

# Dark-ages reionization and galaxy formation simulation – X. The small contribution of quasars to reionization

Yuxiang Qin,<sup>1</sup><sup>★</sup> Simon J. Mutch,<sup>1</sup> Gregory B. Poole,<sup>1</sup> Chuanwu Liu,<sup>1</sup> Paul W. Angel,<sup>1</sup> Alan R. Duffy,<sup>2</sup> Paul M. Geil,<sup>1</sup> Andrei Mesinger<sup>3</sup> and J. Stuart B. Wyithe<sup>1</sup><sup>★</sup>

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

<sup>2</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn VIC 3122, Australia*

<sup>3</sup>*Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy*

Accepted 2017 July 25. Received 2017 July 23; in original form 2017 March 13

## ABSTRACT

Motivated by recent measurements of the number density of faint AGN at high redshift, we investigate the contribution of quasars to reionization by tracking the growth of central supermassive black holes in an update of the MERAXES semi-analytic model. The model is calibrated against the observed stellar mass function at  $z \sim 0.6$ –7, the black hole mass function at  $z \lesssim 0.5$ , the global ionizing emissivity at  $z \sim 2$ –5 and the Thomson scattering optical depth. The model reproduces a Magorrian relation in agreement with observations at  $z < 0.5$  and predicts a decreasing black hole mass towards higher redshifts at fixed total stellar mass. With the implementation of an opening angle of 80 deg for quasar radiation, corresponding to an observable fraction of  $\sim 23.4$  per cent due to obscuration by dust, the model is able to reproduce the observed quasar luminosity function at  $z \sim 0.6$ –6. The stellar light from galaxies hosting faint active galactic nucleus (AGN) contributes a significant or dominant fraction of the UV flux. At high redshift, the model is consistent with the bright end quasar luminosity function and suggests that the recent faint  $z \sim 4$  AGN sample compiled by Giallongo et al. (2015) includes a significant fraction of stellar light. Direct application of this luminosity function to the calculation of AGN ionizing emissivity consequently overestimates the number of ionizing photons produced by quasars by a factor of 3 at  $z \sim 6$ . We conclude that quasars are unlikely to make a significant contribution to reionization.

**Key words:** methods: numerical – galaxies: formation – galaxies: high-redshift – quasars: supermassive black holes.

## 1 INTRODUCTION

The epoch of reionization (EoR) is the phase of the Universe when neutral hydrogen in the intergalactic medium (IGM) was reionized. Star-forming galaxies at high redshift are believed to be one of the dominant sources of ionizing UV photons provided one assumes a high average escape fraction of Lyman continuum radiation ( $f_{\text{esc},*} \gtrsim 10$  per cent) and extends the UV luminosity function to faint dwarf galaxies (Kuhlen & Faucher-Giguère 2012; Duffy et al. 2014; Feng et al. 2016; Mesinger, Greig & Sobacchi 2016). However, the value of  $f_{\text{esc},*}$  is very uncertain. Observations of star-forming galaxies at low redshift usually indicate a much lower escape fraction. For example, by measuring the ratio of Ly  $\alpha$  to H  $\beta$  line emission, Ciardullo et al. (2014) derived an escape fraction of 4.4 per cent, while Matthee et al. (2016) measured a median escape fraction

of 1.6 per cent using stacking of H  $\alpha$ -selected galaxies. Moreover, some theoretical works also suggest a low  $f_{\text{esc},*}$  (Gnedin, Kravtsov & Chen 2008; Hassan et al. 2016; Sun & Furlanetto 2016).

In order to reconcile the difference between low-redshift observations and the photon budget at high redshift, some propose a rapid increase of  $f_{\text{esc},*}$  with redshift (Haardt & Madau 2012; Khaire et al. 2016; Price, Trac & Cen 2016; Sharma et al. 2016) and with decreasing mass (Paardekooper, Khochfar & Dalla 2013; Kimm & Cen 2014; Wise et al. 2014). This is supported by identifying local analogues of high-redshift galaxies and extrapolating the observed  $f_{\text{esc},*}$  using indicators such as the [O III]/[O II] ratio to high redshift (Faisst 2016 and references therein). On the other hand, additional contributors to reionization may also be present. For example, Ma et al. (2016) included a binary population into a set of radiative transfer cosmological simulations and found that they produced significantly more ionizing photons among the old stellar population compared to a model without binaries, reducing the requirement of high escape fraction. In addition, these photons produced at later times can escape from galaxies more easily since the local feedback

\*E-mail: Yuxiang.L.Qin@gmail.com (YQ); swyithe@unimelb.edu.au (SBW)

efficiently clears out nearby gas, leading to a lower required escape fraction on average. With a high escape fraction ( $f_{\text{esc,q}} \sim 1$ , Barkana & Loeb 2001), quasars (active galactic nucleus, AGN) are potential contributors to reionization (Volonteri & Gnedin 2009; Fontanot et al. 2014; Madau & Haardt 2015; Mitra, Roy Choudhury & Ferrara 2015), despite their relatively low number. Recently, Giallongo et al. (2015) identified faint AGN candidates in the CANDELS GOODS-South field and suggested that there is a high number density of faint AGN at  $z = 4-6$ . These faint quasars provide a new source of reionization (Madau & Haardt 2015). However, it is still debated whether there are enough luminous quasars at high redshift to make a significant contribution (Bouwens et al. 2015; Weigel et al. 2015; Manti et al. 2017; Parsa, Dunlop & McLure 2017) and whether the escape fraction of high-redshift, low-luminosity AGN is of the order of unity (Cristiani et al. 2016; Micheva, Iwata & Inoue 2017).

