al. 2012;

Pacucci et al. 2014; Eide et al. 2018). For HMXBs, we use the SEDs of Fragos et al. (2013, 2016). The normalizations scale with the local star-formation rate, and, for HMXBs, additionally with metallicity, following Mineo et al. (2012) and Madau & Fragos (2017).

To couple the radiation field cleanly to the TNG galaxy-formation model, we use the transparent equation of state introduced by Bulichi et al. (2025). Gas denser than the star-formation threshold, $n_{\text{H,SF}}$, remains optically transparent in the RT calculation: radiation is transported through the cell, but its absorption and heating are not coupled to the gas thermochemistry, which instead continues to follow the equilibrium cooling model. This gives a clean numerical interpretation of $f_{\text{esc},\star}$ as the escape fraction from the unresolved ISM. Further details are provided in Zier et al. (2026).

2.4 Structure finding

DM haloes are identified using the FoF algorithm (Davis et al. 1985) with a linking length of 0.2 times the mean inter-particle separation. Gas, stellar, and BH particles are associated with the FoF group of their nearest DM particle if they lie within one linking length. Each FoF group is then processed with the SUBFIND-HBT algorithm (Springel et al. 2001, 2021), which iteratively identifies gravitationally bound substructures (subhaloes) and, unlike the original SUBFIND, tracks subhalo membership across snapshots, improving the identification of infalling systems through mergers (Han et al. 2012). Following common practice, we identify subhaloes as galaxies, designating the most massive subhalo of each FoF group as the central and the remainder as satellites. Because FoF masses can be inflated by percolation via “particle bridges”, we report spherical-overdensity quantities of DM haloes: R_{200c} , defined as the radius enclosing an average density of 200 times the critical density $\rho_c(z)$, and the corresponding enclosed mass M_{200c} . All BH particles in the simulation enter the AGN observable predictions. The host galaxy of a BH is identified as the subhalo to which it is gravitationally bound.

LUMINA records the full set of properties for every BH particle at high time cadence, enabling the construction of AGN light cones alongside other quantities intended for sightline-resolved surveys (Smith et al. 2026). Figure 3 shows a representative $3.6^\circ \times 3.6^\circ$ pencil-beam realization from $z \simeq 3$ to $z \simeq 6$. Each point is placed according to its sightline comoving distance and transverse position, with colour indicating bolometric luminosity and size indicating BH mass. This visualization highlights the wide dynamic range of the simulated AGN population and the ability of LUMINA to connect individual SMBHs to their evolving large-scale environments. Future studies will combine these light-cone catalogues with BH merger trees to study the temporal variability of accretion, feedback, and clustering.

3 AN EMPIRICAL MODEL LINKING SMBHS TO AGN

The BH particles and their mass accretion rates tracked in LUMINA represent the time- and kernel-averaged growth of BHs (plus any unresolved gas or stellar distributions around them), whereas observed AGN are characterized by their multi-band emission along a particular line of sight. Bridging these two descriptions requires a set of choices about how the resolved mean accretion rate is related to the instantaneous bolometric luminosity, how that luminosity is distributed across the electromagnetic spectrum, and how the emerging radiation is post-processed by circumnuclear and host-galaxy material. In this section, we present a compact post-processing framework

that makes these choices explicit, with three ingredients: (i) a log-normal scatter applied to the instantaneous bolometric luminosity of each BH particle to capture unresolved AGN variability; (ii) a pair of empirical SED templates, one for canonical Type-I AGN and one for LRDs, with the mapping between BH particles and SED templates determined by BH mass; and (iii) a gas absorption and dust extinction prescription applied to the canonical AGN population. The framework is deliberately constructed to be portable. Although calibrated and applied here to the SMBH population of LUMINA, none of its ingredients depend on the specifics of the simulation, and the same recipe can be applied to other cosmological simulations, semi-analytic models, or empirical halo–BH mapping schemes to generate high-redshift AGN population predictions.

3.1 Stochasticity of SMBH accretion

We first model an intrinsic stochasticity (variability) of SMBH accretion at the sub-resolution level by convolving the raw bolometric LF computed from binned estimations of simulated SMBHs with a Gaussian kernel of dispersion σ_{bol} . We then reduce the bolometric luminosities systematically by $\Delta \log L_{\text{bol}} = \ln 10 \sigma_{\text{bol}}^2 / 2$ (the mean shift of a log-normal distribution) to ensure conservation of total energy output from AGN. Empirically, we find that a scatter of $\sigma_{\text{bol}} = 0.3$ dex brings the predicted AGN LFs in LUMINA into good agreement with observational constraints of quasars (Shen et al. 2020a). This will be later shown in Section 4.1. We keep it as the fiducial choice and explore variations in Section 5.3. We note that the BH accretion rates are capped at the Eddington limit in LUMINA, but the empirical scatter we add here can effectively boost the accretion rates to moderately super-Eddington.

Similar phenomenological treatments have been adopted in other large-volume cosmological hydrodynamic simulations to account for unresolved AGN variability as well. For example, Rosas-Guevara et al. (2016) explored similar log-normal scatters of 0.3 – 0.5 dex for the AGN LFs in the EAGLE simulation. Ding et al. (2026) introduced a larger scatter of 0.75 dex for quasar LFs in the FLAMINGO simulation (Schaye et al. 2023) to bracket the impact of unresolved small-scale variability. Such a scatter is motivated by the observed variability of AGN across a broad hierarchy of timescales, from hours in the X-ray corona (e.g., Vaughan et al. 2003; González-Martín & Vaughan 2012), through months to years in the optical/UV “damped-random-walk” (DRW) regime (e.g., Kelly et al. 2009; MacLeod et al. 2010), to the $\sim 0.1 - 100$ Myr timescales inferred from e.g. broad Eddington-ratio distribution function (ERDF) measured at fixed host-galaxy properties (e.g., Hickox et al. 2014; Aird et al. 2018). Most large-volume cosmological simulations cannot directly resolve the clumpy multiphase gas structures, angular momentum transport, accretion-disc-state transitions, and other relevant small-scale physical processes that underlie these observational phenomena. Recent ultra-high-resolution zoom-in simulations reveal that gas inflow towards SMBHs becomes increasingly intermittent on progressively smaller scales, driven by nuclear gas structure, gravitational torques, stellar/AGN feedback, magnetic fields, and disc-scale instabilities (e.g., Hopkins & Quataert 2010, 2011; Anglés-Alcázar et al. 2017a,b, 2021; Hopkins et al. 2024; Hopkins 2025). We defer a more thorough discussion of this prescription to Section 5.

3.2 Canonical AGN SED model

For Type-I and Type-II AGN, we adopt a universal model for the intrinsic SED constructed in Shen et al. (2020a). Here we briefly

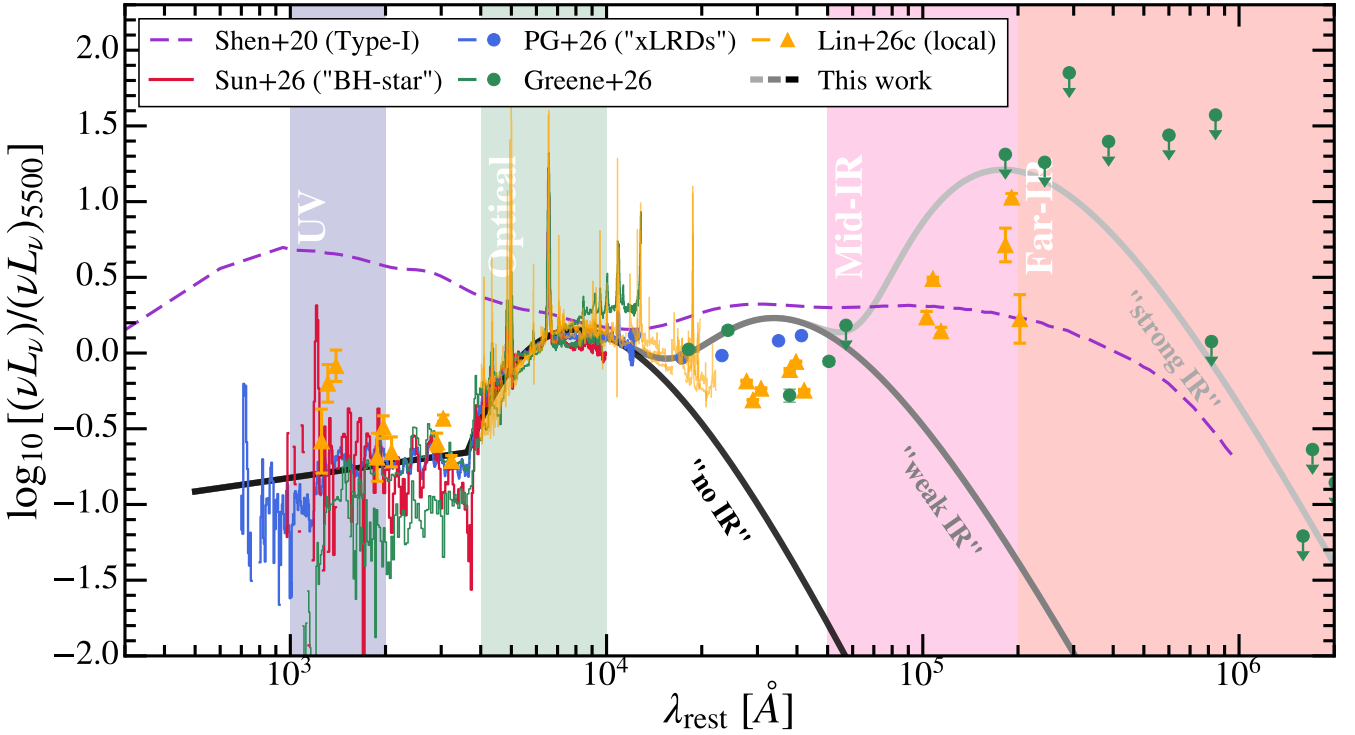


Figure 4. An overview of the AGN SED model adopted. We compile the stacked LRD spectra from Pérez-González et al. (2026) (the “xLRDs” in their sample) and Sun et al. (2026) (“BH-star”, LRD spectra subtracting host galaxy) as well as individual source spectra from Greene et al. (2026) and Lin et al. (2026c) (the three local LRD analogues), including their IR constraints. They are normalized at 5500Å. A universal SED of LRDs appears in optical-to-near-IR wavelengths, where we find that a modified blackbody with temperature $T \sim 4000$ K and $\beta \sim 0.5$ gives a reasonable fit. In the UV, we approximate it as a power-law, $\nu L_\nu \sim \nu^{-0.3}$. In the mid- and far-IR, we attempt adding blackbody components with temperature $T \sim 1000$ K and 200 K (which can be produced by e.g. warm and cold dust) and tune the normalization to the maximum allowed by current observational constraints. In the background, we also show the universal quasar SED assumed in Shen et al. (2020a). The LRDs are distinct from unobscured Type-I AGN with weaker UV and mid-IR emission and more dominant optical-to-near-IR emission. We also assume LRDs are X-ray dark. These together form the basis of the bolometric corrections we derive in Table 2.

summarize it. For the optical/UV/IR wavelengths, we adopt the SED template in Krawczyk et al. (2013), which was based on $> 100\,000$ luminous broad-line quasars observed at $0.064 < z < 5.46$ that are classified as not strongly affected by dust reddening and obscuration. The dust extinction corrections on the AGN luminosities will be considered separately in the following section. This SED template starts at $\sim 30\mu\text{m}$ and truncates at 912Å . We extend the SED to the EUV ($\lambda < 912\text{Å}$) using the power-law model $f_\nu = \nu^{\alpha_\nu}$ with index $\alpha_\nu = -1.70$ (Lusso et al. 2015). We truncate this extension at 600Å , where the measurements of Lusso et al. (2015) ended, and directly connect the flux at 600Å with the X-ray template, which will be discussed below. The EUV SED here is consistent with what we adopted in the radiative transfer calculations in LUMINA. In the IR, we extend the template to $100\mu\text{m}$ using the Richards et al. (2006b) SED. Dust emission is effectively included in this IR SED template, so we do not explicitly model any dust emission.

The X-ray SED template is modelled as a cut-off power law, $f(E) \propto E^{1-\Gamma} \exp(-E/E_c)$, with the photon index $\Gamma = 1.9$ and the cut-off energy $E_c = 300$ keV (e.g., Dadina 2008; Ueda et al. 2014; Aird et al. 2015). An additional reflection component is added using the PEXRAV model (Magdziarz & Zdziarski 1995) implemented in XSPEC (Arnaud 1996), assuming the reflection relative strength $R = 1$, the inclination angle $i = 60^\circ$, and Solar abundances. The normalization of the X-ray SED is determined according to the well-

known correlation between X-ray and optical luminosities

$$\log L_\nu(2\text{ keV}) = \beta \log L_\nu(2500\text{Å}) + C, \quad (9)$$

where the unit of L_ν is $\text{erg s}^{-1} \text{Hz}^{-1}$. One can define α_{ox} accordingly as

$$\alpha_{\text{ox}} \equiv \frac{\log L_\nu(2\text{ keV}) - \log L_\nu(2500\text{Å})}{\log \nu(2\text{ keV}) - \log \nu(2500\text{Å})} = 0.384 \log \left(\frac{L_\nu(2\text{ keV})}{L_\nu(2500\text{Å})} \right). \quad (10)$$

Then Equation (9) can be rewritten as

$$\alpha_{\text{ox}} = -A \log \left(\frac{L_\nu(2500\text{Å})}{\text{erg s}^{-1} \text{Hz}^{-1}} \right) + C', \quad (11)$$

where $A = 0.384(1 - \beta)$ and $C' = 0.384C$. These prefactors have been measured through observations. We choose the parameters measured by Steffen et al. (2006), $\beta = 0.721 \pm 0.011$ and $C = 4.531 \pm 0.688$, which have been adopted in Hopkins et al. (2007) and Shen et al. (2020a), and are consistent with more recent observations. The X-ray SED is then scaled according to α_{ox} with respect to the optical SED.

Following Shen et al. (2020a), we define the bolometric luminosity as the integrated luminosity from $30\mu\text{m}$ to 500 keV, which represents the total energy budget produced by accretion onto the SMBH. Extending the integration range to $\geq 100\mu\text{m}$, as done in e.g. Richards et al. (2006b), has a negligible impact on the bolometric luminosity.

Some studies instead truncate the integration at $1\mu\text{m}$ (e.g., Runnoe et al. 2012a; Duras et al. 2020). For our SED template, we find that this leads to a factor of 1.3 – 1.4 reduction in the bolometric correction, though this remains a subdominant effect for the purposes of this paper. The bolometric correction is defined as the ratio between the bolometric luminosity, L_{bol} , and the observed luminosity in a given band, L_{band} .

Even Type-I AGN do not share a perfectly universal SED. Real variations in the spectral shape translate into scatter in the bolometric corrections and influence the observed AGN LFs. We follow the same approach as Shen et al. (2020a), generating an ensemble of SEDs by varying the optical/FUV/EUV spectral slopes, the X-ray photon index Γ , and α_{ox} , and evaluating the resulting dispersion in the bolometric corrections. Adopting the fitting functions and best-fit parameters from Shen et al. (2020a), we compute the dispersion of bolometric corrections $\sigma_{\text{corr}}(\log L_{\text{bol}})$, which corresponds to a ~ 0.1 dex uncorrelated dispersion in AGN SEDs, consistent with observations. We also adopt the double power-law fit for the dependence of the bolometric corrections on L_{bol} from Shen et al. (2020a), along with the best-fit parameters therein. We note that the optical/UV and X-ray luminosities derived from these bolometric corrections do not yet account for gas absorption and dust extinction, which we discuss in the following sections.

The SED model above is empirical and cannot self-consistently capture systematic variations of AGN spectra with the physical state of the accretion flow. At low Eddington ratios, the thin disc transitions to a radiatively inefficient ADAF (e.g., Narayan & Yi 1994), with a weak UV “big blue bump”, a harder X-ray spectrum, reduced obscuration, and, at the lowest accretion rates, a disappearing broad-line region (e.g., Ho 1999, 2008; Elitzur & Ho 2009; She et al. 2018). Su et al. (2026) recently showed that an accretion-disc-based SED model predicts substantially lower hard X-ray luminosities at $L_{\text{bol}} \lesssim 10^{38} \text{ erg s}^{-1}$ than the Shen et al. (2020a) bolometric corrections calibrated on luminous quasars. In our framework, we treat it as electromagnetically silent when $\lambda_{\text{Edd}} < 2 \times 10^{-3}$ (Section 2.2) and it has almost no impact in the luminosity range of high-redshift AGN we study in LUMINA. At the opposite extreme, near or above the Eddington limit, photon advection through a slim disc reduces the radiative efficiency, saturates the UV/optical continuum, and radiation-pressure drives outflows (e.g., Paczyński & Wiita 1980; Abramowicz et al. 1988; Sądowski 2011; Kubota & Done 2019; Jiang et al. 2019). This regime is plausibly vital for the rapid growth of seed BHs into $z \gtrsim 6$ quasars (Section 5.1) and for many observed properties of LRDs (e.g., Pacucci & Narayan 2024; Inayoshi et al. 2025; Madau & Maiolino 2026), for which we adopt a separate empirical SED in the following section. The σ_{corr} absorbs intra-population variations within the canonical regime but not the qualitatively distinct SEDs at either λ_{Edd} extreme. A self-consistent λ_{Edd} -dependent treatment is deferred to future work.

3.3 LRD SED model

Compared to the canonical AGN SED model described above, LRDs are notably faint in X-rays given their optical luminosities and often show weak mid-IR and far-IR emission, which is challenging to explain by simple dust-reddening of a standard Type-I AGN (e.g., Yue et al. 2024; Ananna et al. 2024; Maiolino et al. 2025; Casey et al. 2025). Recent spectroscopic observations have motivated a picture in which many LRDs host accreting SMBHs embedded in extremely dense gas cocoons (e.g., Naidu et al. 2025; de Graaff et al. 2025c; Inayoshi & Maiolino 2025; Ji et al. 2025; Rusakov et al. 2026), which generate “star-like” (but not identical to stars) blackbody SEDs

in the optical featuring strong Balmer breaks and Balmer absorption features. Many recent works (e.g., de Graaff et al. 2025a; Umeda et al. 2026; Sun et al. 2026) have revealed a surprising quasi-universality of LRD SEDs at least in the optical-to-near-IR. In this section, we will attempt to build an effective universal SED model for LRDs. We note that our approach is deliberately agnostic about the underlying physical configuration of LRDs, whether they are “star-like” objects, dust-obscured AGN, or something else entirely. We restrict ourselves to constructing an empirical SED template from current observations and to bracketing its associated uncertainties.

We first take stacked LRD SEDs from Pérez-González et al. (2026) and Sun et al. (2026). Pérez-González et al. (2026) classify the LRD population into several spectroscopic flavours based on *JWST*/NIRSpec follow-up. We adopt their stack of extreme LRDs (“xLRDs”), which is the AGN-dominated subset with comparatively minor host-galaxy continuum contamination and interpreted as LRD engines therein. Sun et al. (2026) present a complementary stacked SED after subtracting their best-estimate host-galaxy contribution under the “BH-star” picture. In addition, we compile the spectra of individual sources from Greene et al. (2026) (see also Setton et al. 2025a) and the three local LRD analogues of Lin et al. (2026c), including their mid- and far-IR constraints. These compiled SEDs are shown in Figure 4. A degree of universality emerges across these LRD SEDs at optical-to-near-IR wavelengths, with the remaining diversity concentrated in the mid- and far-IR where current observational constraints are weakest.

Inspired by this universality, we model the optical part of the SED as a modified blackbody

$$L_{\nu} \propto \nu^{3+\beta} \frac{1}{e^{h\nu/k_{\text{B}}T} - 1}, \quad (12)$$

with temperature $T = 4000 \text{ K}$ and $\beta = 0.5$, broadly consistent with the best-fits of de Graaff et al. (2025a) and with the $\sim 3000 - 6000 \text{ K}$ colour temperatures inferred in other works (e.g., Naidu et al. 2025; Setton et al. 2025a; Akins et al. 2025b; Umeda et al. 2026). In the UV, we approximate it as a power law, $\nu L_{\nu} \sim \nu^{-0.3}$. In the mid- and far-IR, we attempt to add blackbody components with $T \simeq 1000 \text{ K}$ and $T \simeq 200 \text{ K}$, with normalisations tuned to the maximum allowed by current observational constraints. The IR components could be produced by warm and cold dust, but are not limited to these sources.

We consider three scenarios, starting from no mid- and far-IR components at all and then progressively adding the two. They are shown in Figure 4, labelled as “no IR”, “weak IR”, and “strong IR”, respectively. We caution that the mid- and far-IR portion of the LRD SED remains genuinely uncertain: stacking of deep MIRI, ALMA, and Spitzer/Herschel imaging of LRDs yields strong constraints on the hot and cold dust emission (e.g., Casey et al. 2025; Setton et al. 2025b; Akins et al. 2025b; Xiao et al. 2025), while evidence of hot dust emission from a rising near-IR continuum in stacked MIRI SEDs has also been reported (Delvecchio et al. 2025). Our three SED models are designed to bracket this observed diversity, so effectively we allow for a potential mixture of scenarios like “BH-star”, dust-obscured, and others. We then integrate the SED from 500 \AA to $1000 \mu\text{m}$ (assuming that LRDs are completely dark in X-rays, but see also Hviding et al. 2026) to obtain the bolometric luminosity. We intentionally include the entire far-IR component in integration for LRDs in order to fully bracket the uncertainties, although cold dust in the host galaxies could be the major source of far-IR radiation, like in canonical AGN. In Table 2, we summarize the bolometric corrections derived for LRDs in three different scenarios, and they are compared to the average values computed in Greene et al. (2026) and bolometric corrections of canonical AGN. In the rest of the paper,

	UV (1450Å)	B band (~ 4400Å)	“Optical” (~ 5100Å)	H α (~ 6563Å)	Hard X-ray (2-10 keV)	Mid-IR (~ 15 μ m)
LRDs						
Greene et al. (2026) (no cold dust)	(26.8) ^a	(8.5)	5.4	19	-	-
Greene et al. (2026) (with cold dust)	(183.6)	(58.6)	37	131	-	-
This work (no IR)	12.0	3.8	2.4	-	∞	∞
This work (weak IR)	24.5	7.7	4.9	-	∞	39.6
This work (strong IR)	148.1	46.8	29.5	-	∞	1.6
Canonical AGN						
Shen et al. (2020a) (at $L_{\text{bol}} = 10^{45} \text{ erg s}^{-1}$)	5.4	10.8	11.8		34.2	12.6
Greene & Ho (2005)				170		
Richards et al. (2006b)			10.3			
Runnoe et al. (2012a,b)	5.7[4.2] ^b		11.0[8.1]			11.6[8.5]
Duras et al. (2020)		8.5[5.13]			34.7[20.9]	

Table 2. Bolometric corrections (defined as $L_{\text{bol}}/(\nu L_{\nu})$ for single wavelengths and $L_{\text{bol}}/L_{\text{band}}$ for integrated luminosity in a band) computed for LRDs. We compare the corrections in three scenarios described in Section 3.3 labelled as: “no IR”, “weak IR”, and “strong IR”, to reflect the uncertainties in the IR. We compare them to the “average” bolometric corrections computed in Greene et al. (2026). The two scenarios highlighted in bold form are the lower and upper limits of bolometric corrections considered in this work. In the bottom, we show the bolometric corrections of canonical AGN at $L_{\text{bol}} = 10^{45} \text{ erg s}^{-1}$ from literature for comparison.

^a The values in parentheses are not directly reported in their original paper. We take their bolometric corrections at 5100Å, and convert them to other bands using our LRD SED. ^b These works assume different integration ranges to compute bolometric luminosities. We use our quasar SED template to correct the values to be consistent with our assumption. Their original values are displayed in the brackets.

we choose the “weak IR” and “strong IR” models to make predictions for the lower and upper limits of AGN LFs.

The scatters of the bolometric corrections of LRDs are inferred from the dispersions of the scaling relations between optical, UV, and blackbody luminosities (de Graaff et al. 2025a). We adopt $\sigma_{\text{corr}} \approx 0.34$ dex in the UV and $\sigma_{\text{corr}} \approx 0.23$ dex in the optical, although these are subdominant to the systematic IR uncertainty discussed above. We do not assign an additional dispersion to the IR component, as that uncertainty is already captured by the IR SED variants.

3.4 Gas absorption and dust extinction

The absorption and scattering of surrounding gas and dust further modify the intrinsic emission of AGN. The photoelectric absorption of gas influences X-ray emission, while dust plays a key role in the optical and UV. Here, we first introduce a model for the neutral hydrogen column density (N_{H}) distribution, which determines the absorption in the X-ray. N_{H} is then converted to the column density of dust, assuming a dust-to-gas ratio, which determines the dust extinction in the optical and UV. We note that the absorption/extinction model described below is applied to the canonical AGN. For LRDs, we do not apply any additional correction on top of the SED we adopt, since we assume that it has already been reprocessed by the gas (and potentially dust) around the central BH.

