

# The host galaxies of $z = 7$ quasars: predictions from the BLUE TIDES simulation

Madeline A. Marshall <sup>1,2</sup>★ Yueying Ni,<sup>3</sup> Tiziana Di Matteo,<sup>3</sup> J. Stuart B. Wyithe,<sup>1,2</sup> Stephen Wilkins <sup>4</sup>,  
Rupert A. C. Croft<sup>3</sup> and Jussi K. Kuusisto<sup>4</sup>

<sup>1</sup>*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

<sup>2</sup>*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*

<sup>3</sup>*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA*

<sup>4</sup>*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*

Accepted 2020 September 23. Received 2020 September 23; in original form 2019 December 4

## ABSTRACT

We examine the properties of the host galaxies of  $z = 7$  quasars using the large volume, cosmological hydrodynamical simulation BLUE TIDES. We find that the 10 most massive black holes and the 191 quasars in the simulation (with  $M_{\text{UV,AGN}} < M_{\text{UV,host}}$ ) are hosted by massive galaxies with stellar masses  $\log(M_*/M_\odot) = 10.8 \pm 0.2$ , and  $10.2 \pm 0.4$ , which have large star formation rates of  $513_{-351}^{+1225} M_\odot/\text{yr}$  and  $191_{-120}^{+288} M_\odot/\text{yr}$ , respectively. The hosts of the most massive black holes and quasars in BLUE TIDES are generally bulge-dominated, with bulge-to-total mass ratio  $B/T \simeq 0.85 \pm 0.1$ ; however, their morphologies are not biased relative to the overall  $z = 7$  galaxy sample. We find that the hosts of the most massive black holes and quasars are compact, with half-mass radii  $R_{0.5} = 0.41_{-0.14}^{+0.18}$  kpc and  $0.40_{-0.09}^{+0.11}$  kpc, respectively; galaxies with similar masses and luminosities have a wider range of sizes with a larger median value,  $R_{0.5} = 0.71_{-0.25}^{+0.28}$  kpc. We make mock James Webb Space Telescope (JWST) images of these quasars and their host galaxies. We find that distinguishing the host from the quasar emission will be possible but still challenging with JWST, due to the small sizes of quasar hosts. We find that quasar samples are biased tracers of the intrinsic black hole–stellar mass relation, following a relation that is 0.2 dex higher than that of the full galaxy sample. Finally, we find that the most massive black holes and quasars are more likely to be found in denser environments than the typical  $M_{\text{BH}} > 10^{6.5} M_\odot$  black hole, indicating that minor mergers play at least some role in growing black holes in the early Universe.

**Key words:** galaxies: quasars: supermassive black holes – galaxies: evolution – galaxies: high-redshift.

## 1 INTRODUCTION

High-redshift quasars ( $z \gtrsim 6$ , Fan et al. 2000, 2001, 2003, 2004) are some of the most extreme systems in the Universe, with intense accretion at or even above the Eddington limit (Willott et al. 2010b; De Rosa et al. 2011, 2014; Trakhtenbrot, Volonteri & Natarajan 2017b) forming black holes with masses of  $10^8$ – $10^9 M_\odot$  (Barth et al. 2003; Jiang et al. 2007; Kurk et al. 2007; De Rosa et al. 2011, 2014) in less than a billion years. These luminous systems are invaluable probes of the early Universe, providing constraints on black hole seed theories (e.g. Mortlock et al. 2011; Volonteri 2012; Bañados et al. 2017), the Epoch of Reionization (e.g. Fan et al. 2006b; Mortlock et al. 2011; Greig & Mesinger 2017; Davies et al. 2018; Greig, Mesinger & Bañados 2019), and the relation between the growth of black holes and their host galaxies (e.g. Shields et al. 2006; Wang et al. 2013; Valiante et al. 2014; Schulze & Wisotzki 2014; Willott, Bergeron & Omont 2017). The space density of typical Sloan Digital Sky Survey (SDSS) quasars ( $M_{\text{UV,AGN}} < -26$ ) is less than 1 per Gpc<sup>3</sup> at  $z \gtrsim 6$  (Willott et al. 2010a; Kashikawa et al. 2015; Jiang et al. 2016; Wang et al. 2019). Their rarity and extreme properties raise many questions such as ‘Are the biggest black holes found in the rarest, most overdense regions, i.e. the biggest

haloes and galaxies (e.g. Springel et al. 2005b; Shen et al. 2007; Fanidakis et al. 2013; Ren, Trenti & Matteo 2020)?’ and ‘Is this rapid growth driven by galaxy mergers, with hosts that are highly star forming, or are their host galaxies more discy and quiet (e.g. Mor et al. 2012; Netzer et al. 2014; Trakhtenbrot et al. 2017a)?’ For further discussion, see, for example, the recent reviews of Valiante et al. (2017), Mayer & Bonoli (2018), and Inayoshi, Visbal & Haiman (2020).

Understanding the host galaxies of high-redshift quasars is essential for addressing these questions. However, this requires detection and ideally accurate measurements of quasar host galaxies, which is extremely challenging with current telescopes (see e.g. Mechtley et al. 2012). In the rest-frame ultraviolet (UV)/optical, which traces the emission from the accretion disc and the host stellar component, the quasars often outshine their hosts, entirely concealing the host galaxy emission (Mechtley et al. 2012). The detection of  $z \gtrsim 6$  quasar host galaxies has indeed eluded the *Hubble Space Telescope* (HST). The only current detections of high-redshift quasar hosts are instead in the rest-frame far-infrared, observed in the sub-mm (e.g. Bertoldi et al. 2003; Walter et al. 2003, 2004; Riechers et al. 2007; Wang et al. 2010, 2011; Venemans et al. 2019), which traces cold dust in the host galaxy.

Observations in sub-mm and mm wavelengths with the Atacama Large Millimeter Array (ALMA) and the IRAM Plateau de Bure Interferometer (PdBI), for example, imply a diverse population

\* E-mail: [madelinem1@student.unimelb.edu.au](mailto:madelinem1@student.unimelb.edu.au)

of quasar hosts, with inferred dynamical masses of  $10^{10}$ – $10^{11} M_{\odot}$  (Walter et al. 2009; Wang et al. 2013; Venemans et al. 2016, 2017; Willott et al. 2017; Trakhtenbrot et al. 2017a; Izumi et al. 2018; Pensabene et al. 2020; Nguyen et al. 2020), dust masses of  $10^7$ – $10^9 M_{\odot}$  (Venemans et al. 2016; Izumi et al. 2018; Nguyen et al. 2020), sizes of 1–5 kpc (Wang et al. 2013; Venemans et al. 2016; Willott et al. 2017; Izumi et al. 2018), and a wide range of star formation rates (SFRs) of 10–3000  $M_{\odot}/\text{yr}$  (Venemans et al. 2016, 2017; Willott et al. 2017; Trakhtenbrot et al. 2017a; Izumi et al. 2018; Shao et al. 2019; Nguyen et al. 2020). The hosts are found in a variety of dynamical states, with some having nearby companions, which may suggest a merger system (e.g. Trakhtenbrot et al. 2017a), while some show signatures of a rotating disc (e.g. Willott et al. 2017; Trakhtenbrot et al. 2017a), or even no ordered motion (Venemans 2017). However, since cold dust may not trace the stellar distribution, there may be significant biases in stellar properties inferred through these observations (e.g. Narayanan et al. 2009; Valiante et al. 2014; Lupi et al. 2019).

Upcoming facilities will provide the next frontier for understanding high-redshift quasars and their host galaxies. Infrared surveys with *Euclid* (Amiaux et al. 2012) and the Nancy Grace Roman Space Telescope (RST, formerly the Wide Field Infrared Survey Telescope or WFIRST; Spergel et al. 2015) will significantly increase the known sample of  $z \gtrsim 6$  quasars. The improved resolution of the James Webb Space Telescope (JWST; Gardner et al. 2006) will allow the first detections of the stellar component of their host galaxies, which will be invaluable for accurately determining the properties of quasar hosts. Making detailed theoretical predictions for the results of these groundbreaking instruments is thus a current priority.

Due to the rarity of high-redshift quasars, comprehensive theoretical predictions require high-resolution simulations with large computational volumes. Cosmological hydrodynamical simulations such as Massive Black (Di Matteo et al. 2012), with a volume of  $(0.76 \text{ Gpc})^3$ , and BLUE TIDES (Feng et al. 2015), with a volume of  $(0.57 \text{ Gpc})^3$ , have pioneered this area. These simulations have been used to investigate the rapid growth of black holes (Di Matteo et al. 2012; DeGraf et al. 2012b; Feng et al. 2014; Di Matteo et al. 2017) and their relationship to their host galaxies (Khandai et al. 2012; DeGraf et al. 2015; Huang et al. 2018) and make predictions for the highest redshift quasars that are observed (DeGraf et al. 2012a; Tenneti et al. 2018; Ni et al. 2018).

Previous BLUE TIDES analyses were performed with the phase I simulation, which reached a minimum redshift of  $z = 8.0$ , and BLUE TIDES-II, the second phase of the simulation, which had been run to  $z = 7.5$  when last analysed (Tenneti et al. 2018). In this paper, we use the BLUE TIDES-II simulation extended further to  $z = 7.0$  to make predictions for the properties of quasar host galaxies. At  $z = 7.5$ , there is one quasar analogue in the BLUE TIDES simulation, as studied by Tenneti et al. (2018). Extending the simulation from  $z = 7.5$  to  $7.0$ , a period of only 58 Myr, results in a considerable increase in the number of observable quasar analogues to the order of 100, since this is such an intense growth period for black holes in the Universe. This statistical sample allows us to make predictions for the broader quasar population and not just for individual, extreme systems as was possible previously.

The paper is outlined as follows. In Section 2, we describe the simulation and the post-processing used to obtain mock spectra of the quasars and their host galaxies. In Section 3, we consider the intrinsic galaxy properties of the hosts of black holes and quasars. We consider observable properties in Section 4, making spectra and mock JWST images. In Section 5, we examine the black hole–stellar mass relation,

showing how observations of these quasars will lead to a biased measurement. We explore the environments of quasars in Section 6 before concluding in Section 7. The cosmological parameters used throughout are from the 9-yr Wilkinson Microwave Anisotropy Probe (WMAP; Hinshaw et al. 2013):  $\Omega_M = 0.2814$ ,  $\Omega_{\Lambda} = 0.7186$ ,  $\Omega_b = 0.0464$ ,  $\sigma_8 = 0.820$ ,  $\eta_s = 0.971$ , and  $h = 0.697$ .

## 2 SIMULATION

### 2.1 BLUE TIDES

The BLUE TIDES simulation<sup>1</sup> (Feng et al. 2015) is a cosmological hydrodynamical simulation, which uses the Pressure Entropy Smoothed Particle Hydrodynamics (SPH) code MP-Gadget to model the evolution of  $2 \times 7040^3$  particles in a cosmological box of volume  $(400/h \text{ cMpc})^3$ . The mass resolution of the simulation is  $1.2 \times 10^7/h M_{\odot}$  for dark matter particles and  $2.4 \times 10^6/h M_{\odot}$  for gas particles (in the initial condition). Star particles are converted from gas particles with sufficient SFRs, and each has a stellar mass of  $6 \times 10^5/h M_{\odot}$ . The gravitational softening length of  $\epsilon_{\text{grav}} = 1.5/h \text{ ckpc}$  is the effective spatial resolution. From the initial conditions at  $z = 99$ , BLUE TIDES evolved the box to  $z = 8$  in phase I (Feng et al. 2015). Phase II of the simulation continued the evolution of the box from  $z = 8$  to lower redshifts, with the first results from this phase given in Tenneti et al. (2018). Here, we focus on the lowest redshift currently reached by phase II,  $z = 7.0$ . From the simulation, we consider the 108 000 most massive haloes, with masses  $M_{\text{vir}} > 10^{10.8} M_{\odot}$ , which contain galaxies with  $M_* > 10^{5.9} M_{\odot}$  and black holes with  $M_{\text{BH}} > 10^{5.8} M_{\odot}$  (the seed mass).

BLUE TIDES implements a variety of sub-grid physics to model galaxy and black hole formation and their feedback processes. Here, we briefly list some of its basic features and refer the reader to the original paper (Feng et al. 2015) for more detailed descriptions. In the BLUE TIDES simulation, gas cooling is performed through both primordial radiative cooling (Katz, Hernquist & Weinberg 1999) and metal line cooling (Vogelsberger et al. 2014). Star formation is based on the multiphase star formation model originally from Springel & Hernquist (2003) with modifications following Vogelsberger et al. (2013). We also implement the formation of molecular hydrogen and model its effects on star formation using the prescription from Krumholz & Gnedin (2011), where we self-consistently estimate the fraction of molecular hydrogen gas from the baryon column density, which in turn couples the density gradient into the SFR. For stellar feedback, we apply a type-II supernova wind feedback model from Okamoto et al. (2010), assuming wind speeds proportional to the local one-dimensional dark matter velocity dispersion. The large volume of BLUE TIDES also allows the inclusion of a model of ‘patchy reionization’ (Battaglia et al. 2013), yielding a mean reionization redshift  $z \simeq 10$  and incorporating the UV background estimated by Faucher-Giguère et al. (2009).

The black hole sub-grid model associated with black hole growth and active galactic nuclei (AGN) feedback is the same as that in the MASSIVEBLACK I & II simulations, originally developed in Springel, Di Matteo & Hernquist (2005a) and Di Matteo, Springel & Hernquist (2005), with modifications consistent with ILLUSTRIS; see DeGraf et al. (2012b) and DeGraf et al. (2015) for full details. Black holes are seeded with a mass of  $M_{\text{BH,seed}} = 5 \times 10^5/h M_{\odot}$  in dark matter haloes above a threshold mass of

<sup>1</sup><http://BlueTides-project.org/>

$M_{\text{Halo}} = 5 \times 10^{10}/h M_{\odot}$ . The simulation makes no direct assumption of the black hole formation mechanism, although this mass is most consistent with seed masses predicted by direct collapse scenarios (e.g. Begelman, Volonteri & Rees 2006; Shang, Bryan & Haiman 2010; Volonteri 2010; Latif et al. 2013). Black holes grow by merging with other black holes, and via gas accretion at the Bondi–Hoyle accretion rate (Hoyle & Lyttleton 1939; Bondi & Hoyle 1944; Bondi 1952),  $\dot{M}_{\text{BH}} = \alpha 4\pi G^2 M_{\text{BH}}^2 \rho_{\text{BH}} (c_s^2 + v^2)^{-3/2}$ , where  $\rho_{\text{BH}}$  is the local gas density,  $c_s$  is the local sound speed,  $v$  is the velocity of the black hole relative to the surrounding gas, and  $\alpha$  is a dimensionless parameter. Mildly super-Eddington accretion is permitted, with the accretion rate limited to two times the Eddington limit. In this sub-grid model, the black hole mass grows smoothly at this accretion rate. In order to account for the discrete nature of gas particles, once a black hole has grown by an amount equivalent to the mass of a gas particle, a gas particle is removed and its mass transferred to the black hole’s ‘dynamical mass’. This discrete dynamical mass allows the particle dynamics to be calculated correctly within the simulation; however, the continuous black hole mass is always used in any analyses. Finally, we assume that black holes radiate with bolometric luminosity  $L_{\text{AGN}} = \eta \dot{M}_{\text{BH}} c^2$ , with a radiative efficiency  $\eta$  of 0.1.

Throughout this work, we generally consider only black holes with  $M_{\text{BH}} > 10^{6.5} M_{\odot}$ , in order to minimize any possible influence of the seeding prescription on our analysis. We also consider only the  $z = 7.0$  snapshot. This is in contrast to Tenneti et al. (2018), which explored a range of snapshots around  $z \simeq 7.5$  in order to find the brightest quasar in the simulation, as quasar luminosity varies significantly due to the time variability of black hole accretion. Using only one snapshot is more representative of an observational sample, in which galaxies are observed at a random phase in their growth history, and not, for example, only when their black hole is at its peak luminosity.

To extract the properties of galaxies from the simulation, we run a friends-of-friends algorithm (Davis et al. 1985). The galaxy properties, such as the star formation density, stellar mass function, and UV luminosity function, have been shown to match current observational constraints at  $z = 8, 9$ , and 10 (Feng et al. 2015; Waters et al. 2016; Wilkins et al. 2017).