To explore the consequences for galaxy formation and reionization from these faint quasars, this paper describes the addition of a population of evolving black holes to the MERAXES<sup>1</sup> semi-analytic model of galaxy formation and reionization (Mutch et al. 2016). This new model enables a detailed exploration of the relative role of quasars during the EoR. The paper is organized as follows. We begin with a description of the semi-analytic model in Section 2, in which the black hole growth model is introduced in detail. We present the black hole properties in Section 3 and explore reionization from quasars in Sections 4 and 5. Conclusions are given in Section 6. In this work, we adopt cosmological parameters from the Planck 2015 results ( $\Omega_m, \Omega_b, \Omega_\Lambda, h, \sigma_8, n_s = 0.308, 0.0484, 0.692, 0.678, 0.815, 0.968$ ; Planck Collaboration XIII 2016).

## 2 MODELLING BLACK HOLE GROWTH

Built on halo<sup>2</sup> merger trees constructed from the *Tiamat* collisionless  $N$ -body simulation (Poole et al. 2016), the MERAXES semi-analytic model (Mutch et al. 2016) was specifically designed to study galaxy formation and reionization at high redshift. The model computes galaxy properties according to different astrophysical processes including gas infall, cooling, star formation, supernova feedback, metal enrichment, mergers and reionization. In order to properly track the evolution of galaxies during reionization, *Tiamat* provide 100 snapshots between  $z = 35$  and 5 with a time interval of  $\sim 11.1$  Myr and 64 additional snapshots between  $z = 5$  and 2 separated equally in units of Hubble time. The mass resolution of *Tiamat* is  $\sim 2.64 \times 10^6 h^{-1} M_\odot$  and the box size is  $67.8 h^{-1}$  Mpc. Additionally, in order to obtain information of more massive objects and lower redshifts, we also take advantages of the dark matter halo merger trees generated from the *Tiamat-125-HR* simulation (Poole et al. in prep). The *Tiamat-125-HR* simulation shares identical cosmology with *Tiamat* but has a lower mass resolution of  $0.12 \times 10^9 h^{-1} M_\odot$  in a bigger simulation volume, with side lengths equal to 125  $h^{-1}$  Mpc. The *Tiamat-125-HR* trees are constructed down to  $z = 0.56$  with the same snapshot separation strategy as the *Tiamat* trees.

In the following subsections, we briefly describe the galaxy formation in MERAXES and introduce the new implementation of black

hole growth and feedback in detail. More details about the implemented galaxy formation physics can be found in Mutch et al. (2016).

### 2.1 Galaxy evolution

In the MERAXES semi-analytic model, cooling and star formation are assumed to be negligible in haloes below the atomic cooling mass threshold,  $\sim 10^8 M_\odot$ . Thus, once a halo grows larger than the atomic cooling limit it is designated as a galaxy, which has three baryonic components: gas, stars and a central black hole. During one time step,  $\Delta t$ , additional gas falls into the hot gas component of a galaxy from the IGM when the mass fraction of baryons in the halo is lower than the cosmic mean value,  $f_b = \Omega_b/\Omega_m$ :

$$\Delta m_{\text{hot}} = \max \left[ 0, \chi_r f_b M_{\text{vir}} - (m_* + m_{\text{cold}} + m_{\text{hot}} + m_{\text{eject}}) \right], \quad (1)$$

where  $m_*$ ,  $m_{\text{cold}}$ ,  $m_{\text{hot}}$  and  $m_{\text{eject}}$  are the masses of the stellar component, cold gas, hot gas and ejected gas, respectively, and  $M_{\text{vir}}$  is the virial mass of the host halo in which the galaxy forms.<sup>3</sup>  $\chi_r$  is a baryon fraction modifier to take account of reionization feedback and will be introduced later in Section 4.1.

Some of the hot gas,  $m_{\text{cool}}$ , cools and collapses on to the cold disc. Assuming the cooling process is in quasi-static thermal equilibrium, one can calculate the cooling time by

$$t_{\text{cool}}(r) = \frac{3\bar{\mu}m_p k T_{\text{hot}}}{2\rho_{\text{hot}}(r)\Lambda(T_{\text{hot}}, Z_{\text{hot}})}, \quad (2)$$

where  $\mu = 0.59$ ,  $m_p$  and  $k$  are the mean molecular weight for fully ionized gas, the mass of a proton and the Boltzmann constant, respectively.  $T_{\text{hot}}$  and  $\rho_{\text{hot}}$  are the temperature and density of the hot gas.  $\Lambda$  is the cooling function (Sutherland & Dopita 1993) determined by the temperature and metallicity,  $Z_{\text{hot}}$ . Assuming the hot gas shares the same temperature as the host halo ( $T_{\text{hot}} = T_{\text{vir}}$ ) due to shock heating, and follows a singular isothermal sphere density profile, equation (2) becomes