Large-volume cosmological simulations cannot resolve the circumnuclear gas distribution and underestimate the gas column density along the line of sight (see the measurement done in Ni et al. 2022). Therefore, an observation-driven empirical model is required. Following Shen et al. (2020a), we adopt the N_{H} distribution from Ueda et al. (2014), based on measurements of N_{H} and intrinsic hard X-ray luminosity for individual objects in their sample. The model provides the probability distribution of N_{H} as a function of intrinsic hard X-ray luminosity and redshift. We refer readers to Ueda et al. (2014) for the detailed functional form and parameters. In this model, the abundance of CTK ($\log(N_{\text{H}}/\text{cm}^{-2}) \geq 24$) AGN is assumed to be the same as absorbed

Compton-thin (CTN) ($\log(N_{\text{H}}/\text{cm}^{-2}) = 22 - 24$), with both fractions decreasing at higher hard X-ray luminosities and increasing at higher redshift with a plateau at $z \geq 2$. A more recent X-ray AGN synthesis model (Ananna et al. 2019) that includes ultra-hard X-ray data from *NuSTAR* suggests an enhanced CTK AGN fraction compared to previous measurements, which is remarkably consistent with Carnall et al. (2023) using an independent method. As shown in Figure 5, we find that boosting the CTK AGN fraction relative to the obscured CTN population (f_{CTK}) by a factor of 2–3 in the Ueda et al. (2014) model can mimic these new constraints at $z \approx 4$. The same agreement is consistently found at $3 \lesssim z \lesssim 5$ where the Ananna et al. (2019) model covers. In this work, the fiducial boost factor we adopt is 2.5. This update is mainly motivated by the observational knowledge about heavily gas-obscured AGN. However, its impact on the resulting LFs is quite limited as we will show in Section 5.3.

Given N_{H} , we next apply a multiplicative factor $\exp(-N_{\text{H}}[\sigma_{\text{pe}}(E) + \sigma_{\text{T}}])$ to the intrinsic luminosity, where $\sigma_{\text{pe}}(E)$ is the photoelectric absorption cross-section from Morrison & McCammon (1983), and σ_{T} is the Thomson scattering cross-section, the non-relativistic limit of the Compton (Klein-Nishina) cross-section, which is accurate to within a few percent throughout the 2–10 keV hard X-ray band of interest. We note that this screen-absorption treatment is a simplification of the full radiative transfer of X-ray photons through circumnuclear material, and can be inaccurate deep in the CTK regime. It does not capture e.g. multi-scattered photons that return to the line of sight in optically thick media, variations of the Compton-reflection continuum produced by scattering off circumnuclear material, and the fluorescent Fe $K\alpha$ line that contributes a small but non-zero fraction of the integrated hard X-ray luminosity. Taken together, these approximations imply that our predicted 2–10 keV AGN luminosities are robust for the CTN population but are approximate for the heavily CTK sources. Nevertheless, the impact on the LF is limited since the latter population contributes little to the post-absorption LF. A future treatment based on Monte Carlo radiative-transfer torus models, such as MYTorus (Murphy & Yaqoob 2009), BNTorus/borus02

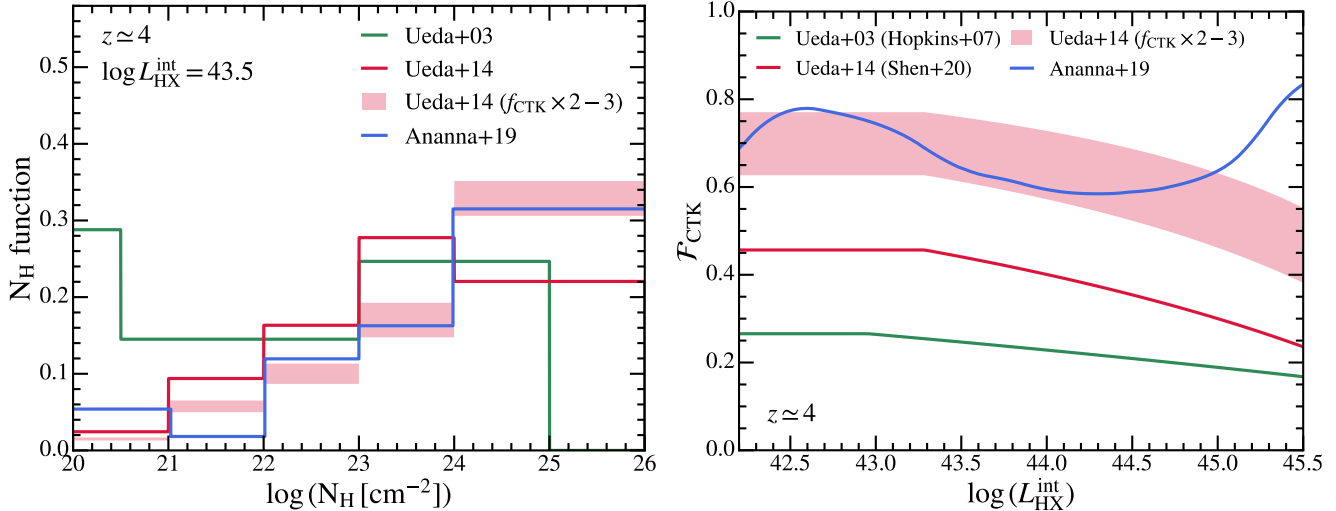


Figure 5. *Left:* Hydrogen column density distribution function at $z \approx 4$ at intrinsic hard X-ray luminosity $L_{\text{HX}}^{\text{int}} = 10^{43.5} \text{ erg s}^{-1}$. The distribution is normalised over $\log(N_{\text{H}} [\text{cm}^{-2}]) = [20, 26]$. We show the N_{H} distribution from Ueda et al. (2003) (used in Hopkins et al. 2007), Ueda et al. (2014) (used in Shen et al. 2020a), and more recent constraints from Ananna et al. (2019) leveraging the *NuSTAR* data. We show a modified Ueda et al. (2014) distribution, increasing the CTK AGN fraction relative to the obscured CTN population (f_{CTK}) by a factor of 2–3, which roughly agrees with Ananna et al. (2019) at these luminosities at $3 \leq z \leq 5$. *Right:* CTK AGN fraction (f_{CTK} , relative to the whole population, different from f_{CTK}) versus $L_{\text{HX}}^{\text{int}}$ at $z \approx 4$. We show the inferred fractions from Ueda et al. (2003, 2014) and Ananna et al. (2019) for comparison. The modified Ueda et al. (2014) model by boosting CTK fraction matches roughly the model in Ananna et al. (2019) until reaching $L_{\text{HX}}^{\text{int}} \sim 10^{45} \text{ erg s}^{-1}$, where the statistics of the observational sample become poor. In this work, the fiducial f_{CTK} boost factor we take is 2.5.

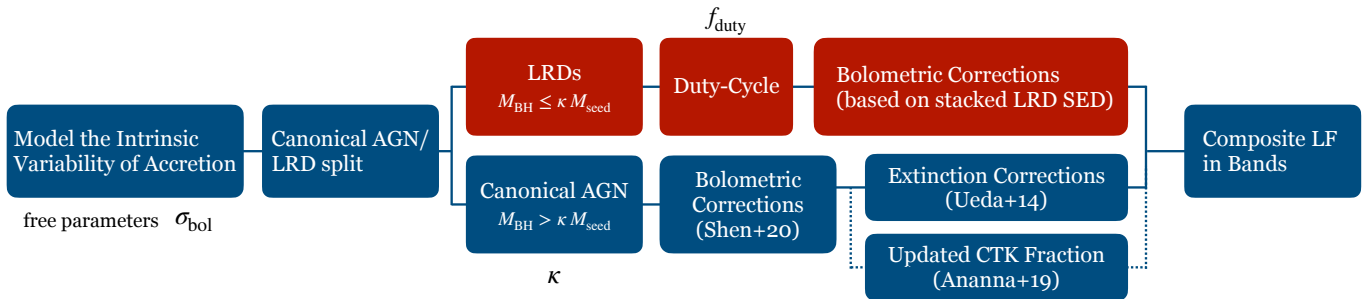


Figure 6. A schematic view of the empirical model for AGN LF. We start from the raw intrinsic bolometric luminosities of SMBHs informed from LUMINA and apply a scatter in bolometric luminosities due to unresolved variability of accretion (see Section 3.1). Then we perform a split of canonical AGN and the ones in the LRD phase based on the BH particle mass in LUMINA. The former undergoes the standard AGN bolometric and extinction/absorption corrections to obtain the LFs in observed bands, while the latter relies on empirical bolometric corrections based on observed stacked LRD SEDs and has an additional duty-cycle parameter. The three free parameters of this model are highlighted in black.

(Brightman & Nandra 2011; Baloković et al. 2018), or the more recent clumpy-torus calculations (Tanimoto et al. 2019; Buchner et al. 2019), would refine these predictions.

To determine the dust abundance, a dust-to-gas ratio is required along with the extinction curve as observed in the Small Magellanic Cloud (Pei 1992; commonly assumed for high-redshift metal-poor environments, note that this is different from the choice in Shen et al. 2020a but consistent with Hopkins et al. 2007). Observations have revealed that the gas-phase mass-metallicity relation of galaxies overall decreases with redshift (e.g., Lilly et al. 2013; Zahid et al. 2013; Sanders et al. 2021). Assuming that the dust-to-metal ratio remains constant, the decrement in the gas-phase metallicity of AGN host galaxies will lead to a decrement in the dust-to-gas ratio at higher redshift. We choose to adopt a redshift-dependent dust-to-gas ratio that scales as the gas-phase metallicity given by the fit in Ma et al. (2016), which is consistent with recent *JWST* constraints on high-redshift mass-metallicity relations (e.g., Nakajima et al. 2023;

Curti et al. 2024; Sarkar et al. 2025). The value of the dust-to-gas ratio in the local Universe still follows Hopkins et al. (2007) with $(A_{\text{B}}/N_{\text{H}}) = 8.47 \times 10^{-22} \text{ cm}^2$.

The absorption and extinction model and the bolometric corrections allow us to link the bolometric LF with the observed one in a certain band. Unless otherwise specified, the AGN LFs presented in this paper include both the obscured and unobscured AGN, and the observed AGN LFs account for gas absorption and dust extinction.

3.5 An empirical split of LRDs and canonical AGN

We assume that a subset of LRDs traces the earliest, still-unresolved growth phase of SMBHs from their seeds (e.g., Inayoshi 2025). This interpretation is physically motivated by a broad class of massive-seed scenarios, such as quasi-stars as a phase of direct-collapse BHs (e.g., Begelman et al. 2006, 2008; Volonteri & Begelman 2010; Begelman & Dexter 2026), the collapse of supermassive stars (e.g.,

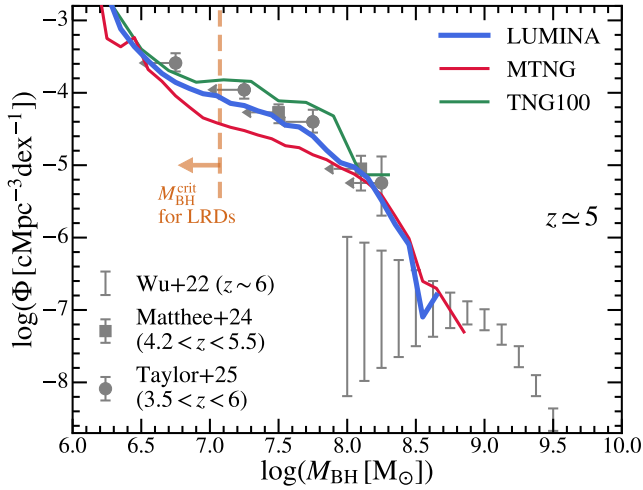


Figure 7. SMBH mass function at $z \approx 5$ in LUMINA, TNG100, and MTNG. We compare the results with observational constraints of BLAGN from Wu et al. (2022) for quasars and Matthee et al. (2024); Taylor et al. (2025) for LRDs. The latter should be considered as upper limits due to the large uncertainties in single-epoch BH mass estimates using local scaling relations. Specifically, we highlight the upper mass threshold ($M_{\text{BH}}^{\text{crit}} = 10 M_{\text{seed}}$) we set for the LRD population with the orange dashed line.

Begelman 2010; Hosokawa et al. 2013; Woods et al. 2017; Pacucci et al. 2025), and DM-seeded BHs (e.g., Xiao et al. 2021; Shen et al. 2025a; Jiang et al. 2026). We introduce a simple empirical mass cut: for BHs with $M_{\text{BH}} \leq M_{\text{BH}}^{\text{crit}} \equiv \kappa M_{\text{seed}}$, we assume that the object is in an early-growth phase that is not properly resolved in the simulation. In this regime, the BH particle should be interpreted not as an isolated, fully exposed SMBH, but as a sub-grid tracer of an unresolved nuclear system consisting of the seed BH, its dense ambient gas reservoir, and possibly a surrounding nuclear star cluster in a proto-galactic nucleus. We defer a more detailed discussion of this interpretation to Section 5. We then assume that this unresolved early-growth population appears observationally as LRDs and assign it the representative LRD SED described above in Section 3.3. We adopt the LRD bolometric correction factors in Table 2 and do not apply any additional absorption or extinction corrections, since the assumed LRD SED already represents radiation reprocessed by the surrounding material. We also introduce a duty cycle, f_{duty} , such that only a fraction of these low-mass SMBHs are active and observable as LRDs at any given time. The remaining $(1 - f_{\text{duty}})$ population is treated as electromagnetically silent in the bands we study in this work. In Section 5.3, we discuss the impact of this assumption. BHs above the mass threshold are instead treated as canonical AGN and follow the bolometric correction and absorption/extinction correction described in the Sections above. A summary of this pipeline is shown in Figure 6. Unless otherwise stated, we adopt $\kappa = 10$ (which results in $M_{\text{BH}}^{\text{crit}} \sim 10^7 M_{\odot}$) and $f_{\text{duty}} = 0.3$ as our fiducial values, motivated by comparisons to the observed AGN LFs as will be shown in Section 4.

In Figure 7, we show the SMBH mass function at $z \approx 5$ in LUMINA, compared to results from TNG100 (Springel et al. 2018) and MTNG (Pakmor et al. 2023), which adopt the same galaxy formation model but differ in both mass resolution and simulation volume. The TNG100 volume is $\sim 1/92$ that of LUMINA, but it has ~ 2.6 times better mass resolution and includes magnetic fields. In contrast, MTNG has a volume ~ 3.2 times larger than LUMINA, but with ~ 8.6 times

poorer mass resolution. We illustrate the fiducial $M_{\text{BH}}^{\text{crit}}$ below which we model BHs as LRDs. For comparison, we also show the SMBH mass function from observations for luminous quasars (Wu et al. 2022) and LRDs (Matthee et al. 2024; Taylor et al. 2025). The LRD constraints should be interpreted as upper limits, given the substantial uncertainties in the broadening mechanism of the Balmer lines (e.g., Naidu et al. 2025; Rusakov et al. 2026, but see also Brazzini et al. 2025, 2026; Scholtz et al. 2026) and in single-epoch virial mass estimators calibrated using normal AGN more generally. We note that we do not interpret the simulated BH particle masses as a one-to-one tracer of the physical SMBH mass, so the comparison here is only to give a qualitative sense of the BH particle mass distribution in the simulation. In the remaining sections of the paper, we will focus on the AGN LFs rather than the SMBH mass function as the primary point of comparison with observations.

4 RESULTS

4.1 Bolometric LFs

We begin with the bolometric LF of AGN. This quantity is particularly well suited for comparison with theoretical models because it is directly tied to the BH mass accretion rates and total energy output predicted by simulations (see Equation 7). In Figure 8, we show the bolometric LFs of AGN in LUMINA at $z \approx 3 - 6$. For comparison, we also show results from TNG100 and MTNG. At the luminous end, LUMINA and MTNG show good convergence across redshift. At the faint end, LUMINA is converged with TNG100 at $z \lesssim 4$, whereas a mild resolution dependence persists at higher redshift, with TNG100 predicting a higher abundance of low-luminosity AGN. Given this residual resolution dependence, we remain conservative and avoid drawing overly precise quantitative conclusions at the faint end at $\lesssim 0.2$ dex level. Another notable feature is the upturn at the very faint end ($L_{\text{bol}} \lesssim 10^{44} \text{ erg s}^{-1}$). A similar feature is also visible in the SMBH mass functions shown in Figure 7. These features are likely associated with the relatively large seed BH mass adopted in the simulations, since the Eddington luminosity of our seed BHs happens to be $L_{\text{Edd}} \sim 10^{44.2} \text{ erg s}^{-1}$. Newly seeded BHs are inserted directly at M_{seed} , and the unresolved early growth phase of lower-mass BHs is therefore not explicitly followed. It is one of the key reasons for us to develop the empirical model for low-mass BHs to encapsulate physics in the unresolved phase. The feature can be amplified by the duty-cycle of LRDs in our model, which suppresses the LRD contribution to the bolometric LFs at the faint end.

We compare the simulation predictions to the pre-*JWST* observational constraints compiled by Shen et al. (2020a), which include measurements in the UV, optical, soft and hard X-ray, and mid-IR bands, and the best-fit model therein. These multi-wavelength LFs are mapped onto the bolometric plane using bolometric corrections therein, while preserving the offset in $\log \Phi$ between each observed data point and the best-fit model LF. We choose the ‘‘Global fit A’’ from Shen et al. (2020a), which has a relatively steep faint-end slope driven mainly by the UV data in that compilation. We note that, in preparing these observational data, we keep the original Shen et al. (2020a) model without any modifications (e.g., the change in f_{CTR}) introduced above. With our fiducial choice of $\sigma_{\text{bol}} = 0.3$ dex, the bolometric LF in LUMINA agrees well with the observational constraints for moderately luminous quasars with $L_{\text{bol}} \gtrsim 10^{46} \text{ erg s}^{-1}$. At even higher luminosities, the finite simulation volume becomes insufficient to sample the rarest bright quasars. In Section 5.3, we compare the fiducial LF to versions with different levels of scat-

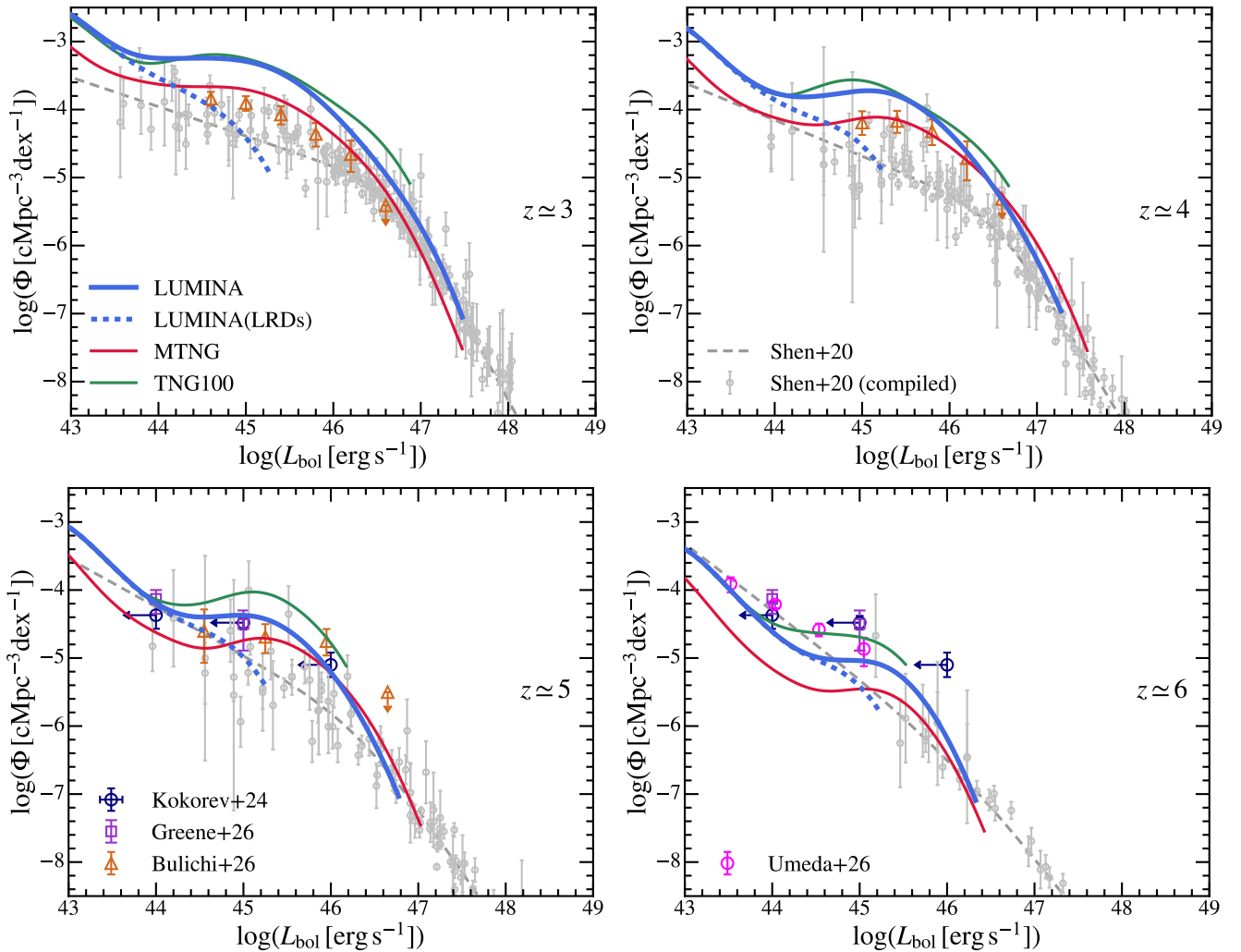


Figure 8. Bolometric LFs of AGN at $z = 3 - 6$. We present the bolometric LFs of SMBHs in LUMINA, TNG100, and MTNG. The blue dashed line shows the contribution by LRDs. These simulation predictions are compared to the best-fit quasar LF in Shen et al. (2020a) (the “Global fit A”) and the compiled observational data there (moved to the bolometric plane by preserving the offsets to the best-fit model). We also compare our results to the bolometric LF of LRDs (e.g., Kokorev et al. 2024; Greene et al. 2026; Umeda et al. 2026) and obscured AGN in the MEOW survey (Bulichi et al. 2026).

ter, and also place our results in the context of other simulations in Section 5.4.

At the faint end, we first include recent bolometric LF constraints inferred from observed LRDs. Kokorev et al. (2024) inferred bolometric luminosities of LRDs from their rest-frame optical luminosities assuming standard AGN bolometric corrections and dust corrections (e.g., Greene & Ho 2005; Richards et al. 2006b). However, more recent work by Greene et al. (2026) suggests that LRDs have substantially different bolometric corrections from classical unobscured AGN (see our discussion in Section 3.3 above). We therefore treat the luminosities inferred by Kokorev et al. (2024) as conservative upper limits. Updated constraints on the LRD bolometric LF have been presented by Greene et al. (2026). Similar conclusions have also been reached by Umeda et al. (2026), who estimated bolometric luminosities using a BH-envelope model. In addition to LRDs, we further include the mid-IR-based constraints from the *JWST* MIRI Early Obscured AGN Wide Survey (MEOW; Leung in prep., Bulichi et al. 2026), which provide an independent census of obscured AGN activity at high redshifts.

Overall, recent *JWST* observations probing the faint end of the

AGN LFs ($L_{\text{bol}} \lesssim 10^{46} \text{ erg s}^{-1}$) suggest an enhanced AGN abundance relative to previous UV and X-ray surveys. These new measurements are broadly consistent with the predictions from LUMINA. Meanwhile, relative to the pre-*JWST* compilation of Shen et al. (2020a), LUMINA predicts a higher abundance of low-luminosity AGN. This discrepancy is expected, since the faint-end LF in Shen et al. (2020a) is largely constrained by X-ray observations, which are likely incomplete if a substantial fraction of LRDs and other faint high-redshift AGN are intrinsically X-ray weak or heavily obscured by gas (e.g., Yue et al. 2024; Maiolino et al. 2025). We revisit this issue in Section 4.4, where we compare the X-ray LFs directly.

4.2 Optical and $\text{H}\alpha$ LFs

We next consider AGN LFs in specific rest-frame bands, which are less impacted by the varying approaches observational works use to derive bolometric luminosities. We begin with the rest-frame optical, where LRDs are most prominent and where their bolometric corrections are expected to be substantially lower than those of canonical unobscured quasars (see Section 3.3). We choose the rest-frame B

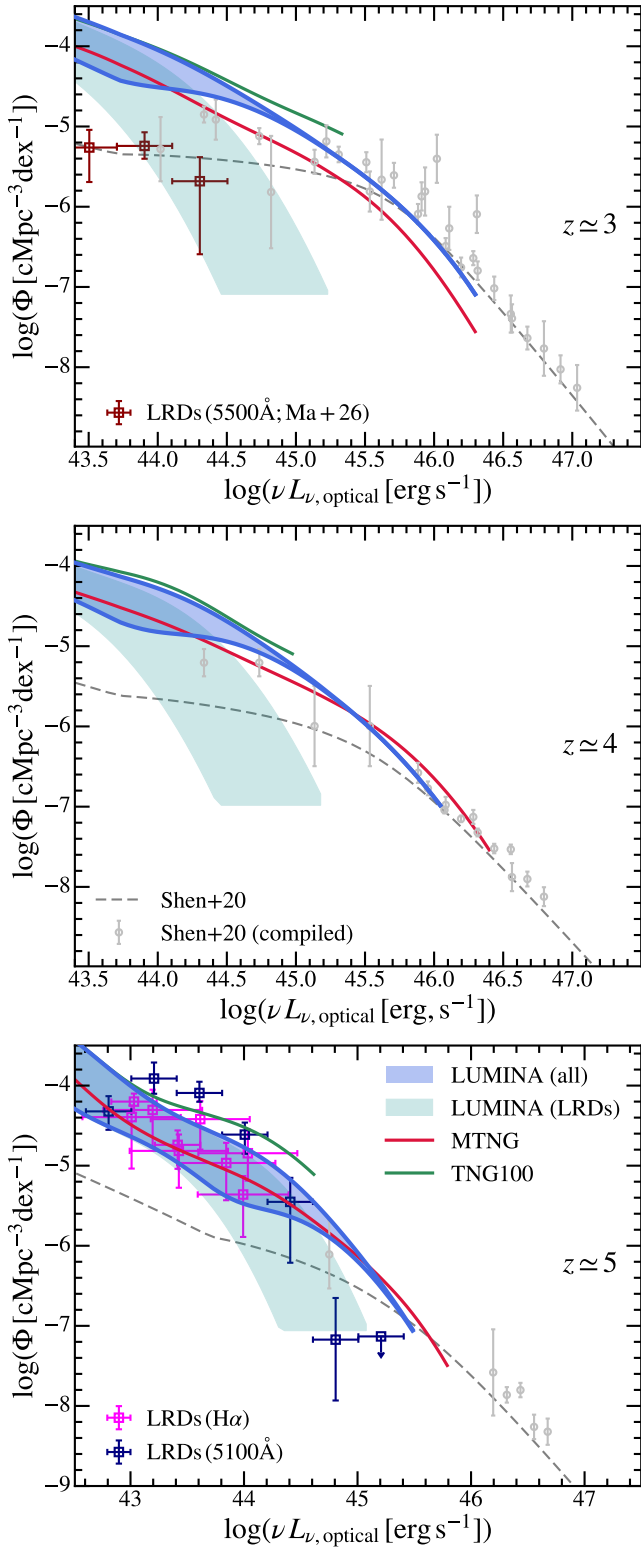


Figure 9. Rest-frame optical (B band, $\sim 4400\text{\AA}$) LFs of AGN at $z \approx 3 - 5$ from LUMINA, TNG100, and MTNG. The shaded region represents the lower and upper limits of the LFs given uncertainties in bolometric corrections of the LRDs. We compare the results to optical observations compiled in Shen et al. (2020a) and LRD constraints from Kokorev et al. (2024); Ma et al. (2025) (at 5100\AA) and Matthee et al. (2024); Lin et al. (2024, 2026a) based on $H\alpha$ at $z \sim 5 - 6$ and the ones from Ma et al. (2026) (at 5500\AA) at $z \sim 3$. The LRD constraints are all corrected to the B band luminosity using the SED model in Section 3.3 and the scaling relation in Equation 13.

band ($\lambda \approx 4400\text{\AA}$) to do the comparison, mainly to follow the conventions of observations compiled in Hopkins et al. (2007); Shen et al. (2020a). This band is also where our bolometric corrections and dispersions for canonical AGN are derived. Alternative optical LFs at 5100\AA are shown in Appendix A1.