To determine the stellar mass of the galaxies from the total stellar mass contained in their host dark matter haloes, we calculate the galaxy  $R_{200}$ , the radius containing 200 times the critical stellar mass density (the critical density of the Universe multiplied by the baryon fraction and star formation efficiency of the simulation). We define the stellar mass of a galaxy as the mass contained within  $R_{200}$ . This generally includes the inner dense core of the galaxy and also the more diffuse outer regions, ensuring that we include the majority of particles truly associated with each galaxy. For determining the sizes of galaxies, we calculate the half-mass radius inside this  $R_{200}$ ,  $R_{0.5}$ .

We determine the morphology of the galaxy by its bulge-to-total ratio, calculated using the bulge-to-disc decomposition method of Scannapieco et al. (2009). We first construct a circularity parameter  $\epsilon = j_z / j_{\text{circ}}(r)$  for each star particle in the galaxy within  $R_{200}$ , where  $j_z$  is the projection of the specific angular momentum of the star particle in the direction of the total angular momentum of the galaxy, and  $j_{\text{circ}}(r)$  is the angular momentum expected for a circular orbit at the radius  $r$ :  $j_{\text{circ}} = r v_{\text{circ}}(r) = \sqrt{GM(< r)r}$ . We identify star particles with  $\epsilon > 0.7$  as disc stars and define the bulge-to-total ratio as  $B/T = 1 - f_{\epsilon > 0.7}$ , where  $f_{\epsilon > 0.7}$  is the fraction of disc stars in the galaxy.

## 2.2 Mock spectra

### 2.2.1 Galaxy SEDs

To determine the spectral energy distribution (SED) of a galaxy, we assign a SED from a simple stellar population to each star particle within  $R_{200}$ , based on its mass, age, and metallicity. We do this using the Binary Population and Spectral Population Synthesis model (BPASS, version 2.2; Stanway & Eldridge 2018), assuming a modified Salpeter initial mass function with a high-mass cutoff of  $300 M_{\odot}$ . The SED of the galaxy is taken as the sum of the SEDs of each of its star particles. To determine the relative contribution of the stellar and nebular emission, we assume an escape fraction of 0.9.

### 2.2.2 Quasar spectra

To assign spectra to each of our quasars, we use the CLOUDY spectral synthesis code (Ferland et al. 2017), as in Tenneti et al. (2018).

The continuum is given by

$$f_{\nu} = \nu^{\alpha_{\text{UV}}} \exp\left(\frac{-h\nu}{kT_{\text{BB}}}\right) \exp\left(\frac{-kT_{\text{IR}}}{h\nu}\right) + a\nu^{\alpha_{\text{X}}}, \quad (1)$$

where  $\alpha_{\text{UV}} = -0.5$ ,  $\alpha_{\text{X}} = -1$ ,  $kT_{\text{IR}} = 0.01\text{Ryd}$ , and  $T_{\text{BB}}$  are the temperature of the accretion disc, which is determined by the black hole mass and its accretion rate

$$T_{\text{BB}} = \left(\frac{3c^6}{8\pi 6^3 \sigma_{\text{SB}} G^2} \frac{\dot{M}_{\text{BH}}}{M_{\text{BH}}^2}\right)^{1/4} = 2.24 \times 10^9 \left(\frac{\dot{M}_{\text{BH}}}{M_{\odot}/\text{yr}}\right)^{1/4} \left(\frac{M_{\text{BH}}}{M_{\odot}}\right)^{-1/2} \text{K}. \quad (2)$$

The normalization of the continuum is set by the bolometric luminosity of the quasar.

The emission lines are calculated with CLOUDY assuming a hydrogen density of  $10^{10}\text{cm}^{-3}$  at the face of the cloud, which has inner radius  $10^{18}\text{cm}$ , and a total hydrogen column density of  $10^{22}\text{cm}^{-2}$ .

We also implement Lyman-forest extinction on the redshifted spectra (Madau 1995; STScI Development Team 2018) for both the quasars and the host galaxies.

### 2.2.3 Dust attenuation and extinction

As in Wilkins et al. (2017), we model the dust attenuation of galaxies by relating the density of metals along a line of sight to the UV-band dust optical depth  $\tau_{\text{UV}}$ . For each star particle in the galaxy, we calculate  $\tau_{\text{UV},*}$  as

$$\tau_{\text{UV},*} = -\kappa \Sigma(x, y, z) \left(\frac{\lambda}{5500\text{\AA}}\right)^{\gamma}, \quad (3)$$

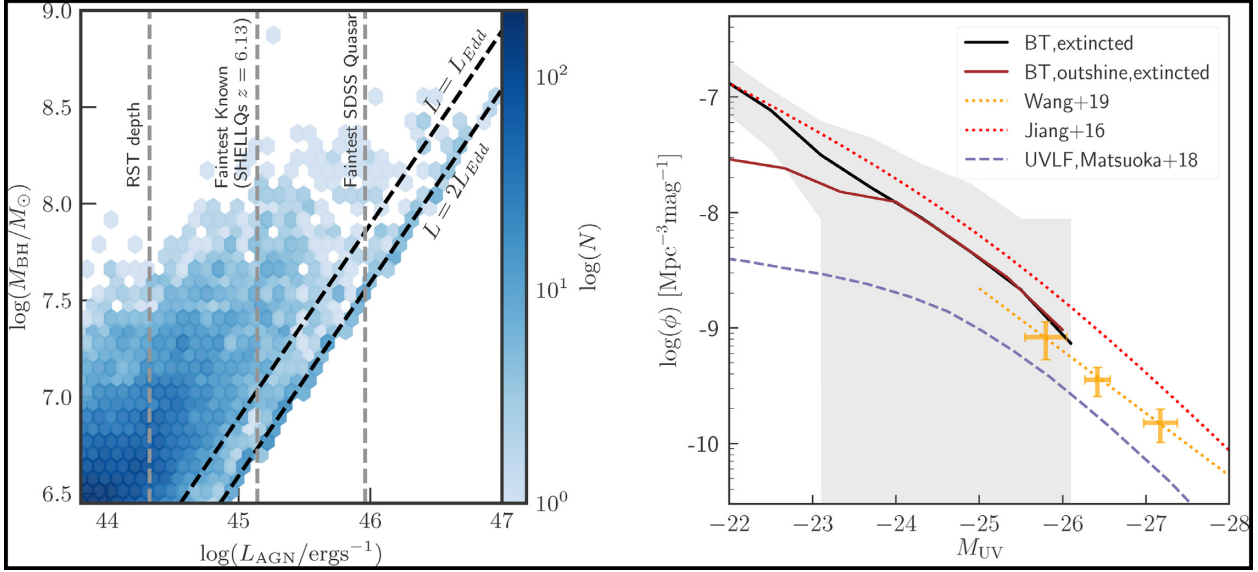
where  $\Sigma(x, y, z) = \int_{z'=0}^z \rho_{\text{metal}}(x, y, z') dz'$  is the metal surface density at the position of the star particle, along the  $z$ -direction line of sight, and  $\kappa$  and  $\gamma$  are free parameters. Here, we use  $\kappa = 10^{4.6}$  and  $\gamma = -1.0$ , which are calibrated against the observed galaxy UV luminosity function at redshift  $z = 7$ . The total dust-attenuated galaxy luminosity is the sum of the extinguished luminosities of each individual star particle.

We apply the same technique to determine the dust attenuation of the AGN, with the dust optical depth calculated using the metal column density integrated along a line of sight to the quasar:

$$\tau_{\text{UV,AGN}} = \kappa \int_{\text{ray}} \rho_{\text{metal}}(l) dl \left(\frac{\lambda}{5500\text{\AA}}\right)^{\gamma}, \quad (4)$$

with the same values of  $\kappa$  and  $\gamma$  (see also Ni et al. 2019). Dust attenuation of the AGN mainly traces the regions of high gas





**Figure 1.** *Left:* The distribution of black hole masses and AGN bolometric luminosities for BLUETIDES galaxies at  $z = 7$  (blue density plot). The Eddington limit is shown for reference (dashed black line), as well as twice the Eddington limit, which is the upper limit of the black hole accretion rate set in the simulation. The luminosity of the faintest SDSS quasar, the faintest currently known high-redshift quasar, and the RST detection limit are shown for reference (grey dashed lines). *Right:* The UV-band luminosity function of AGN at  $z = 7$ . The black solid line is the dust-extinguished quasar luminosity function from BLUETIDES. The brown solid line is the luminosity function including only the AGN that outshine their host galaxies. The grey shaded area gives the error estimate by considering the dust extinction through all lines of sight of the AGN population. The orange solid symbols with error bars show the measured binned quasar luminosity function from Wang et al. (2019) at  $z \simeq 6.7$ . The red dotted line is the  $z \simeq 6$  fitted quasar luminosity function measured by Jiang et al. (2016). The purple dashed line gives the luminosity function from Matsuoka et al. (2018b), based on the population of  $5.7 < z < 6.5$  quasars and extrapolated to  $z = 7$ .

density near the centre of the galaxy, with the gas metallicity  $Z$  only modulating the dust extinction at a sub-dominant level (see fig. 14 in Ni et al. 2019, for illustration). Because of the angular variation in the density field surrounding the central black hole, the dust extinction for AGN is sensitive to the choice of line of sight, unlike for galaxies, whose dust attenuation is accumulated over the extended source. For each quasar, we therefore calculate  $\tau_{UV,AGN}$  along approximately 1000 lines of sight. See Ni et al. (2019) for full details.

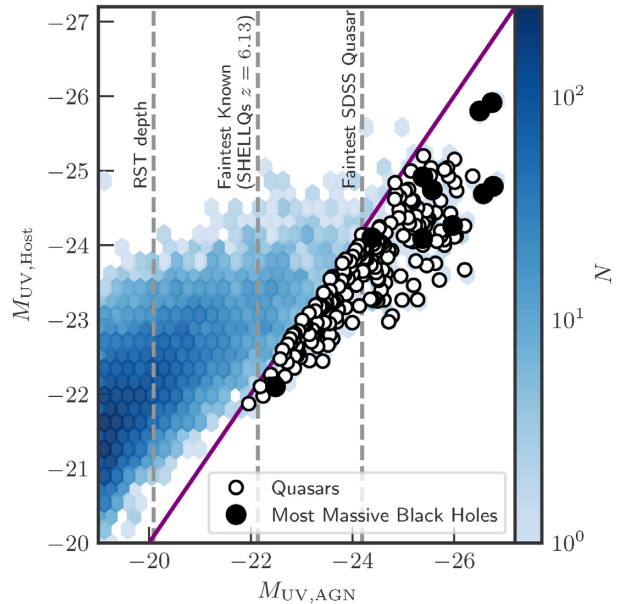
### 3 PROPERTIES OF BLACK HOLE HOST GALAXIES

#### 3.1 Sample selection

The black hole population in BLUETIDES at  $z = 7.0$  is presented in Fig. 1, which shows the distributions of their masses and AGN luminosities, as well as the AGN UV luminosity function. By considering the UV-band dust extinction as described in Section 2.2.3 and Ni et al. (2019), BLUETIDES produces a quasar luminosity function that is in good agreement with the high-redshift observations of Jiang et al. (2016) and Wang et al. (2019) – BLUETIDES predicts the expected number density of high-redshift quasars. The most massive black holes in BLUETIDES at  $z = 7$  have masses of  $M_{BH} \simeq 10^{8.5} M_{\odot}$ , and the most luminous AGN have intrinsic bolometric luminosities,  $L_{bol} \simeq 10^{47} \text{erg s}^{-1} \simeq 10^{13} L_{\odot}$ , equivalent to those of currently observed high-redshift quasars.

From this black hole population, we select three samples; the most massive black holes, ‘quasars’, and ‘hidden quasars’.

**Massive black hole sample:** We consider the 10 most massive black holes at  $z = 7$ , which have masses  $M_{BH} = 10^{8.44} - 10^{8.89} M_{\odot}$ , to be our ‘massive black hole’ sample.



**Figure 2.** The distribution of host and AGN intrinsic UV absolute magnitudes for BLUETIDES galaxies at  $z = 7$  (blue density plot). We classify quasars (white circles) as those with  $M_{UV,AGN} < M_{UV,Host}$ , since the AGN outshines the host galaxy. The 10 most massive black holes (black circles) are also shown. The luminosity of the faintest SDSS quasar, the faintest currently known high-redshift quasar, and the RST detection limit are shown for reference (grey dashed lines).



**Quasar sample:** To select ‘quasars’ from the simulation, we consider the galaxy and AGN UV-band absolute magnitudes, as shown in Fig. 2. We make the simple assumption that every bright ( $L_{\text{AGN}} > 10^{44} \text{erg s}^{-1}$ ) black hole with  $M_{\text{UV,AGN}} < M_{\text{UV,Host}}$  would be classified as a quasar, since the AGN outshines the host galaxy. This results in a sample of 205 BLUE TIDES quasars, which is only 2.1 per cent of the black holes with  $M_{\text{BH}} > 10^{6.5} M_{\odot}$ , and 2.6 per cent of black holes with  $L_{\text{AGN}} > 10^{44} \text{erg s}^{-1}$  (see Fig. 2). We note that the assumption that  $M_{\text{UV,AGN}} < M_{\text{UV,Host}}$  for a galaxy to be classified as a quasar is not an accurate representation of the true observational quasar selection techniques and may underestimate the number of galaxies in our sample that would be observed as quasars.

**Hidden quasar sample:** Within our simulation, 70.5 per cent of black holes brighter than the faintest currently known high-redshift quasar,  $m_{\text{UV}} = 24.85$  ( $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$  at  $z = 7$ ; Matsuoka et al. 2018a), have host luminosities that outshine the AGN. These 488 black holes are experiencing significant black hole growth, with high AGN luminosities, but are simply ‘hidden’ by their luminous host galaxies. We consider all black holes with intrinsic luminosities  $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$  and with  $M_{\text{UV,AGN}} > M_{\text{UV,Host}}$  as ‘hidden’ quasars, i.e. those outshined by their host galaxy.

### 3.2 Galaxy properties

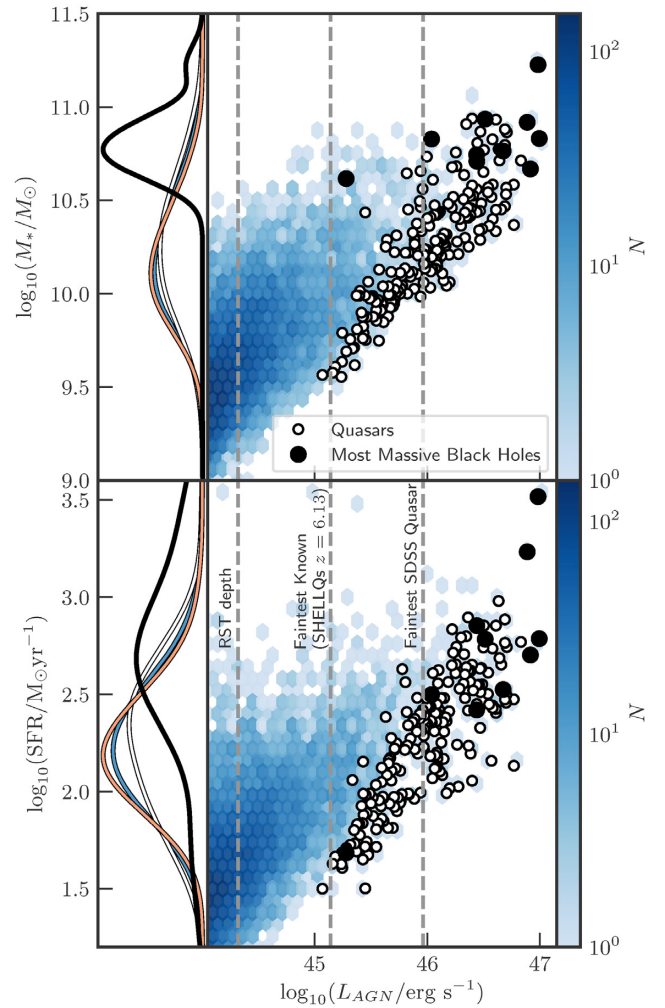
We now investigate the properties of the hosts of the most massive black holes and quasars in BLUE TIDES at  $z = 7$ .

In Fig. 3, we show the relation between AGN luminosity and both stellar mass and SFR at  $z = 7$ . The most massive black holes are in massive galaxies with stellar masses  $\log(M_*/M_{\odot}) = 10.80^{+0.20}_{-0.16}$ , which have large SFRs,  $513^{+1225}_{-351} M_{\odot}/\text{yr}$ .<sup>2</sup> We find that the quasar hosts also have large but lower stellar masses of  $\log(M_*/M_{\odot}) = 10.25^{+0.40}_{-0.37}$  and lower SFRs of  $191^{+288}_{-120} M_{\odot}/\text{yr}$ .