$$t_{\text{cool}}(r) = \frac{6\pi\bar{\mu}m_p k T_{\text{vir}} R_{\text{vir}} r^2}{m_{\text{hot}} \Lambda(T_{\text{vir}}, Z_{\text{hot}})}, \quad (3)$$

where  $R_{\text{vir}}$  is the virial radius. Following Croton et al. (2006), we calculate the cooling radius,  $r_{\text{cool}}$ , at which the cooling time is equal to the halo dynamical time, through

$$r_{\text{cool}} = \sqrt{\frac{m_{\text{hot}} \Lambda(T_{\text{vir}}, Z_{\text{hot}})}{6\pi\bar{\mu}m_p k T_{\text{vir}} V_{\text{vir}}}}, \quad (4)$$

where  $V_{\text{vir}}$  is the virial velocity. Cooling is sufficient within  $r_{\text{cool}}$  and we estimate the cooling mass by

$$m_{\text{cool}} \equiv m_{\text{hot}} \times \min \left[ 1, \min \left( 1, \frac{r_{\text{cool}}}{R_{\text{vir}}} \right) \times \frac{\Delta t}{t_{\text{cool}}} \right], \quad (5)$$

which is removed from the hot gas reservoir and redistributed into the cold gas disc,  $\Delta m_{\text{hot}} = -\Delta m_{\text{cool}} = -m_{\text{cool}}$ . From equation (5), we see that depending on the ratio of  $r_{\text{cool}}$  to  $R_{\text{vir}}$ , cooling is separated into two regimes: *static hot halo* ( $r_{\text{cool}} > R_{\text{vir}}$ ) and *rapid cooling* ( $r_{\text{cool}} < R_{\text{vir}}$ , see more in Croton et al. 2006). In the case of a static hot halo, cooling is in thermal equilibrium and the cooling

<sup>1</sup> <http://dragons.ph.unimelb.edu.au>

<sup>2</sup> Note that in this work, haloes are defined as the substructures of FOF groups. Central haloes (the most massive halo in a FOF group) and satellites co-evolve in MERAXES and both contribute to reionization.

<sup>3</sup> A central halo consist of the majority of its FOF particles and dominates the gravitational potential of the entire system. In the code, the infall gas is added into the central halo based on the baryon fraction of its FOF group while the satellites do not get any fresh gas (see more details in Mutch et al. 2016).

rate is determined by the continuity equation. On the other hand, if the cooling time is shorter than the dynamical time this leads to a rapid cooling process, with a free-fall of all the hot gas on to the cold gas disc. We note that  $m_{\text{cool}}$  is set to be zero when the mass of the hot halo drops below the atomic cooling due to stripping.

When the galaxy collects enough cold gas,  $m_{\text{cold}} > m_{\text{crit}}$ , based on Kennicutt (1998) and Kauffmann (1996), it forms new stars,  $\Delta m_*$ , through a burst

$$\Delta m_* = \min \left[ \frac{\alpha_{\text{sf}} \max(0, m_{\text{cold}} - m_{\text{crit}}) \Delta t}{t_{\text{dyn, disc}}}, m_{\text{cold}} \right], \quad (6)$$

where  $t_{\text{dyn, disc}} = 1.5\sqrt{2}\lambda R_{\text{vir}}/V_{\text{vir}}$  is the dynamical time of the cold gas disc ( $\lambda$  is the spin parameter of the host halo) and  $\alpha_{\text{sf}}$  is a free parameter corresponding to the star formation efficiency. This mass is removed from the cold gas reservoir,  $\Delta m_{\text{cold}} = -\Delta m_*$ .

The new stellar mass is assumed to form following a Salpeter (1955) initial mass function (IMF). Ultimately some of these stars,  $\eta_{\text{SNII}} \Delta m_*$ , recycle their mass back to the interstellar medium (ISM) through type-II supernovae explosion.<sup>4</sup> The energy produced by these supernovae heats the ISM and converts some of the cold gas to the hot phase or, if there is sufficient energy, even ejects a fraction of hot gas from the galaxy. Assuming that the efficiency for supernovae energy coupling with the ISM scales with mass and is in proportion to a free parameter (Guo et al. 2011),  $\alpha_{\text{energy}}$ , then the total energy released by supernovae that couples to the ISM is

$$e_{\text{total}} = \min \left\{ \alpha_{\text{energy}} \left[ 0.5 + \left( \frac{V_{\text{max}}}{70 \text{ km s}^{-1}} \right)^{-2.0} \right], 1 \right\} \times 10^{51} \text{ erg} \times \eta_{\text{SNII}} \Delta m_*, \quad (7)$$

where  $V_{\text{max}}$  is the maximum circular velocity of the host halo. Since a massive star ( $\gtrsim 8 M_{\odot}$ ) takes  $\sim 40$  million years, or 4 snapshots, before reaching its type-II supernova stage, MERAXES accounts for supernovae not only from the current snapshot,  $j$ , but also from the stars formed in the previous 4 snapshots. Therefore, the total energy released during one snapshot is

$$E_{\text{total}} = \sum_{i=j-4}^{i=j} e_{\text{total}, i} (\Delta m_{*, i}, V_{\text{max}, i}, \eta_{\text{SNII}, i}). \quad (8)$$