In Figure 9, we show the rest-frame optical LFs of AGN at $z \approx 3 - 5$ in LUMINA. Predictions from TNG100 and MTNG are shown here for reference. For the LRDs in the simulation, we adopt the bolometric correction factors listed in Table 2. Specifically, we use the values inferred for the “weak IR” and “strong IR” cases as lower and upper limits, respectively, and present the resulting LFs as a shaded region. To avoid overcrowding the figure, the TNG100 and MTNG LFs assume the “weak IR” model. For comparison, we include the pre-*JWST* observational datasets compiled in Shen et al. (2020a), which all have been converted to the B band. We also include the LRD LF measurements from Kokorev et al. (2024); Ma et al. (2025, 2026). These measurements are typically reported at rest-frame 5100\AA or 5500\AA , and we convert them to the B band using the LRD SED in Section 3.3. The corresponding difference between $\log L_{5100}$ ($\log L_{5500}$) and $\log L_{4400}$ is 0.20 dex (0.27 dex). In addition, we include the broad $H\alpha$ emitter LFs from Matthee et al. (2024); Lin et al. (2024, 2026a). We convert these measurements to rest-frame optical continuum luminosities using the empirical relation

$$L_{5100} = 10^{44} \text{ erg s}^{-1} \left(\frac{L_{H\alpha}}{10^{43.30} \text{ erg s}^{-1}} \right)^{1/1.04}, \quad (13)$$

as found in de Graaff et al. (2025a) for their full LRD sample. It is a tight correlation with scatter ~ 0.2 dex. The resulting $H\alpha$ -based constraints are broadly consistent with the continuum luminosity constraints in observations.

The predictions from LUMINA simultaneously reproduce several key features of the observed LRD LFs: the high abundance of sources at $L_{\text{opt}} \sim 10^{43} - 10^{44} \text{ erg s}^{-1}$, the rapid decline in number density when approaching $L_{\text{opt}} \sim 10^{45} \text{ erg s}^{-1}$ at $z \sim 5$ (Ma et al. 2026), and the lower number density of LRDs at $z \sim 3$ (Ma et al. 2025), except for their faintest luminosity bin. At the same time, the total AGN LF connects smoothly to the pre-*JWST* constraints on the canonical AGN population compiled by Shen et al. (2020a). The faint-end optical LF is sensitive to our choices of κ and f_{duty} . For example, reducing κ and $M_{\text{BH}}^{\text{crit}}$ will reduce the brightest LRDs we predict, and reducing f_{duty} results in a change in the normalization of the LF. Therefore, the agreement with the observed optical LFs should be interpreted partly as a calibration of these effective parameters. In Section 5, we revisit this point by exploring alternative parameter choices, the associated degeneracies, and their physical implications. In the following sections, we test whether the same empirical model and fiducial parameter choices remain consistent with observations in other wavelength bands.

4.3 UV LFs

The rest-frame UV emission in LRDs remains one of their most puzzling features. Several scenarios have been proposed, including leaked or reprocessed AGN light (e.g., Greene et al. 2024; Labbe et al. 2024; Kocevski et al. 2025; Lin et al. 2026c; Madau & Maiolino 2026), emission from young stars in the host galaxy or compact nuclear starbursts (e.g., Killi et al. 2024; Akins et al. 2025a; Asada et al. 2026; Inayoshi et al. 2026; Baggen et al. 2026; Naidu et al. 2025; Sun et al. 2026), nebular emission from low-density ionized gas around the central SMBH (e.g., Chen et al. 2025a,b), or a mixture of them (Cloonan et al. 2026). The UV emission is crucial for the photometric identification of LRDs, since the population is often

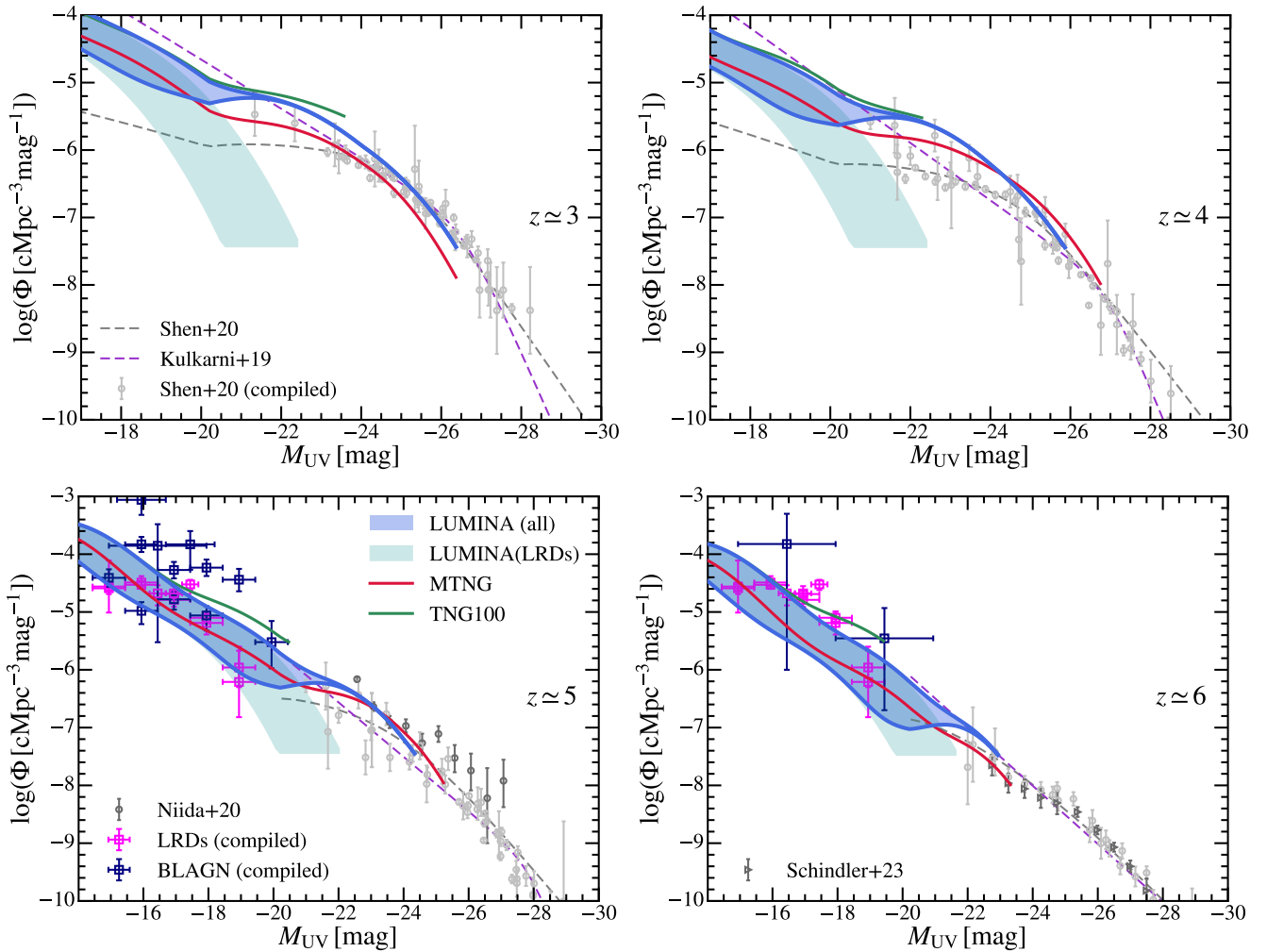


Figure 10. Rest-frame UV LFs of AGN at $z = 3 - 6$ from LUMINA, TNG100, and MTNG. We compare them with the UV observations compiled in Shen et al. (2020a) and additional constraints from Niida et al. (2020); Schindler et al. (2023). For reference, we also show the best-fit AGN UV LF from Shen et al. (2020a) and Kulkarni et al. (2019). For LRDs, we show two groups of observational constraints: the BLAGN LFs from Matthee et al. (2024); Maiolino et al. (2024); Harikane et al. (2023b); Taylor et al. (2025) and the LRD LFs from Kokorev et al. (2024); Greene et al. (2024); Kocevski et al. (2025), where the latter have additional explicit/implicit “V-shape” SED selection criteria. We assume that AGN account for only 15% of the UV luminosity in LRDs (Sun et al. 2026) and shift them to the fainter end accordingly.

selected through its characteristic “V-shaped” SED, with a blue UV continuum and a red optical continuum. It is therefore important to ask whether a coherent picture can simultaneously explain the UV and optical luminosities of LRDs.

In Figure 10, we show the rest-frame UV LFs of AGN at $z \approx 3 - 6$ in LUMINA, again comparing our predictions to TNG100 and MTNG. At the bright end, we include the pre-*JWST* observational constraints compiled by Shen et al. (2020a), together with the measurements from Niida et al. (2020) at $z \approx 5$ and Schindler et al. (2023) at $z \approx 6$. The best-fit quasar UV LFs in Kulkarni et al. (2019) and Shen et al. (2020a) are also shown for reference. Similar to the optical LFs discussed above, the simulations show reasonable agreement with these pre-*JWST* observations at moderately bright luminosities ($M_{\text{UV}} \lesssim -22$).

At the faint end, we compare our predictions to two categories of observational constraints at $z \approx 5 - 6$. The first category is based on spectroscopically selected BLAGN, without an explicit colour or morphological pre-selection (Matthee et al. 2024; Harikane et al. 2023b; Maiolino et al. 2024; Taylor et al. 2025). The second category

consists of LRDs selected with an additional “V-shape” SED criterion, applied either explicitly via photometry (Kokorev et al. 2024; Kocevski et al. 2025) or implicitly through spectroscopic follow-up of photometric LRD samples (Greene et al. 2024). The BLAGN abundance from the first category is generally larger than the LRD abundance from the second category at fixed M_{UV} , although the flux-limited blind Grism selection in Matthee et al. (2024) gives results close to the LRD LFs. Recent works (e.g., Brazzini et al. 2026; Madau & Maiolino 2026) interpret this difference as evidence that LRDs are a $\sim 10 - 30\%$ sub-population of the broader low-luminosity BLAGN, with the bluer majority named “Little Blue Dots” (LBDs). The potential selection bias and the physical origin of UV emission in LRDs/LBDs are yet to be understood.

Here, we adopt one interpretation recently developed by Sun et al. (2026), in which the rest-frame UV emission of LRDs is dominated by their host galaxies, with only a $\sim 15\%$ contribution from the AGN at $\lambda \sim 1500\text{\AA}$. We therefore multiply the reported total UV luminosities of LRDs by this AGN fraction before comparing them to the AGN-only LF, corresponding to a shift of -0.8 dex in luminosity

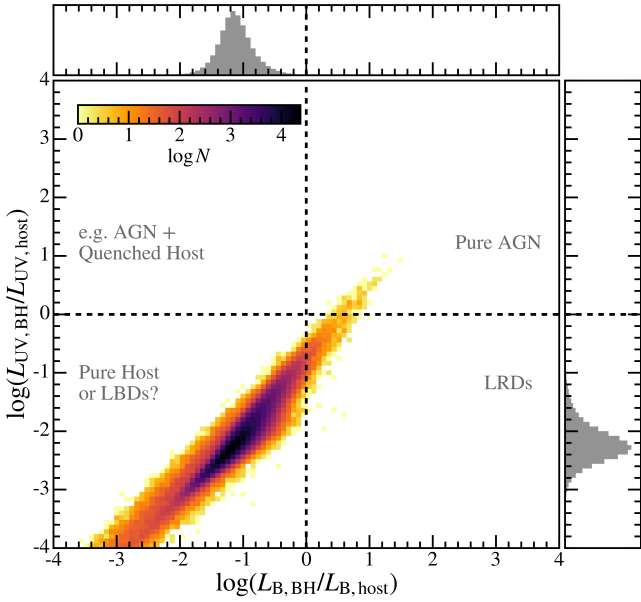


Figure 11. The luminosity ratio between BH and host galaxy in the UV versus in the optical B band at $z \approx 5$. We show the distribution of BHs below the LRD selection mass threshold $M_{\text{BH}}^{\text{crit}}$ on this plane. The majority of the BHs are dominated by their host galaxies in the UV and optical, which raises the concern that the fraction of sources that would pass observational selection criteria of LRDs is smaller than the f_{duty} implied. Therefore, a major remaining issue for the simulation is the overbright and overmassive host galaxies of LRDs.

(+2 mag in M_{UV}). This treatment is also consistent with the LRD SED model adopted in Section 3.3, which is based on stacked LRD spectra either after subtracting the host-galaxy component (Sun et al. 2026) or of sources most strongly dominated by the optical component (Pérez-González et al. 2026). After applying this empirical correction, we find that the faint-end predictions from LUMINA agree reasonably well with the inferred AGN contribution to the UV LFs of LRDs. We restrict our quantitative comparison to the LRD observations (second category mentioned above), while showing the BLAGN LFs only for context (applied the same correction factor in M_{UV} just to keep the relative position to LRD constraints). This is because the host-galaxy contribution in BLAGN (that are not classified as LRDs) could be much higher than inferred in Sun et al. (2026), and it is still possible that the UV emission is produced by leaked or post-processed radiation from the accreting SMBHs (e.g., Chen et al. 2025a; Madau & Maiolino 2026), which our empirical LRD SED model would fail to represent. In both cases, we underestimate the UV luminosities. It could be the reason that the observed BLAGN LFs are systematically above our model predictions as well as the LRD LFs.

Another caveat of our model is that the simulated host galaxies of the selected LRD-like BHs are typically too UV luminous. Their host-galaxy intrinsic UV luminosities are more than an order of magnitude higher than what is required to explain the total observed UV luminosities of LRDs. Dust attenuation is likely not sufficient to reconcile this tension: given the relatively blue UV slopes $\beta_{\text{UV}} \sim -2$ of LRDs (e.g., Greene et al. 2024; Kocevski et al. 2025), $A_{\text{UV}} \lesssim 0.5$ mag if one assumes a typical Meurer et al. (1999) IRX- β relation of host galaxies. At the same time, their optical luminosities are also too large for the systems to be identified as LRDs in real observations. For example, the median stellar mass and intrinsic UV magnitude of host

galaxies, for LRDs above $L_{\text{bol}} > 10^{43} \text{ erg s}^{-1}$, are around $10^9 M_{\odot}$ and -21 mag in LUMINA, compared to $10^{8.3} M_{\odot}$ and -18.5 mag reported in Sun et al. (2026), $5 \times 10^7 M_{\odot}$ reported in Matthee et al. (2025), and $10^{8.5} M_{\odot}$ reported in Lin et al. (2026a) using clustering measurements. This issue persists even when we select BHs with the largest optical luminosity ratios relative to their host galaxies.

We illustrate this tension in Figure 11, where we show the distribution of luminosity ratios between BHs and hosts in the UV versus in the B band for the LRDs in LUMINA. We use the LRD bolometric corrections of the “weak IR” model here, while the “strong IR” model will only make BHs even dimmer in both bands. We roughly divide the space into four regions. If BHs or hosts dominate in both bands, the source should appear as a pure AGN or galaxy. In the picture we described above, only the ones that are dominated by the BH in the optical and by the host galaxy in the UV will appear as LRDs. There are AGN in LUMINA that lie in the LRD regime, but the fraction of them is too small compared to the $f_{\text{duty}} = 0.3$ implied by our comparisons of LFs. Meanwhile, a substantial population of AGN is moderately dominated by their host galaxies, which in fact could appear as the LBDs. The conclusion here is that host galaxies of LRDs in LUMINA are overmassive and overbright to self-consistently explain the observed LRDs with the “V-shape” SED selection criteria. This remains a major unresolved issue in cosmological simulations such as LUMINA. As already highlighted in e.g. Matthee et al. (2025); Habouzit (2025), the host galaxy stellar masses of LRDs are significantly lower than predictions in a range of modern cosmological simulations and the values inferred from local scaling relations between SMBH and host masses. The tensions we find here are independent of the BH mass measurements, yet consistent with these findings.

4.4 X-ray LFs

Another critical validation of our model comes from the X-ray LFs of AGN. Hard X-ray emission in canonical AGN is expected to trace the compact corona associated with accretion onto BHs and is relatively less affected by contamination from host-galaxy starlight than the UV or optical bands (e.g., Osterbrock & Ferland 2006; Brandt & Alexander 2015; Hickox & Alexander 2018). At the same time, the X-ray band provides an important consistency check on our LRD interpretation. Most LRDs and *JWST*-selected faint AGN appear to be X-ray weak or undetected relative to expectations from standard AGN scaling relations (e.g. Yue et al. 2024; Ananna et al. 2024; Maiolino et al. 2025, see α_{ox} discussed in Section 3.2), although the recently identified “X-ray dot” (Hviding et al. 2026) may represent an important exception in a transitional phase. Nevertheless, we assume that all our selected LRD-like objects are X-ray dark. The key test of our model is whether the remaining canonical AGN population in the simulation can reproduce the observed X-ray LFs.

In Figure 12, we show the hard X-ray (2 – 10 keV) LFs of AGN at $z \approx 3 - 6$ in LUMINA, TNG100, and MTNG. For comparison, we show the observational constraints compiled in Shen et al. (2020a) based on pre-*JWST* X-ray surveys. These constraints predate the improved characterization of heavily obscured and CTK AGN uncovered by *NuSTAR* (e.g., Marchesi et al. 2018; Ananna et al. 2019; Tanimoto et al. 2022). We have updated our column density distribution model in Section 3.4 accordingly. For the comparisons here, we include new observational constraints from Ananna et al. (2019) and Pouliaxis et al. (2024). Ananna et al. (2019) constructed an AGN population-synthesis model designed to reproduce X-ray AGN number counts, the cosmic X-ray background, and the observed CTK AGN fraction, yielding an X-ray LF in the $(z, L_{\text{HX}}, N_{\text{H}})$ space. Pouliaxis et al. (2024)

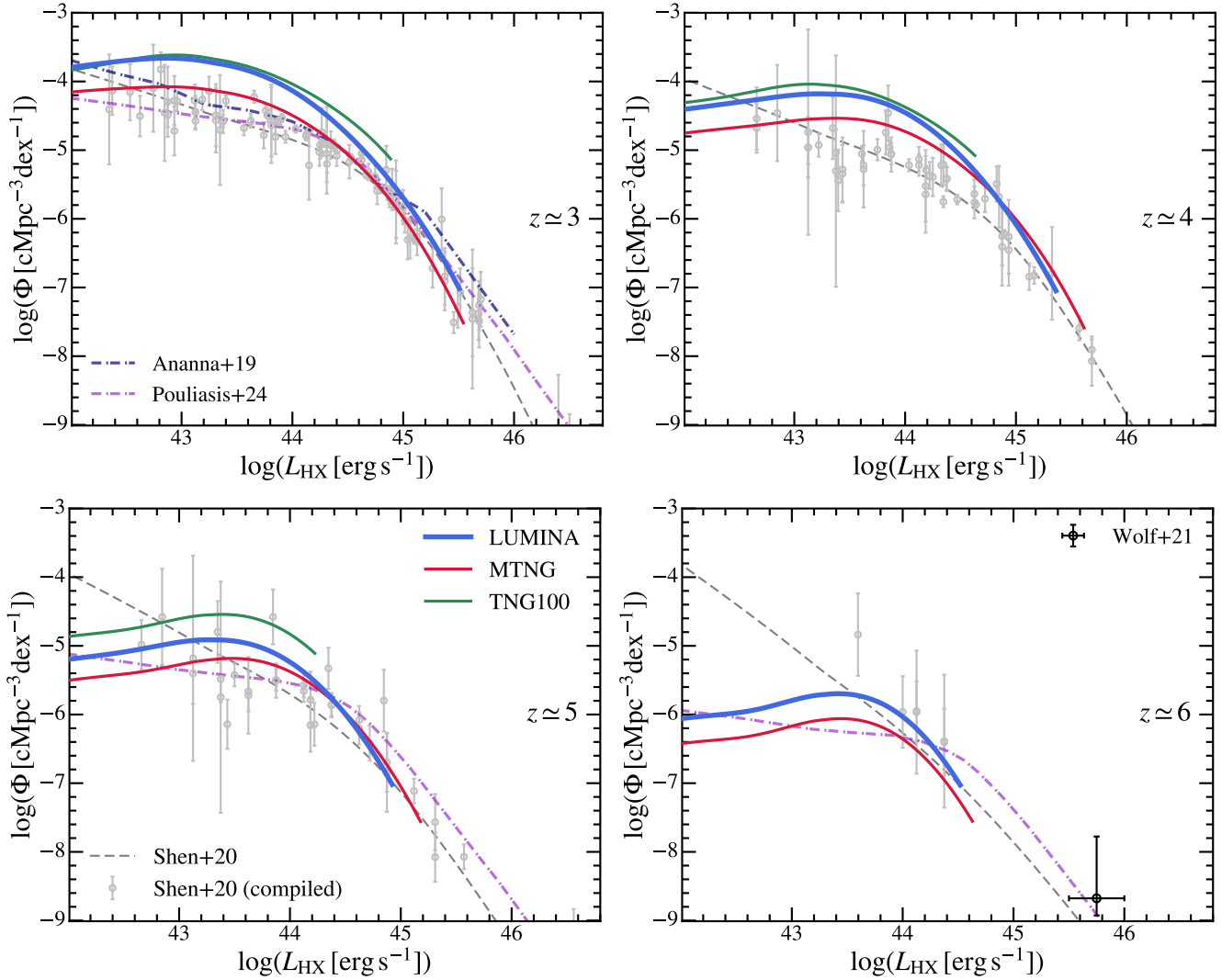


Figure 12. Hard X-ray (2 – 10 keV) LFs of AGN at $z = 3 - 6$ from LUMINA, TNG100, and MTNG. We note that the X-ray LFs here are only contributed by the non-LRD population we identify in the simulations, so there is no dependence on the f_{duty} and bolometric corrections of LRDs. We compare the simulation results with X-ray observations compiled in Shen et al. (2020a) along with more recent constraints from Ananna et al. (2019); Wolf et al. (2021); Pouliaisis et al. (2024). At $z \approx 6$, all BHs in TNG100 are below $M_{\text{BH}}^{\text{crit}}$ and therefore no prediction is presented.

derived the X-ray luminosity and absorption functions at $3 \leq z \leq 6$ using a large sample of soft X-ray-selected AGN from the Chandra Deep Fields, the Chandra COSMOS-Legacy survey, and the XMM-XXL northern field. We also include the $z \approx 6$ constraint from Wolf et al. (2021), based on an eROSITA-detected quasar, although this point lies close to or beyond the effective dynamical range sampled by our simulations.

As in many X-ray luminosity-function studies, the published L_{HX} values in these works are intrinsic, absorption-corrected luminosities. To compare them to our post-absorption predictions, we first generate the intrinsic model LF of Shen et al. (2020a) by setting $N_{\text{H}} = 0$. We then record the offset in $\log \Phi$ between the observational constraints and this intrinsic reference model, and apply the same offset to the post-absorption LF from Shen et al. (2020a). This procedure preserves the relative information carried by the observational measurements while placing them on the same luminosity definition as our model predictions.

Encouragingly, the simulation predictions reproduce the observed

X-ray LFs across the redshift range considered here. At the bright end, the agreement is similar to what we found in the optical and UV bands. Towards the faint end, the predicted X-ray LF flattens, broadly following the trend inferred from observations. The X-ray comparison supports a picture in which the canonical AGN population in LUMINA is broadly consistent with existing hard X-ray constraints, while the LRD-like population is largely hidden from current X-ray surveys. We overpredict the number density of faint X-ray AGN with $L_{\text{HX}} \lesssim 10^{44} \text{ erg s}^{-1}$ at $z \lesssim 4$. A similar excess was already seen in the bolometric LFs discussed in Section 4.1. However, this discrepancy should not necessarily be interpreted as a failure of the model. As will become clearer from the IR LFs below, current X-ray surveys can remain incomplete for canonical obscured AGN even before accounting for the LRD population.