These SFRs are broadly consistent with those observed in the hosts of luminous high-redshift quasars with the PdBI of  $\simeq 1700 M_{\odot}/\text{yr}$  (Walter et al. 2009), and with ALMA: 100–1600  $M_{\odot}/\text{yr}$  (Venemans et al. 2016), 200–3500  $M_{\odot}/\text{yr}$  (Trakhtenbrot et al. 2017a), 30–3000  $M_{\odot}/\text{yr}$  (Decarli et al. 2018), 50–2700  $M_{\odot}/\text{yr}$  (Venemans et al. 2018), 900–3200  $M_{\odot}/\text{yr}$  (Nguyen et al. 2020), and  $\simeq 2500 M_{\odot}/\text{yr}$  (Shao et al. 2019). However, the simulation does not contain quasar hosts with extreme SFRs of  $\gtrsim 1000 M_{\odot}/\text{yr}$ , as are observed. This is most likely because by  $z = 7$ , BLUE TIDES has not yet produced a population of extremely luminous quasars, which are those generally found in such extreme hosts. Note, however, that SFRs derived from far-infrared observations can have uncertainties of a factor of  $\sim 2$ –3 (e.g. Venemans et al. 2018). A comparison of BLUE TIDES at lower redshift with more precise SFRs measured in the rest-frame UV using JWST, for example, would allow for a deeper understanding of quasar host SFRs.

From Fig. 3, we see that, on average, lower luminosity quasars have less extreme host galaxies, with lower masses and SFRs. The hosts of lower luminosity high-redshift quasars are indeed observed to have lower SFRs:  $\lesssim 10 M_{\odot}/\text{yr}$  (Willott et al. 2017), 100–500  $M_{\odot}/\text{yr}$  (Trakhtenbrot et al. 2017a), 23–40  $M_{\odot}/\text{yr}$  (Izumi et al. 2018), and 200–500  $M_{\odot}/\text{yr}$  (Nguyen et al. 2020). While our  $z = 7$  predictions do not extend to such low SFRs, we expect that by  $z \simeq 6$ , there could be more scatter in the relation, alongside more ‘quenched’ quasars, where feedback has significantly reduced the star formation in the host galaxy.

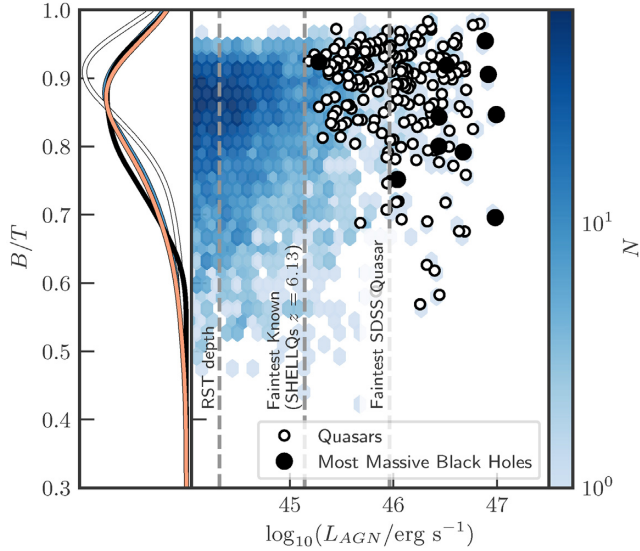
<sup>2</sup>Errors presented in this manner correspond to the 16th and 84th percentiles of the distributions, relative to the median value.



**Figure 3.** The relations between AGN luminosity and stellar mass (upper right-hand panel) and star formation rate (lower right-hand panel). The blue density plot shows the distribution for all BLUE TIDES galaxies, with the most massive black holes and quasars also plotted (see legend). The left-hand panels show the distributions of the host properties for the most massive black holes (black line), quasars (white line), hidden quasars (salmon line), and for all black holes with  $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$  (blue line). The luminosity of the faintest SDSS quasar, the faintest currently known high-redshift quasar ( $L_{\text{AGN}} = 10^{45.1} \text{erg s}^{-1}$ ), and the RST detection limit are shown in the right-hand panels for reference (grey dashed lines).

Fig. 3 also shows the one-dimensional distributions of stellar mass and SFR for these samples, alongside all black holes brighter than  $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$ , and ‘hidden’ quasars. This shows that the most massive black holes live in more massive galaxies, with higher SFRs, than the total sample of bright black holes ( $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$ ), which have  $\log(M_*/M_{\odot}) = 10.18^{+0.37}_{-0.32}$  and SFRs of  $170^{+219}_{-90} M_{\odot}/\text{yr}$ . The hidden quasars are hosted by galaxies with  $\log(M_*/M_{\odot}) = 10.16^{+0.34}_{-0.30}$  and SFRs of  $166^{+189}_{-83} M_{\odot}/\text{yr}$ .

At a fixed AGN luminosity, the quasars are hosted by less massive galaxies with lower SFRs than the hidden quasars. This is expected due to the  $M_{\text{UV,AGN}} < M_{\text{UV,Host}}$  selection: the quasar sample contains galaxies with lower  $M_{\text{UV,Host}}$  for fixed  $M_{\text{UV,AGN}}$ , which is produced by having a lower SFR. Galaxies with higher SFRs have higher luminosities, which outshine their quasar, resulting in ‘hidden’ quasars of the same quasar luminosity. As stellar mass is an integrated quantity, the selection effect is weakened slightly.



**Figure 4.** The relation between AGN luminosity and the bulge-to-total mass ratio ( $B/T$ ). The blue density plot shows the distribution for all BLUE TIDES galaxies, with the most massive black holes and quasars also plotted (see legend). The left-hand panel shows the distribution of  $B/T$  for the most massive black holes (black line), quasars (white line), hidden quasars (salmon line), and all black holes brighter than  $L_{\text{AGN}} > 10^{45.1} \text{ erg s}^{-1}$  (blue line). The luminosity of the faintest SDSS quasar, the faintest currently known high-redshift quasar ( $L_{\text{AGN}} = 10^{45.1} \text{ erg s}^{-1}$ ), and the RST detection limit are shown in the right-hand panel for reference (grey dashed lines).

In Fig. 4, we show the relation between AGN luminosity and the ratio of stellar mass contained in a galaxy’s bulge to its total stellar mass ( $B/T$ ). The hosts of the most massive black holes and quasars all show bulge-dominated morphologies, although there is a large tail to lower  $B/T$ , with  $B/T = 0.85^{+0.09}_{-0.10}$ , and  $0.89^{+0.07}_{-0.10}$  for the two samples, respectively. Their morphologies have a similar distribution to that of the total sample of bright black holes ( $L_{\text{AGN}} > 10^{45.1} \text{ erg s}^{-1}$ ), with  $B/T = 0.85^{+0.09}_{-0.12}$ , and hidden quasars, with  $B/T = 0.84^{+0.10}_{-0.14}$ . The hosts of the most massive black holes and quasars in BLUE TIDES are generally bulge-dominated but are not biased in morphology relative to the overall galaxy sample at  $z = 7$ .

Lupi et al. (2019) performed a high-resolution cosmological zoom-in simulation of a halo containing a black hole with mass  $M_{\text{BH}} = 10^{8.9} M_{\odot}$  at  $z = 7$ , similar to that of the most massive black hole in BLUE TIDES. They found its host galaxy to have a mass of  $M_{*} \simeq 10^{11} M_{\odot}$  and a large SFR of  $\sim 10^{2.5} M_{\odot}/\text{yr}$  at  $z = 7$ , equivalent to the most massive and star-forming galaxies in BLUE TIDES. Their quasar host is less bulge-dominated than those in our quasar sample, with a bulge-to-total mass ratio of  $B/T \simeq 0.45$ . This is potentially due to the increased resolution of their simulation, which has the ability to better resolve the disc structure. The results of Lupi et al. (2019) are therefore reasonably consistent with the BLUE TIDES simulation, given their sample of only one quasar host.

Fig. 5 shows the relation between half-mass radius  $R_{0.5}$  and stellar mass, halo mass, and host UV magnitude. For comparison, we also show a range of observations of  $z \simeq 6-7$  Lyman-break galaxies with  $M_{\text{UV}} < -20$ . These observed galaxies have a wide range of sizes, consistent with the sizes of the general BLUE TIDES galaxy sample. The most massive black hole hosts in BLUE TIDES have small radii of  $R_{0.5} = 0.41^{+0.18}_{-0.14} \text{ kpc}$ , as do the quasar hosts, with  $R_{0.5} = 0.40^{+0.11}_{-0.09} \text{ kpc}$ . The total sample of bright black holes

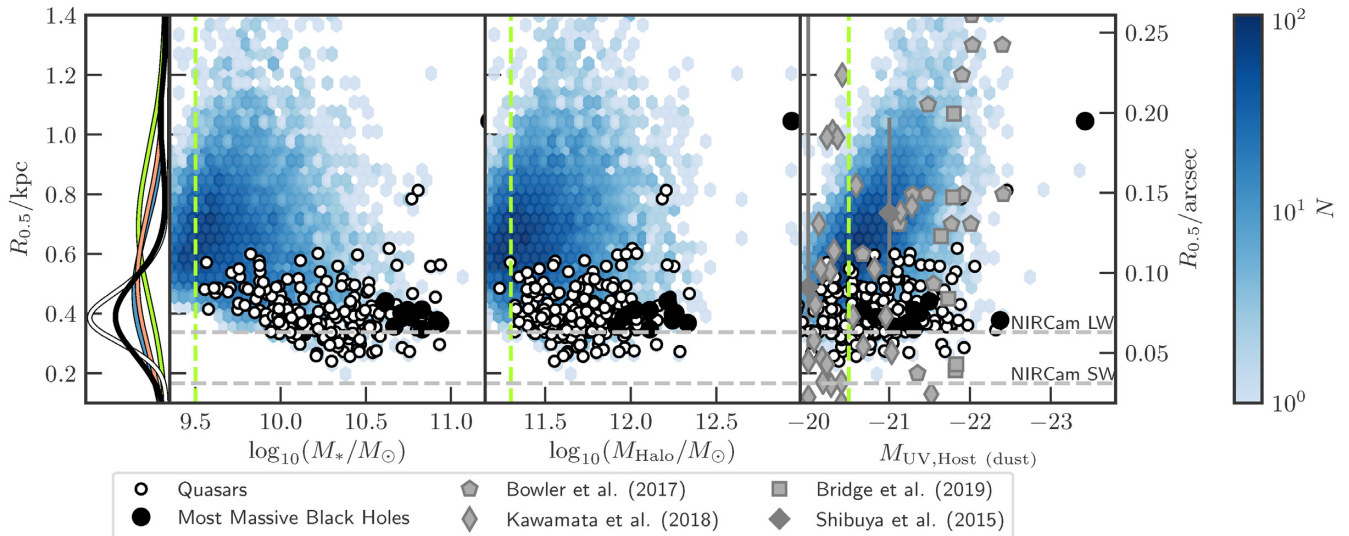
( $L_{\text{AGN}} > 10^{45.1} \text{ erg s}^{-1}$ ) has a wider distribution of galaxy sizes with a larger median value,  $R_{0.5} = 0.51^{+0.27}_{-0.21} \text{ kpc}$ , as do hidden quasars, which have  $R_{0.5} = 0.56^{+0.30}_{-0.23} \text{ kpc}$ . For comparison, galaxies with similar masses and luminosities ( $M_{\text{Halo}} > 10^{11.3} M_{\odot}$ ,  $M_{*} > 10^{9.5} M_{\odot}$ ,  $M_{\text{UV,Host (dust)}} < -20.5$ ) have sizes of  $R_{0.5} = 0.71^{+0.28}_{-0.25} \text{ kpc}$ . A feature of massive black hole and quasar hosts is therefore that they are generally very compact and are some of the smallest galaxies present at  $z = 7$ . This is consistent with the lower redshift conclusions of Bornancini & Lambas (2020), who observe that  $1.4 \leq z \leq 2.5$  AGN and quasar hosts are more compact than star-forming galaxies. Silverman et al. (2019) find that  $1.2 \leq z \leq 1.7$  AGN hosts have intermediate sizes, between those of star-forming, disc-dominated galaxies, and more compact quiescent, bulge-dominated galaxies.

Observations of high-redshift quasar host galaxies at sub-mm wavelengths generally measure larger extents of their gas and dust distributions than our predictions for the sizes of their stellar distributions. For example, studies of [CII] line emission in  $z \gtrsim 6$  quasars using ALMA find radii of 1.7–3.5 kpc (Wang et al. 2013), 2.1–4.0 kpc (Venemans et al. 2016), a median radius of 2.25 kpc (Willott et al. 2017), and for low-luminosity quasars, radii of 2.6–5.2 kpc (Izumi et al. 2018) and 2.1–4.0 kpc (Izumi et al. 2018). In a larger study, Decarli et al. (2018) measured the [CII] emission of a sample of 27  $z > 5.94$  quasars with ALMA, finding radii of 1.2–4.1 kpc. These sizes are much larger than our predictions for the stellar distributions of quasar host galaxies, of  $R_{0.5} = 0.40^{+0.11}_{-0.09} \text{ kpc}$ , and even the general galaxy distribution of  $R_{0.5} = 0.71^{+0.28}_{-0.25} \text{ kpc}$ . Note, however, that FIR continuum and [CII] line observations generally find that [CII] emission is more extended than the dust continuum emission (Wang et al. 2013; Venemans et al. 2016; Willott et al. 2017; Izumi et al. 2018), with, in some cases, the continuum radius smaller than the [CII] radius by even a factor of  $\sim 3$  (Izumi et al. 2018). As the different observations trace different components, this suggests that the gas and dust, and likewise potentially the stars in galaxies, may follow different distributions (see also Khandai et al. 2012; Lupi et al. 2019).

Using the FIRE simulation and a radiative transfer code to study galaxies in  $1 < z < 5$ , Cochrane et al. (2019) find that emission from the stellar component is generally less extended than that of dust continuum emission, cool gas, and dust; an example galaxy is quoted as having  $R_{*} \simeq 0.8 \text{ kpc}$ ,  $R_{\text{cool gas}} \simeq 2.7 \text{ kpc}$ , and  $R_{\text{dust}} \simeq 2.4 \text{ kpc}$ . Observations of main-sequence star-forming galaxies at  $z \simeq 4-6$  with both ALMA and *HST* find that the [CII] radii exceed the rest-frame UV radii by factors of  $\sim 2-3$ , with median sizes  $R_{[\text{CII}]} = 2.1 \pm 0.16 \text{ kpc}$ ,  $R_{\text{F160W}} = 0.91 \pm 0.06 \text{ kpc}$ , and  $R_{\text{F814W}} = 0.66 \pm 0.04 \text{ kpc}$ , where  $R_{\text{F160W}}$  and  $R_{\text{F814W}}$  are radii measured from the *HST* F160W and F814W filters, respectively (Fujimoto et al. 2020). Thus, it seems reasonable that while ALMA observations find extended emission in quasar hosts from dust and cold gas, our predictions expect the stellar emission to be much more compact. We will use BLUE TIDES to make predictions for the gas and dust properties of quasar hosts and compare these to the stellar properties in future work.

#### 4 QUASAR OBSERVATIONS

In Section 3, we considered the properties of the host galaxies of the most massive black holes and intrinsically bright quasars in the BLUE TIDES simulation. We now consider the effects of dust attenuation and survey magnitude limits to mimic true quasar



**Figure 5.** The relation between half-mass radius and stellar mass (left), halo mass (centre), and dust-attenuated galaxy UV magnitude (right). The blue density plots show the distribution for all BLUEFIELDS galaxies, with the most massive black holes and quasars also plotted (see legend). Also shown in the right-hand panel are a range of observations of individual  $z \simeq 6 - 7$  Lyman-break galaxies (Bowler et al. 2016; Kawamata et al. 2018; Bridge et al. 2019) and the size–luminosity relation for Lyman-break galaxies at  $z = 7$  derived by Shibuya, Ouchi & Harikane (2015). Horizontal grey dashed lines show the pixel scales of the JWST NIRCam short-wavelength (SW;  $0.6\text{--}2.3 \mu\text{m}$ ) and long-wavelength (LW;  $2.4\text{--}5.0 \mu\text{m}$ ) detectors of  $0''.031$  and  $0''.063$ , respectively, for reference. The left-most panel shows the distributions of half-mass radius for the most massive black holes (black line), quasars (white line), hidden quasars (salmon line), all black holes brighter than  $L_{\text{AGN}} > 10^{45.1} \text{erg s}^{-1}$  (blue line), and the total sample of galaxies with  $M_* > 10^{9.5} M_\odot$ ,  $M_{\text{Halo}} > 10^{11.3} M_\odot$ , and  $M_{\text{UV,Host (dust)}} < -20.5$  (green line). These limits are shown in the corresponding panels for reference (green dashed lines).

observations to make predictions for upcoming observations with JWST.