Since the hot gas shares the same virial temperature as the host halo, assuming the mass loading factor for reheating cold gas is a free parameter,  $\alpha_{\text{mass}}$ , the energy utilized in gas heating can be calculated by

$$E_{\text{reheat}} = \frac{1}{2} \alpha_{\text{mass}} \Delta m_* V_{\text{vir}}^2, \quad (9)$$

where  $V_{\text{vir}}$  is the virial velocity. Depending on the available energy,  $E_{\text{total}}$  and the required energy for re-heating,  $E_{\text{reheat}}$ , the reheated mass,  $m_{\text{reheat}}$ , is

$$m_{\text{reheat}} = \min \left[ m_{\text{cold}}, \frac{\min(E_{\text{total}}, E_{\text{reheat}})}{0.5 V_{\text{vir}}^2} \right]. \quad (10)$$

This mass is removed from the cold gas reservoir and redistributed in the hot gas component,  $\Delta m_{\text{cold}} = -\Delta m_{\text{hot}} = -m_{\text{reheat}}$ . If there is still some energy left after reheating, the supernovae feedback

will further remove hot gas from the galaxy, adding it to the ejected component

$$\Delta m_{\text{eject}} = \min \left[ m_{\text{hot}}, \frac{\max(0, E_{\text{total}} - 0.5 m_{\text{reheat}} V_{\text{vir}}^2)}{0.5 V_{\text{vir}}^2} \right], \quad (11)$$

which is removed from the hot gas reservoir,  $\Delta m_{\text{hot}} = -\Delta m_{\text{eject}}$ .

In addition, the metals produced by supernovae enrich the environment, which then enhances the cooling rate (see equation 2). Moreover, mergers drive strong turbulence and hence increase the possibility of star formation. Major mergers generally induce more energetic bursts than minor mergers since they induce strong inflows and easily trigger bar-like instabilities in the cloud (Somerville, Primack & Faber 2001). Therefore, following mergers, MERAXES also includes a starburst mechanism. Reionization feedback will be introduced in Section 4.1. There are more details of the semi-analytic model in Mutch et al. (2016). This work extends the current model with a detailed black hole growth prescription based on Croton et al. (2016) and is described in the following subsection.

## 2.2 Black hole growth

In the updated model, every newly formed galaxy is seeded with a central black hole<sup>5</sup> of mass  $1000 h^{-1} M_{\odot}$ . The two gas reservoirs in the galaxy (i.e. hot and cold) lead to two different black hole growth scenarios, termed radio and quasar modes (Croton et al. 2016). In the normal quiescent state, black holes only accrete mass from the hot gas reservoir, resulting in radio emission in the centre of galaxy. However, mergers trigger rapid accretion on to the black hole from the cold gas disc, causing them to radiate as quasars. In this work, we do not distinguish AGN with different types and refer to them all as quasars unless specified otherwise.

### 2.2.1 Accretion of hot gas

Whenever there is a static hot gas reservoir,  $m_{\text{hot}}$ , around the galaxy, some of it will cool,  $m_{\text{cool}}$ , and form a cold gas disc,  $m_{\text{cold}}$  (see the previous section), while some will be directly accreted by the central black hole. We adopt the Bondi–Hoyle accretion model proposed in Croton et al. (2016) to describe this hot gas accretion. The Bondi–Hoyle accretion rate is given by

$$\dot{m}_{\text{Bondi}} = \frac{2.5\pi G^2 m_{\text{bh}}^2 \rho_{\text{hot}}}{c_s^3}, \quad (12)$$

where  $G$ ,  $m_{\text{bh}}$ ,  $c_s$  and  $\rho_{\text{hot}}$  are the gravitational constant, the black hole mass, the speed of sound and the density of the hot gas reservoir, respectively. We define the Bondi radius as  $r_{\text{Bondi}} = \frac{2Gm_{\text{bh}}}{c_s^2}$ , and by equating the sound crossing travel time from the centre to the Bondi radius with the local cooling time (see equation 2), we obtain

$$\frac{\rho_{\text{hot}}}{c_s^3} = \frac{3\mu m_p kT}{4G\Lambda m_{\text{bh}}}. \quad (13)$$

Then, assuming the accretion rate does not change during one time step,<sup>6</sup>  $\Delta t$ , the accretion mass is

$$\Delta m_{\text{bh, hot}} = \min(m_{\text{hot}}, m_{\text{Edd}}, k_h \dot{m}_{\text{Bondi}} \Delta t), \quad (14)$$

<sup>5</sup> Dual or multiple AGN are not considered.

<sup>6</sup> It is generally true for the radio mode. However, the quasar mode introduced in Section 2.2.2 can induce significant increases in black hole mass during a single time step, leading to a non-negligible change of the accretion rate.