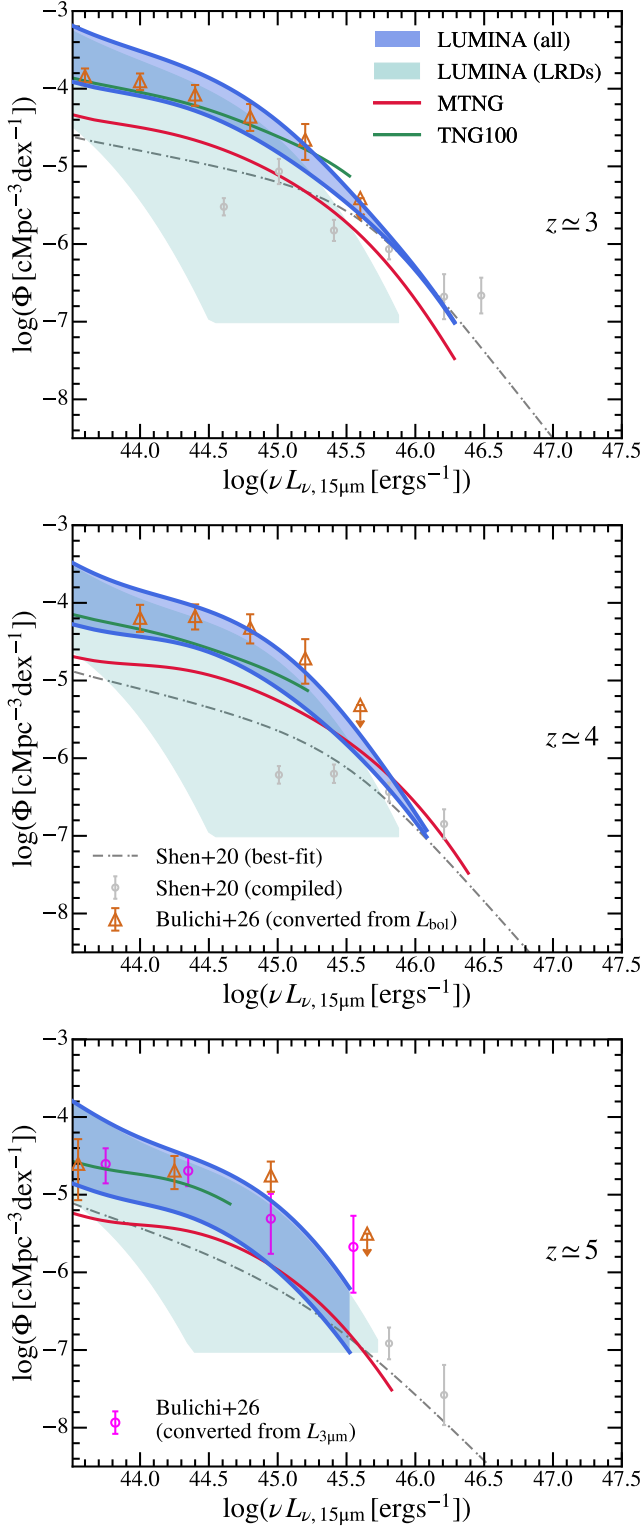


Figure 13. Rest-frame mid-IR band ($\sim 15\mu\text{m}$) LFs of AGN at $z = 3 - 5$ in LUMINA. Results from TNG100 and MTNG are shown for reference. For comparison, we include observational constraints compiled in Shen et al. (2020a). We also include *JWST* MIRI constraints from Bulichi et al. (2026), with data converted from their bolometric LFs or from their $3\mu\text{m}$ LFs assuming the canonical AGN SED (Krawczyk et al. 2013).

4.5 Mid-IR LFs

The final band we consider is the mid-IR. For canonical AGN, the near- to mid-IR emission is generally interpreted as accretion-disc radiation absorbed and reprocessed by circumnuclear dust, often associated with the dusty torus or more extended dusty polar structures (e.g., Antonucci 1993; Urry & Padovani 1995; Fritz et al. 2006; Netzer 2015; Ramos Almeida & Ricci 2017; Hickox & Alexander 2018), while the far-IR can receive substantial contributions from colder dust heated by star-formation in the host galaxy. As discussed in Section 3.3, the mid- to far-IR SEDs of LRDs remain highly uncertain, and even the physical origin of their IR emission is still debated. Some studies find little evidence for hot or cold dust emission in individual or stacked LRD samples, challenging simple dust-reddened AGN or dusty-starburst interpretations (e.g., Setton et al. 2025a; Chen et al. 2025a; Xiao et al. 2025). However, evidence for warm or hot dust emission in LRDs has been reported in stacking analyses (Delvecchio et al. 2025) and in individual sources (Ronayne et al. 2025; Barro et al. 2026). Therefore, the IR LF provides a complementary consistency check, similar in spirit to the X-ray LF. If LRDs are relatively faint in the mid-IR compared to canonical obscured AGN, then the IR LF should mainly validate whether the canonical AGN population in the simulation remains consistent with observations.

Observational constraints in this regime have recently become possible with deep *JWST*/MIRI surveys. Early MIRI results showed that mid-IR selection can reveal a population of heavily obscured high-redshift AGN missed by X-ray surveys (e.g., Yang et al. 2023; Pérez-González et al. 2024; Lyu et al. 2024). More recently, the MEOW survey (Leung in prep., Bulichi et al. 2026) used deep MIRI imaging, together with NIRCam, spectroscopic, and HST data, to identify obscured AGN out to $z \sim 6$ and infer a rapidly increasing obscured fraction towards high redshift. We explore two ways of converting the MEOW constraints to $15\mu\text{m}$ LF. The first is taking their bolometric LFs and applying a typical mid-IR bolometric correction of ~ 10 (see Table 2). The second is taking their $3\mu\text{m}$ LFs and converting that to $15\mu\text{m}$ with luminosity differences inferred from the Krawczyk et al. (2013) template we used for canonical AGN SED. They yield broadly consistent results.

In Figure 13, we show the mid-IR ($\sim 15\mu\text{m}$) LFs of AGN at $z \approx 3 - 5$ in LUMINA, compared to TNG100 and MTNG. We compare these predictions to the observational constraints compiled by Shen et al. (2020a), as well as to the recent MIRI-selected constraints from Bulichi et al. (2026). We find that the faint-end IR LF lies above the extrapolation of the best-fitting model from Shen et al. (2020a), but is more consistent with the higher AGN abundance inferred by Bulichi et al. (2026). This agreement is particularly encouraging because Bulichi et al. (2026) found a substantially larger obscured AGN fraction than previous UV-, optical-, X-ray-, or pre-*JWST* IR-selected surveys. We emphasize that this faint-end enhancement is not driven solely by the LRD-like population. In our “weak IR” model (the lower limit of predicted IR LF), the faint-end mid-IR LF remains substantially contributed, and in some luminosity ranges dominated, by canonical AGN in the simulation. Thus, the mid-IR comparison supports the interpretation that LUMINA contains a significant population of obscured but otherwise canonical AGN, in addition to the empirical LRD component introduced above.

4.6 The large-scale environment of AGN

So far, we have focused primarily on the LFs, and hence the abundances, of AGN. The large-scale environment of AGN provides an important and independent probe of their host DM haloes (e.g., Eilers

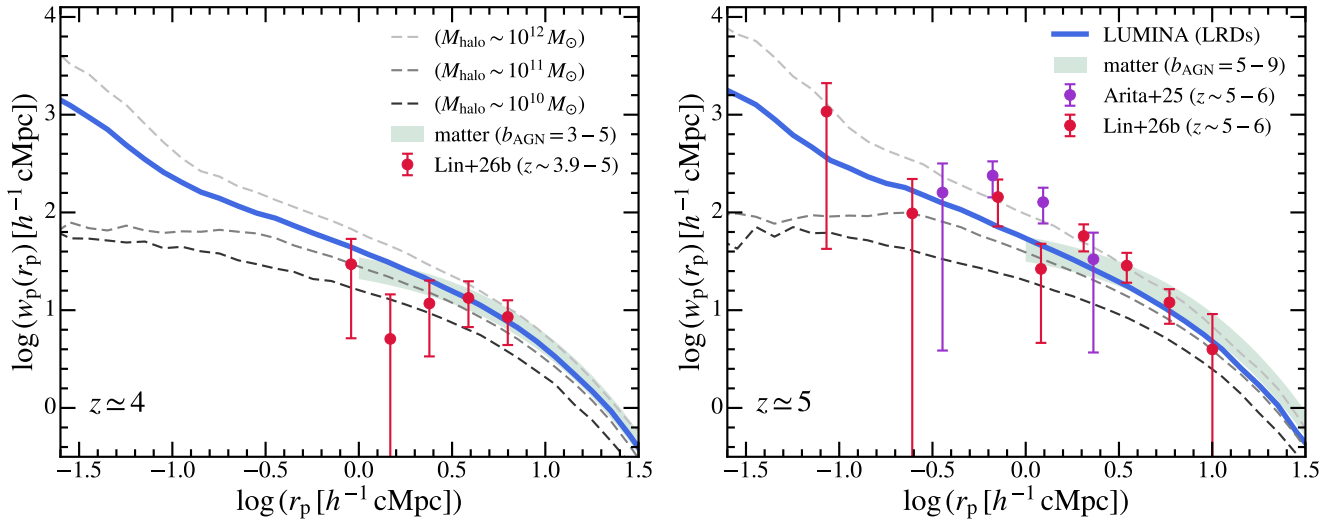


Figure 14. Projected cross-correlation function (w_p) of LRDs with galaxies at $z \approx 4$ (left panel) and $z \approx 5$ (right panel) in LUMINA. As discussed in the main text, we select galaxies with $\text{SFR} > 1.7 M_\odot \text{ yr}^{-1}$ to mimic the tracer galaxy sample used in observations. We further measure the bias factor of the selected galaxy sample and apply a $b_{\text{gal}}^{\text{sim}}/b_{\text{gal}}^{\text{obs}}$ correction to the observed w_p in Lin et al. (2026b) and Arita et al. (2025). For reference, we show the cross-correlation function of different mass haloes with the same galaxy sample in LUMINA. We also show the biased matter auto-correlation function assuming different bias factors of LRD. At $z \approx 4$, our LRD clustering signal agrees well with Lin et al. (2026b), suggesting host halo mass mildly above $10^{11} M_\odot$ and a bias factor ~ 4 . At $z \approx 5$, our LRD sample has a bias factor of ~ 5 and still lies in haloes slightly above $10^{11} M_\odot$, while the observed LRDs (Arita et al. 2025; Lin et al. 2026b) have larger bias factors and correlation functions close to our $10^{12} M_\odot$ haloes. Notably, the data point from Lin et al. (2026b) at the smallest r_p may suggest the one-halo term from clustering of satellite star-forming galaxies within massive haloes that only shows up when halo mass gets above $10^{11} M_\odot$ and is consistent with our results.

et al. 2024; Pizzati et al. 2024, 2025). This is particularly valuable for LRDs. Unlike luminosity- or SED-based arguments, which depend sensitively on bolometric corrections, dust attenuation, host-galaxy subtraction, and the physical origin of the broad-line emission, clustering directly traces the connection between LRDs and the underlying matter density field. Massive haloes originate from rare peaks in the Gaussian random density field and are expected to be more clustered compared to the matter density field (Kaiser 1984; Bardeen et al. 1986). Therefore, the large-scale autocorrelation or cross-correlation amplitude of a tracer population and the inferred bias (e.g., Mo & White 1996; Sheth et al. 2001; Cooray & Sheth 2002; Tinker et al. 2010) provides information about the characteristic host halo mass of AGN. This test is especially useful for distinguishing between different physical interpretations of LRDs.

Recent *JWST* observations have begun to make this test feasible. For instance, Matthee et al. (2025) used the deep NIRCcam Grism ALT survey to study the environments of faint broad $\text{H}\alpha$ emitters at $z \approx 4 - 5$. By comparing the number of neighbouring emission-line galaxies around LRDs to that around a reference sample of star-forming galaxies, they inferred host stellar masses of order $M_* \sim 5 \times 10^7 M_\odot$. This is much lower than what would be inferred from galaxy-only SED fits and implies large BH-to-stellar-mass ratios. Arita et al. (2025) measured the cross-correlation functions of low-luminosity BLAGN found by *JWST* with photometrically-selected galaxies at $z \approx 5 - 6$. They found that the typical host halo mass of these AGN is $\log(M_{\text{halo}}/M_\odot) \approx 11.6 - 11.7$. In a complementary analysis, Lin et al. (2026b) measured the large-scale environments of LRDs and their cross-correlation functions with $\text{H}\alpha$ emitters at $3.9 < z < 6$ in the GOODS-N field using the CONGRESS and FRESCO survey data. Their inferred characteristic host halo masses are $\log(M_{\text{halo}}/M_\odot) \approx 11.0 - 11.2$, with corresponding host stellar masses of $\log(M_*/M_\odot) \approx 8.4 - 8.6$.

In LUMINA, we compute the cross-correlation functions of LRDs

and tracer galaxies as

$$\xi_{\text{ag}}(r) = \frac{D_a D_g(r) - D_a R_g(r) - D_g R_a(r) + R_a R_g(r)}{R_a R_g(r)}, \quad (14)$$

where “a” and “g” represent AGN and tracer galaxies, respectively. $D_a D_g(r)$ is the number of cross pairs between the two types of sources at a distance r , while $D_a R_g(r)$, $D_g R_a(r)$, and $R_a R_g(r)$ are the corresponding pair counts involving randomly distributed sources. The auto-correlation function of LRDs or galaxies can be obtained in a similar way using the same source type for all terms. One can further define the projected cross-correlation function as

$$w_p(r_p) = 2 \int_0^{r_{\pi, \text{max}}} \xi(r_p, r_\pi) \text{d}r_\pi, \quad (15)$$

where r_p and r_π are the transverse and line-of-sight separations, respectively. We adopt the canonical choice in literature, $r_{\pi, \text{max}} = 8 h^{-1} \text{ cMpc}$. The code `CORRFUNC` (Sinha & Garrison 2019; Sinha & Garrison 2020) is used to facilitate this calculation.

The selection of the tracer galaxies is critical, as the cross-correlation function of AGN and galaxies in the linear perturbation theory can be decomposed as (see the review Desjacques et al. 2018)

$$\xi_{\text{ag}}(r) \sim \xi_{\text{mm}}(r) b_{\text{AGN}} b_{\text{gal}}, \quad (16)$$

where ξ_{mm} is the matter auto-correlation function, b_{AGN} and b_{gal} are the bias factors of AGN and tracer galaxies, respectively. Lin et al. (2026b) used $\text{H}\alpha$ emitters above the luminosity threshold $10^{41.5} \text{ erg s}^{-1}$ as the tracer galaxies. We convert this to a cut in SFR of $1.7 M_\odot \text{ yr}^{-1}$, assuming the conversion factors from Murphy et al. (2011); Kennicutt & Evans (2012); Shen et al. (2020b). We first compute the auto-correlation function of galaxies above this SFR threshold from LUMINA and obtain b_{gal} by comparing it to ξ_{mm} at $r = 8 h^{-1} \text{ cMpc}$ at the redshift of interest. The calculation of ξ_{mm}

is done using `COLOSSUS` (Diemer 2018), assuming the Eisenstein & Hu (1998) transfer function. The b_{gal} of tracer galaxies in LUMINA is 4.8 at $z \approx 5$ and 3.7 at $z \approx 4$. For comparison, the galaxy sample in Lin et al. (2026b) has $b_{\text{gal}} = 5.9$ at $5 < z < 6$ and $b_{\text{gal}} = 4.11$ at $3.9 < z < 5$. The galaxy sample selected in Arita et al. (2025) has $b_{\text{gal}} = 5.2$ at $5 < z < 6$. Therefore, to account for the remaining differences in tracers and obtain a clean comparison of the clustering power of AGN, we will apply a correction factor $b_{\text{gal}}^{\text{sim}}/b_{\text{gal}}^{\text{obs}}$ to the correlation functions reported in observations.

In Figure 14, we compare the projected cross-correlation functions between LRDs (with $L_{\text{bol}} > 10^{43} \text{ erg s}^{-1}$) and selected star-forming galaxies in LUMINA at $z \approx 4$ and $z \approx 5$ with the measurements from Lin et al. (2026b)¹ and Arita et al. (2025)². The simulation predictions agree well with the observed amplitudes of LRD-galaxy clustering at both redshifts, which provides an important validation of the model beyond LFs alone. To translate the clustering amplitude into an approximate LRD host halo mass scale, we further compare the LRD-galaxy cross-correlation to the cross-correlations between DM haloes of different masses and the same star-forming-galaxy sample. The observed and simulated LRD clustering amplitudes are best matched by haloes with characteristic masses slightly above $10^{11} M_{\odot}$. This inferred halo mass is consistent with the estimate from Lin et al. (2026b) and Arita et al. (2025), supporting a picture in which LRDs are associated with relatively common, moderate-mass haloes rather than rare massive quasar host haloes with $M_{\text{halo}} \sim 10^{12} - 10^{13} M_{\odot}$ (e.g., Arita et al. 2023; Eilers et al. 2024; Pizzati et al. 2024). It is also worth noting that the data point from Lin et al. (2026b) at the smallest r_p may reflect the one-halo term arising from the clustering of satellite star-forming galaxies within massive haloes, a feature that only emerges once the halo mass exceeds $10^{11} M_{\odot}$, and which is consistent with our results. We note that our LRD-galaxy cross-correlation is almost identical to the galaxy auto-correlation function in our sample, so we show only the former in the figure. This is consistent with what has been found observationally (Arita et al. 2025; Lin et al. 2026b), although it is not too surprising in our case, since the BH seeding and the LRD model depend only on halo/BH mass with no additional requirement on the halo environment.

In observational studies, halo mass is usually inferred from clustering by comparing it to ξ_{mm} , measuring the bias, and applying an analytical bias-mass relation (e.g., Tinker et al. 2010). Motivated by this, we show the LRD-galaxy cross-correlation function by multiplying ξ_{mm} with the b_{gal} measured in our galaxy sample and different values of b_{AGN} . The LRDs selected in LUMINA have bias factor ≈ 4 at $z \approx 4$ and ≈ 5 at $z \approx 5$. For comparison, Lin et al. (2026b) found $b_{\text{AGN}} \approx 3$ at $3.9 < z < 5$ and ≈ 8 at $5 < z < 6$. Arita et al. (2025) found $b_{\text{AGN}} \approx 6.6$ at $5 < z < 6$. In general, these values are mildly higher than our predictions. It is interesting that the host halo masses they infer are consistent with ours, suggesting that the discrepancy may arise from inaccuracies in the tracer galaxy bias estimates or in the bias-halo mass relation they adopt.

At the same time, clustering should not be interpreted purely as a one-to-one measure of halo mass. The clustering of galaxies and haloes is not determined by halo mass alone: halo assembly bias (Gao et al. 2005; Gao & White 2007) can modify the clustering amplitude

¹ Lin et al. (2026b) originally measured the volume-weighted projected cross-correlation function. We find that it can be converted to w_p with a simple multiplication factor 16.

² Arita et al. (2025) used $r_{\pi, \text{max}} = 50 h^{-1} \text{ cMpc}$ when measuring w_p . Empirically, based on LUMINA, we find that increasing $r_{\pi, \text{max}}$ from $8 h^{-1} \text{ cMpc}$ to $50 h^{-1} \text{ cMpc}$ has < 0.1 dex impact on w_p at $r_p \lesssim 2 h^{-1} \text{ cMpc}$ where Arita et al. (2025) data reach.

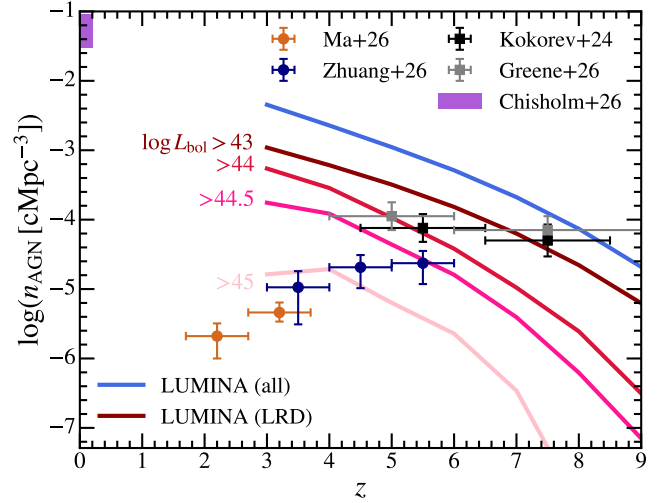


Figure 15. Number density of AGN versus redshift. We show the number density of all AGN with $L_{\text{bol}} \geq 10^{43} \text{ erg s}^{-1}$, and LRDs with different thresholds of L_{bol} as labelled. For comparison, we show observational constraints for LRDs from Kokorev et al. (2024); Greene et al. (2026) at $z \geq 5$ and Ma et al. (2026); Zhuang et al. (2026) at lower redshifts. The constraints from Kokorev et al. (2024); Greene et al. (2026) reach down to $L_{\text{bol}} \sim 10^{43} - 10^{43.5} \text{ erg s}^{-1}$ with updated bolometric corrections, while Ma et al. (2026); Zhuang et al. (2026) reach $L_{\text{bol}} \sim 10^{44} - 10^{44.5} \text{ erg s}^{-1}$. At $z = 0$, we show the LRD remnant estimation from Chisholm et al. (2026) (we apply the same f_{duty} factor just for comparison with our high-redshift LRD predictions), which considers LRDs as globular clusters in formation.

at fixed mass if LRD formation preferentially selects haloes with particular secondary properties (Wang et al. 2026), such as early formation time, high concentration, or low spin. This issue is especially relevant because several proposed LRD formation channels are tied to unusual halo environments or assembly histories, such as the metal-free haloes exposed to strong Lyman-Werner radiation from nearby companions in direct-collapse BH scenarios (e.g., Bromm & Loeb 2003; Begelman et al. 2006), and concentrated early-forming haloes in DM-seeding scenarios (e.g., Xiao et al. 2021; Shen et al. 2025a; Jiang et al. 2026; Roberts et al. 2026). Therefore, the agreement in clustering should be interpreted with some caution. We cannot rule out the possibility that LRDs reside in somewhat less massive DM haloes with secondary properties that enhance their large-scale bias. In that case, the duty cycle required to match the luminosity-function constraints would likely be smaller than the value inferred under a mass-only halo-bias interpretation.

4.7 AGN in the cosmological context

The comparisons above focus on LFs in individual bands and on the large-scale environments of AGN. We now place the model in a broader cosmological context by examining the redshift evolution of several integrated properties of SMBHs.

4.7.1 Redshift evolution of AGN number density

In Figure 15, we show the comoving number density of all AGN with $L_{\text{bol}} \geq 10^{43} \text{ erg s}^{-1}$ as a function of redshift. We show the LRD population selected by our empirical model with different L_{bol} thresholds as labelled. On the observational side, Ma et al. (2026); Zhuang

et al. (2026) reported low-redshift searches of LRDs using ground-based and NIRCam/WFSS Grism surveys to an observational limit of $M_{5500} \sim -20.5$, which corresponds to $L_{\text{bol}} \approx 10^{44.5} - 10^{45} \text{ erg s}^{-1}$ assuming our LRD bolometric corrections. In addition, for LRD searches at $z \geq 5$, we take the sum of the binned estimations in Kokorev et al. (2024); Greene et al. (2026) to obtain LRD number densities. These high-redshift observations typically have lower detection limits. In Greene et al. (2026), the lowest L_{bol} bin extends to $\sim 10^{43.5} \text{ erg s}^{-1}$. In Kokorev et al. (2024), the lowest L_{bol} bin centers on $10^{44} \text{ erg s}^{-1}$ and would reduce by a factor of ~ 10 to $10^{43} \text{ erg s}^{-1}$ if one adopts updated bolometric corrections (see the comparison in Greene et al. 2026).

In LUMINA, the total AGN number density increases monotonically towards lower redshift. The LRD population follows a similar evolution pattern, while the brightest LRD abundance declines faster. At $z \approx 3 - 5$, the predicted number density of observable LRDs is consistent with the measurements from Zhuang et al. (2026) and higher than the ones from Ma et al. (2026) assuming a consistent L_{bol} limit. At $z \geq 5$, the LRD abundance reported in other *JWST* surveys is more consistent with the LRDs in LUMINA, assuming a consistent L_{bol} limit of $10^{43} - 10^{44} \text{ erg s}^{-1}$ (Kokorev et al. 2024; Greene et al. 2026). We do not find the rise-and-decline trend of LRD number density versus redshift, as found in observational studies and predicted in other theoretical works (e.g., Pacucci & Loeb 2025; Inayoshi 2025). From our analysis, it is possible that observational bias (deeper detection limits of *JWST* compared to ground-based surveys targeting LRDs at lower redshift) contributes to the decline of LRD abundance at $z \leq 4$, and the observed samples may represent only the bright tail of a larger underlying population. However, the full story is likely more complicated since LRD analogues are rare in the local Universe (e.g., Lin et al. 2025, 2026c; Ji et al. 2026b), while the number density of SMBHs below $M_{\text{BH}}^{\text{crit}}$ should keep increasing dramatically towards lower redshifts.