## 4.1 Observable quasar sample selection

### 4.1.1 Magnitude limits of quasar observations

To select our observable quasar samples, we first consider the magnitude limits of various observational surveys.

The most well-known sample of high-redshift quasars is that of the SDSS (e.g. Fan et al. 2003, 2006a; Jiang et al. 2016). The faintest quasar in this sample is SDSS J0129–0035, with  $m_{1450} = 22.8$ , or  $M_{1450} = -23.89$  at  $z = 5.78$  (Wang et al. 2013; Bañados et al. 2016). This is of similar luminosity to the brightest quasars in the BLUEFIELDS simulation at  $z = 7$ .

The faintest high-redshift quasars observed to date are those discovered in the Subaru High- $z$  Exploration of Low-Luminosity Quasars (SHELLQs) project (Matsuoka et al. 2018a), which uses imaging from the Subaru Hyper Suprime-Cam and follow-up spectroscopy using the Gran Telescopio Canarias and the Subaru Telescope. This sample includes  $5.7 < z < 6.8$  quasars down to magnitudes of  $m_{1450} = 24.85$ , or  $M_{1450} = -21.93$  (HSC J1423–0018 at  $z = 6.13$ ).

Surveys with upcoming facilities will significantly increase the sample of known high-redshift quasars. The *Euclid* spacecraft, expected to launch in the latter half of 2022, will perform a wide survey of 15 000 square degrees to a magnitude of 24.0 in the  $Y$ ,  $J$ , and  $H$  bands and a deep survey of 40 square degrees to a magnitude of 26.0 (Laureijs et al. 2011). The Vera C. Rubin Observatory will perform the Legacy Survey of Space and Time (LSST), a survey over 18 000 square degrees reaching depths of magnitude 26.1 in the  $z$  band and 24.9 in the  $y$  band, commencing in 2023 (LSST Science Collaboration 2009). These large surveys will discover a large number of high-redshift quasars, complementing the smaller, existing sample of SHELLQs quasars, which found quasars to a

similar depth. The deep *Euclid* survey will discover even fainter quasars, although its much smaller area will result in a smaller quasar sample.

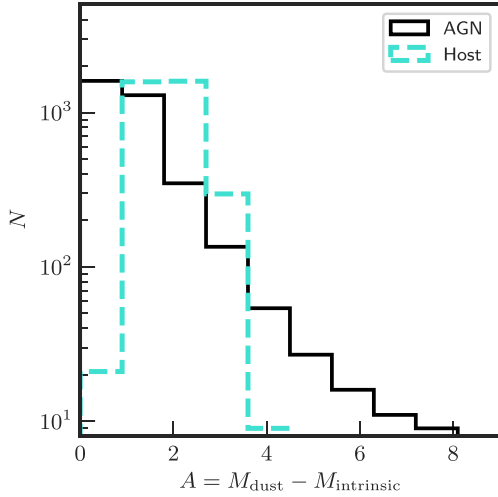
At the forefront of upcoming high-redshift quasar discovery surveys is the RST High Latitude Survey, which will cover 2000 square degrees to a magnitude of 26.9 in the  $Y$ ,  $J$ , and  $H$  bands (Spergel et al. 2015). While RST will not launch until at least 2025, potentially beyond the 5-yr mission plan of JWST, the large volume and significant depth of this survey will result in the largest sample of faint quasars in the foreseeable future. We assume that there will be some bluer comparison data of significant depth, which can be used to select dropouts, so take  $m_{1450} = 26.9$  as the faintest  $z = 7$  AGN luminosity that could be detected by RST.

### 4.1.2 Observable quasars

In Section 3, we selected our quasar sample based on the intrinsic UV-band magnitudes of the black holes and host galaxies (Fig. 2), with ‘quasars’ defined as black holes that had intrinsic magnitudes brighter than their hosts. In Fig. 6, we show the difference between dust-attenuated and intrinsic magnitudes for both AGN and host galaxies,  $A = M_{\text{dust}} - M_{\text{intrinsic}}$ , as calculated following the procedures outlined in Section 2.2.3. Here, and throughout the remainder of this paper, we take the dust extinction for the AGN as that along the line of sight with the minimum  $\tau_{\text{UV,AGN}}$ , as an optimistic estimate of the AGN dust extinction. This generally corresponds to the face-on direction (see Ni et al. 2019).

Fig. 6 shows that AGN and host galaxies experience a similar level of dust attenuation in the majority of cases. The AGN population, however, exhibits a small tail in the distribution extending to large  $A$ . These AGN with extreme dust attenuation are a mixed population of black holes, with a variety of masses and accretion rates. This extinction results in some of the ‘intrinsic quasars’ having dust-attenuated AGN magnitudes that no longer outshine





**Figure 6.** The distributions of dust extinction  $A = M_{\text{dust}} - M_{\text{intrinsic}}$  applied to galaxies and AGN. We take the dust extinction for the AGN as that along the line of sight with the least dust extinction, which is typically the face-on direction, as an optimistic assumption.

their host galaxy. It is therefore important when making mock observational samples to select them based on their dust-attenuated magnitudes.

In Fig. 7, we show the relation between galaxy and AGN dust-attenuated UV magnitudes for the BLUE TIDES galaxies. Since surveys are limited by apparent and not absolute magnitude, we convert our AGN magnitudes using  $m - M = \text{Distance Modulus} - 2.5 \log(1 + z) = 46.99$ . This uses a  $k$ -correction of  $2.5 \log(1 + z)$ , which accounts for the flux per unit wavelength changing with redshift by a factor of  $(1 + z)$  (Oke & Sandage 1968; Hogg et al. 2002). We show these AGN apparent magnitudes in Fig. 7 and overplot the observational selection limits described in Section 4.1.1.

As in Section 3, we make the simple assumption that AGN, which outshine their host galaxy in the UV band, are classified as ‘quasars’. In contrast to Section 3, however, we perform this classification using dust-attenuated magnitudes: ‘quasars’ are black holes with  $M_{\text{UV,AGN(dust)}} < M_{\text{UV,Host(dust)}}$ . Using the limiting magnitudes from SDSS, SHELLQs, and RST, we define our three observable quasar samples as:

- (i) SDSS quasars:  $M_{\text{UV,AGN(dust)}} < M_{\text{UV,Host(dust)}}$  and  $m_{\text{UV,AGN(dust)}} < 22.8$
- (ii) Currently observable quasars:  $M_{\text{UV,AGN(dust)}} < M_{\text{UV,Host(dust)}}$  and  $22.8 < m_{\text{UV,AGN(dust)}} < 24.85$
- (iii) RST quasars:  $M_{\text{UV,AGN(dust)}} < M_{\text{UV,Host(dust)}}$  and  $24.85 < m_{\text{UV,AGN(dust)}} < 26.9$

The ‘SDSS’, ‘currently observable’, and ‘RST’ quasar samples contain 23, 177, and 498 quasars, respectively, which is all black holes in the simulation with  $m_{\text{UV,AGN(dust)}} < 22.8$ , 99 per cent of black holes with  $22.8 < m_{\text{UV,AGN(dust)}} < 24.85$ , and 38 per cent of black holes with  $24.85 < m_{\text{UV,AGN(dust)}} < 26.9$  (see Fig. 2). Defining the observable quasar samples using the dust-attenuated magnitudes therefore selects a different, larger sample of black holes.

We define the black hole with the median  $M_{\text{BH}}$  in the most massive black hole sample as the ‘median massive black hole’ and the quasars with the median  $M_{\text{UV,AGN(dust)}}$  from each of these three quasar samples as the ‘median quasars’. We show the spectra of these objects in Fig. 8 as examples.

## 4.2 Images of quasar hosts

To visualize the host galaxies of the most massive black holes and ‘SDSS’, ‘currently observable’, and ‘RST’ quasars, we first make images of their mass distributions using GAEPSI2,<sup>3</sup> a suite of routines for visualizing SPH simulations. We select four galaxies from each sample with a representative range of morphologies ( $B/T \simeq 0.65, 0.75, 0.85$  and  $0.95$ ) to image. We note that the minimum  $B/T$  for the most massive black hole sample is  $\simeq 0.70$  and so we select only three galaxies from that sample. The mass distributions of these sample galaxies from a face-on and edge-on perspective are shown in Figs 9 and 10, respectively, with colours depicting stellar age. These images show a variety of sizes, shapes, and ages of the black hole and quasar host galaxies.

We also construct a matched sample for comparison with the three representative most massive black hole hosts. These matched galaxies are chosen to have small black holes,  $10^{6.5} < M_{\text{BH}}/M_{\odot} < 10^7$ , but the most similar stellar mass and  $B/T$  to those of the three most massive black hole hosts. Images of these representative most massive black holes and the matched galaxy sample can be seen in Fig. 11. The galaxies in the matched sample are more diffuse and slightly more extended than the hosts of the most massive black holes. This is consistent with our findings from Fig. 5, which shows that the hosts of massive black holes are more likely to be compact than other galaxies of equivalent stellar mass.

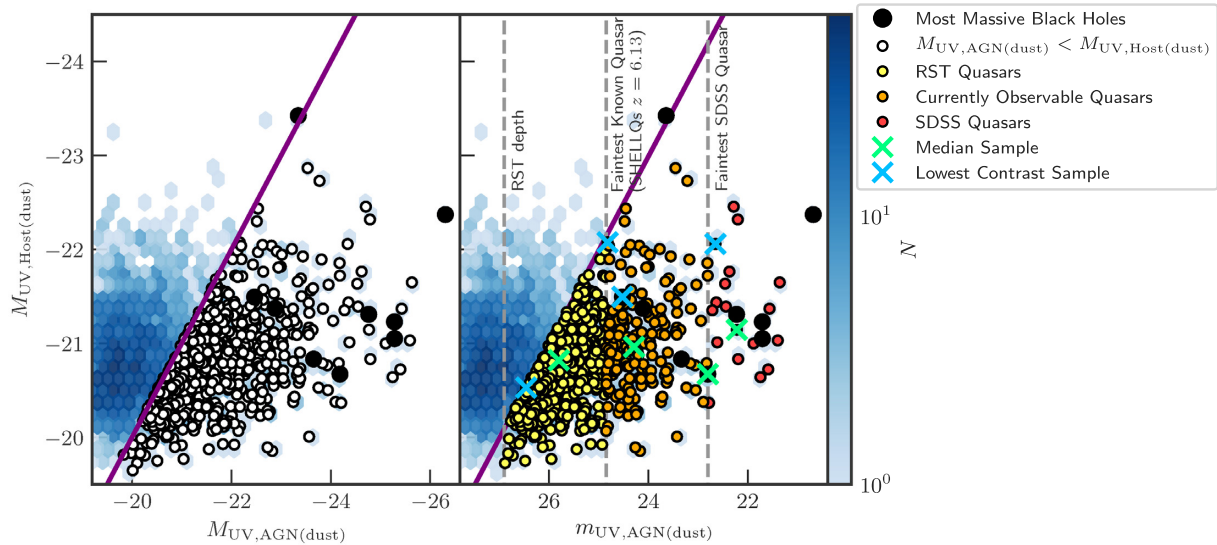
We produce mock JWST images using SYNTHOBS,<sup>4</sup> a package for producing synthetic observations from SPH simulations. SYNTHOBS takes the flux of each stellar particle, applies the specified photometric filter, and convolves this emission with the corresponding JWST point-spread function (PSF). We assume that the quasar emission comes from a single point, with the quasar thus appearing as a point source convolved with the PSF in the images. By applying the appropriate smoothing, SYNTHOBS produces a mock image with the pixel scale of the instrument. We include the effects of noise by adding a random background noise map to the SYNTHOBS images, with noise  $\sigma$  from the predicted  $10\sigma$  sensitivity of JWST, using a circular photometric aperture 2.5 pixels in radius (STSci 2017). Dust attenuation is applied following the methodology described in Section 2.2.3, and we use the minimum AGN dust attenuation of the various sightlines.

In Fig. 12, we show mock JWST imaging in the NIRCcam F200W filter of the galaxies hosting the median massive black hole and the median quasars (i.e. those whose spectra are shown in Fig. 8). These images show the combined (dust attenuated) quasar and host emission with varying exposure times: 1 ks, 5 ks, 10 ks, and an image with no noise background for comparison. This shows that deep exposure times of  $\gtrsim 5$  ks are required to observe the detailed structure present in the noise-less images, which is likely to be necessary for detecting the underlying host galaxy emission, which generally has low surface brightness and is hidden by the bright quasar emission (see discussion below). We therefore choose to adopt an exposure time of 10 ks for our mock images herein. We choose to show the F200W filter as an example, as it has the highest sensitivity of the NIRCcam wide-band filters, resulting in the least background noise for a given exposure time.

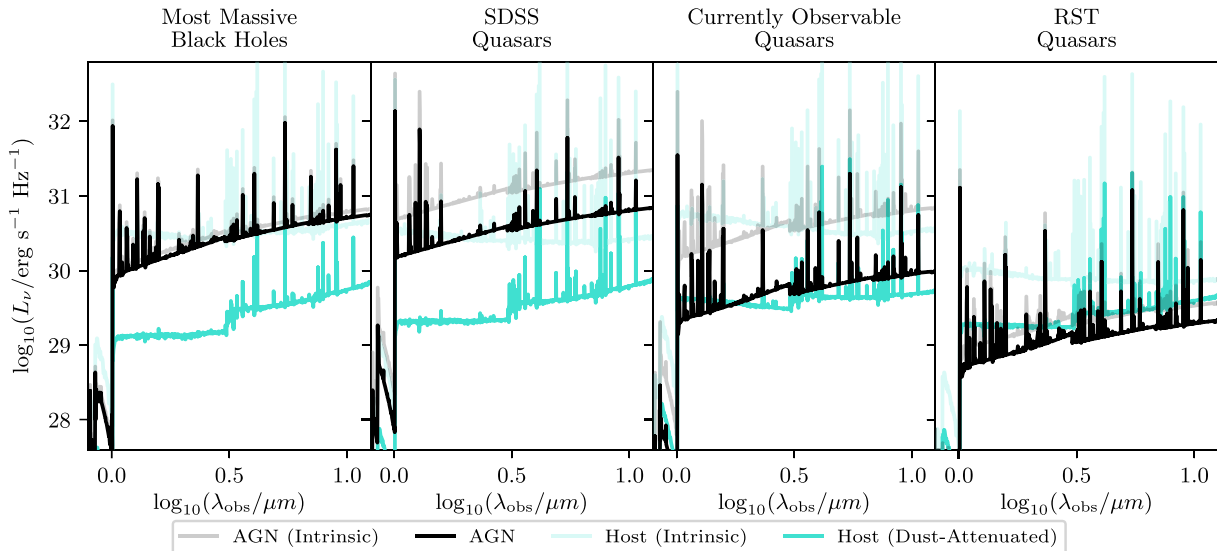
In Fig. 13, we show mock JWST imaging in the NIRCcam F200W filter of the hosts of the median massive black hole and the median quasars, with and without dust attenuation, and with and without

<sup>3</sup><https://github.com/rainwoodman/gaepsi2>

<sup>4</sup><https://github.com/stephenwilkins/SynthObs>



**Figure 7.** The distribution of galaxy and AGN dust-attenuated UV magnitudes for BLUE TIDES galaxies at  $z = 7$  (blue density plot). The left-hand panel shows the absolute host and AGN magnitudes, and the right-hand panel shows the apparent AGN magnitude for comparison with observational detection limits. Using the left-hand panel, we classify quasars as those with  $M_{UV,AGN(dust)} < M_{UV,Host(dust)}$ , since the AGN outshines the host galaxy. Using the observational detection limits shown in the right-hand panel (grey dashed lines), we further split the quasars into three samples: SDSS quasars, currently observable quasars, and RST quasars. We mark with green crosses the median mass black hole in the most massive black hole sample, and the black hole with the median  $M_{UV,AGN(dust)}$  in the SDSS, currently observable, and RST quasar samples (‘median sample’). We also mark the black hole in each sample, which has the lowest contrast between the quasar and the host galaxy brightness (‘lowest contrast sample’, blue crosses). The most massive black holes are also shown.