<sup>4</sup> MERAXES assumes the Salpeter IMF with a mass range of  $0.1\text{--}120 M_{\odot}$ . We use the lifetime-mass relation of stars (Portinari, Chiosi & Bressan 1997) to calculate the fraction ( $\eta_{\text{SNII}}$ ) of stellar mass in a single stellar population that have reached the supernova stage after a certain period of time. For example, assuming stars more massive than  $8 M_{\odot}$  will reach the type-II supernova stage after  $\sim 40$  Myr,  $\eta_{\text{SNII}} = 7.432 \times 10^{-3}$ .

where  $k_h$  is a free parameter, used to adjust the efficiency of black hole growth in the radio mode since black holes may not be accreting at the full Bondi–Hoyle accretion rate.  $m_{\text{Edd}}$  is the Eddington limit, which will be introduced later. This mass is removed from the hot gas reservoir,  $\Delta m_{\text{hot}} = -\Delta m_{\text{bh, hot}}$ . Assuming a fraction,  $\eta$ , of the accreted mass is radiated, the black hole only grows by

$$\Delta m_{\text{bh}} = (1 - \eta)\Delta m_{\text{bh, hot}}. \quad (15)$$

Moreover, the radiation acting outward is limited by the Eddington luminosity

$$\eta \dot{m}_{\text{bh}}(t)c^2 = \epsilon \frac{4\pi G m_{\text{bh}}(t)m_p c}{\sigma_T}, \quad (16)$$

where  $\epsilon$  is the Eddington ratio ( $\epsilon = 1$  in this work),  $m_{\text{bh}}(t)$  and  $\dot{m}_{\text{bh}}(t)$  are the black hole mass and growth rate at time  $t$ , respectively,  $\sigma_T$  is the Thomson cross-section area and  $c$  is the speed of light. By integrating equation (16) through one time step, this provides the second limitation in equation (14)

$$m_{\text{Edd}} = m_{\text{bh}} \left[ \exp\left(\frac{\epsilon \Delta t}{\eta t_{\text{Edd}}}\right) - 1 \right], \quad (17)$$

where  $t_{\text{Edd}} \equiv \frac{\sigma_T c}{4\pi G m_p} \approx 450 \text{ Myr}$  is the Eddington accretion time-scale and  $m_{\text{bh}}$  is the black hole mass at the beginning of this time step.

Assuming adiabatic heating and that a fraction of the radiated energy,  $\kappa_r$ , is coupled to the surrounding gas and therefore can contribute to feedback, then the heated mass due to black hole feedback can be calculated through

$$m_{\text{heat}} = \frac{\kappa_r \eta \Delta m_{\text{bh, hot}} c^2}{0.5 V_{\text{vir}}^2}, \quad (18)$$

which is subtracted from the cooling flow,  $\Delta m_{\text{cool}} = -m_{\text{heat}}$ . However, if the heating from black hole feedback is too strong,  $m_{\text{heat}} > m_{\text{cool}}$  (see equation 5), it will significantly suppress the cooling flow, which will consequently restrain the black hole accretion of hot gas. This suppression is referred to as radio mode feedback. In this case, following<sup>7</sup> Croton et al. (2016),  $\Delta m_{\text{bh, hot}}$  is rescaled to be the amount of mass within the cooling radius,  $\Delta m_{\text{bh, hot}} = \Delta m_{\text{bh, hot}} \times \frac{m_{\text{cool}}}{m_{\text{heat}}}$  and the heated mass consequently shrinks to be the cooling mass,  $m_{\text{heat}} = m_{\text{cool}}$ , resulting in a complete quenching of cooling,  $\Delta m_{\text{cool}} = -m_{\text{heat}} \rightarrow m_{\text{cool}} = 0$ . Radio mode feedback limits gas condensation in cluster cooling flows and regulates star formation in massive galaxies (Croton et al. 2006). We note that this does not have a significant impact on the results from *Tiamat* at  $z \geq 2$  due to the limited number of massive objects in the box. However, black hole feedback is important in more massive objects at lower redshifts, which can be observed using the *Tiamat-125-HR* halo merger trees.

We note that at high redshift, accreted hot gas does not contribute significantly to black hole growth, with the accreted mass during the radio mode typically being only  $\sim 0.1$  per cent of the mass from the quasar mode (see the Appendix) that is introduced in the next section. Additionally, because the radio mode accretion rate is approximately 3 orders of magnitude smaller than the Eddington accretion rate, we ignore the energy due to the radio mode for the calculation of the quasar luminosity.

<sup>7</sup> We do not limit the heating radius (see Croton et al. 2016) to only move outwards.

### 2.2.2 Accretion of cold gas

Many binary AGN have been detected in merging galaxies (Shields et al. 2012; Comerford et al. 2013; Comerford & Greene 2014; Comerford et al. 2015; Müller-Sánchez et al. 2015). This suggests that galaxy mergers might trigger AGN activity, which is also supported by hydrodynamic simulations of galaxy mergers (Capelo et al. 2015; Volonteri et al. 2015a,b; Steinborn et al. 2016). Following Croton et al. (2016), when a merger occurs between two galaxies, the central black holes coalesce and their masses combine.<sup>8</sup> As mergers drive strong gas inflows towards the central region (Capelo et al. 2015), cold gas is funnelled to the central black hole of the resulting merged galaxy, significantly increasing its mass. The amount of accretion can be estimated by (Bonoli et al. 2009; Croton et al. 2016)

$$\Delta m_{\text{bh, max}} = \min \left[ m_{\text{cold}}, \frac{k_c \gamma m_{\text{cold}}}{\left(1 + \frac{280 \text{ km s}^{-1}}{V_{\text{vir}}}\right)^2} \right] + \Delta m'_{\text{bh, max}}, \quad (19)$$

where  $m_{\text{cold}}$  is the amount of mass available in the cold gas disc,  $k_c$  is a free parameter used to modulate the strength of black hole accretion and  $\gamma \leq 1$  is the mass ratio between the two merging galaxies, respectively. The term  $\Delta m'_{\text{bh, max}}$  corresponds to the accretion mass left from the quasar mode in the previous snapshot, which will be introduced later.