This apparent tension indicates several key missing ingredients in our BH seeding prescription. In LUMINA, BHs continue to be seeded in low-mass haloes once they cross the seeding threshold, even at relatively low redshifts. This may be somewhat unphysical if BH seeding occurs only under special physical conditions at high redshifts, such as the extremely metal-poor environment that is necessary for the formation of direct-collapse BHs (e.g., Rees 1984; Bromm & Loeb 2003; Begelman et al. 2006; Lodato & Natarajan 2006) and/or supermassive stars (e.g., Baumgarte & Shapiro 1999; Volonteri & Begelman 2010; Hosokawa et al. 2013; Woods et al. 2017). Meanwhile, the LRD phenomenon itself, i.e., the activation of the LRD phase in low-mass SMBHs, may be directly tied to the early-Universe environment. Failing to capture these conditions in the model could lead to an overprediction of low-mass, faint AGN at later times. A recent work (Chisholm et al. 2026) links LRDs to the formation of globular clusters and estimates a local LRD remnant number density of $\sim 0.1 - 0.3 \text{ cMpc}^{-3}$, which we include in the figure for comparison, applying the same $f_{\text{duty}} = 0.3$ as for our high-redshift LRDs. Their inferred number density lies above even the extrapolation of our results. The key difference here is that they associate LRDs with the early-formation phase of globular clusters, which implies that f_{duty} should evolve to a substantially lower value than we adopt.

4.7.2 The Soltan argument

The Soltan argument attributes the SMBH mass density in the local Universe to the integrated radiatively efficient accretion (manifest as AGN) over cosmic time (e.g., Soltan 1982; Yu & Tremaine 2002;

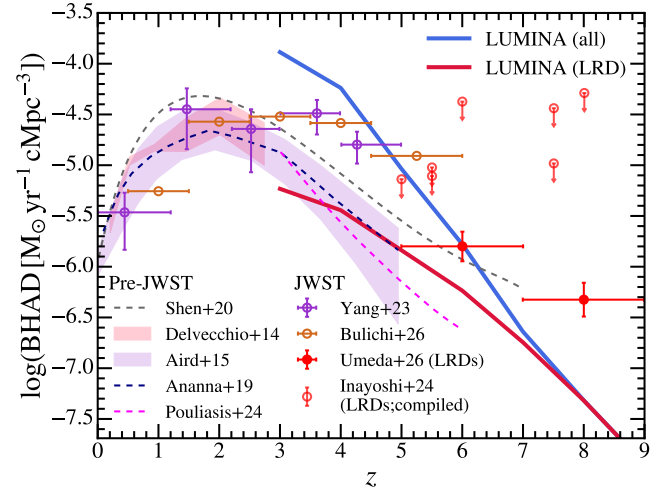


Figure 16. SMBH accretion rate density (BHAD) versus redshift, obtained through a summation of accretion rates of BHs in LUMINA in the bolometric luminosity range $L_{\text{bol}} = 10^{43} - 10^{48} \text{ erg s}^{-1}$. We show the BHAD contributed by all AGN and by LRDs in LUMINA and compare them to LRD constraints compiled in Inayoshi & Ichikawa (2024) (originally taken from Matthee et al. 2024; Greene et al. 2024; Kokorev et al. 2024; Akins et al. 2025b, treated as upper limits because canonical quasar bolometric corrections have been applied) and Umeda et al. (2026). At lower redshifts, X-ray constraints from Aird et al. (2015); Ananna et al. (2019); Pouliaxis et al. (2024) and mid-IR constraints from Delvecchio et al. (2014); Yang et al. (2023); Bulichi et al. (2026) are shown for reference. Many of these works assume $\epsilon_{\text{rad}} = 0.1$, and we have corrected them to be consistent with $\epsilon_{\text{rad}} = 0.2$ assumed in LUMINA.

Marconi et al. 2004; Hopkins et al. 2007; Shankar et al. 2009; Shen et al. 2020a). It is a critical integral constraint for SMBH models. In Figure 16, we show the SMBH accretion rate density (BHAD) as a function of redshift. Commonly, in LF studies, this quantity can be computed by integrating the bolometric LF and converting luminosity to BH mass growth as

$$\dot{\rho}_{\text{BH}} = \frac{1 - \epsilon_{\text{rad}}}{\epsilon_{\text{rad}} c^2} \int_{L_{\text{min}}}^{L_{\text{max}}} L_{\text{bol}} \Phi(L_{\text{bol}}) d \log L_{\text{bol}}, \quad (17)$$

where L_{min} and L_{max} are usually taken to be 10^{43} and $10^{48} \text{ erg s}^{-1}$. From simulations, however, we can directly compute BHAD as the summation of the accretion rates of BHs within the bolometric luminosity range. A finite-volume correction would be required at the bright end, where the LUMINA box cannot directly sample the rarest quasars up to $L_{\text{bol}} = 10^{48} \text{ erg s}^{-1}$. We estimate this missing contribution by integrating the bolometric quasar LF of Shen et al. (2020a) from the maximum BH luminosity in LUMINA to the integration limit. We find that it accounts for less than 10% of the total BHAD even at $z \approx 7$ and is smaller at lower redshifts. Neglecting it, therefore, does not impact our conclusions.

We find that the total BHAD in LUMINA rises rapidly from $z \geq 8$ to $z \approx 3$. Over the redshift range where comparisons are available, the predicted total BHAD is higher than constraints derived from X-ray (Aird et al. 2015; Ananna et al. 2019; Pouliaxis et al. 2024) and pre-*JWST* mid-IR surveys (Delvecchio et al. 2014). Our predictions are more consistent with more recent *JWST*/MIRI survey results (Yang et al. 2023; Bulichi et al. 2026). Many of these works assume $\epsilon_{\text{rad}} = 0.1$. Since the inferred $\dot{\rho}_{\text{BH}} \propto (1 - \epsilon_{\text{rad}})/\epsilon_{\text{rad}}$, we have applied a correction factor to these observationally inferred BHAD to be consistent with the $\epsilon_{\text{rad}} = 0.2$ assumed in LUMINA. The LRD contri-

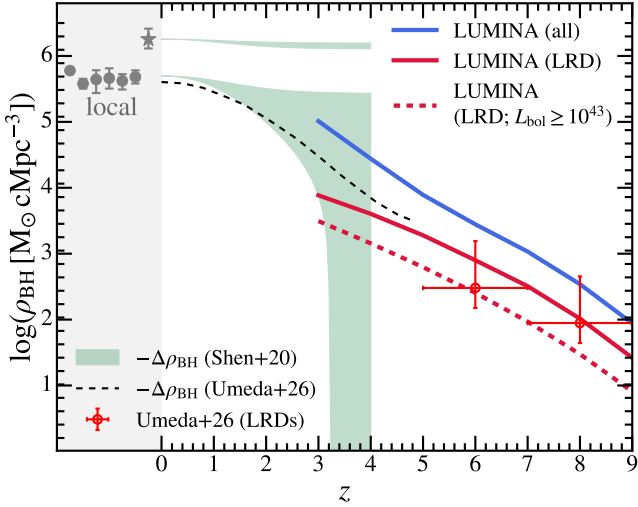


Figure 17. SMBH mass density versus redshift. We show the mass density of all SMBHs, LRDs, and LRDs with $L_{\text{bol}} \geq 10^{43}$ erg s^{-1} . The Local Universe constraints are taken from Shankar et al. (2004, 2009); Marconi et al. (2004); Graham & Driver (2007); Hopkins et al. (2007); Yu & Lu (2008) (grey circles) and Liepold & Ma (2024) from the MASSIVE survey (grey star), which suggests a substantially higher local BH mass density. The green band shows a back propagation of the BH mass density by subtracting the contribution by quasars in Shen et al. (2020a) through redshifts from the local value, assuming $\epsilon_{\text{rad}} = 0.1 - 0.2$. We show both the results assuming the canonical local $\rho_{\text{BH}} \approx 5 \times 10^5 M_{\odot} \text{cMpc}^{-3}$ and assuming the updated $\rho_{\text{BH}} \approx 1.8 \times 10^6 M_{\odot} \text{cMpc}^{-3}$ in Liepold & Ma (2024). We show the results derived in a similar way from Umeda et al. (2026) using X-ray LFs for reference. For LRDs, we show the observational estimates from Umeda et al. (2026).

bution is lower than the total AGN BHAD but remains non-negligible at $z \gtrsim 4$, indicating that the LRD phase can contribute appreciably to early SMBH growth without dominating the entire accretion budget. We also compare to observational estimates of the LRD accretion rate density obtained in Inayoshi & Ichikawa (2024), which are based on measurements from Matthee et al. (2024); Greene et al. (2024); Kokorev et al. (2024); Akins et al. (2025b). These estimates typically adopt canonical quasar bolometric corrections and should be regarded as an approximate upper limit. Umeda et al. (2026) provides an updated estimate using a BH-envelope model, with a similar resulting LRD SED to us. The LRD BHAD in LUMINA lies below the upper-limit compilation and closer to the lower values inferred by Umeda et al. (2026). This suggests that LRDs can be numerous and optically prominent without requiring an unrealistically large global SMBH growth rate or creating a severe tension with the BH mass density budget (Inayoshi & Ichikawa 2024).

In Figure 17, we show the corresponding SMBH mass density as a function of redshift. We compute the total SMBH mass density in LUMINA, the mass density of BHs in LRD phase, and the mass density in actively visible LRDs with $L_{\text{bol}} \geq 10^{43}$ erg s^{-1} . We examine the consistency of the model with the Soltan argument. We start from the local SMBH mass-density estimates from Shankar et al. (2004); Marconi et al. (2004); Graham & Driver (2007); Hopkins et al. (2007); Yu & Lu (2008); Shankar et al. (2009), although the recent MASSIVE survey suggests a substantially higher local SMBH mass density than past measurements (Liepold & Ma 2024), which is independently supported by the large amplitude of the nanohertz gravitational-wave background reported by the pulsar timing arrays (Agazie et al. 2023; Sato-Polito et al. 2025). We first take a canonical

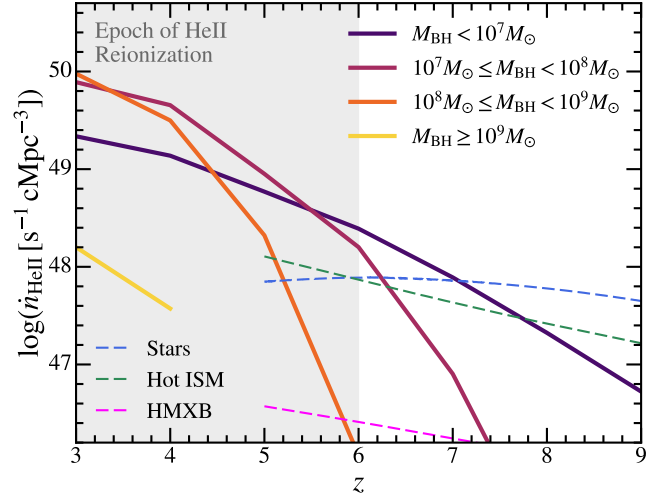


Figure 18. Comoving production rate density of He II-ionizing photons at 54.42 – 100 eV, $\dot{n}_{\text{ion}}^{\text{HeII}}$, as a function of redshift in LUMINA. We show decomposed emissivity from SMBHs in four mass bins. The vertical shaded band marks the He II-reionization epoch in LUMINA, $z \approx 3 - 6$. We also show the contribution from other sources to $z \approx 5$, which are subdominant for He II reionization. They are therefore not tracked at $z < 4.75$ (see Section 2.3). The main drivers for He II reionization in LUMINA are SMBHs in the mass range $10^7 - 10^9 M_{\odot}$.

estimate of $\rho_{\text{BH}} \sim 5 \times 10^5 M_{\odot} \text{cMpc}^{-3}$, and back-propagate this local SMBH mass density by subtracting the integrated quasar accretion history from Shen et al. (2020a), assuming $\epsilon_{\text{rad}} = 0.1$ and 0.2 to obtain lower and upper limits of ρ_{BH} at high redshifts. A similar exercise was carried out by Umeda et al. (2026) using X-ray LFs, and they obtained similar results to us. The total SMBH mass density in LUMINA remains below the local value and evolves in a manner broadly compatible with the back-propagated constraints. The LRD population contributes a substantial fraction of the high-redshift SMBH mass density, but itself lies comfortably below any SMBH mass budget constraints. The predicted LRD mass density is consistent with the observational estimates in Umeda et al. (2026).

However, if we take the updated local SMBH mass density $\rho_{\text{BH}} \sim 1.8 \times 10^6 M_{\odot} \text{cMpc}^{-3}$ from Liepold & Ma (2024), the back-propagated mass density at high redshifts is substantially higher than the value from LUMINA, and is in fact too large to be reconciled with existing AGN LF constraints. Closing this gap may require a substantial fraction of SMBH mass to be assembled through obscured or radiatively inefficient accretion that is missed by the optical/UV/X-ray AGN census at low redshift. This is qualitatively consistent with our findings in Section 4.5 that both the observed mid-IR LF and the predictions from LUMINA are higher than the pre-*JWST* census, but our SMBH density comparison here shows that it is still not sufficient to explain the deficit. One cannot rule out the possibility that the Liepold & Ma (2024) result itself is biased by selection effects in the dynamically-measured SMBH samples that anchor the massive end of the local scaling relations.

4.7.3 The role of AGN in He II reionization

While stars in low-mass galaxies are widely believed to dominate the reionization of hydrogen (e.g., Faucher-Giguère et al. 2008a; Kuhlen & Faucher-Giguère 2012; Robertson et al. 2015; Onoue et al. 2017; Parsa et al. 2018), AGN have long been considered the principal

drivers of the reionization of singly ionized helium, owing to their hard EUV spectra extending well beyond $4\text{Ryd} \approx 54.42\text{ eV}$ (e.g., Miralda-Escudé 1998; Madau et al. 1999; Schirber & Bullock 2003; Faucher-Giguère et al. 2008b, 2009; McQuinn et al. 2009; Haardt & Madau 2012; Worseck et al. 2014, 2016; Basu et al. 2024). The large abundance of LRDs discovered by *JWST* at $z \gtrsim 4$ has revived the long-standing question of whether AGN could also contribute appreciably to hydrogen reionization (e.g., Madau & Haardt 2015; Haardt & Salvaterra 2015; Madau et al. 2024; Asthana et al. 2025), although models in which AGN drive both phase transitions remain in tension with the observed $\text{Ly}\alpha$ forest and IGM thermal history (Asthana et al. 2025).

In *LUMINA*, the joint cosmic radiation field of stars and AGN is followed self-consistently with on-the-fly radiative transfer (Section 2.3). As described in Zier et al. (2026), *LUMINA* predicts a He II reionization history that begins around $z \approx 6$ and completes near $z \approx 3$, broadly consistent with existing constraints from the He II $\text{Ly}\alpha$ forest (e.g., Worseck et al. 2014, 2016, 2019) and the IGM thermal history (Lidz et al. 2010; Gaikwad et al. 2021). Here, we leverage this self-consistent picture to dissect the AGN contribution to He II reionization in greater detail, asking how SMBHs of different masses and evolutionary phases share the photon budget.

We compute the comoving production rate density of He II-ionizing photons by summing the bolometric luminosities of all accreting SMBHs in the simulation volume,

$$\dot{n}_{\text{ion}}^{\text{HeII}} \approx \frac{\mathcal{F}_{\text{HeII}}}{\langle \epsilon_{\text{HeII}} \rangle V_{\text{box}}} \sum_i f_{\text{esc}}(L_{\text{bol}}^i) L_{\text{bol}}^i, \quad (18)$$

where $\mathcal{F}_{\text{HeII}} = 0.0455$ is the fraction of the bolometric luminosity emerging at the He II-ionizing band (54.42 – 100 eV) given the EUV spectral shape adopted in our canonical AGN SED model (Section 3.2; Lusso et al. 2015; Shen et al. 2020a), $\langle \epsilon_{\text{HeII}} \rangle \approx 74.83\text{ eV}$ is the mean energy in the band, V_{box} is the comoving volume of the simulation, and the sum runs over active SMBH particles. $f_{\text{esc}}(L_{\text{bol}})$ is the luminosity-dependent escape fraction introduced in Section 2.3.

In Figure 18, we show the resulting He II-ionizing photon emissivity in *LUMINA*, decomposed into four BH mass bins. We also show the contribution from stars, hot ISM, and HMXBs at $z \gtrsim 5$ for comparison. It is clear that AGN dominate the production of He II-ionizing photon since $z \gtrsim 7$ and justifies our choice to deactivate other sources at $z < 4.75$ (see a more thorough discussion in Zier et al. 2026). During the bulk of He II reionization at $z \lesssim 6$, the photon budget is dominated by SMBHs in the range $10^7 - 10^9 M_{\odot}$, with only a modest contribution from BHs above $10^9 M_{\odot}$ that reflects their rarity at these epochs. At the beginning of He II reionization, emissivity is dominated by BHs with mass $10^7 - 10^8 M_{\odot}$ while more massive BHs at $10^8 - 10^9 M_{\odot}$ become equally important closer to the end of reionization. Only at $z \gtrsim 6$, when the most massive SMBHs have not yet assembled, do BHs below $10^7 M_{\odot}$ contribute an appreciable share of the ionizing emissivity. The characteristic bolometric luminosity of $\sim 10^8 M_{\odot}$ BHs is $L_{\text{bol}} \sim L_{\text{Edd}} \sim 10^{46}\text{ erg s}^{-1} \sim 10^{12.5} L_{\odot}$, which is aligned with the conventional wisdom that BHs at the knee of the quasar LF (Shen et al. 2020a) drive the He II reionization (e.g., Sokasian et al. 2002; Furlanetto & Oh 2008; McQuinn et al. 2009; Yung et al. 2021).

This mass dichotomy has a useful consequence for our empirical AGN model. The mass scale below which BHs do not contribute substantially to the ionizing emissivity coincides with our adopted LRD threshold, $M_{\text{BH}}^{\text{crit}} \sim 10^7 M_{\odot}$ (Section 3.5). The EUV spectra of LRDs are observationally poorly constrained: most LRDs are remarkably X-ray faint or undetected, and the dense gas inferred from their optical spectra suggests that their intrinsic EUV continuum

may differ substantially from the canonical quasar template adopted in Section 2.3. Reassuringly, Figure 18 demonstrates that LRD-mass BHs contribute negligibly to the He II-ionizing photon budget over the redshift range where He II reionization actually proceeds in *LUMINA*. Our conclusion that AGN drive He II reionization is therefore essentially independent of how the EUV emission of LRDs is modelled and rests on the well-characterized population of $\gtrsim 10^7 M_{\odot}$ canonical AGN.

5 DISCUSSION

5.1 AGN variability is essential for early SMBH growth

As discussed in Section 3.1 above, one fundamental limitation of any cosmological simulation is that many relevant physical processes related to gas inflow and BH accretion operate many orders of magnitude below the resolution limit. The Bondi–Hoyle prescription effectively delivers a “smooth” accretion rate set by the kpc-scale gas properties, while real AGN are observed to vary over timescales of hours to decades or longer. Quantitatively, in large optical quasar samples, the asymptotic structure-function amplitudes (which quantify the total light-curve variability) are typically $\sim 0.1 - 0.4$ mag on rest-frame timescales of $\sim 10^2 - 10^3$ days (e.g., Vanden Berk et al. 2004; Kelly et al. 2009; MacLeod et al. 2010; Caplar et al. 2017; Burke et al. 2021). In hard X-rays, where variability traces the corona on much shorter timescales, fractional rms amplitudes of a few tens of percent are commonly observed on timescales of hours to days (e.g., Vaughan et al. 2003; González-Martín & Vaughan 2012; Ponti et al. 2012), corresponding to ~ 0.1 dex scatter in luminosity. On the much longer timescales relevant to the lifetime of an individual quasar ($\sim 0.1 - 100$ Myr), the observed ERDF can span several orders of magnitude in λ_{Edd} (e.g., Hopkins & Hernquist 2009; Hickox et al. 2014; Aird et al. 2018; Sartori et al. 2018), with effective log-normal dispersions of $\sim 0.3 - 0.4$ dex (Kollmeier et al. 2006; Kauffmann & Heckman 2009; Kelly et al. 2010; Kelly & Shen 2013; Lusso et al. 2012). Our adopted $\sigma_{\text{bol}} = 0.3$ dex therefore sits naturally between the short-timescale optical/X-ray variability amplitude and the long-timescale ERDF dispersion.

The $\sigma_{\text{bol}} = 0.3$ dex log-normal scatter applied in our empirical model (Section 3.1) is an essential and physically motivated ingredient. It captures unresolved accretion variability and translates the simulated mean accretion rates into a realistic distribution of observed AGN luminosities. However, we caution that this treatment remains simplified in several aspects: (i) The scatter is applied to the bolometric luminosity, so it does not just reflect intrinsic accretion-rate variability but may also absorb fluctuations in the obscuration properties. (ii) The assumed log-normal kernel is itself a simplification, whereas the observed ERDFs of AGN are often better described by a Schechter form, with a power-law tail towards low Eddington ratios and an exponential cutoff near the Eddington limit (e.g., Hopkins & Hernquist 2009; Schulze & Wisotzki 2010; Aird et al. 2018). (iii) We adopt a constant scatter across the full luminosity range, whereas a luminosity- or mass-dependent kernel may provide a better description (see Figure 19 for a hint).

Variability is also likely vital for early SMBH growth and quenching galaxies. Recent zoom-in simulations have shown that allowing bursty super-Eddington episodes, physically motivated by photon-trapping in slim discs, can grow $\sim 10^4 M_{\odot}$ seeds to $\sim 10^{8-9} M_{\odot}$ within several hundred Myr, in regimes where Eddington-limited Bondi–Hoyle accretion would fail by orders of magnitude (e.g., Lupi et al. 2024; Pacucci & Narayan 2024; Huško et al. 2025). The duty

cycle of these super-Eddington bursts is short (often $\lesssim 10$ Myr), but the integrated mass growth is dominated by them. It offers a more efficient pathway to assembling the most massive quasars known at $z \gtrsim 6-7$ with $M_{\text{BH}} \gtrsim 10^9 M_{\odot}$ (e.g., Mortlock et al. 2011; Bañados et al. 2018; Wang et al. 2018; Wu et al. 2022), whose existence has long challenged growth from $\lesssim 10^{4-5} M_{\odot}$ seeds under standard Eddington-limited accretion (e.g., Volonteri & Begelman 2010; Smith & Bromm 2019; Inayoshi et al. 2020). The recent *JWST* detections of BLAGN out to $z > 10$ (Bogdán et al. 2024; Maiolino et al. 2025) only sharpen this issue. In addition, in many modern AGN feedback prescriptions (including the IllustrisTNG model on which LUMINA is based), AGN feedback transitions from a radiatively-coupled thermal mode to a substantially more efficient kinetic mode once the BH mass exceeds a characteristic threshold of a few times $10^8 M_{\odot}$ (Sijacki et al. 2007; Weinberger et al. 2017; Shen et al. 2025b), which quenches massive galaxies in these models (e.g., Terrazas et al. 2020; Piotrowska et al. 2022). A stochastic early growth could let BHs enter the low-accretion rate kinetic mode earlier and more frequently (even below the mass threshold), and accelerate the quenching of massive host galaxies and potentially help to explain the population of quiescent massive galaxies now identified by *JWST* at $z \gtrsim 4$ (e.g., Carnall et al. 2023; Looser et al. 2024; Glazebrook et al. 2024; de Graaff et al. 2025b). Exploring these possible consequences of a bursty early growth phase of SMBHs will require more self-consistent adjustments to the BH accretion and feedback recipes in simulations rather than the simple empirical scatter of luminosity, and is deferred for future work.

The amplification of bright-quasar abundances by intrinsic variability is closely analogous to the amplification of the high-redshift galaxy UV LF by bursty star-formation. *JWST* surveys at $z \gtrsim 10$ revealed an excess of UV-bright galaxies relative to pre-*JWST* extrapolations (e.g., Finkelstein et al. 2023; Harikane et al. 2023a; Finkelstein et al. 2024; Naidu et al. 2026), and a leading explanation is that the rest-UV luminosities of low-mass high-redshift galaxies vary stochastically by $\sigma_{\text{UV}} \simeq 1-2$ mag driven by the short feedback-modulated SFR cycles seen in zoom-in simulations (e.g., Mirocha & Furlanetto 2023; Shen et al. 2023, 2024; Mason et al. 2023; Sun et al. 2023; Semenov et al. 2025a,b). Despite potentially very different relevant astrophysical timescales, the conceptual lesson is the same: in regimes where the underlying source population is steeply declining at the bright end and where short-timescale fluctuations are large, time variability is a central component of the observable LF.

5.2 Summary of explicit & implicit assumptions about LRDs

We have described above how LUMINA, paired with an empirical model connecting SMBHs to observable AGN, can reproduce several key properties of high-redshift quasars and LRDs. Here, we provide a more direct summary of the explicit and implicit assumptions entering this framework. The most important explicit assumptions are:

- (i) In LUMINA, we place $8 \times 10^5 h^{-1} M_{\odot}$ seed BH particles in every halo that exceeds $5 \times 10^{10} h^{-1} M_{\odot}$ and does not already contain a BH. It provides the endpoint of the unresolved BH growth below the seed mass.
- (ii) For BH particles below $M_{\text{BH}}^{\text{crit}} = \kappa M_{\text{seed}} = 1.18 \times 10^7 M_{\odot}$ ($\kappa = 10$), a fraction $f_{\text{duty}} = 0.3$ is assumed to appear observationally as LRDs. This intuitively associates LRDs with the early growth phase of SMBHs.
- (iii) We assume Bondi–Hoyle accretion onto BH particles, with-

out an additional boost factor, limited by the Eddington rate, and adopt a constant radiative efficiency $\epsilon_{\text{rad}} = 0.2$ for all BHs.

- (iv) We assume an intrinsic unresolved scatter in instantaneous BH luminosities of $\sigma_{\text{bol}} = 0.3$ dex, which phenomenologically accounts for accretion variability below the resolution scale of the simulation.