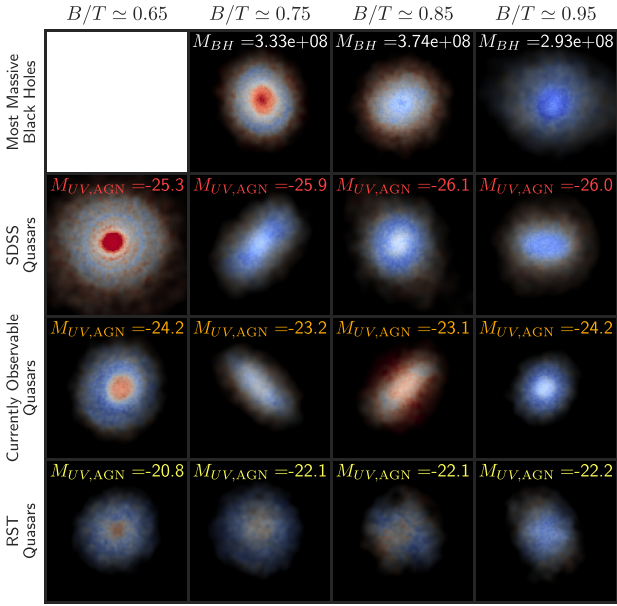


**Figure 8.** The spectra of an AGN and its host galaxy for the median mass black hole in the most massive black hole sample (‘median massive black hole’, far left), and the quasar with the median  $M_{UV,AGN(dust)}$  in the SDSS (middle left), currently observable (middle right), and RST (far right) quasar samples (‘median quasars’). The upper grey curves show the intrinsic AGN spectra, while the lower black curves show the AGN spectra with our dust extinction law applied. The upper turquoise curves show the intrinsic host galaxy spectra, while the lower turquoise curves show the host galaxy spectra with our dust extinction law applied.

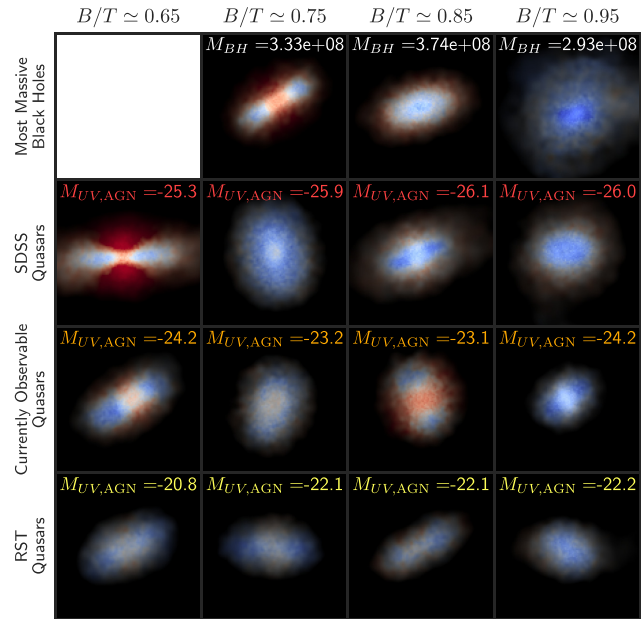
the quasar emission. With a resolution of 0.031 arcsec, JWST only partially resolves the host galaxies, with diameters of  $\sim 0.8$  kpc or  $\sim 0.15$  arcsec at  $z = 7$ . Their emission is centrally concentrated, and so the hosts appear as a smeared PSF at this resolution. However, as the density of dust is highest in the central regions, the dust-attenuated images show more interesting, asymmetrical features.

The limited resolution of these small galaxies makes it difficult to distinguish the host galaxy once the point-source quasar emission is included in the images. For the intrinsic images, the image is broader

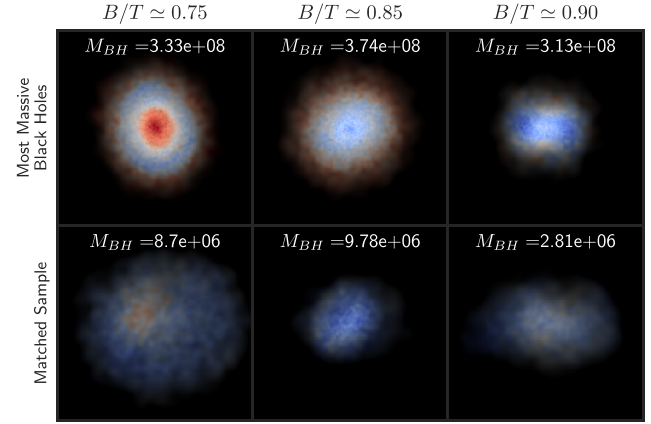
than the quasar image (i.e. the PSF of the telescope), suggesting that an accurate modelling technique should be able to detect the host emission despite the presence of the quasar. However, including the effect of dust attenuation makes the host more difficult to distinguish from the quasar, as its emission becomes fainter and less extended, particularly for the median massive black hole, SDSS, and currently observable quasars. For these three systems, the brightness contrast between the host and the quasar is  $\sim 1.5$  orders of magnitude at the centre, decreasing with distance from the quasar, with the two having



**Figure 9.** The stellar mass distribution of four BLUETIDES galaxies from each of the samples: the 10 most massive black holes, SDSS quasars, currently observable quasars, and RST quasars, selected to show a range of morphologies ( $B/T \simeq 0.65, 0.75, 0.85$  and  $0.95$ ; left to right shows more disc-like to more bulge-dominated galaxies). The minimum  $B/T$  for the most massive black hole sample is  $\sim 0.70$  and so only three galaxies from that sample are shown. Each galaxy is viewed face-on, with a field of view of  $3 \times 3$  kpc. The colour depicts the age of the stellar population, from bluest ( $\leq 20$  Myr) to reddest ( $\geq 220$  Myr), with a linear scale.



**Figure 10.** The stellar mass distribution of four BLUETIDES galaxies from each of the samples: the 10 most massive black holes, SDSS quasars, currently observable quasars, and RST quasars, selected to show a range of morphologies ( $B/T \simeq 0.65, 0.75, 0.85$  and  $0.95$ ; left to right shows more disc-like to more bulge-dominated galaxies). The minimum  $B/T$  for the most massive black hole sample is  $\sim 0.70$  and so only three galaxies from that sample are shown. Each galaxy is viewed edge-on, with a field of view of  $3 \times 3$  kpc. The colour depicts the age of the stellar population, from bluest ( $\leq 20$  Myr) to reddest ( $\geq 220$  Myr), with a linear scale.



**Figure 11.** The stellar mass distribution of three BLUETIDES galaxies from the most massive black hole sample, alongside a matched sample of galaxies with a similar stellar mass and bulge-to-total ratio, but low black hole masses. Each galaxy is viewed face-on, with a field of view of  $3 \times 3$  kpc. The colour depicts the age of the stellar population, from bluest ( $\leq 20$  Myr) to reddest ( $\geq 220$  Myr), with a linear scale.

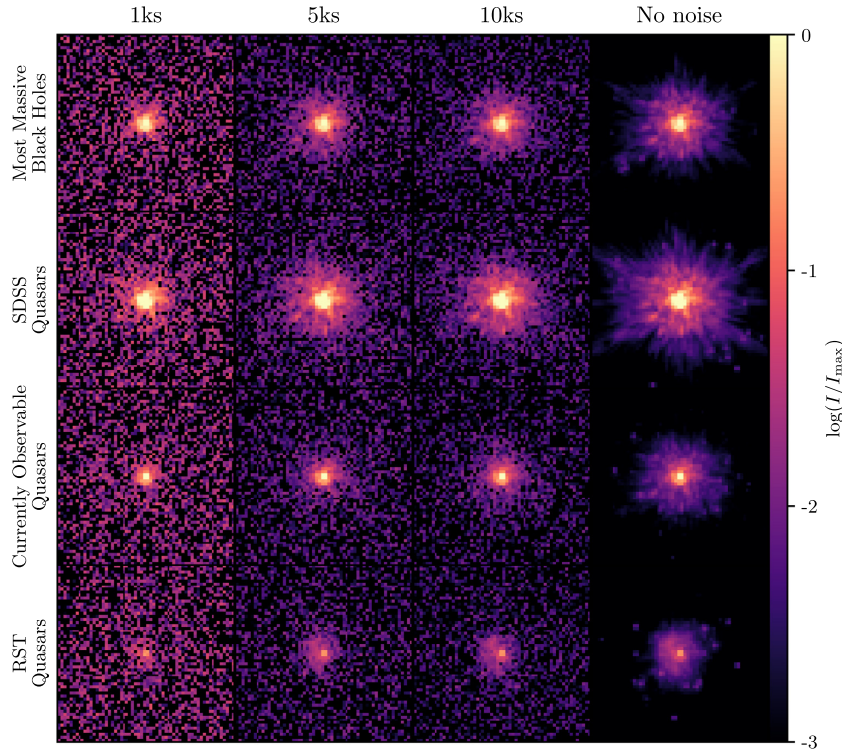
similar brightness towards the edge of the host galaxy at  $\sim 0.5''$ . The median RST quasar has a lower contrast between the quasar and the host, resulting in the host being more easily visible around the bright, central emission from the quasar. Distinguishing the host galaxy from the quasar emission will therefore still be challenging with JWST, even with its improved resolution over *HST*.

While it appears that the host galaxies are more easily detected in the fainter quasar samples, this is an effect of the contrast ratio between the host and the quasar and not the quasar's total brightness. For the median massive black hole and the median SDSS, currently observable, and RST quasars, the difference in AGN and host magnitude is 3.51, 3.62, 1.74, and 0.37, respectively (see Fig. 7). Thus, for the median massive black hole and SDSS quasar, as the AGN magnitude is much brighter than the host, it completely obscures any host emission. For the median currently observable quasar, the bright point source still dominates; however, the total emission is somewhat broader than the quasar PSF, which may allow for a host detection with an accurate modelling technique. For the median RST quasar, the lower contrast ratio results in an easily distinguishable host galaxy.

To investigate this effect further, we consider the black hole in each sample, which has the lowest contrast ratio between the AGN and the host luminosity,  $M_{UV,Host(dust)} - M_{UV,AGN(dust)}$  (with  $M_{UV,AGN(dust)} < M_{UV,Host(dust)}$ ). Mock JWST images of these systems are shown in Fig. 14. The host galaxies of these black holes are much easier to distinguish from the quasar point source emission, particularly for those that have the lowest contrast ratios and for hosts that have more extended emission. Thus, while the host galaxies of quasars will still be difficult to detect with JWST in general, it should be possible to detect the hosts of quasars with low contrast ratios.

We show mock JWST images of the median currently observable quasar in all of the NIRCcam wide-band filters red-ward of the Lyman break in Fig. 15. Most filters show a combined image that is slightly broader than the quasar PSF; however, no filter makes the host clearly more detectable. As the wavelength increases, the resolution of the telescope decreases. Thus, while the contrast ratio of the quasar and its host should be lower at larger wavelengths due to the spectral shapes of quasars and host galaxies (see Fig. 8), redder NIRCcam filters do not particularly make the host more easily distinguishable.





**Figure 12.** Simulated face-on images of BLUEtIDES galaxies in the JWST NIRCcam F200W filter, showing the median massive black hole and the median SDSS, currently observable, and RST quasars (i.e. the galaxies whose spectra are shown in Fig. 8). The combined quasar and host galaxy emission including dust attenuation is shown. In the first three panels from left to right, the images assume an exposure time of 1 ks, 5 ks, and 10 ks, which are predicted to achieve a  $10\sigma$  detection of 63.4, 18.2, and 13.2 nJy point sources, using a circular photometric aperture 2.5 pixels in radius (STSci 2017). The right-most panel shows the images with no noise background. The field of view is  $10 \times 10$  kpc, or  $1''.86 \times 1''.86$ . Note that all panels are shown with the same intensity scale.

The instrument sensitivity increases from the F090W to F200W filters and then decreases for the higher wavelength filters (STSci 2017). The highest sensitivity F200W filter results in the clearest image of the quasar system for a given exposure time (here 10 ks) and thus may offer the best results for detecting quasar host galaxies with JWST.

These conclusions are based only on examining the resulting images by eye. In future work, we will make more detailed and robust predictions for the detectability of quasar hosts with JWST by running an observational technique used to detect quasar host galaxies on these simulated images.

## 5 BIASES IN THE OBSERVED SCALING RELATIONS

We now consider the  $z = 7$  black hole–stellar mass and black hole–bulge mass relations predicted by BLUEtIDES, shown in Fig. 16. The best-fitting relations for black holes with  $M_{\text{BH}} > 10^{6.5} M_{\odot}$  and galaxies with  $M_* > 10^{9.7} M_{\odot}$  are

$$\log(M_{\text{BH}}/M_{\odot}) = (1.30 \pm 0.02) \log(M_*/M_{\odot}) - (5.9 \pm 0.2), \quad (5)$$

and

$$\log(M_{\text{BH}}/M_{\odot}) = (1.50 \pm 0.02) \log(M_{\text{Bulge}}/M_{\odot}) - (7.7 \pm 0.2), \quad (6)$$

with errors calculated from 10 000 bootstrap realizations. The standard deviation of the residuals, or scatter, is 0.2 dex for both relations, so the simulation shows no preference for a tighter correlation of black hole mass with either total or bulge stellar mass. We note that these relations are unlikely to be sensitive to the black hole

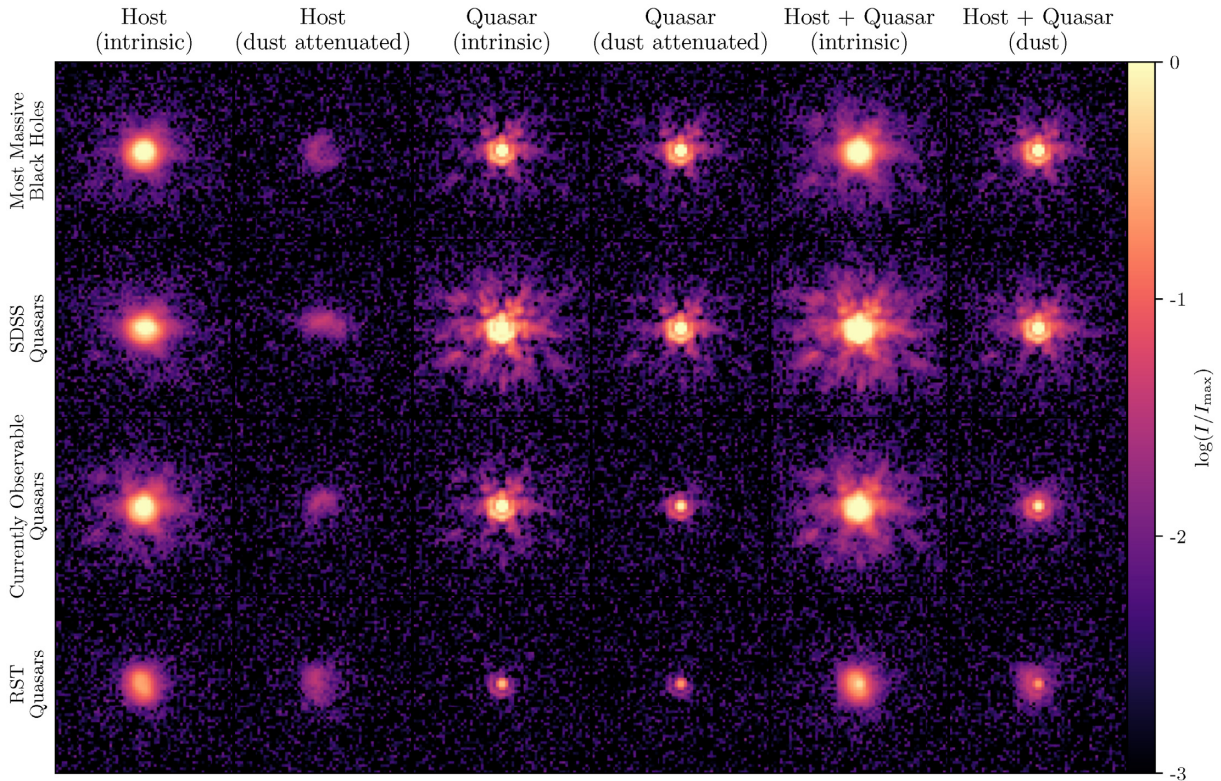
seeding prescription, as we consider only black holes that have grown significantly above the seed mass of  $10^{5.8} M_{\odot}$ .