Unlike the radio mode, black holes grow dramatically during the quasar mode. The AGN activity lifetime (a few  $10^7$  yr up to a Gyr Fiore et al. 2012) is much longer than the 11 Myr time step at  $z > 5$ . In some cases, this leaves the central black hole insufficient time to consume all of the newly accreted gas,  $\Delta m_{\text{bh, max}}$ , at the Eddington limit. Therefore, the mass actually accreted by the black hole is

$$\Delta m_{\text{bh, cold}} = \min(m_{\text{Edd}}, \Delta m_{\text{bh, max}}), \quad (20)$$

which is removed from the cold gas reservoir,  $\Delta m_{\text{cold}} = -\Delta m_{\text{bh, cold}}$ . In our model, during the quasar mode black holes are assumed to either accrete and radiate at the Eddington rate or stay quiescent if the accretion mass is not sufficient. Therefore, depending on the total available mass brought in,  $\Delta m_{\text{bh, max}}$ , and the Eddington limit,  $m_{\text{Edd}}$ , there are two possible scenarios when the central black hole is undergoing a merger:

(i)  $\Delta m_{\text{bh, max}} < m_{\text{Edd}}$ : In this case, there is inadequate mass to feed the central black hole at the Eddington rate for the entire time step. The duration of accretion can be calculated through

$$t_{\text{acc}} = \ln \left( \frac{\Delta m_{\text{bh, cold}}}{m_{\text{bh}}} + 1 \right) \times \frac{\eta t_{\text{Edd}}}{\epsilon}. \quad (21)$$

When the quasar is observed at a random time  $t_{\text{obs}}$ , the bolometric luminosity is

$$L_{\text{bol}} \equiv \epsilon m_{\text{bh}}|_{t=t_{\text{obs}}} \frac{c^2}{t_{\text{Edd}}} = \epsilon m_{\text{bh}} \exp\left(\frac{\epsilon t_{\text{obs}}}{\eta t_{\text{Edd}}}\right) \times \frac{c^2}{t_{\text{Edd}}}, \quad (22)$$

when  $t_{\text{obs}} \leq t_{\text{acc}}$ , otherwise  $L_{\text{bol}} = 0$ .

(ii)  $\Delta m_{\text{bh, max}} \geq m_{\text{Edd}}$ : In this case, the merger event delivers sufficient cold gas into the accretion disc. In the case of  $\Delta m_{\text{bh, max}} > m_{\text{Edd}}$ , instead of consuming this instantaneously, causing a super-Eddington accretion event, some of the mass is accreted by the central black hole, limited by the Eddington rate, while the rest,  $\Delta m'_{\text{bh, max}} = \Delta m_{\text{bh, max}} - m_{\text{Edd}}$ , is stored in the accretion disc to be consumed in the next time step. Similarly, when this quasar is

<sup>8</sup> Ignoring the loss due to gravitational wave emission.

**Table 1.** A list of relevant parameters in the model with description and adopted value. The redshift varying  $f_{\text{esc},*}$  model (Mutch et al. 2016, hereafter M16) is shown for comparison. The M16BH model includes black hole feedback compared to the M16 model but without reionization from quasars. We refer the interested reader to M16.

Parameter	Section	Equation	Description	Fiducial	M16BH	M16
$\alpha_{\text{sf}}$	2.1	6	Star formation efficiency	0.08	0.03	0.03
$\alpha_{\text{energy}}$	2.1	7	Energy coupling efficiency normalization	1.0	0.5	0.5
$\alpha_{\text{mass}}$	2.1	9	Mass loading normalization	15.0	9.0	9.0
$\eta$	2.2.1	15	Black hole efficiency of converting mass to energy	0.06	0.06	-
$\epsilon$	2.2.1	16	Eddington ratio	1.0	1.0	-
$k_{\text{h}}$	2.2.1	14	Black hole growth efficiency for the radio mode	0.3	0.1	-
$\kappa_{\text{r}}$	2.2.2	16	Black hole feedback efficiency for the radio mode	1.0	1.0	-
$k_{\text{c}}$	2.2.2	19	Black hole growth efficiency for the quasar mode	0.05	0.05	-
$\kappa_{\text{q}}$	2.2.2	-	Black hole feedback efficiency for the quasar mode	0.0005	0.0005	-
$N_{\gamma,*}$	4.1	28	Mean number of ionizing photons produced per stellar baryon	4,000	4,000	4,000
$f_{\text{esc},\text{q}}$	4.1	28	Ionizing photon escape fraction for quasars	1.0 <sup>a</sup>	0.0	-
$f_{\text{esc},*}$	4.1	28	Ionizing photon escape fraction for the stellar component	$f_{\text{esc},*,z}{}^{b,c}$	$f_{\text{esc},*,\text{M16}}{}^d$	$f_{\text{esc},*,\text{M16}}$

<sup>a</sup>  $f_{\text{esc},\text{q}} = 0$  is adopted for the StellarReion model in which stars are the only reionization source.