The remaining ingredients are mainly observationally motivated, including the SEDs and bolometric corrections of canonical AGN and LRDs, and the column-density distribution of canonical AGN, which determines absorption and dust extinction corrections. In this sense, our model separates the problem into two parts: the simulation provides the abundance, growth, and environments of BH particles, while the empirical model specifies how these BH particles appear in different wavelength bands.

The physical interpretation of f_{duty} is broader than its implementation suggests. For example, because LUMINA seeds BHs in every halo above the adopted threshold, f_{duty} effectively absorbs the true occupation fraction of BH seeds in early haloes. Alternatively, a timescale-based interpretation of f_{duty} emerges if we assume that essentially every newly seeded BH passes through a single LRD-like episode and is observable while in that phase. In that limit, $f_{\text{duty}} \sim t_{\text{LRD}}/t_{\text{low}}$, where t_{LRD} is the duration of the LRD-emitting phase and t_{low} is the total time a BH spends below $M_{\text{BH}}^{\text{crit}}$. Both time scales can be set from physical models. For example, in the late-stage quasi-star scenario (Begelman et al. 2008; Begelman & Dexter 2026), the LRD phase coincides when a central BH swallowing a convectively powered envelope and lasts of order the Salpeter time $t_{\text{Sal}} \equiv \epsilon_{\text{rad}} M_{\text{BH}} c^2 / [(1 - \epsilon_{\text{rad}}) L_{\text{Edd}}] \simeq 112.5$ Myr (assuming $\epsilon_{\text{rad}} = 0.2$), modulo a factor that depends on the envelope-to-BH mass ratio at envelope dispersal (Begelman 2010; Volonteri & Begelman 2010; Coughlin & Begelman 2024). The lifetime t_{low} is set by the time required for the BH to grow from M_{seed} to κM_{seed} at its average accretion rate. For fast but Eddington-limited growth with $\langle \lambda_{\text{Edd}} \rangle \sim 0.5-1$ and our fiducial $\kappa = 10$, we have $t_{\text{low}} \sim t_{\text{Sal}} \ln \kappa / \langle \lambda_{\text{Edd}} \rangle \simeq 260$ Myr in LUMINA. Therefore, one predicts $f_{\text{duty}} \sim 0.2-0.4$, comfortably bracketing our fiducial choice. Alternative scenarios, in which LRDs correspond to recurrent super-Eddington accretion bursts each lasting only a few Myr (Inayoshi 2025), are also consistent with $f_{\text{duty}} = 0.3$ provided that the bursts repeat a few times during the LRD-eligibility window, as those models in fact predict.

The BH particle in LUMINA should be viewed as a sub-grid tracer of an unresolved proto-nuclear ecosystem that may include the seed BH itself, a surrounding gas reservoir, and possibly a compact nuclear star cluster. Several physical models have been proposed for what such systems look like, including but not limited to the quasi-star picture mentioned above, and an intermediate-mass BH growing rapidly inside a dense nuclear star cluster (or a cluster of clusters) fed by runaway stellar collisions and mergers (e.g., Devecchi & Volonteri 2009; Devecchi et al. 2012; Davies et al. 2011; Katz et al. 2015; Shi et al. 2024; Dekel et al. 2025; Pacucci et al. 2025). LUMINA resolves the large-scale environment, halo assembly, and global gas supply common to all of these scenarios, but does not resolve the internal radiative and dynamical processes that distinguish them. Our empirical model for AGN is deliberately agnostic to which picture applies. The broad consistency with the quasi-star timescale estimates discussed above is encouraging in this context, but is not by itself sufficient to discriminate between the underlying scenarios.

Several implicit assumptions also enter the empirical model. The transition between the LRD-like and canonical AGN phases is treated as a sharp mass cut, although in reality it likely depends on additional properties such as gas density, metallicity, and halo assembly

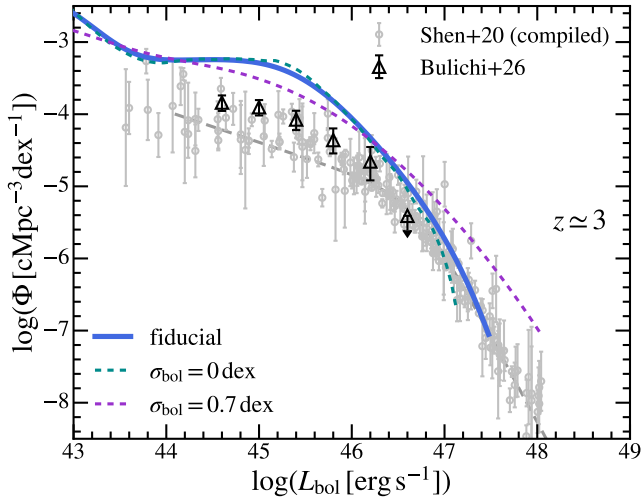


Figure 19. Bolometric LFs of AGN at $z \approx 3$ in LUMINA, assuming $\sigma_{\text{bol}} = 0$ and 0.7 dex, compared to the fiducial model. We choose $z \approx 3$ for this comparison because it is the redshift where LUMINA predictions have most overlap with the bright-end observational constraints. A lower (higher) σ_{bol} results in under- (over-) prediction of the bright-end LF, while the faint end is less affected. A luminosity- or mass-dependent σ_{bol} may result in a better agreement at both the bright end and $L_{\text{bol}} \sim 10^{45} \text{ erg s}^{-1}$, where LUMINA overpredicts the AGN number density.

history. The discrepancy with the observed decline of the LRD number density at lower redshifts (Section 4.7) may be one consequence. We further assume that the LRD SED is approximately universal, independent of redshift, BH mass, luminosity, and host-galaxy properties, and that the same effective parameters apply across $z \approx 3 - 6$. These choices are deliberately simple as the main goal of this work is to test whether a plausible LRD-like phase can be embedded in a large-volume cosmological simulation while remaining consistent with a wide range of AGN observables and constraints. The agreement found above demonstrates that such a region of parameter space exists, although it may not be unique.

5.3 The impact of model parameters and their degeneracies

The empirical model presented in Section 3 introduces a small set of free parameters: the unresolved bolometric luminosity scatter σ_{bol} , the LRD mass threshold $\kappa = M_{\text{BH}}^{\text{crit}}/M_{\text{seed}}$, and the LRD duty cycle f_{duty} . In this subsection, we examine how each of these parameters individually shapes the predicted multi-band AGN LFs, and discuss potential degeneracies between them and with other implicit assumptions of the simulation. For brevity, we focus on the bolometric LF at $z \approx 3$, where LUMINA predictions overlap most strongly with the bright-end observational constraints, and on the optical and hard-X-ray LFs at $z \approx 5$, where most LRD observational constraints are.

In Figure 19, we show the bolometric LFs at $z \approx 3$ for three values of σ_{bol} bracketing our fiducial value of 0.3 dex. The bright end of the LF responds strongly to σ_{bol} . Removing the scatter ($\sigma_{\text{bol}} = 0$) underpredicts the abundance of luminous AGN at $L_{\text{bol}} \sim 10^{47} \text{ erg s}^{-1}$. This is a minor effect here for LUMINA, but becomes critical for larger-volume simulations such as MTNG, which we included in the comparisons above. On the other hand, adopting a higher $\sigma_{\text{bol}} = 0.7$ dex severely overpredicts the bright-end LF. The faint end, however, is much less sensitive to σ_{bol} . This asymmetry is anticipated as a log-normal scatter applied to a steeply declining bright-end LF pref-

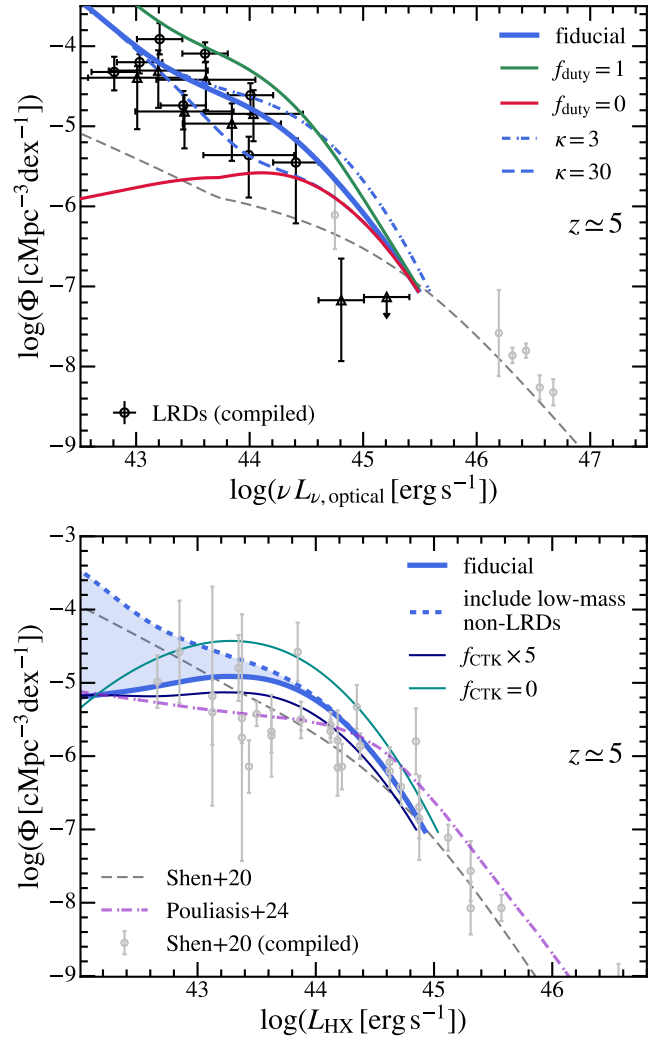


Figure 20. *Top:* AGN B-band optical LF at $z \approx 5$ in LUMINA, for several combinations of f_{duty} and κ . For clarity, we show only the predictions using the “weak IR” bolometric correction of LRDs. The observational constraints of LRDs compiled in Section 4.2 are shown here for comparison. f_{duty} controls the normalization of the faint-end LF that is dominated by LRDs, while κ mainly influences the transitional luminosity between LRD and canonical AGN LFs. *Bottom:* AGN hard X-ray LF at $z \approx 5$ in LUMINA. In contrast to the optical, f_{duty} has no impact on the X-ray LF, because LRDs are intrinsically X-ray weak in our model. The dominant uncertainty here is whether the $(1 - f_{\text{duty}})$ population of BHs below $M_{\text{BH}}^{\text{crit}}$ is included as canonical AGN or treated as electromagnetically silent; this choice substantially affects the predicted X-ray LF at $L_{\text{HX}} \lesssim 10^{43} \text{ erg s}^{-1}$. We also show the predictions assuming different values of f_{CTK} , which mildly affects the normalization of the X-ray LF.

entially scatters faint, abundant sources up to high luminosities, while a shallow faint-end LF is essentially invariant under the same convolution.

In the top panel of Figure 20, we show the rest-frame B-band LF at $z \approx 5$ for several variations of κ and f_{duty} around our fiducial values. For clarity, we show only the predictions using the “weak IR” model. The two parameters are physically distinct and shape the predicted LF in complementary ways. The threshold κ controls which simulated BHs are eligible for staying in the LRD phase. Through the LRD bolometric correction at roughly fixed Eddington ratio, it sets the

maximum LRD luminosity in the optical, as well as the transition luminosity of the LRD- and canonical-AGN-dominated branches of the LF. Our fiducial value $\kappa = 10$ ($M_{\text{BH}}^{\text{crit}} \sim 10^7 M_{\odot}$) is chosen to coincide with the sharp cut-off of LRD number densities at brighter luminosities (e.g., Ma et al. 2025). Raising κ to 30 – 100 extends the LRD branch into the bright-AGN regime and produces a high-luminosity bump, while lowering κ truncates the LRD branch at lower luminosities, both of which are less consistent with the data. On the other hand, the duty cycle f_{duty} controls the fraction of eligible BHs that are in the LRD phase at any given epoch and therefore modulates the normalization of the LRD-dominated faint end without affecting the maximum LRD luminosity. Variations in either parameter leave the bright end of the LF, dominated by canonical AGN, essentially unchanged.

In the bottom panel of Figure 20, we show the hard X-ray LF at $z \approx 5$. Because LRDs in our model are assumed X-ray dark (Section 3.3), f_{duty} itself has no impact on the X-ray LF. The major uncertainty in X-rays is whether the $(1 - f_{\text{duty}})$ population of BHs below $M_{\text{BH}}^{\text{crit}}$ that are not in the LRD phase at a given epoch is assigned the canonical AGN SED or is instead treated as electromagnetically silent. This treatment of the off-LRD duty cycle low-mass BHs can substantially alter the predicted X-ray LFs at $L_{\text{HX}} \lesssim 10^{43} \text{ erg s}^{-1}$. Assigning them the canonical AGN SED produces a significantly larger faint-end abundance. Nevertheless, the current X-ray observational constraints cannot distinguish the two scenarios. We note that this population has almost no influence on the optical/UV LFs since LRDs strongly dominate over dust-obscured canonical AGN at the faint end. The physical nature of these off-LRD duty cycle BHs in our model is highly uncertain. Future deep, high-angular-resolution X-ray surveys, in particular with the AXIS (Reynolds et al. 2023) and NewAthena (Cruise et al. 2025) mission concepts, will provide the most direct route to constraining the transition between the LRD phase and the broader low-mass AGN population at $z \gtrsim 4$. In this figure, we also show the X-ray LFs assuming different CTK AGN fraction (f_{CTK}) in our N_{H} distribution model. Removing the CTK population ($f_{\text{CTK}} = 0$) or increasing the f_{CTK} boost to 5 has a moderate impact on the X-ray LF at $\sim 0.2 - 0.5$ dex level.

Beyond the parameters of the empirical model, the predicted AGN LFs also depend on the BH accretion physics built into LUMINA itself. In particular, the Bondi–Hoyle accretion rate adopted in the simulation does not include the multiplicative “boost factor” α that is commonly applied in lower-resolution cosmological simulations, e.g., $\alpha \sim 100$ in Illustris (Springel et al. 2005; Sijacki et al. 2015) and ASTRID (Bird et al. 2022; Ni et al. 2022). An ad-hoc accretion boost above unity (along with the removal or relaxation of the Eddington cap) would shift the predicted BH mass and bolometric LFs to the more massive and more luminous end (e.g., Booth & Schaye 2009; Negri & Volonteri 2017; Bennett et al. 2024). To preserve a match to the observed LRD LFs, both $M_{\text{BH}}^{\text{crit}}$ and f_{duty} would then need to be reduced, and more luminous BHs would correspondingly be placed in lower-mass DM haloes and host galaxies. As noted in Section 4.3, the simulated host galaxies of LRD-mass BHs in LUMINA are often too optically luminous to satisfy the observational LRD selection criteria, and also too UV luminous, overshooting the UV LF. Therefore, this re-calibration may help relieve the AGN-host relation tension. It would also reduce the clustering bias of LRDs, but only modestly, and is unlikely to spoil the clustering comparison in Section 4.6 given the large statistical uncertainties in current observational measurements (e.g., Matthee et al. 2024; Arita et al. 2025; Lin et al. 2026b). A more complete future treatment of high-redshift AGN demographics will likely require a physically motivated seeding and accretion model

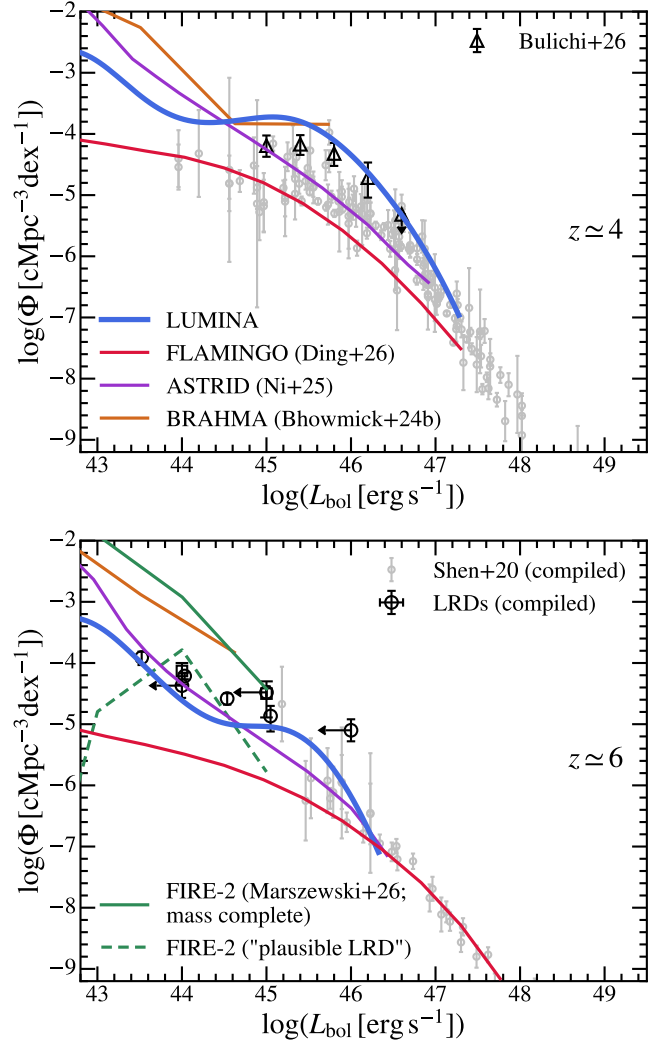


Figure 21. Comparison of AGN bolometric LFs at $z \approx 4$ and 6 between LUMINA and ASTRID (Ni et al. 2022, 2025), FLAMINGO (Schaye et al. 2023; Ding et al. 2026), BRAHMA (Bhowmick et al. 2024b), and FIRE-2 (a “passive” BH model; Marszewski et al. 2026). We include the observational data compiled in Section 4.1 for comparison.

combined with a self-consistent treatment of super-Eddington phases (as we discussed in Section 5.1).

5.4 Comparison to other cosmological simulations

Cosmological simulations adopt substantially different sub-grid prescriptions for SMBH seeding, accretion, and feedback, yielding distinct predictions for properties of SMBHs and their host galaxies both in the local Universe and at higher redshifts (see e.g. the comparisons in Habouzit et al. 2021, 2022). In this section, we compare the bolometric AGN LF predicted by LUMINA with results from a representative set of recent cosmological simulations covering the high-redshift AGN population: ASTRID (Bird et al. 2022; Ni et al. 2022, 2025), FLAMINGO (Schaye et al. 2023; Ding et al. 2026), BRAHMA (Bhowmick et al. 2024a,b), and the post-processing FIRE-2 analysis of Marszewski et al. (2026). These simulations differ substantially in volume, mass resolution, BH seeding, accretion, and feedback pre-

criptions. We briefly summarize each model below before turning to the comparison.

ASTRID (Bird et al. 2022; Ni et al. 2022) is a $(250 h^{-1} \text{cMpc})^3$ smoothed-particle hydrodynamics simulation with 2×5500^3 resolution elements run with MP-Gadget (Feng et al. 2018). BHs are seeded in haloes with $M_{\text{halo}} \gtrsim 5 \times 10^9 h^{-1} M_{\odot}$ and $M_{\star} \gtrsim 2 \times 10^6 h^{-1} M_{\odot}$, and the BH seed masses are drawn probabilistically from a power-law distribution between $3 \times 10^4 h^{-1} M_{\odot}$ and $3 \times 10^5 h^{-1} M_{\odot}$. BHs grow via Bondi–Hoyle accretion with an $\alpha = 100$ boost factor to account for the underestimation of the accretion rate due to the unresolved cold and hot phase of the sub-grid ISM nearby. Mildly super-Eddington accretion is allowed, but the accretion rate is still capped at two times the Eddington rate. AGN feedback is modelled as an isotropic thermal feedback until $z \sim 2.3$ when the kinetic mode feedback is turned on. The LFs shown in Figure 21 are taken from the most recent ASTRID release (Ni et al. 2025), from which we take the results assuming $\epsilon_{\text{rad}} = 0.2$.

FLAMINGO (Schaye et al. 2023) is a suite of large-volume SPH simulations run with swIFT (Schaller et al. 2024). The flagship run covers a $(2.8 \text{Gpc})^3$ volume with 5040^3 baryonic particles and a baryon mass resolution of $1.07 \times 10^9 M_{\odot}$, with smaller boxes at higher resolution. The sub-grid prescriptions for stellar and AGN feedback are calibrated using machine learning to reproduce the low-redshift galaxy stellar mass function and cluster gas fractions. BHs are seeded with $M_{\text{seed}} = 10^5 M_{\odot}$ in FoF haloes with $M_{\text{halo}} \gtrsim 2.757 \times 10^{11} M_{\odot}$. BHs accrete at a Bondi–Hoyle rate with a density-dependent boost factor (Booth & Schaye 2009) and are capped at the Eddington rate. In their fiducial model, AGN feedback is modelled thermally. They assume $\epsilon_{\text{rad}} = 0.1$. Ding et al. (2026) apply an additional log-normal scatter of 0.75 dex in post-processing to the bolometric luminosities before constructing the AGN LFs shown here, motivated by the comparison to the Shen et al. (2020a) compilation at $z \lesssim 3$.

BRAHMA (Bhowmick et al. 2024a) is a suite of small-volume cosmological simulations run with AREPO using a sub-grid galaxy formation model close to that of IllustrisTNG. The recent addition (Bhowmick et al. 2024b) is designed specifically to systematically explore heavy-seed (direct-collapse-like) channels for the formation of the first massive BHs. Heavy seeds with $M_{\text{seed}} \approx 10^5 h^{-1} M_{\odot}$ are placed in haloes that satisfy combinations of gas-density, metallicity, spin, Lyman-Werner-flux, and environmental criteria, in $(18 \text{cMpc})^3$ boxes. To be specific, we take their “SM5_LW10” model for comparison. The small simulation volume restricts the BRAHMA LFs to $L_{\text{bol}} \lesssim 10^{45} \text{erg s}^{-1}$ in Figure 21.

The FIRE-2 high-redshift suite (Ma et al. 2018, 2019) consists of zoom-in simulations of individual haloes at $z \gtrsim 5$ using the FIRE-2 model (Hopkins et al. 2018). Active BHs are not followed self-consistently. Instead, in Marszewski et al. (2026), hypothetical BHs are placed at the centres of resolved galaxies in post-processing (“passive” BH model), and their accretion rates are computed from the central gas distribution using either a gravitational-torque-driven accretion prescription (Hopkins & Quataert 2011; Anglés-Alcázar et al. 2017a,b) or a simple free-fall accretion model. Bolometric luminosities are derived assuming $\epsilon_{\text{rad}} = 0.1$. We take two illustrative LFs from their gravitational-torque-driven accretion model: a “mass-complete” LF that places a BH at the centre of every resolved host galaxy above the simulation completeness threshold, and a “plausible LRD” scenario that further restricts to BHs that are super-Eddington-accreting, Eddington-luminosity-limited in host galaxies with $M_{\star} > 2 \times 10^7 M_{\odot}$.

This paper based on LUMINA differs from these efforts in the following aspects. First, LUMINA strikes a favourable balance between simulation volume and numerical resolution: it resolves low-mass

galaxies and faint AGN more reliably than the FLAMINGO flagship run, while sampling rare luminous quasars more completely than BRAHMA, FIRE-2 zoom-in volumes, and even ASTRID. Second, the on-the-fly multi-frequency radiative transfer through H I and He II reionization couples the AGN ionizing emissivity to the IGM thermal and ionization state self-consistently. The He II reionization is driven by the same AGN population studied here. Third, rather than restricting the comparison to BH accretion rates and bolometric LFs, we develop an empirical model that maps simulated SMBHs onto multi-band AGN observables, and demonstrate that a coherent picture of the high-redshift AGN, including LRDs, can be established within a single framework.

Returning to the bolometric LFs in Figure 21, as discussed in earlier sections, LUMINA agrees well with the Shen et al. (2020a) compilation at the bright end and with the *JWST* constraints at the faint end. ASTRID also reproduces the Shen et al. (2020a) compilation at $z \approx 4$ and matches LUMINA closely at $z \approx 6$, but underpredicts compared to the *JWST*/MIRI constraints at $z \approx 4$. We note that the ASTRID results plotted here adopt a post-processing radiative efficiency $\epsilon_{\text{rad}} = 0.2$ rather than the $\epsilon_{\text{rad}} = 0.1$ used in the original simulation, which improves the bright-end agreement at the expense of internal consistency. FLAMINGO underpredicts the bolometric LF across the full luminosity range at $z \approx 4$, plausibly reflecting the resolution limitations highlighted in our earlier LUMINA–MIRI comparison. At $z \approx 6$, FLAMINGO recovers the bright end but still underpredicts at the faint-end. BRAHMA is restricted by its small simulation volume and only gives predictions at the faint end. A large abundance of low-luminosity AGN at both redshifts is found in this simulation, which sits above the observational constraints of LRDs at $L_{\text{bol}} \sim 10^{44} \text{erg s}^{-1}$. Finally, the mass-complete FIRE-2 LF overpredicts the faint-AGN abundance by roughly a decade, similar to BRAHMA, which indicates a rapid fuelling of gas to central BHs even in bursty and clumpy star-forming galaxies in FIRE-2. Their subsample restricted to super-Eddington-accreting, Eddington-luminosity-capped BHs in relatively massive hosts (the “plausible LRD” sample) reproduces the *JWST* LRD LF at around $L_{\text{bol}} \sim 10^{44} - 10^{45} \text{erg s}^{-1}$.