The BLUEtIDES black hole–bulge mass relation at  $z = 7$  is steeper than the observed local relation (Kormendy & Ho 2013), which is also shown in Fig. 16. However, the simulations and local observations are reasonably consistent, particularly at the highest masses where the observed relation is best measured. Fig. 16 also shows a range of observations of  $5 \lesssim z \lesssim 7$  quasars, assuming their stellar mass is equal to their measured dynamical mass (Willott et al. 2017; Izumi et al. 2018; Pensabene et al. 2020). Quasars observed with  $M_{\text{BH}} > 10^{8.5} M_{\odot}$  show a wide range of dynamical masses and generally lie above the local relation. These observed black holes are larger than those present in the BLUEtIDES simulation at  $z = 7$ , so a comparison cannot be made. Observations of lower luminosity quasars with  $M_{\text{BH}} < 10^{8.5} M_{\odot}$  are consistent with the BLUEtIDES relation.

In Fig. 17, we plot the black hole–stellar mass relation for the most massive black holes, SDSS quasars, currently observable quasars, and RST quasars. This shows that the hosts of the most massive black holes and quasars have large black hole masses for their stellar mass. To investigate the effect of this bias on the observed black hole–stellar mass relation, we make fits to the SDSS, currently observable, and RST quasar samples, constraining the slope to be equal to that of the total sample:

$$\log(M_{\text{BH}}/M_{\odot}) = 1.32 \log(M_*/M_{\odot}) + b, \quad (7)$$

where for the full sample,  $b = -6.06$  (equation 5). The SDSS, currently observable, and RST quasar samples have a normalization of  $b = -5.85$ ,  $-5.87$ , and  $-5.90$ , respectively,  $\sim 0.2$  dex higher



**Figure 13.** Simulated face-on images of BLUE TIDES galaxies in the JWST NIRCам F200W filter, showing the median massive black hole and the median SDSS, currently observable, and RST quasars (i.e. the galaxies whose spectra are shown in Fig. 8). The host galaxy emission is shown with and without dust attenuation in the two left-most panels. The emission from the quasar is shown with and without dust attenuation in the middle panels, with the combined quasar and host galaxy image shown with and without dust attenuation (applied to both the host and the quasar) in the right-most panels. These images assume an exposure time of 10 ks, which is predicted to achieve a  $10\sigma$  detection of 13.2 nJy (AB mag  $\sim 28.8$ ) point sources, using a circular photometric aperture 2.5 pixels in radius (STScI 2017). The field of view is  $10 \times 10$  kpc, or  $1''.86 \times 1''.86$ . Note that all panels are shown with the same intensity scale.

than that of the full galaxy sample. Quasar samples are therefore biased samples of the intrinsic black hole–stellar mass relations, consistent with expectations from observations (e.g. Lauer et al. 2007; Salviander et al. 2007; Schulze & Wisotzki 2014; Willott et al. 2017). Our simulation provides a calibration of this systematic effect. A similar bias to larger black hole masses is theoretically expected to occur when observing the black hole–velocity dispersion relation (see e.g. Volonteri & Stark 2011).

Interestingly, we find that the fainter quasar samples are as biased as the bright quasar samples in measuring the intrinsic black hole–stellar mass relation. To investigate this further, we remove the ‘quasar’ constraint that  $M_{UV,AGN(dust)} < M_{UV,Host(dust)}$  and instead consider all AGN brighter than each survey limit. We make fits to these SDSS, currently observable, and RST AGN samples, again constraining the slope to be equal to that of the total sample (equation 7). The SDSS, currently observable, and RST AGN samples have a normalization of  $b = -5.85$ ,  $-5.87$  and  $-5.97$ , respectively, relative to the total sample with  $b = -6.06$ . Here, the bias decreases with survey depth. This is qualitatively consistent with the expectations of Lauer et al. (2007), which found that for a survey limited only by  $L_{AGN}$  and not the host properties, the bias will decrease, but is not removed entirely, as the survey goes to fainter luminosity limits.

The constraint that  $M_{UV,AGN(dust)} < M_{UV,Host(dust)}$  for quasars therefore results in the RST quasar sample measuring a larger bias in the black hole–host mass relation than for all RST AGN. We consider black hole samples with  $M_{UV,AGN(dust)} - M_{UV,Host(dust)} < -2, 0$ , and  $2$ , with no magnitude limit, and find the best fits to equation (7).

These fits have normalizations of  $b = -5.79$ ,  $-5.89$ , and  $-5.99$  respectively; the brighter the AGN relative to the host, the larger the black hole mass at fixed stellar mass. Thus, to reduce the bias in the measured black hole–stellar mass relation, observations should target AGN that are faint relative to their host galaxies. There are more of these objects in fainter AGN samples; hence, the sample of all RST AGN has a reduced bias compared with the brighter SDSS and currently observable AGN samples.

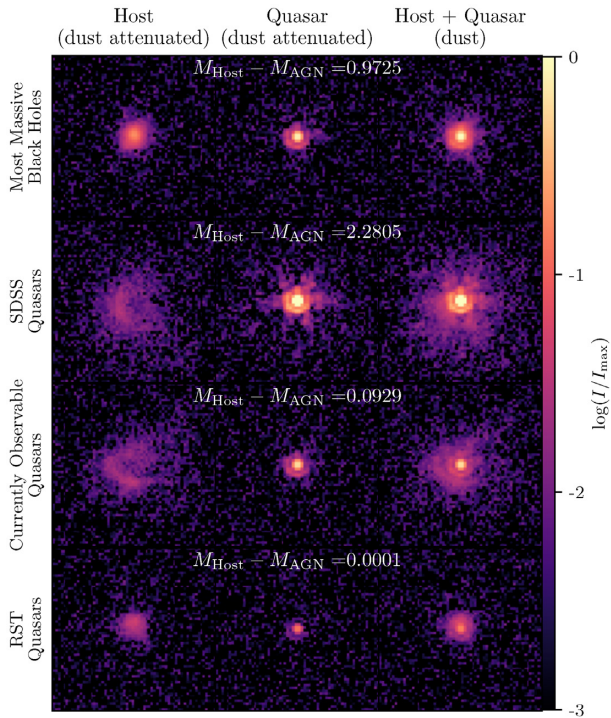
## 6 THE ENVIRONMENTS OF HIGH-REDSHIFT QUASARS

We now study the environments of high-redshift quasars in the BLUE TIDES simulation by investigating neighbouring galaxies that host a black hole. The requirement for a companion to host a black hole is due to no halo sub-finding algorithm being implemented on the simulation; it cannot identify multiple galaxies within an individual dark matter halo, which are precisely the systems we are interested in. As an approximation, we identify neighbouring galaxies via their black holes and assign all particles within  $R_{0.5}$  of the black hole to that galaxy.

### 6.1 The number of nearby galaxies

We first consider the number of nearby galaxies around each black hole ( $M_{BH} > 10^{5.8} M_{\odot}$ ). We compare the most massive black hole





**Figure 14.** Simulated face-on images of BLUETIDES galaxies in the JWST NIRCам F200W filter, showing the galaxy from each black hole sample that has the lowest contrast ratio between the quasar and the host  $M_{UV,Host(dust)} - M_{UV,AGN(dust)}$  (with  $M_{UV,AGN(dust)} < M_{UV,Host(dust)}$ ). The host galaxy emission is shown in the left-most panels, with the emission from the quasar shown in the middle panels. The combined quasar and host galaxy image is shown in the right-most panels. All images include the effect of dust attenuation. These images assume an exposure time of 10 ks, which is predicted to achieve a  $10\sigma$  detection of 13.2 nJy (AB mag  $\sim 28.8$ ) point sources, using a circular photometric aperture 2.5 pixels in radius (STSci 2017). The field of view is  $10 \times 10$  kpc, or  $1''.86 \times 1''.86$ . Note that all panels are shown with the same intensity scale.

and quasar samples to the overall sample of black holes with  $M_{BH} > 10^{6.5} M_{\odot}$ .

In Fig. 18, we show the average number of neighbours within a given distance that would be observed at various magnitude limits for black holes in each sample. To a magnitude limit of the faintest SDSS quasar,  $m_{UV} = 22.8$ , no companions are detected within 340 kpc, on average, around any black hole. At a deeper magnitude limit of  $m_{UV} = 24.85$ , the magnitude of the faintest known high-redshift quasar, no black hole samples are predicted to have nearby companions within  $\sim 80$  kpc, on average. The average number of companions increases slightly at larger distances for the massive black hole and quasar samples. At distances  $\gtrsim 150$  kpc, the most massive black holes have an average of  $\sim 1$  companion with  $m_{UV} < 24.85$ , more than expected for the overall sample. However, this enhancement is not significant, given the uncertainties.

Many more companions would be observable with RST, with a survey depth of  $m_{UV} = 26.9$ . The average number of  $m_{UV} < 26.9$  companions for each sample increases with distance from the black hole, with the most massive black holes having an average of  $\sim 2$  companions within  $\sim 100$  kpc and  $\sim 4$  companions within  $\sim 300$  kpc. This sample shows the largest number of companions, with the SDSS quasars, currently observable quasars, RST quasars, and all black holes having progressively less companion galaxies, with all black holes having an average of  $\sim 0.5$  companions within  $\sim 100$

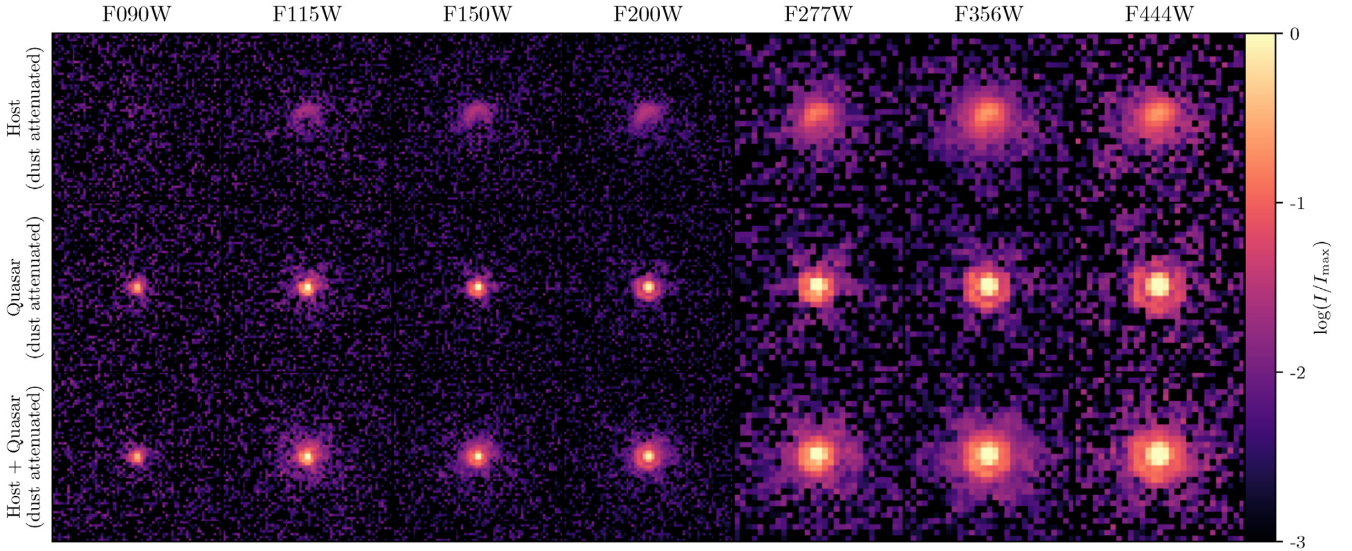
kpc and  $\sim 1.5$  companions within  $\sim 300$  kpc. A similar enhancement is seen when considering all companion galaxies, with no magnitude limit. On average, the most massive black holes have the most companion galaxies within 50–340 kpc, followed by the quasar samples from brightest to faintest, with the enhancement above the overall black hole sample largest at larger distances ( $> 150$  kpc). As we predict more neighbouring galaxies at larger separations ( $> 50$  kpc), the majority of companion galaxies are too distant to be detected in the small field of view of ALMA.

The most massive black holes are more likely to be found in denser environments than the typical  $M_{BH} > 10^{6.5} M_{\odot}$  black hole, with quasars showing a weaker enhancement. However, the increased average number of companions found around the most massive black holes and quasars relative to the general sample is statistically insignificant, with a large variation seen in the number of galaxies around each black hole. Our conclusions are consistent with the more comprehensive analysis of Habouzit et al. (2019), who investigated black hole environments in the Horizon-AGN simulation at  $z \simeq 4-6$ . Habouzit et al. (2019) found that, on average, massive black holes live in regions with more nearby galaxies, with an excess of up to 10 galaxies within 1 cMpc at  $z \simeq 4-5$ . The enhancement is larger for more massive black holes. Habouzit et al. (2019) found a diversity in number counts, with some massive black holes having similar numbers of nearby neighbours to the average number counts, consistent with our expectations.

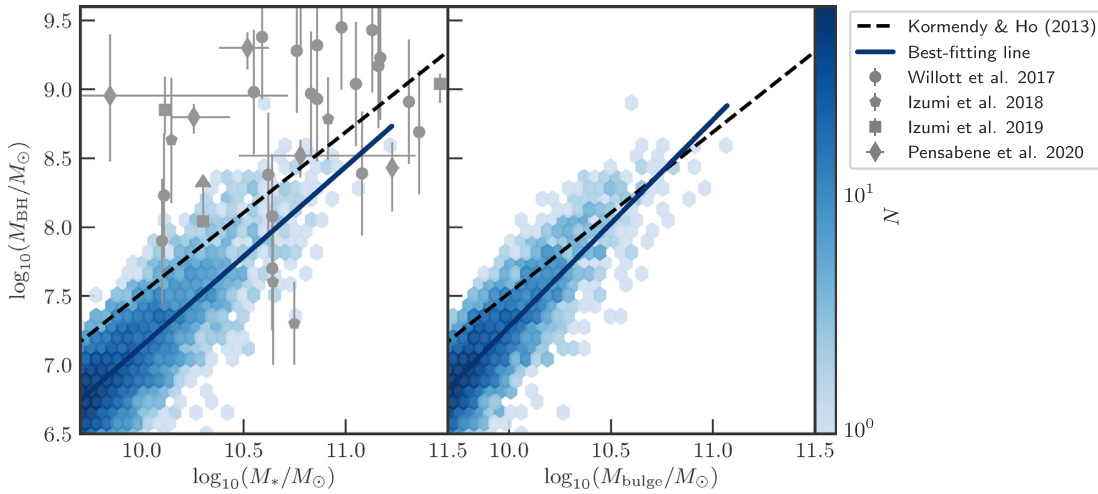
Companion galaxies have been observed near high-redshift quasars, particularly in sub-mm observations (e.g. Wagg et al. 2012; McGreer et al. 2014; Decarli et al. 2017; Willott et al. 2017; Neeleman et al. 2019). In the rest-frame UV/optical, McGreer et al. (2014) detected companion galaxies around two quasars, although they found that bright companion galaxies within 20 kpc are uncommon, with an incidence of  $\lesssim 2/29$  for  $\gtrsim 5L^*$  galaxies and  $\lesssim 1/6$  for  $2 \lesssim L \lesssim 5L^*$  galaxies. In the sub-mm, Trakhtenbrot et al. (2017a) observed six  $z \simeq 4.8$  quasars with ALMA and found nearby companions around three of the quasars at distances of 14–45 kpc. Given that these companions are not detected in the infrared with Spitzer, Trakhtenbrot et al. (2017a) conclude that there must be significant dust attenuation in these galaxies. Continuing this study, Nguyen et al. (2020) observed an additional 12 quasars, finding nearby companions around five of the 18 quasars. Decarli et al. (2017) detected a companion galaxy around four of 25  $z \gtrsim 6$  quasars with ALMA, and similarly Willott et al. (2017) found one quasar companion in a sample of five  $z \gtrsim 6$  quasars. Mazzucchelli et al. (2019) took follow-up observations of these companions in the optical/IR, detecting the emission from only one of the companions, finding that the remaining three must be ‘highly dust-enshrouded’. Willott et al. (2005) also hypothesize that the lack of companions observed in rest-frame UV observations, relative to sub-mm observations, is a result of dust attenuation.