6 CONCLUSIONS

In this work, we use the large-volume cosmological radiation-hydrodynamic simulation LUMINA to study the high-redshift AGN population. The balance of large volume and high resolution offered by LUMINA allows us to simultaneously model some rare luminous quasars, the more abundant moderate-luminosity AGN, and the low-mass SMBH population that may be connected to the recently discovered LRDs. We construct an empirical mapping between the simulated SMBH population and observed AGN across multiple wavelength bands, and test whether the resulting AGN population is consistent with the available observational constraints from LFs, large-scale environments, and integrated properties of AGN. As described in Section 3, this empirical model starts from the SMBH population in LUMINA and applies a modest unresolved log-normal scatter to the instantaneous bolometric luminosity, with fiducial value $\sigma_{\text{bol}} = 0.3$ dex (Section 3.1). The simulated BHs are then split into two classes by mass: those above a critical mass threshold $M_{\text{BH}}^{\text{crit}} = \kappa M_{\text{seed}}$ are assigned a canonical quasar SED with luminosity-dependent absorption/extinction corrections (Sections 3.2 and 3.4), while those below the threshold may instead enter an LRD-like early-growth phase characterized by an empirical universal LRD SED (Section 3.3) and a

duty-cycle parameter f_{duty} (Section 3.5). Our fiducial model adopts $\kappa = 10$ and $f_{\text{duty}} = 0.3$. Our main results are as follows.

(i) **Bolometric LFs** (Section 4.1). The bolometric AGN LFs predicted by LUMINA agree well with pre-*JWST* multi-wavelength constraints at $z \simeq 3 - 6$ for moderately luminous quasars. The bright end is broadly converged between LUMINA and MTNG, while the faint end remains mildly resolution-dependent when compared to TNG100. The stochastic luminosity scatter is essential for matching the observed abundance of luminous AGN and can be interpreted as an effective treatment of unresolved accretion variability below the resolution scale of the simulation. The faint-end bolometric LF in LUMINA matches more recent *JWST* constraints that include the contribution from LRDs and MIRI-selected dust-obscured AGN.

(ii) **Optical and UV LFs** (Sections 4.2 and 4.3). LUMINA reproduces the LFs of both the canonical AGN population and the LRDs in the rest-frame optical and UV. In the optical, the model prediction agrees with the large abundance of LRDs at $L_{\text{opt}} \sim 10^{43} - 10^{44} \text{ erg s}^{-1}$ while smoothly connecting to the pre-*JWST* constraints on canonical AGN at the bright end. The comparison of optical LFs motivates our fiducial choice of κ and f_{duty} . In the UV, the simulation prediction again agrees with pre-*JWST* constraints at the bright end, while the comparison to LRDs depends sensitively on the interpretation of the physical origin of the UV emission. We explore one scenario in which the LRD UV emission is significantly contributed by the host galaxies. An empirical decomposition of AGN and host-galaxy light in UV inferred from recent observations results in good agreement between LUMINA predictions and observed LRD UV LFs. However, the simulated hosts of LRD-mass BHs are often too UV- and optically luminous to satisfy observational LRD selection criteria, highlighting an important limitation of theoretical models in reproducing the AGN-host galaxy scaling relations at high redshifts.

(iii) **X-ray and mid-IR LFs** (Sections 4.4 and 4.5). The hard X-ray LFs of canonical AGN (assuming no LRD contribution) in LUMINA broadly reproduce the observed LFs at $z \simeq 3 - 6$, including the flattening towards the faint end. The model slightly overpredicts faint X-ray AGN at $z \lesssim 4$, although this should be interpreted with caution given that current X-ray surveys may still miss a larger population of heavily obscured or intrinsically X-ray weak AGN than previously accounted for. In the mid-IR, LUMINA appears to contain a more complete sample of AGN compared to extrapolations of pre-*JWST* LF constraints, in agreement with recent *JWST*/MIRI constraints that capture even the X-ray weak and previously missed sources.

(iv) **Large-scale environments of LRDs** (Section 4.6). The projected cross-correlation between LRDs and star-forming galaxies in LUMINA at $z \simeq 4$ and $z \simeq 5$ agrees well with recent *JWST* survey results. The clustering amplitudes are consistent with characteristic LRD host halo masses slightly above $10^{11} M_{\odot}$, similar to those inferred observationally. Therefore, the LRD-like population appears to be associated with relatively common, moderate-mass haloes rather than the rare massive haloes hosting luminous quasars. This serves as an important cross-check of our empirical LRD model, which is largely calibrated based on optical LRD LFs. However, halo assembly bias may complicate a direct mapping between clustering amplitude and halo mass if LRD formation preferentially selects haloes with unusual secondary properties.

(v) **AGN in the cosmological context** (Section 4.7). The predicted number density of observable LRDs is broadly consistent with current constraints at $z \simeq 4 - 6$. The predicted LRD abundance continues to rise towards lower redshifts while the observed number densities show an apparent decline at $z \lesssim 4$. Although we show

that observational bias and selection effects can play a role, this may indicate limitations of our BH seeding or LRD model, which could miss additional criteria uniquely associated with the high-redshift environment. The total BHAD in LUMINA is broadly consistent with recent *JWST*/MIRI constraints at $z \lesssim 5$ and is higher than the value inferred from previous X-ray and mid-IR surveys. The LRD contribution is non-negligible at $z \simeq 4 - 6$ but remains below the total AGN contribution and is closer to estimates based on updated LRD bolometric corrections than to upper limits derived from canonical quasar corrections. The accumulated SMBH mass density at high redshifts remains compatible with canonical local SMBH mass density constraints when we subtract the integrated mass gain from low-redshift quasars (based on the Soltan argument). The full AGN population in LUMINA, including both canonical AGN and the LRD phase, is therefore consistent with the global SMBH growth budget.

(vi) **AGN-driven He II reionization** (Section 4.7). With multi-frequency radiative transfer followed on the fly, LUMINA predicts a He II reionization history that begins around $z \simeq 6$ and completes by $z \simeq 3$, driven essentially by AGN and broadly consistent with observational constraints. The He II-ionizing photon budget during the bulk of this epoch is dominated by SMBHs with $M_{\text{BH}} \simeq 10^7 - 10^9 M_{\odot}$. Crucially, BHs below our adopted LRD mass threshold contribute negligibly to the He II-ionizing emissivity over this redshift range.

Taken together, these results show that LUMINA provides a coherent cosmological framework for interpreting the high-redshift AGN population. The simulation can simultaneously reproduce the abundance of classical quasars alongside the emerging population of faint AGN revealed by *JWST*, the large-scale environments of AGN, the integrated SMBH growth rates, and the reionization of He II in the IGM driven by the same SMBH population, all within a single self-consistent framework. In this picture, LRDs are considered as a possible early-growth phase of SMBHs. The observational evidence collected for LRDs through LFs at multiple wavelengths and large-scale clustering is mutually consistent, with minimal assumptions motivated empirically from observations. Meanwhile, the remaining canonical AGN component stays essential for matching the X-ray, mid-IR, and bright-end AGN optical/UV LF constraints, as well as supplying the bulk of the He II-ionizing photon budget. Several uncertainties remain, as discussed in detail in Section 5.2 and 5.3. The few free parameters in the AGN empirical model and the BH seeding and accretion prescriptions are simplified treatments of unresolved physics with some level of degeneracies. The role of LUMINA in this work is to provide a cosmological backbone within which a physically plausible, effective LRD phase can be tested against the abundance, environments, and other observables of SMBHs simultaneously. Future progress will require a larger and more complete observational sample and more predictive theoretical models.

ACKNOWLEDGEMENTS

We thank Pablo G. Pérez-González, Xiaojing Lin, Jenny Greene, David Setton, and Wendy Sun for sharing their stacked SED data with us and for insightful discussions. We thank Claude-André Faucher-Giguère, Philip Hopkins, and Gordon Richards for their valuable feedback in discussions. We also thank Zihao Wang for sharing the scripts of clustering measurements. An award of computer time was provided by the INCITE program. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Advanced Scientific Computing Research programs in the Office of Science of the U.S.

Department of Energy under Contract No. DE-AC05-00OR22725. Support for programs JWST-AR-04814 (XS, MV) and JWST-AR-08709 (AS) was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127. Support for OZ was provided by Harvard University through the Institute for Theory and Computation Fellowship. RK acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery Grant and a Discovery Launch Supplement, funding reference numbers RGPIN-2024-06222 and DGECR-2024-00144, and from York University's Global Research Excellence Initiative. MV acknowledges support through NASA ATP Grant 23-ATP23-149 and NSF AAG Grant AST-2307699. VS and LH acknowledge support from the Simons Foundation through the "Learning the Universe" initiative. LCH was supported by the National Science Foundation of China (12233001), the National Key R&D Program of China (2022YFF0503401), and the China Manned Space Program (CMS-CSST-2025-A09).

DATA AVAILABILITY

All LUMINA data will be made publicly available in 2028. Some of the data presented in the figures of this article are already available at <https://www.lumina-simulation.com>. Additional data are available from the corresponding author upon reasonable request.

REFERENCES

- Abramowicz M. A., Czerny B., Lasota J. P., Szuszkiewicz E., 1988, *ApJ*, **332**, 646
- Agazie G., et al., 2023, *ApJ*, **951**, L8
- Aird J., et al., 2010, *MNRAS*, **401**, 2531
- Aird J., Coil A. L., Georgakakis A., Nandra K., Barro G., Pérez-González P. G., 2015, *MNRAS*, **451**, 1892
- Aird J., Coil A. L., Georgakakis A., 2018, *MNRAS*, **474**, 1225
- Akins H. B., et al., 2025a, *ApJ*, **980**, L29
- Akins H. B., et al., 2025b, *ApJ*, **991**, 37
- Alexander D. M., Hickox R. C., 2012, *New Astron. Rev.*, **56**, 93
- Alexander D. M., et al., 2025, *New Astron. Rev.*, **101**, 101733
- Ananna T. T., et al., 2019, *ApJ*, **871**, 240
- Ananna T. T., Bogdán Á., Kovács O. E., Natarajan P., Hickox R. C., 2024, *ApJ*, **969**, L18
- Anglés-Alcázar D., Davé R., Faucher-Giguère C.-A., Özel F., Hopkins P. F., 2017a, *MNRAS*, **464**, 2840
- Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017b, *MNRAS*, **472**, L109
- Anglés-Alcázar D., et al., 2021, *ApJ*, **917**, 53
- Antonucci R., 1993, *ARA&A*, **31**, 473
- Arita J., et al., 2023, *ApJ*, **954**, 210
- Arita J., Kashikawa N., Onoue M., Yoshioka T., Takeda Y., Hoshi H., Shimizu S., 2025, *MNRAS*, **536**, 3677
- Arnauk K. A., 1996, in Jacoby G. H., Barnes J., eds, *Astronomical Society of the Pacific Conference Series Vol. 101, Astronomical Data Analysis Software and Systems V*. p. 17
- Asada Y., Inayoshi K., Fei Q., Fujimoto S., Willott C., 2026, *arXiv e-prints*, p. [arXiv:2601.10573](https://arxiv.org/abs/2601.10573)
- Assef R. J., et al., 2011, *ApJ*, **728**, 56
- Asthana S., Haehnelt M. G., Kulkarni G., Bolton J. S., Gaikwad P., Keating L. C., Puchwein E., 2025, *MNRAS*, **542**, 2968
- Bañados E., et al., 2018, *Nature*, **553**, 473
- Baggen J. F. W., Scoggins M. T., van Dokkum P., Haiman Z., Torralba A., Matthee J., 2026, *ApJ*, **1002**, L4
- Bagla J. S., 2002, *Journal of Astrophysics and Astronomy*, **23**, 185
- Bagla J. S., Ray S., 2003, *New Astronomy*, **8**, 665
- Baloković M., et al., 2018, *ApJ*, **854**, 42
- Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, *ApJ*, **304**, 15
- Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., Wang W.-H., Steffen A. T., Capak P., 2005, *AJ*, **129**, 578
- Barro G., et al., 2026, *ApJ*, **997**, 48
- Basu A., Garaldi E., Ciardi B., 2024, *MNRAS*, **532**, 841
- Baumgarte T. W., Shapiro S. L., 1999, *ApJ*, **526**, 941
- Begelman M. C., 2010, *MNRAS*, **402**, 673
- Begelman M. C., Dexter J., 2026, *ApJ*, **996**, 48
- Begelman M. C., Volonteri M., Rees M. J., 2006, *MNRAS*, **370**, 289
- Begelman M. C., Rossi E. M., Armitage P. J., 2008, *MNRAS*, **387**, 1649
- Bennett J. S., Sijacki D., Costa T., Laporte N., Witten C., 2024, *MNRAS*, **527**, 1033
- Bhowmick A. K., et al., 2024a, *MNRAS*, **531**, 4311
- Bhowmick A. K., et al., 2024b, *MNRAS*, **534**, 1907
- Bird S., Ni Y., Di Matteo T., Croft R., Feng Y., Chen N., 2022, *MNRAS*, **512**, 3703
- Bogdán Á., et al., 2024, *Nature Astronomy*, **8**, 126
- Bondi H., 1952, *MNRAS*, **112**, 195
- Bondi H., Hoyle F., 1944, *MNRAS*, **104**, 273
- Bonoli S., Izquierdo-Villalba D., Spinoso D., Colpi M., Sesana A., Polkas M., Springel V., 2025, *arXiv e-prints*, p. [arXiv:2509.12325](https://arxiv.org/abs/2509.12325)
- Booth C. M., Schaye J., 2009, *MNRAS*, **398**, 53
- 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
- Boyle B. J., Shanks T., Peterson B. A., 1988, *MNRAS*, **235**, 935
- Boyle B. J., Griffiths R. E., Shanks T., Stewart G. C., Georgantopoulos I., 1993, *MNRAS*, **260**, 49
- Brandt W. N., Alexander D. M., 2015, *A&A Rev.*, **23**, 1
- Brazzini M., D'Eugenio F., Maiolino R., Juodžbalis I., Ji X., Scholtz J., Chang S.-J., 2025, *MNRAS*, **544**, L167
- Brazzini M., et al., 2026, *arXiv e-prints*, p. [arXiv:2601.22214](https://arxiv.org/abs/2601.22214)
- Brightman M., Nandra K., 2011, *MNRAS*, **413**, 1206
- Bromm V., Loeb A., 2003, *ApJ*, **596**, 34
- Brown M. J. I., et al., 2006, *ApJ*, **638**, 88
- Buchner J., Brightman M., Nandra K., Nikutta R., Bauer F. E., 2019, *A&A*, **629**, A16
- Bulich T.-E., et al., 2025, *MNRAS*, **544**, 355
- Bulich T.-E., et al., 2026, *arXiv e-prints*, p. [arXiv:2603.22393](https://arxiv.org/abs/2603.22393)
- Burke C. J., et al., 2021, *Science*, **373**, 789
- Burlon D., Ajello M., Greiner J., Comastri A., Merloni A., Gehrels N., 2011, *ApJ*, **728**, 58
- Caplar N., Lilly S. J., Trakhtenbrot B., 2017, *ApJ*, **834**, 111
- Carnall A. C., et al., 2023, *Nature*, **619**, 716
- Carroll C. M., Hickox R. C., Masini A., Lanz L., Assef R. J., Stern D., Chen C.-T. J., Ananna T. T., 2021, *ApJ*, **908**, 185
- Casey C. M., et al., 2025, *ApJ*, **990**, L61
- Chabrier G., 2003, *PASP*, **115**, 763
- Chang S.-J., Gronke M., Matthee J., Mason C., 2026, *MNRAS*, **545**, staf2131
- Chen C.-H., Ho L. C., Li R., Zhuang M.-Y., 2025a, *ApJ*, **983**, 60
- Chen C.-H., Ho L. C., Li R., Inayoshi K., 2025b, *ApJ*, **989**, L12
- Chisholm J., et al., 2026, *arXiv e-prints*, p. [arXiv:2602.15935](https://arxiv.org/abs/2602.15935)
- Cicone C., et al., 2014, *A&A*, **562**, A21
- Cloonan A. P., et al., 2026, *arXiv e-prints*, p. [arXiv:2603.24700](https://arxiv.org/abs/2603.24700)
- Cooray A., Sheth R., 2002, *Phys. Rep.*, **372**, 1
- Coughlin E. R., Begelman M. C., 2024, *ApJ*, **970**, 158
- Cowie L. L., Binney J., 1977, *ApJ*, **215**, 723
- Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, *AJ*, **112**, 839
- Crain R. A., et al., 2015, *MNRAS*, **450**, 1937
- Croton D. J., et al., 2006, *MNRAS*, **365**, 11
- Cruise M., et al., 2025, *Nature Astronomy*, **9**, 36
- Curti M., et al., 2024, *A&A*, **684**, A75
- Dadina M., 2008, *A&A*, **485**, 417
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, *MNRAS*, **486**, 2827
- Davies M. B., Miller M. C., Bellovary J. M., 2011, *ApJ*, **740**, L42
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, **292**, 371
- Dayal P., Ferrara A., Dunlop J. S., Pacucci F., 2014, *MNRAS*, **445**, 2545

- Dayal P., Rossi E. M., Shiralilou B., Piana O., Choudhury T. R., Volonteri M., 2019, *MNRAS*, **486**, 2336
- Dekel A., Stone N. C., Chowdhury D. D., Gilbaum S., Li Z., Mandelker N., van den Bosch F. C., 2025, *A&A*, **695**, A97
- Delvecchio I., et al., 2014, *MNRAS*, **439**, 2736
- Delvecchio I., et al., 2025, *A&A*, **704**, A313
- Desjacques V., Jeong D., Schmidt F., 2018, *Phys. Rep.*, **733**, 1
- Devecchi B., Volonteri M., 2009, *ApJ*, **694**, 302
- Devecchi B., Volonteri M., Rossi E. M., Colpi M., Portegies Zwart S., 2012, *MNRAS*, **421**, 1465
- Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, **433**, 604
- Diemer B., 2018, *ApJS*, **239**, 35
- Ding B., Pizzati E., Schaye J., Hennawi J. F., McDonald W., Schaller M., 2026, *MNRAS*, **547**, stg308
- Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, **463**, 3948
- Dubois Y., et al., 2021, *A&A*, **651**, A109
- Dubroca B., Feugeas J.-L., 1999, *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, **329**, 915
- Duras F., et al., 2020, *A&A*, **636**, A73
- Durodola E., Pacucci F., Hickox R. C., 2025, *ApJ*, **985**, 169
- Ebrero J., et al., 2009, *A&A*, **493**, 55
- Eide M. B., Graziani L., Ciardi B., Feng Y., Kakiichi K., Di Matteo T., 2018, *MNRAS*, **476**, 1174
- Eilers A.-C., et al., 2024, *ApJ*, **974**, 275
- Eisenstein D. J., Hu W., 1998, *ApJ*, **496**, 605
- Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, *Publ. Astron. Soc. Australia*, **34**, e058
- Elitzur M., Ho L. C., 2009, *ApJ*, **701**, L91
- Fabian A. C., 1994, *ARA&A*, **32**, 277
- Fabian A. C., 2012, *ARA&A*, **50**, 455
- Fabian A. C., Nulsen P. E. J., 1977, *MNRAS*, **180**, 479
- Fabian A. C., Nulsen P. E. J., Canizares C. R., 1984, *Nature*, **310**, 733
- Fan X., et al., 2001, *AJ*, **122**, 2833
- Fan X., Narayanan V. K., Strauss M. A., White R. L., Becker R. H., Pentericci L., Rix H.-W., 2002, *AJ*, **123**, 1247
- Fan X., et al., 2004, *AJ*, **128**, 515
- Fan X., et al., 2006, *AJ*, **132**, 117
- Fan X., Bañados E., Simcoe R. A., 2023, *ARA&A*, **61**, 373
- Faucher-Giguère C.-A., Lidz A., Hernquist L., Zaldarriaga M., 2008a, *ApJ*, **682**, L9
- Faucher-Giguère C.-A., Lidz A., Hernquist L., Zaldarriaga M., 2008b, *ApJ*, **688**, 85
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, **703**, 1416
- Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, *MNRAS*, **455**, 2778
- Feng Y., Bird S., Anderson L., Font-Ribera A., Pedersen C., 2018, MP-Gadget/MP-Gadget: A tag for getting a DOI, doi:10.5281/zenodo.1451799
- Ferrarese L., Merritt D., 2000, *ApJ*, **539**, L9
- Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, *A&A*, **518**, L155
- Fialkov A., Barkana R., Tseliakhovich D., Hirata C. M., 2012, *MNRAS*, **424**, 1335
- Finkelstein S. L., Bagley M. B., Ferguson H. C., Wilkins S. M., Kartaltepe J. S., Papovich C., et al., 2023, *ApJ*, **946**, L13
- Finkelstein S. L., et al., 2024, *ApJ*, **969**, L2
- Fixsen D. J., 2009, *ApJ*, **707**, 916
- Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2013, *ApJ*, **776**, L31
- Fragos T., Lehmer B. D., Naoz S., Zezas A., Basu-Zych A., 2016, *ApJ*, **827**, L21
- Fritz J., Franceschini A., Hatziminaoglou E., 2006, *MNRAS*, **366**, 767
- Fujimoto S., et al., 2022, *Nature*, **604**, 261
- Furlanetto S. R., Oh S. P., 2008, *ApJ*, **681**, 1
- Gaikwad P., Srianand R., Haehnelt M. G., Choudhury T. R., 2021, *MNRAS*, **506**, 4389
- Gao L., White S. D. M., 2007, *MNRAS*, **377**, L5
- Gao L., Springel V., White S. D. M., 2005, *MNRAS*, **363**, L66
- Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, *MNRAS*, **512**, 4909
- Gebhardt K., et al., 2000, *ApJ*, **539**, L13
- Giallongo E., et al., 2015, *A&A*, **578**, A83
- Glazebrook K., et al., 2024, *Nature*, **628**, 277
- Glikman E., Djorgovski S. G., Stern D., Dey A., Jannuzi B. T., Lee K.-S., 2011, *ApJ*, **728**, L26
- Gnedin N. Y., Abel T., 2001, *New Astronomy*, **6**, 437
- Godunov S. K., Bohachevsky I., 1959, *Matematičeskij sbornik*, **47**, 271
- González-Martín O., Vaughan S., 2012, *A&A*, **544**, A80
- Graham A. W., Driver S. P., 2007, *ApJ*, **655**, 77
- Green P. J., et al., 1995, *ApJ*, **450**, 51
- Greene J. E., Ho L. C., 2005, *ApJ*, **630**, 122
- Greene J. E., et al., 2024, *ApJ*, **964**, 39
- Greene J. E., et al., 2026, *ApJ*, **996**, 129
- Grimes J. A., Rawlings S., Willott C. J., 2004, *MNRAS*, **349**, 503
- Gültekin K., et al., 2009, *ApJ*, **698**, 198
- Haardt F., Madau P., 2012, *ApJ*, **746**, 125
- Haardt F., Salvaterra R., 2015, *A&A*, **575**, L16
- Habouzit M., 2025, *MNRAS*, **537**, 2323
- Habouzit M., et al., 2021, *MNRAS*, **503**, 1940
- Habouzit M., et al., 2022, *MNRAS*, **509**, 3015
- Han J., Jing Y. P., Wang H., Wang W., 2012, *MNRAS*, **427**, 2437
- Hao L., et al., 2005, *AJ*, **129**, 1795
- Harikane Y., et al., 2023a, *ApJS*, **265**, 5
- Harikane Y., et al., 2023b, *ApJ*, **959**, 39
- Hasinger G., Miyaji T., Schmidt M., 2005, *A&A*, **441**, 417
- Hayes M. J., et al., 2024, *ApJ*, **971**, L16
- Heckman T. M., Best P. N., 2014, *ARA&A*, **52**, 589
- Hickox R. C., Alexander D. M., 2018, *ARA&A*, **56**, 625
- Hickox R. C., Mullaney J. R., Alexander D. M., Chen C.-T. J., Civano F. M., Goulding A. D., Hainline K. N., 2014, *ApJ*, **782**, 9
- Hill G. J., Goodrich R. W., Depoy D. L., 1996, *ApJ*, **462**, 163
- Ho L. C., 1999, *ApJ*, **516**, 672
- Ho L. C., 2008, *ARA&A*, **46**, 475
- Hopkins P. F., 2025, *The Open Journal of Astrophysics*, **8**, 56
- Hopkins P. F., Hernquist L., 2009, *ApJ*, **698**, 1550
- Hopkins P. F., Quataert E., 2010, *MNRAS*, **407**, 1529
- Hopkins P. F., Quataert E., 2011, *MNRAS*, **415**, 1027
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, *ApJS*, **163**, 1
- Hopkins P. F., Richards G. T., Hernquist L., 2007, *ApJ*, **654**, 731
- Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, *ApJS*, **175**, 356
- Hopkins P. F., et al., 2018, *MNRAS*, **480**, 800
- Hopkins P. F., et al., 2024, *The Open Journal of Astrophysics*, **7**, 18
- Hosokawa T., Yorke H. W., Inayoshi K., Omukai K., Yoshida N., 2013, *ApJ*, **778**, 178
- Howlett C., Lewis A., Hall A., Challinor A., 2012, *J. Cosmol. Astropart. Phys.*, **2012**, 027
- Huško F., Lacey C. G., Roper W. J., Schaye J., Briggs J. M., Schaller M., 2025, *MNRAS*, **537**, 2559
- Hviding R. E., et al., 2025, *A&A*, **702**, A57
- Hviding R. E., et al., 2026, *ApJ*, **1000**, L18
- Inayoshi K., 2025, *ApJ*, **988**, L22
- Inayoshi K., Ho L. C., 2025, *arXiv e-prints*, p. arXiv:2512.03130
- Inayoshi K., Ichikawa K., 2024, *ApJ*, **973**, L49
- Inayoshi K., Maiolino R., 2025, *ApJ*, **980**, L27
- Inayoshi K., Visbal E., Haiman Z., 2020, *ARA&A*, **58**, 27
- Inayoshi K., Kimura S. S., Noda H., 2025, *PASJ*, **77**, 811
- Inayoshi K., Murase K., Kashiyama K., 2026, *ApJ*, **1000**, 90
- Ji X., et al., 2025, *MNRAS*, **544**, 3900
- Ji X., et al., 2026a, *arXiv e-prints*, p. arXiv:2604.03370
- Ji X., et al., 2026b, *MNRAS*, **545**, staf2235
- Jiang L., et al., 2016, *ApJ*, **833**, 222
- Jiang Y.-F., Stone J. M., Davis S. W., 2019, *ApJ*, **880**, 67