In Fig. 18, we consider the effect of dust attenuation on the observed number of companion galaxies by showing the number of companions that are intrinsically brighter than each magnitude limit and the fraction of these companions that have dust-attenuated magnitudes fainter than the limit and thus would be ‘missed’ by observations in the rest-frame UV. We find that at a depth of  $m_{UV} = 22.8$ , almost 100 per cent of companions with intrinsic magnitudes of  $m_{UV} < 22.8$  are missed due to dust attenuation (with  $m_{UV,Dust} > 22.8$ ). However, the overall number of these intrinsic companions is low (an average of  $< 1$  within 300 kpc). At a magnitude limit of  $m_{UV} = 24.85$ , around 50 per cent of the companions of the most massive black holes are missed, while for the quasar samples, this is around 75 per cent. RST, at a depth of  $m_{UV} = 26.9$ ,





**Figure 15.** Simulated face-on images of the median currently observable quasar in the JWST NIRCcam wide-band filters red-ward of the  $z = 7$  Lyman break. The host galaxy emission is shown in the top panels, the emission from the quasar in the middle panels, with the combined quasar and host galaxy image shown in the bottom panels. All images include dust extinction of both the quasar and the host galaxy. These images assume an exposure time of 10 ks, with  $10\sigma$  detection sensitivities as predicted by STSci (2017). The field of view is  $12 \times 12$  kpc, or  $2''.23 \times 2''.23$ . Note that all panels are shown with the same intensity scale.



**Figure 16.** The relation between black hole mass and stellar mass (*left*) and black hole mass and bulge mass (*right*) for BLUETIDES galaxies at  $z = 7$  and their best-fitting relations as given in equations (5) and (6). We plot a range of observations of  $5 \lesssim z \lesssim 7$  quasars from the literature (Willott et al. 2017; Izumi et al. 2018; Pensabene et al. 2020), assuming that their stellar mass is equal to their measured dynamical mass. We also plot the observed black hole–bulge mass relation at  $z = 0$  (Kormendy & Ho 2013). This relation is also shown in the left (stellar mass) panel for comparison, assuming that the hosts are pure elliptical galaxies with  $M_* = M_{\text{bulge}}$ .

will be able to detect the majority of companions, with less than 10 per cent of intrinsic companions missed due to dust attenuation.

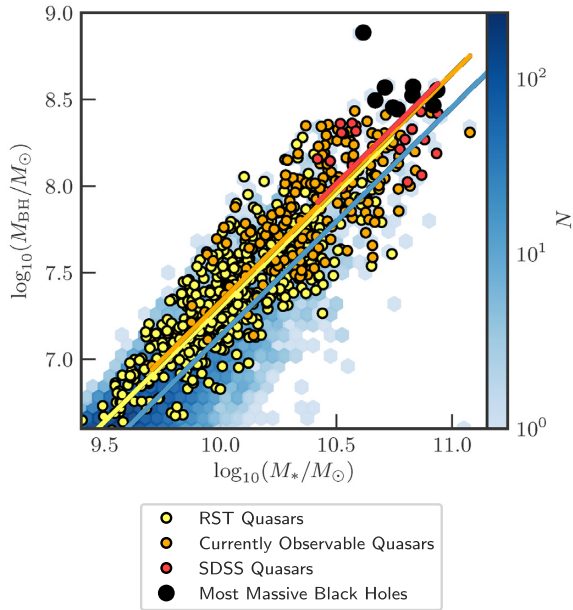
Overall, our predictions expect that a large fraction (up to 75 per cent at  $m_{\text{UV}} < 24.85$ ) of quasar companions will be ‘missed’ in current rest-frame UV observations due to dust obscuration. These dusty galaxies are likely to be observable in the sub-mm, and so our predictions are consistent with expectations (e.g Willott et al. 2005).

## 6.2 Properties of nearby neighbours

We now restrict our investigation to the nearest neighbour to each black hole, with distances less than 200 kpc.

We find that 90 per cent of the most massive black holes have their nearest neighbour within 200 kpc, compared with 87 per cent of SDSS quasars, 80 per cent of currently observable quasars, and 67 per cent of RST quasars. For comparison, 63 per cent of all black holes with  $M_{\text{BH}} > 10^{6.5} M_{\odot}$  have their nearest neighbour within 200 kpc.

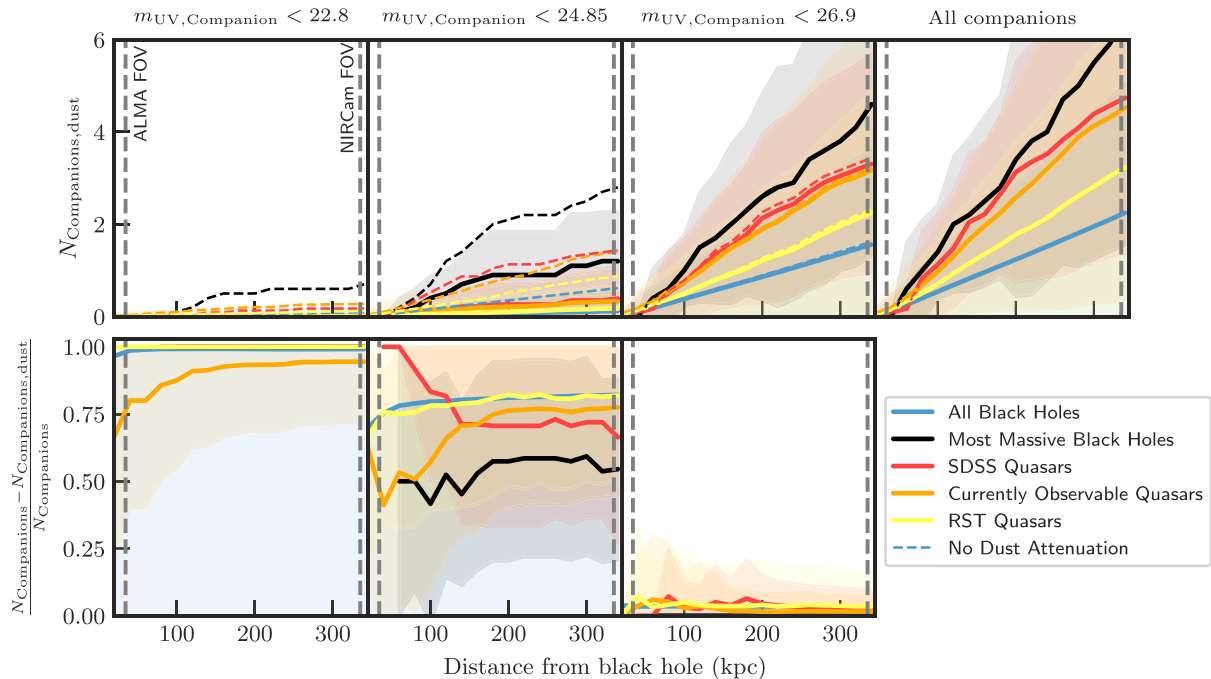
Fig. 19 shows various properties of the nearest neighbours: their distance, UV magnitude (both with and without dust attenuation), stellar mass and black hole mass, and the differences between the properties of the neighbour and those of the black hole host. Most of the nearest neighbours lie within 100 kpc or 20 arcsec of the black hole host galaxy. The vast majority of these neighbours are



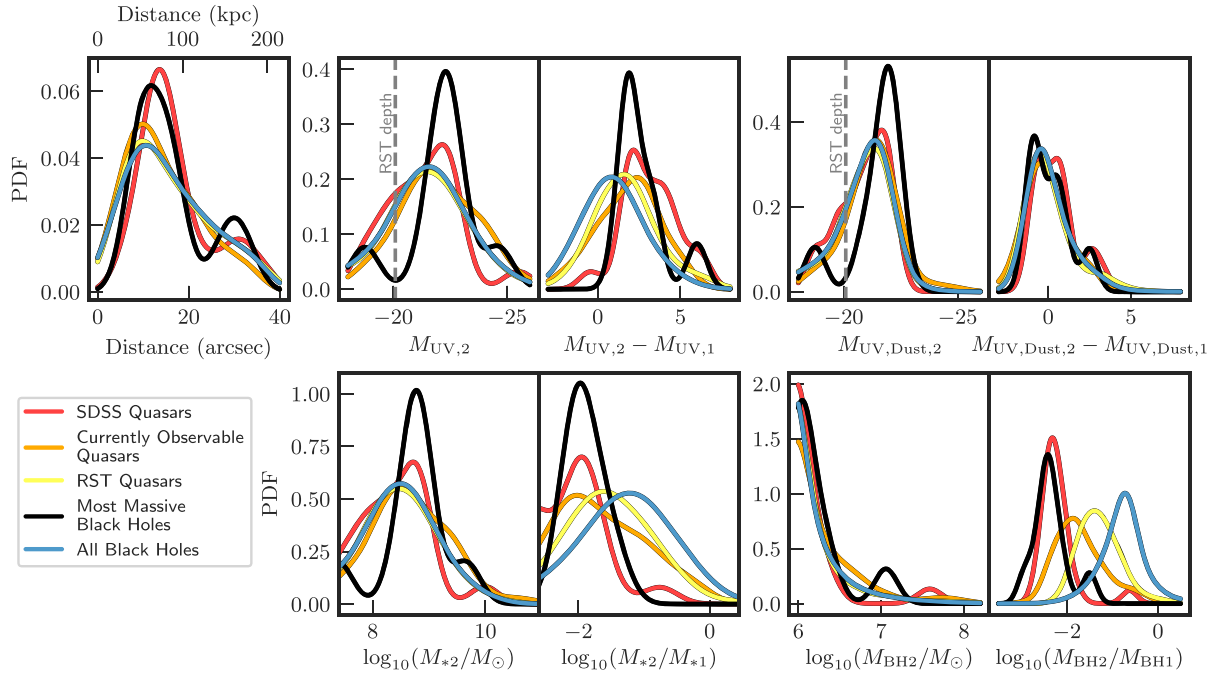
**Figure 17.** The relation between black hole mass and stellar mass. The blue density plot shows the distribution for all BLUETIDES galaxies, while the circles show the most massive black holes and SDSS, currently observable, and RST quasar samples (see legend). The solid lines are fits to the total (blue), SDSS (red), currently observable (orange), and RST quasar (yellow) samples, constraining the slope to be the same as that for the total sample (equation 7).

brighter than  $M_{UV} = -20$  and so should be readily detectable by RST. Some companions are fainter than the black hole host by up to five magnitudes, although most are of similar brightness. The stellar mass and black hole mass distributions of the neighbouring galaxies are consistent between the various black hole samples. However, as the most massive black holes and quasars are hosted by massive galaxies, the stellar mass ratios between the neighbour and the black hole host  $M_{*2}/M_{*1}$  are lower. More than 75 per cent of the neighbours of quasars, and 90 per cent of the neighbours of the most massive black holes, have less than 1/10th of the stellar mass of the black hole host and so would be classified as only minor mergers; this is compared with around 60 per cent of the neighbours of all  $M_{BH} > 10^{6.5} M_{\odot}$  black holes. More than 75 per cent of the neighbours of quasars and most massive black holes have black hole mass ratios  $M_{BH2}/M_{BH1}$  that are also less than 1/10, compared with 26 per cent of neighbours of all  $M_{BH} > 10^{6.5} M_{\odot}$  black holes.

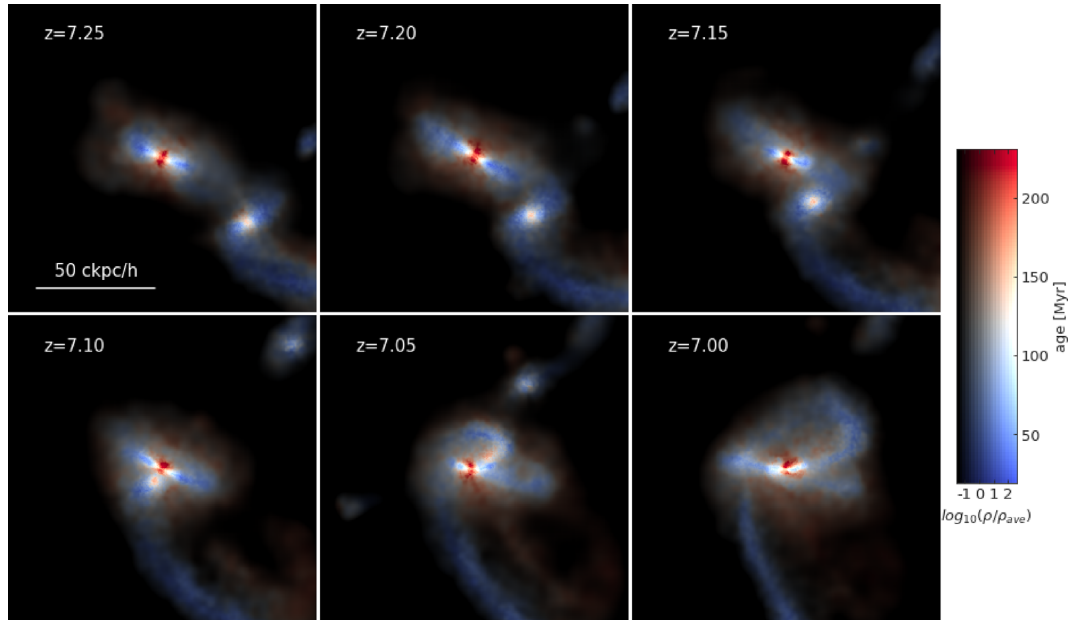
The quasar companions observed by Trakhtenbrot et al. (2017a) have dynamical masses  $M_{dyn} = (2.1 - 10.7) \times 10^{10} M_{*}$  relative to  $M_{dyn} = (3.7 - 7.4) \times 10^{10} M_{*}$  for the quasar hosts, and so these interactions would be classified as major mergers. The quasar companions found by McGreer et al. (2014) are also likely to be major mergers. These companions are also found at projected distances of 5 and 12 kpc, which, depending on the angle of projection, are much smaller than the average distance of companions in the BLUETIDES simulation, as are the majority of ALMA-discovered companions, due to its small field of view. This may be a result of the companion classification used in the simulation being ineffective at low separations.



**Figure 18.** *Top row:* The average number of companion galaxies around each black hole in the sample, as a function of distance from the black hole. Each panel shows companions brighter than a given magnitude limit. From left to right: companions with  $m_{UV} < 22.8$ , i.e. the magnitude of the faintest SDSS quasar; companions with  $m_{UV} < 24.85$ , i.e. the magnitude of the faintest currently known quasar; companions observable in RST, with  $m_{UV} < 26.9$ ; and all companion galaxies. The solid coloured lines show the average number of companions at each magnitude limit, including the effect of dust attenuation, and shaded regions show the  $\pm 1\sigma$  range. The dashed coloured lines show the average number of companions that are intrinsically brighter than the magnitude limit (i.e. without the effect of dust attenuation). Vertical grey dashed lines show the ALMA (Trakhtenbrot et al. 2017a) and JWST NIRCcam fields of view. *Bottom row:* The fraction of companions around a black hole that are missed due to dust attenuation at each magnitude limit. The solid lines show the average fraction, and the shaded regions show the  $\pm 1\sigma$  range. The ‘all black hole’ sample refers to black holes with  $M_{BH} > 10^{6.5} M_{\odot}$ .



**Figure 19.** The properties of the nearest neighbour to each black hole in the various samples that have a distance of less than 200 kpc. The various panels show probability distribution functions for the distance to the nearest neighbour in arcseconds and kpc, the UV magnitude of the neighbouring galaxy (with and without dust attenuation), the difference between the host and companion’s magnitude, the companion’s stellar and black hole mass, and the stellar and black hole mass ratio of the neighbour’s mass to that of the sample black hole’s host. The ‘all black hole’ sample refers to black holes with  $M_{\text{BH}} > 10^{6.5} M_{\odot}$ . These probability distribution functions are calculated using kernel density estimation with Gaussian kernels and a covariance factor of 0.4 for appropriate smoothing.



**Figure 20.** The host galaxies of two merging black holes from  $z = 7.25$  to  $z = 7.0$ . At  $z = 7.3$ , the primary black hole (at the centre of the images) has a black hole mass of  $\log(M_{\text{BH}}/M_{\odot}) = 8.25$ , while the companion has a mass of  $\log(M_{\text{BH}}/M_{\odot}) = 6.37$  and is at a distance of 82 ckpc or 9.9 kpc from the primary black hole. The black hole that results from this merger is one of the 10 most massive black holes in the simulation at  $z = 7$ , with a mass of  $\log(M_{\text{BH}}/M_{\odot}) = 8.56$ , in a galaxy with a stellar mass of  $\log(M_{*}/M_{\odot}) = 11.11$ . Each panel is 120/h ckpc per side, showing the stellar density colour coded by the age of star (from blue to red indicating young to old populations, respectively).



We examine one of the most massive black holes, which has its nearest neighbour within 200 kpc, more closely. This black hole has a mass of  $\log(M_{\text{BH}}/M_{\odot}) = 8.56$  and is hosted by a galaxy of mass  $\log(M_{*}/M_{\odot}) = 11.11$ . We find that at  $z = 7.0$ , its dark matter halo contains five additional black holes, with masses  $\log(M_{\text{BH}}/M_{\odot}) = (8.20, 7.80, 6.53, 7.06, 5.97)$ , in galaxies with stellar masses of  $\log(M_{*}/M_{\odot}) = (10.54, 10.58, 9.24, 9.66, 8.57)$ . Imaging this system at various redshifts shows that this galaxy has been involved in a recent merger between  $z = 7.3$  and  $z = 7.0$ , with the central black hole (of mass  $\log(M_{\text{BH}}/M_{\odot}) = 8.25$  at  $z = 7.3$ ) merging with another black hole of mass  $\log(M_{\text{BH}}/M_{\odot}) = 6.37$  (Fig. 20).

## 7 CONCLUSIONS

In this paper, we use the BLUE TIDES simulation to make predictions for the host galaxies of the most massive black holes and quasars at  $z = 7$ . Our main findings are as follows.

(i) The 10 most massive black holes are in massive galaxies with stellar masses  $\log(M_{*}/M_{\odot}) = 10.80^{+0.20}_{-0.16}$ , which have large SFRs,  $513^{+1225}_{-351} M_{\odot}/\text{yr}$ . Quasar hosts are less massive,  $\log(M_{*}/M_{\odot}) = 10.25^{+0.40}_{-0.37}$ , with lower SFRs,  $191^{+288}_{-120} M_{\odot}/\text{yr}$ . Lower luminosity quasars are hosted by less extreme host galaxies.

(ii) The hosts of the most massive black holes and quasars in BLUE TIDES are generally bulge-dominated, with  $B/T \simeq 0.85 \pm 0.1$ ; however, their morphologies are not biased relative to the overall  $z = 7$  galaxy sample.

(iii) The hosts of the most massive black holes and quasars are compact, with half-mass radii of  $R_{0.5} = 0.41^{+0.18}_{-0.14}$  and  $0.40^{+0.11}_{-0.09}$  kpc, respectively. Galaxies of similar mass and luminosity have a wider range of sizes with a larger median value,  $R_{0.5} = 0.71^{+0.28}_{-0.25}$  kpc.

(iv) Despite its increased resolution over HST, distinguishing the compact host galaxies from the quasar emission will still be challenging with JWST, as shown through our mock images. This will be more successful for galaxies that have the lowest contrast ratio between the host and the AGN.

(v) The  $z = 7$  sample has a black hole–stellar mass relation that is steeper than the local Kormendy & Ho (2013) relation, but the two are reasonably consistent, particularly at the highest masses where the observations are most robust. Sub-mm observations of  $5 \lesssim z \lesssim 7$  quasars with  $M_{\text{BH}} < 10^{8.5} M_{\odot}$  are consistent with our predicted relation.

(vi) Observations of quasars are biased to measure a higher black hole–stellar mass relation than the intrinsic relation. The SDSS, currently observable, and RST quasar samples have black hole–stellar mass relations 0.2 dex higher than the total galaxy sample, providing an estimate of the systematic offset of quasar observations of the  $M_{*}$ – $M_{\text{BH}}$  relation from the true population. To reduce this bias in the measured black hole–stellar mass relation, observations should target AGN that are faint relative to their host galaxies, which are more likely to be found in deep surveys.

(vii) The most massive black holes and quasars have more nearby companions than the typical  $M_{\text{BH}} > 10^{6.5} M_{\odot}$  black hole. The majority of their nearest neighbours have stellar mass ratios  $M_{*2}/M_{*1} < 0.1$  and thus would be classified as minor mergers. A large fraction of these nearby companion galaxies will be missed by rest-frame UV observations due to dust attenuation.

## ACKNOWLEDGEMENTS

We thank the referee for valuable and constructive comments that helped to improve the quality of this paper. We also thank Bram

Venemans for a useful discussion on the measurement of star formation rates from sub-mm observations. This research was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. The BLUE TIDES simulation was run on the BlueWaters facility at the National Center for Supercomputing Applications. Part of this work was performed on the OzSTAR national facility at Swinburne University of Technology, which is funded by Swinburne University of Technology and the National Collaborative Research Infrastructure Strategy (NCRIS). MAM acknowledges the support of an Australian Government Research Training Program (RTP) Scholarship, and a Postgraduate Writing-Up Award sponsored by the Albert Shimmins Fund. TDM acknowledges funding from NSF ACI-1614853, NSF AST-1517593, NSF AST-1616168, and NASA ATP 19-ATP19-0084. TDM and RAC also acknowledge ATP 80NSSC18K101 and NASA ATP 17-0123. This paper makes use of version 17.00 of CLOUDY, last described by Ferland et al. (2017).

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Amiaux J. et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann J. M., eds, *Euclid Mission: building of a reference survey*. Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE
- Bañados E. et al., 2016, *ApJS*, 227, 11
- Bañados E. et al., 2017, An 800-million-solar-mass black hole in a significantly neutral Universe at a redshift of 7.5. *Nature*, Springer Science and Business Media LLC, p. 473
- Barth A. J., Martini P., Nelson C. H., Ho L. C., 2003, *ApJ*, 594, L95
- Battaglia N., Trac H., Cen R., Loeb A., 2013, *ApJ*, 776, 81
- Begelman M. C., Volonteri M., Rees M. J., 2006, *MNRAS*, 370, 289
- Bertoldi F. et al., 2003, *A&A*, 409, L47
- Bondi H., 1952, *MNRAS*, 112, 195
- Bondi H., Hoyle F., 1944, *MNRAS*, 104, 273
- Bornancini C. G., Lambas D. G., 2020, *MNRAS*, 494, 1189
- Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2016, *MNRAS*, 466, 3612
- Bridge J. S. et al., 2019, *ApJ*, 882, 42
- Cochrane R. K. et al., 2019, *MNRAS*, 488, 1779
- Davies F. B. et al., 2018, *ApJ*, 864, 142
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- De Rosa G. et al., 2014, *ApJ*, 790, 145
- De Rosa G., Decarli R., Walter F., Fan X., Jiang L., Kurk J., Pasquali A., Rix H. W., 2011, *ApJ*, 739, 56
- Decarli R. et al., 2017, *Nat*, 545, 457
- Decarli R. et al., 2018, *ApJ*, 854, 97
- DeGraf C., Matteo T. D., Khandai N., Croft R., Lopez J., Springel V., 2012a, *MNRAS*, 424, 1892
- DeGraf C., Matteo T. D., Khandai N., Croft R., 2012b, *ApJ*, 755, L8
- DeGraf C., Matteo T. D., Treu T., Feng Y., Woo J.-H., Park D., 2015, *MNRAS*, 454, 913
- Di Matteo T., Springel V., Hernquist L., 2005, *Nat*, 433, 604
- Di Matteo T., Khandai N., DeGraf C., Feng Y., Croft R. A. C., Lopez J., Springel V., 2012, *ApJ*, 745, L29
- Di Matteo T., Croft R. A. C., Feng Y., Waters D., Wilkins S., 2017, *MNRAS*, 467, 4243
- Fan X. et al., 2000, *AJ*, 120, 1167
- Fan X. et al., 2001, *AJ*, 122, 2833
- Fan X. et al., 2003, *AJ*, 125, 1649
- Fan X. et al., 2004, *AJ*, 128, 515

- Fan X. et al., 2006a, *AJ*, 131, 1203
- Fan X. et al., 2006b, *AJ*, 132, 117
- Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, *MNRAS*, 436, 315
- Faucher-Giguère C.-A., Lidz A., Zaldarriaga M., Hernquist L., 2009, *ApJ*, 703, 1416
- Feng Y., Matteo T. D., Croft R., Khandai N., 2014, *MNRAS*, 440, 1865
- Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2015, *MNRAS*, 455, 2778
- Ferland G. J. et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385
- Fujimoto S. et al., 2020, *ApJ*, 900, 1
- Gardner J. P. et al., 2006, *Space Sci. Rev.*, 123, 485
- Greig B., Mesinger A., 2017, *MNRAS*, 465, 4838
- Greig B., Mesinger A., Bañados E., 2019, *MNRAS*, 484, 5094
- Habouzit M., Volonteri M., Somerville R. S., Dubois Y., Peirani S., Pichon C., Devriendt J., 2019, *MNRAS*, 489, 1206
- Hinshaw G. et al., 2013, *ApJS*, 208, 19
- Hogg D. W., Baldry I. K., Blanton M. R., Eisenstein D. J., 2002, The K-correction, preprint ([astro-ph/0210394v1](https://arxiv.org/abs/astro-ph/0210394v1))
- Hoyle F., Lyttleton R. A., 1939, *Math. Proc. Cambridge Philos. Soc.*, 35, 405
- Huang K.-W., Matteo T. D., Bhowmick A. K., Feng Y., Ma C.-P., 2018, *MNRAS*, 478, 5063
- Inayoshi K., Visbal E., Haiman Z., 2020, *ARA&A*, 58,
- Izumi T. et al., 2018, *PASJ*, 70, 36
- Jiang L. et al., 2016, *ApJ*, 833, 222
- Jiang L., Fan X., Vestergaard M., Kurk J. D., Walter F., Kelly B. C., Strauss M. A., 2007, *AJ*, 134, 1150
- Kashikawa N. et al., 2015, *ApJ*, 798, 28
- Katz N., Hernquist L., Weinberg D. H., 1999, *ApJ*, 523, 463
- Kawamata R., Ishigaki M., Shimasaku K., Oguri M., Ouchi M., Tanigawa S., 2018, *ApJ*, 855, 4
- Khandai N., Feng Y., DeGraf C., Matteo T. D., Croft R. A. C., 2012, *MNRAS*, 423, 2397
- Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511
- Krumholz M. R., Gnedin N. Y., 2011, *ApJ*, 729, 36
- Kurk J. D. et al., 2007, *ApJ*, 669, 32
- Latif M. A., Schleicher D. R. G., Schmidt W., Niemeyer J. C., 2013, *MNRAS*, 436, 2989
- Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, *ApJ*, 670, 249
- Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- LSST Science Collaboration, 2009, *LSST Science Book, Version 2.0*
- Lupi A., Volonteri M., Decarli R., Bovino S., Silk J., Bergeron J., 2019, *MNRAS*, 488, 4004
- Madau P., 1995, *ApJ*, 441, 18
- Matsuoka Y. et al., 2018a, *PASJ*, 70, S35
- Matsuoka Y. et al., 2018b, *ApJ*, 869, 150
- Mayer L., Bonoli S., 2018, *Rep. Prog. Phys.*, 82, 016901
- Mazzucchelli C. et al., 2019, *ApJ*, 881, 163
- McGreer I. D., Fan X., Strauss M. A., Haiman Z., Richards G. T., Jiang L., Bian F., Schneider D. P., 2014, *AJ*, 148, 73
- Mechtley M. et al., 2012, *ApJ*, 756, L38
- Mor R., Netzer H., Trakhtenbrot B., Shemmer O., Lira P., 2012, *ApJ*, 749, L25
- Mortlock D. J. et al., 2011, *Nat*, 474, 616
- Narayanan D., Cox T. J., Hayward C. C., Younger J. D., Hernquist L., 2009, *MNRAS*, 400, 1919
- Neeleman M. et al., 2019, *ApJ*, 882, 10
- Netzer H., Mor R., Trakhtenbrot B., Shemmer O., Lira P., 2014, *ApJ*, 791, 34
- Nguyen N. H., Lira P., Trakhtenbrot B., Netzer H., Cicone C., Maiolino R., Shemmer O., 2020, *ApJ*, 895, 74
- Ni Y., Di Matteo T., Feng Y., Croft R. A. C., Tenneti A., 2018, *MNRAS*, 481, 4877
- Ni Y., Matteo T. D., Gilli R., Croft R. A. C., Feng Y., Norman C., 2019, *MNRAS*, 495, 2135
- Okamoto T., Frenk C. S., Jenkins A., Theuns T., 2010, *MNRAS*, 406, 208
- Oke J. B., Sandage A., 1968, *ApJ*, 154, 21
- Pensabene A., Carniani S., Perna M., Cresci G., Decarli R., Maiolino R., Marconi A., 2020, *A&A*, 637, A84
- Ren K., Trenti M., Matteo T. D., 2020, *ApJ*, 894, 124
- Riechers D. A., Walter F., Carilli C. L., Bertoldi F., 2007, *ApJ*, 671, L13
- Salviander S., Shields G. A., Gebhardt K., Bonning E. W., 2007, *ApJ*, 662, 131
- Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, *MNRAS*, 396, 696
- Schulze A., Wisotzki L., 2014, *MNRAS*, 438, 3422
- Shang C., Bryan G. L., Haiman Z., 2010, *MNRAS*, 402, 1249
- Shao Y. et al., 2019, *ApJ*, 876, 99
- Shen Y. et al., 2007, *AJ*, 133, 2222
- Shibuya T., Ouchi M., Harikane Y., 2015, *ApJS*, 219, 15
- Shields G. A., Menezes K. L., Massart C. A., Bout P. V., 2006, *ApJ*, 641, 683
- Silverman J. D. et al., 2019, *ApJ*, 887, L5
- Spergel D. et al., 2015, Wide-Field Infrared Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report
- Springel V. et al., 2005b, *Nat*, 435, 629
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V., Di Matteo T., Hernquist L., 2005a, *MNRAS*, 361, 776
- Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 479, 75
- STScI, 2017, NIRCam Sensitivity, <https://jwst-docs.stsci.edu/near-infrared-camera/nircam-predicted-performance/nircam-sensitivity>
- STScI Development Team, 2018, Astrophysics Source Code Library, record ascl:1811.001
- Tenneti A., Wilkins S. M., Matteo T. D., Croft R. A. C., Feng Y., 2018, *MNRAS*, 483, 1388
- Trakhtenbrot B., Lira P., Netzer H., Cicone C., Maiolino R., Shemmer O., 2017a, *ApJ*, 836, 8
- Trakhtenbrot B., Volonteri M., Natarajan P., 2017b, *ApJ*, 836, L1
- Valiante R., Schneider R., Salvadori S., Gallerani S., 2014, *MNRAS*, 444, 2442
- Valiante R., Agarwal B., Habouzit M., Pezzulli E., 2017, *PASA*, 34, e031
- Venemans B. P. et al., 2017, *ApJ*, 837, 146
- Venemans B. P. et al., 2018, *ApJ*, 866, 159
- Venemans B. P., 2017, *Messenger*, 169, 48
- Venemans B. P., Walter F., Zschaechner L., Decarli R., De Rosa G., Findlay J. R., McMahon R. G., Sutherland W. J., 2016, *ApJ*, 816, 37
- Venemans B. P., Neeleman M., Walter F., Novak M., Decarli R., Hennawi J. F., Rix H.-W., 2019, *ApJ*, 874, L30
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
- Volonteri M., 2010, *A&A Rev.*, 18, 279
- Volonteri M., 2012, *Sci*, 337, 544
- Volonteri M., Stark D. P., 2011, *MNRAS*, 417, 2085
- Wagg J. et al., 2012, *ApJ*, 752, L30
- Walter F. et al., 2003, *Nat*, 424, 406
- Walter F., Carilli C., Bertoldi F., Menten K., Cox P., Lo K. Y., Fan X., Strauss M. A., 2004, *ApJ*, 615, L17
- Walter F., Riechers D., Cox P., Neri R., Carilli C., Bertoldi F., Weiss A., Maiolino R., 2009, *Nat*, 457, 699
- Wang R. et al., 2010, *ApJ*, 714, 699
- Wang R. et al., 2011, *AJ*, 142, 101
- Wang R. et al., 2013, *ApJ*, 773, 44
- Wang F. et al., 2019, *ApJ*, 884, 30
- Waters D., Matteo T. D., Feng Y., Wilkins S. M., Croft R. A. C., 2016, *MNRAS*, 463, 3520
- Wilkins S. M., Feng Y., Matteo T. D., Croft R., Lovell C. C., Waters D., 2017, *MNRAS*, 469, 2517
- Willott C. J. et al., 2010a, *AJ*, 139, 906
- Willott C. J. et al., 2010b, *AJ*, 140, 546
- Willott C. J., Percival W. J., McLure R. J., Crampton D., Hutchings J. B., Jarvis M. J., Sawicki M., Simard L., 2005, *ApJ*, 626, 657
- Willott C. J., Bergeron J., Omont A., 2017, *ApJ*, 850, 108