- Jiang F., Jia Z., Zheng H., Ho L. C., Inayoshi K., Shen X., Vogelsberger M., Feng W.-X., 2026, *ApJ*, 996, L19
- Jones L. R., et al., 1997, *MNRAS*, 285, 547
- Just D. W., Brandt W. N., Shemmer O., Steffen A. T., Schneider D. P., Chartas G., Garmire G. P., 2007, *ApJ*, 665, 1004
- Kaiser N., 1984, *ApJ*, 284, L9
- Kannan R., Vogelsberger M., Marinacci F., McKinnon R., Pakmor R., Springel V., 2019, *MNRAS*, 485, 117
- Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022, *MNRAS*, 511, 4005
- Kashikawa N., et al., 2015, *ApJ*, 798, 28
- Katz N., Hernquist L., Weinberg D. H., 1992, *ApJ*, 399, L109
- Katz N., Weinberg D. H., Hernquist L., 1996, *ApJS*, 105, 19
- Katz H., Sijacki D., Haehnelt M. G., 2015, *MNRAS*, 451, 2352
- Kauffmann G., Heckman T. M., 2009, *MNRAS*, 397, 135
- Kelly B. C., Shen Y., 2013, *ApJ*, 764, 45
- Kelly B. C., Bechtold J., Siemiginowska A., 2009, *ApJ*, 698, 895
- Kelly B. C., Vestergaard M., Fan X., Hopkins P., Hernquist L., Siemiginowska A., 2010, *ApJ*, 719, 1315
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kido D., Ioka K., Hotokezaka K., Inayoshi K., Irwin C. M., 2025, *MNRAS*, 544, 3407
- Killi M., et al., 2024, *A&A*, 691, A52
- Kocevski D. D., et al., 2025, *ApJ*, 986, 126
- Kokorev V., et al., 2024, *ApJ*, 968, 38
- Kokubo M., Harikane Y., 2025, *ApJ*, 995, 24
- Kollmeier J. A., et al., 2006, *ApJ*, 648, 128
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Krawczyk C. M., Richards G. T., Mehta S. S., Vogeley M. S., Gallagher S. C., Leighly K. M., Ross N. P., Schneider D. P., 2013, *ApJS*, 206, 4
- Kubota A., Done C., 2019, *MNRAS*, 489, 524
- Kuhlen M., Faucher-Giguère C.-A., 2012, *MNRAS*, 423, 862
- Kulkarni G., Worseck G., Hennawi J. F., 2019, *MNRAS*, 488, 1035
- La Franca F., et al., 2005, *ApJ*, 635, 864
- Labbe I., et al., 2024, *ApJ*, 978, 92
- Lacy M., Ridgway S. E., Sajina A., Petric A. O., Gates E. L., Urrutia T., Storrie-Lombardi L. J., 2015, *ApJ*, 802, 102
- Lambides E., et al., 2024, *ApJ*, 970, 68
- Levermore C. D., 1984, *J. Quant. Spectrosc. Radiative Transfer*, 31, 149
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- Li J., et al., 2025, *ApJ*, 981, 19
- Lidz A., Faucher-Giguère C.-A., Dall'Aglio A., McQuinn M., Fechner C., Zaldarriaga M., Hernquist L., Dutta S., 2010, *ApJ*, 718, 199
- Liebold E. R., Ma C.-P., 2024, *ApJ*, 971, L29
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Lin X., et al., 2024, *ApJ*, 974, 147
- Lin R., et al., 2025, *ApJ*, 980, L34
- Lin X., et al., 2026a, *ApJ*, 996, 93
- Lin X., et al., 2026b, *ApJ*, 997, 61
- Lin X., et al., 2026c, *ApJ*, 997, 364
- Ling C.-T., Goto T., Kim S. J., Wu C. K.-W., Chen A. Y.-A., Kilerci E., Hashimoto T., Phan T. L., 2026, *PASP*, 138, 014102
- Liu Z., et al., 2026, *arXiv e-prints*, p. arXiv:2604.13000
- Lodato G., Natarajan P., 2006, *MNRAS*, 371, 1813
- Looser T. J., et al., 2024, *Nature*, 629, 53
- Lupi A., Trinca A., Volonteri M., Dotti M., Mazzucchelli C., 2024, *A&A*, 689, A128
- Lusso E., Risaliti G., 2016, *ApJ*, 819, 154
- Lusso E., et al., 2010, *A&A*, 512, A34
- Lusso E., et al., 2012, *MNRAS*, 425, 623
- Lusso E., Worseck G., Hennawi J. F., Prochaska J. X., Vignali C., Stern J., O'Meara J. M., 2015, *MNRAS*, 449, 4204
- Lyu J., et al., 2024, *ApJ*, 966, 229
- Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, 456, 2140
- Ma X., et al., 2018, *MNRAS*, 478, 1694
- Ma X., et al., 2019, *MNRAS*, 487, 1844
- Ma Y., et al., 2025, *arXiv e-prints*, p. arXiv:2509.02662
- Ma Y., et al., 2026, *ApJ*, 1000, 59
- MacLeod C. L., et al., 2010, *ApJ*, 721, 1014
- Maccacaro T., della Ceca R., Gioia I. M., Morris S. L., Stocke J. T., Wolter A., 1991, *ApJ*, 374, 117
- Madau P., Fragos T., 2017, *ApJ*, 840, 39
- Madau P., Haardt F., 2015, *ApJ*, 813, L8
- Madau P., Haardt F., 2024, *ApJ*, 976, L24
- Madau P., Maiolino R., 2026, *arXiv e-prints*, p. arXiv:2602.22386
- Madau P., Rees M. J., 2000, *ApJ*, 542, L69
- Madau P., Haardt F., Rees M. J., 1999, *ApJ*, 514, 648
- Madau P., Giallongo E., Grazian A., Haardt F., 2024, *ApJ*, 971, 75
- Magdziarz P., Zdziarski A. A., 1995, *MNRAS*, 273, 837
- Magorrian J., et al., 1998, *AJ*, 115, 2285
- Maiolino R., et al., 2024, *A&A*, 691, A145
- Maiolino R., et al., 2025, *MNRAS*, 538, 1921
- Marchesi S., Ajello M., Marcolutti L., Comastri A., Lanzuisi G., Vignali C., 2018, *ApJ*, 854, 49
- Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, *MNRAS*, 351, 169
- Marszewski A., Faucher-Giguère C.-A., Sun G., Anglés-Alcázar D., Feldmann R., Su K.-Y., Miller T. B., Roy N. C., 2026, *ApJ*, 1002, L13
- Mason C. A., Trenti M., Treu T., 2023, *MNRAS*, 521, 497
- Matsuoka Y., et al., 2018, *ApJ*, 869, 150
- Matthee J., et al., 2024, *ApJ*, 963, 129
- Matthee J., et al., 2025, *ApJ*, 988, 246
- Matute I., La Franca F., Pozzi F., Gruppioni C., Lari C., Zamorani G., 2006, *A&A*, 451, 443
- Mazzucchelli C., et al., 2017, *ApJ*, 849, 91
- McGreer I. D., et al., 2013, *ApJ*, 768, 105
- McQuinn M., Lidz A., Zaldarriaga M., Hernquist L., Hopkins P. F., Dutta S., Faucher-Giguère C.-A., 2009, *ApJ*, 694, 842
- Meurer G. R., Heckman T. M., Calzetti D., 1999, *ApJ*, 521, 64
- Mineo S., Gilfanov M., Sunyaev R., 2012, *MNRAS*, 426, 1870
- Miralda-Escudé J., 1998, *ApJ*, 501, 15
- Mirocha J., Furlanetto S. R., 2023, *MNRAS*, 519, 843
- Miyaji T., Hasinger G., Schmidt M., 2000, *A&A*, 353, 25
- Mo H. J., White S. D. M., 1996, *MNRAS*, 282, 347
- Morrison R., McCammon D., 1983, *ApJ*, 270, 119
- Mortlock D. J., et al., 2011, *Nature*, 474, 616
- Murphy K. D., Yaqoob T., 2009, *MNRAS*, 397, 1549
- Murphy E. J., et al., 2011, *ApJ*, 737, 67
- Naidu R. P., et al., 2025, *arXiv e-prints*, p. arXiv:2503.16596
- Naidu R. P., et al., 2026, *The Open Journal of Astrophysics*, 9, 56033
- Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, *ApJS*, 269, 33
- Napolitano L., et al., 2025, *ApJ*, 989, 75
- Narayan R., Yi I., 1994, *ApJ*, 428, L13
- Negri A., Volonteri M., 2017, *MNRAS*, 467, 3475
- Netzer H., 2015, *ARA&A*, 53, 365
- Ni Y., et al., 2022, *MNRAS*, 513, 670
- Ni Y., Chen N., Zhou Y., Park M., Yang Y., Di Matteo T., Bird S., Croft R., 2025, *ApJ*, 990, 120
- Niida M., et al., 2020, *ApJ*, 904, 89
- Onoue M., et al., 2017, *ApJ*, 847, L15
- Osterbrock D. E., Ferland G. J., 2006, *Astrophysics of gaseous nebulae and active galactic nuclei*
- Pacucci F., Loeb A., 2025, *ApJ*, 989, L19
- Pacucci F., Narayan R., 2024, *ApJ*, 976, 96
- Pacucci F., Mesinger A., Mineo S., Ferrara A., 2014, *MNRAS*, 443, 678
- Pacucci F., Nguyen B., Carniani S., Maiolino R., Fan X., 2023, *ApJ*, 957, L3
- Pacucci F., Hernquist L., Fujii M., 2025, *ApJ*, 994, 40
- Paczynski B., Wiita P. J., 1980, *A&A*, 88, 23
- Page M. J., Mason K. O., McHardy I. M., Jones L. R., Carrera F. J., 1997, *MNRAS*, 291, 324
- Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaal K., Zhu C., 2016, *MNRAS*, 455, 1134
- Pakmor R., et al., 2023, *MNRAS*, 524, 2539

- Parsa S., Dunlop J. S., McLure R. J., 2018, *MNRAS*, 474, 2904
- Pei Y. C., 1992, *ApJ*, 395, 130
- Pérez-González P. G., et al., 2024, *ApJ*, 968, 4
- Pérez-González P. G., et al., 2026, *arXiv e-prints*, p. arXiv:2602.20247
- Pillepich A., et al., 2018a, *MNRAS*, 473, 4077
- Pillepich A., et al., 2018b, *MNRAS*, 475, 648
- Piotrowska J. M., Bluck A. F. L., Maiolino R., Peng Y., 2022, *MNRAS*, 512, 1052
- Pizzati E., et al., 2024, *MNRAS*, 534, 3155
- Pizzati E., Hennawi J. F., Schaye J., Eilers A.-C., Huang J., Schindler J.-T., Wang F., 2025, *MNRAS*, 539, 2910
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Ponti G., Papadakis I., Bianchi S., Guainazzi M., Matt G., Uttley P., Bonilla N. F., 2012, *A&A*, 542, A83
- Pouliasis E., et al., 2024, *A&A*, 685, A97
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, 430, 2427
- Ramos Almeida C., Ricci C., 2017, *Nature Astronomy*, 1, 679
- Reed S. L., et al., 2017, *MNRAS*, 468, 4702
- Rees M. J., 1984, *ARA&A*, 22, 471
- Reynolds C. S., et al., 2023, *arXiv e-prints*, p. arXiv:2311.00780
- Ricci C., Ueda Y., Koss M. J., Trakhtenbrot B., Bauer F. E., Gandhi P., 2015, *ApJ*, 815, L13
- Richards G. T., et al., 2006a, *AJ*, 131, 2766
- Richards G. T., et al., 2006b, *ApJS*, 166, 470
- Roberts M. G., Braff L., Garg A., Profumo S., Jeltema T., 2026, *J. Cosmol. Astropart. Phys.*, 2026, 003
- Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, *ApJ*, 802, L19
- Ronayne K., et al., 2025, *arXiv e-prints*, p. arXiv:2508.20177
- Rosas-Guevara Y., Bower R. G., Schaye J., McAlpine S., Dalla Vecchia C., Frenk C. S., Schaller M., Theuns T., 2016, *MNRAS*, 462, 190
- Ross N. P., Cross N. J. G., 2020, *MNRAS*, 494, 789
- Ross N. P., et al., 2013, *ApJ*, 773, 14
- Runnoe J. C., Brotherton M. S., Shang Z., 2012a, *MNRAS*, 422, 478
- Runnoe J. C., Brotherton M. S., Shang Z., 2012b, *MNRAS*, 426, 2677
- Rusakov V., et al., 2026, *Nature*, 649, 574
- Sanders D. B., Mirabel I. F., 1996, *ARA&A*, 34, 749
- Sanders R. L., et al., 2021, *ApJ*, 914, 19
- Sarkar A., et al., 2025, *ApJ*, 978, 136
- Sartori L. F., Schawinski K., Trakhtenbrot B., Caplar N., Treister E., Koss M. J., Urry C. M., Zhang C. E., 2018, *MNRAS*, 476, L34
- Sato-Polito G., Zaldarriaga M., Quataert E., 2025, *Phys. Rev. D*, 112, 123018
- Sazonov S. Y., Revnivtsev M. G., 2004, *A&A*, 423, 469
- Schaller M., et al., 2024, *MNRAS*, 530, 2378
- Schaye J., et al., 2015, *MNRAS*, 446, 521
- Schaye J., et al., 2023, *MNRAS*, 526, 4978
- Schindler J.-T., et al., 2023, *ApJ*, 943, 67
- Schirber M., Bullock J. S., 2003, *ApJ*, 584, 110
- Schmidt M., 1968, *ApJ*, 151, 393
- Scholtz J., et al., 2026, *arXiv e-prints*, p. arXiv:2603.22277
- Schulze A., Wisotzki L., 2010, *A&A*, 516, A87
- Semenov V. A., Conroy C., Hernquist L., 2025a, *ApJ*, 989, 219
- Semenov V. A., Conroy C., Smith A., Puchwein E., Hernquist L., 2025b, *ApJ*, 990, 7
- Setton D. J., et al., 2025a, *ApJ*, 991, L10
- Setton D. J., et al., 2025b, *ApJ*, 995, 118
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 500, 33
- Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, *MNRAS*, 354, 1020
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, *ApJ*, 690, 20
- She R., Ho L. C., Feng H., Cui C., 2018, *ApJ*, 859, 152
- Shen X., Hopkins P. F., Faucher-Giguère C.-A., Alexander D. M., Richards G. T., Ross N. P., Hickox R. C., 2020a, *MNRAS*, 495, 3252
- Shen X., et al., 2020b, *MNRAS*, 495, 4747
- Shen X., Vogelsberger M., Boylan-Kolchin M., Tacchella S., Kannan R., 2023, *MNRAS*, 525, 3254
- Shen X., Vogelsberger M., Boylan-Kolchin M., Tacchella S., Naidu R. P., 2024, *MNRAS*, 533, 3923
- Shen T., Shen X., Xiao H., Vogelsberger M., Jiang F., 2025a, *arXiv e-prints*, p. arXiv:2504.00075
- Shen X., Zier O., Vogelsberger M., Boylan-Kolchin M., Hernquist L., Tacchella S., Naidu R. P., 2025b, *arXiv e-prints*, p. arXiv:2509.19427
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Shi Y., Kremer K., Hopkins P. F., 2024, *ApJ*, 969, L31
- Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, *MNRAS*, 380, 877
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, 452, 575
- Silverman J. D., et al., 2008, *ApJ*, 679, 118
- Simpson C., Rawlings S., Lacy M., 1999, *MNRAS*, 306, 828
- Sinha M., Garrison L., 2019, in Majumdar A., Arora R., eds, *Software Challenges to Exascale Computing*. Springer Singapore, Singapore, pp 3–20, https://doi.org/10.1007/978-981-13-7729-7_1
- Sinha M., Garrison L. H., 2020, *MNRAS*, 491, 3022
- Sądowski A., 2011, *arXiv e-prints*, p. arXiv:1108.0396
- Smith A., Bromm V., 2019, *Contemporary Physics*, 60, 111
- Smith B., Sigurdsson S., Abel T., 2008, *MNRAS*, 385, 1443
- Smith A., Kannan R., Garaldi E., Vogelsberger M., Pakmor R., Springel V., Hernquist L., 2022, *MNRAS*, 512, 3243
- Smith A., et al., 2026, *arXiv e-prints*, p. arXiv:2605.18939
- Sokasian A., Abel T., Hernquist L., 2002, *MNRAS*, 332, 601
- Soltan 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., 2010, *MNRAS*, 401, 791
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V., Di Matteo T., Hernquist L., 2005, *ApJ*, 620, L79
- Springel V., et al., 2018, *MNRAS*, 475, 676
- Springel V., Pakmor R., Zier O., Reinecke M., 2021, *MNRAS*, 506, 2871
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, *ApJ*, 596, L23
- Steffen A. T., Strateva I., Brandt W. N., Alexander D. M., Koekemoer A. M., Lehmer B. D., Schneider D. P., Vignali C., 2006, *AJ*, 131, 2826
- Strateva I. V., Brandt W. N., Schneider D. P., Vanden Berk D. G., Vignali C., 2005, *AJ*, 130, 387
- Su T., Guo Q., Qiao E., Pei W., Ho L. C., Lacey C. G., 2026, *ApJ*, 999, 258
- Sun G., Faucher-Giguère C.-A., Hayward C. C., Shen X., Wetzel A., Cochrane R. K., 2023, *ApJ*, 955, L35
- Sun W. Q., et al., 2026, *arXiv e-prints*, p. arXiv:2601.20929
- Tabor G., Binney J., 1993, *MNRAS*, 263, 323
- Tanimoto A., Ueda Y., Odaka H., Kawaguchi T., Fukazawa Y., Kawamuro T., 2019, *ApJ*, 877, 95
- Tanimoto A., Ueda Y., Odaka H., Yamada S., Ricci C., 2022, *ApJS*, 260, 30
- Taylor A. J., et al., 2025, *ApJ*, 986, 165
- Tee W. L., Fan X., Wang F., Yang J., 2025, *ApJ*, 983, L26
- Terrazas B. A., et al., 2020, *MNRAS*, 493, 1888
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
- Trebtsch M., Volonteri M., Dubois Y., 2021, *A&A*, 653, A154
- Trinca A., Schneider R., Valiante R., Graziani L., Zappacosta L., Shankar F., 2022, *MNRAS*, 511, 616
- Trinca A., Schneider R., Valiante R., Graziani L., Ferrotti A., Omukai K., Chon S., 2024, *MNRAS*, 529, 3563
- Tselikhovich D., Hirata C., 2010, *Phys. Rev. D*, 82, 083520
- Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, *ApJ*, 598, 886
- Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, *ApJ*, 786, 104
- Umeda H., Inayoshi K., Harikane Y., Murase K., 2026, *ApJ*, 999, 183
- Urry C. M., Padovani P., 1995, *PASP*, 107, 803
- Vanden Berk D. E., et al., 2004, *ApJ*, 601, 692
- Vaughan S., Edelson R., Warwick R. S., Uttley P., 2003, *MNRAS*, 345, 1271
- Venemans B. P., et al., 2013, *ApJ*, 779, 24
- Venemans B. P., et al., 2015, *ApJ*, 801, L11
- Vignali C., Brandt W. N., Schneider D. P., 2003, *AJ*, 125, 433

- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, **436**, 3031
- Vogelsberger M., et al., 2014a, *MNRAS*, **444**, 1518
- Vogelsberger M., et al., 2014b, *Nature*, **509**, 177
- Volonteri M., Begelman M. C., 2010, *MNRAS*, **409**, 1022
- Volonteri M., Dubois Y., Pichon C., Devriendt J., 2016, *MNRAS*, **460**, 2979
- Wang F., et al., 2018, *ApJ*, **869**, L9
- Wang F., et al., 2021, *ApJ*, **907**, L1
- Wang Z., Jiang F., Zheng H., Shen X., Jia Z., Ho L. C., Inayoshi K., Jiang L., 2026, *arXiv e-prints*, p. arXiv:2603.15736
- Warren S. J., Hewett P. C., Osmer P. S., 1994, *ApJ*, **421**, 412
- Weinberger R., et al., 2017, *MNRAS*, **465**, 3291
- Weinberger R., Springel V., Pakmor R., 2020, *ApJS*, **248**, 32
- Wiersma R. P. C., Schaye J., Smith B. D., 2009, *MNRAS*, **393**, 99
- Wilkes B. J., Tananbaum H., Worrall D. M., Avni Y., Oey M. S., Flanagan J., 1994, *ApJS*, **92**, 53
- Willott C. J., Rawlings S., Blundell K. M., Lacy M., 2000, *MNRAS*, **316**, 449
- Willott C. J., et al., 2010, *AJ*, **139**, 906
- Wolf J., et al., 2021, *A&A*, **647**, A5
- Woods T. E., Heger A., Whalen D. J., Haemmerlé L., Klessen R. S., 2017, *ApJ*, **842**, L6
- Worseck G., et al., 2014, *MNRAS*, **445**, 1745
- Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, *ApJ*, **825**, 144
- Worseck G., Davies F. B., Hennawi J. F., Prochaska J. X., 2019, *ApJ*, **875**, 111
- Wu J., et al., 2022, *MNRAS*, **517**, 2659
- Xiao H., Shen X., Hopkins P. F., Zurek K. M., 2021, *J. Cosmol. Astropart. Phys.*, **2021**, 039
- Xiao M., et al., 2025, *A&A*, **700**, A231
- Yang J., et al., 2020, *ApJ*, **897**, L14
- Yang G., et al., 2023, *ApJ*, **950**, L5
- Yeh J. Y.-C., et al., 2023, *MNRAS*, **520**, 2757
- Yenko B., Barger A. J., Trouille L., Winter L. M., 2009, *ApJ*, **698**, 380
- Yu Q., Lu Y., 2008, *ApJ*, **689**, 732
- Yu Q., Tremaine S., 2002, *MNRAS*, **335**, 965
- Yue M., Eilers A.-C., Ananna T. T., Panagiotou C., Kara E., Miyaji T., 2024, *ApJ*, **974**, L26
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Hirschmann M., Davé R., Popping G., Gardner J. P., Venkatesan A., 2021, *MNRAS*, **508**, 2706
- Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, *ApJ*, **771**, L19
- Zhang H., Behroozi P., Volonteri M., Silk J., Fan X., Hopkins P. F., Yang J., Aird J., 2023, *MNRAS*, **518**, 2123
- Zhang Z., et al., 2025a, *arXiv e-prints*, p. arXiv:2512.05180
- Zhang Z., Jiang L., Liu W., Ho L. C., 2025b, *ApJ*, **985**, 119
- Zhuang M.-Y., Li J., Shen Y., Lin X., Shapley A. E., Wang F., Wu Q., Yang Q., 2026, *ApJ*, **999**, 31
- Zier O., Kannan R., Smith A., Vogelsberger M., Verbeek E., 2024, *MNRAS*, **533**, 268
- Zier O., et al., 2026, *arXiv e-prints*, p. arXiv:2605.15310
- de Graaff A., et al., 2025a, *arXiv e-prints*, p. arXiv:2511.21820
- de Graaff A., et al., 2025b, *Nature Astronomy*, **9**, 280
- de Graaff A., et al., 2025c, *A&A*, **701**, A168

APPENDIX A: OPTICAL LF AT 5100 Å

In Section 4.2, we present the AGN LFs in the rest-frame optical B band. Here we show the same LFs instead at $\lambda \sim 5100$ Å, the wavelength at which many observational studies of LRDs report their LFs (Kokorev et al. 2024; Ma et al. 2025). Since both our model predictions and the compiled quasar LF data (Shen et al. 2020a) are defined in the B band, we apply a small constant offset of -0.04 dex to convert them to 5100 Å, derived from our quasar SED template. Figure A1 shows the resulting 5100 Å AGN LF in LUMINA, compared to TNG100 and MTNG. We include the original measurements from

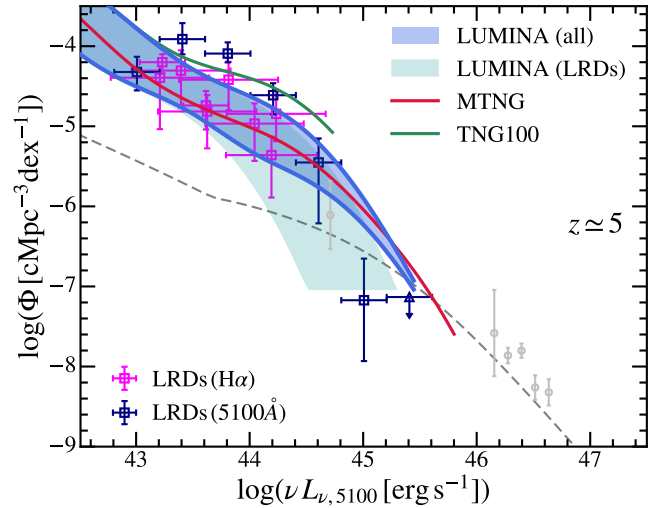


Figure A1. Optical (~ 5100 Å) AGN LFs in LUMINA, TNG100, and MTNG. We compare the simulation predictions with the same set of observations shown in Figure 9 in the main text. Our conclusions about optical LF are essentially unchanged.

Kokorev et al. (2024); Ma et al. (2025) together with the $H\alpha$ -based constraints of Matthee et al. (2024); Lin et al. (2024, 2026a), the latter converted to L_{5100} using the same scaling relation adopted in the main text. The conclusions reached in Section 4.2 remain unchanged.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.