<sup>b</sup>  $f_{\text{esc},*} = 0$  is adopted for the QuasarReion model in which quasars are the only reionization source.

<sup>c</sup>  $f_{\text{esc},*,z} = \min[0.06 \times (\frac{1+z}{6})^{0.5}, 1.0]$  for the fiducial and StellarReion models.

<sup>d</sup>  $f_{\text{esc},*,\text{M16}} = \min[0.04 \times (\frac{1+z}{6})^{2.5}, 1.0]$ .

observed at  $t_{\text{obs}}$ , the bolometric luminosity can be calculated using equation (22).

It is suggested that during mergers, black holes undergo rapid accretion for a certain time period, which is followed by a long quiescent phase (Hopkins et al. 2005a,b,c,d). The assumption that black holes are either accreting at the Eddington rate ( $\epsilon = 1$ , see equation 16) or stay quiescent has been shown to provide a good description of black hole growth for the majority of black holes at high redshift (Bonoli et al. 2009).

The energy injected into galactic gas during the quasar mode is given by  $\kappa_{\text{q}}\eta\Delta m_{\text{bh,cold}}c^2$ , where  $\kappa_{\text{q}}$  represents the mass coupling factor in the quasar mode. Unlike the radio mode, this energy generates a wind, which heats the gas in the cold disc into the hot reservoir. Depending on the amount of energy provided by the quasar, the wind can further unbind and eject the hot gas in a manner similar to the stellar feedback prescription presented in Section 2.1.

### 2.3 Quasar luminosity

In order to compare the predicted black hole population in our model with observations, the intrinsic  $B$  band and UV 1450 Å band luminosities of quasars are calculated as follows:

- (i) We calculate the bolometric magnitude through

$$M_{\text{bol}} = 4.74 - 2.5 \log_{10} \left( \frac{L_{\text{bol}}}{L_{\odot}} \right). \quad (23)$$

- (ii) We calculate the  $B$ -band magnitude in the Vega magnitude system using the bolometric correction proposed by Hopkins, Richards & Hernquist (2007),

$$M_{\text{bol}} - M_{\text{B}} = -2.5 \log_{10} k_{\text{B}}, \quad (24)$$

where

$$k_{\text{B}} \equiv \frac{L_{\text{bol}}}{L_{\text{B}}} = 6.25 \left( \frac{L_{\text{bol}}}{10^{10} L_{\odot}} \right)^{-0.37} + 9.00 \left( \frac{L_{\text{bol}}}{10^{10} L_{\odot}} \right)^{-0.012}. \quad (25)$$

- (iii) We convert the  $B$ -band magnitude from the Vega system to the AB system following Glikman et al. (2010)

$$M_{\text{AB,B}} - M_{\text{B}} = -0.09. \quad (26)$$

- (iv) We extrapolate the  $B$ -band magnitude of which the effective wavelength is 4344 Å (Blanton & Roweis 2007) to the 1450 Å magnitude, assuming the quasar continuum between 1450 and 4344 Å has a power-law slope of  $\alpha_{q,\text{optical}} = 0.44$  relative to wavelength (Schirber & Bullock 2003). Thus,

$$M_{1450} = M_{\text{AB,B}} - 2.5 \log_{10} \left( \frac{1450 \text{ \AA}}{4344 \text{ \AA}} \right)^{\alpha_{q,\text{optical}}} = M_{\text{AB,B}} + 0.524. \quad (27)$$

We will further discuss the quasar luminosity function in Section 3.3.

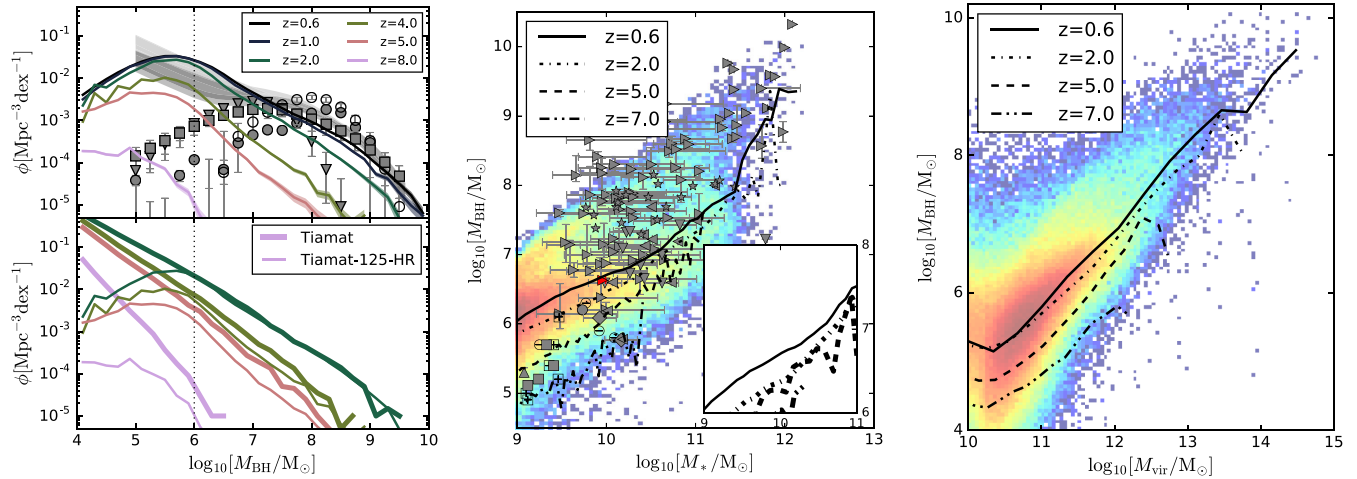
## 3 GALAXY AND BLACK HOLE PROPERTIES

We summarize the relevant model parameters in Table 1 compared to the original value adopted in M16. All other parameters remain the same as M16. In this work, we constrain<sup>9</sup> our model against:

- (i) the observed evolution of the galaxy stellar mass function (Pozzetti et al. 2007; Drory et al. 2009; Marchesini et al. 2009; González et al. 2011; Mortlock et al. 2011; Santini et al. 2012; Ilbert et al. 2013; Muzzin et al. 2013; Duncan et al. 2014; Tomczak et al. 2014; Grazian et al. 2015; Song et al. 2016; Huertas-Company et al. 2016; Stefanon et al. 2017; Davidzon et al. 2017) between  $z \sim 0.6$  and 7. We note that the observed stellar mass functions, based on a diet Salpeter IMF, a Chabrier (2003) IMF or a Kroupa (2001) IMF, were all converted into a standard Salpeter IMF by adding  $-0.15$ ,  $0.22$  or  $0.18$  dex, respectively, to the logarithm of the stellar masses;

- (ii) the observed black hole mass function (Graham et al. 2007; Shankar, Weinberg & Miralda-Escudé 2009; Vika et al. 2009; Davis et al. 2014; Mutlu-Pakdil, Seigar & Davis 2016) and Magorrian

<sup>9</sup> Calibration using the MCMC technique is ongoing. In this work, the calibration is performed by hand.



**Figure 1.** *Left-hand panel:* black hole mass functions from  $z \sim 8-0.6$  in the fiducial model. The results calculated using the *Tiamat* and *Tiamat-125-HR* trees are shown with thick and thin lines, respectively. The shaded regions represent the  $1\sigma$  Poisson uncertainties for the *Tiamat-125-HR* result. Observational data are shown with grey colours:  $\bullet$  Graham et al. (2007),  $\circ$  Vika et al. (2009),  $\nabla$  Davis et al. (2014),  $\blacksquare$  (Mutlu-Pakdil et al. 2016) and (Shankar et al. 2009) at  $z \sim 0$  and  $\blacksquare$   $z \sim 0.5$ . The vertical dotted line represents the resolution limit in *Tiamat-125-HR*. *Middle panel:* correlation between black hole mass and stellar mass. The 2D histogram shows the distribution of galaxies in the fiducial model using the *Tiamat-125-HR* halo merger trees at  $z \sim 0.6$ , while the solid line represents the mean. The Magorrian relations from the model at  $z = 2, 5$  and  $7$  are shown with dash-dotted, dashed and dash-dot-dotted lines, respectively. The results calculated using the *Tiamat* trees are shown with thick lines in the bottom right subplot for comparison with the  $z \sim 0.6$  *Tiamat-125-HR* Magorrian relation. Observations of the local Universe are indicated with different symbols:  $\blacktriangle$  Thornton et al. (2008),  $\circ$  Jiang et al. (2011),  $\blacktriangledown$  Mathur et al. (2012),  $\bullet$  Jiang et al. (2013),  $\blacktriangleright$  Scott et al. (2013),  $\square$  Reines et al. (2013) with BPT AGNs,  $\blacksquare$  Reines et al. (2013) with BPT composites,  $\blacklozenge$  Yuan et al. (2014),  $\blacktriangleleft$  Busch et al. (2014),  $\star$  Sanghvi et al. (2014) at  $z \sim 0.5-1$ , and  $\blacktriangleright$  representing the Milky Way Scott et al. (2013). *Right-hand panel:* correlation between the black hole mass and the virial mass from the fiducial model using the *Tiamat-125-HR* trees. The 2D histogram shows the distribution of galaxies at  $z \sim 0.6$ , while the solid line represents the mean. The scaling relations from the model at  $z = 2, 5$  and  $7$  are shown with dash-dotted, dashed and dash-dot-dotted lines, respectively. [A colour version of this figure is available in the online version.]

relation (Thornton et al. 2008; Jiang, Greene & Ho 2011; Mathur et al. 2012; Jiang et al. 2013; Reines, Greene & Geha 2013; Scott, Graham & Schombert 2013; Busch et al. 2014; Sanghvi et al. 2014; Yuan et al. 2014) at low redshift ( $z \lesssim 0.5$ );

(iii) the latest integrated free electron Thomson scattering optical depth measurement (Planck Collaboration XLVI 2016);

(iv) the predicted global ionizing emissivity from Ly  $\alpha$  opacities (Becker & Bolton 2013).

### 3.1 Black hole properties

The black hole mass functions at  $z \sim 8.0-0.6$  are shown with different colours in the left-hand panel of Fig. 1. The results calculated using the *Tiamat* and *Tiamat-125-HR* trees are shown with thick and thin lines, respectively. The shaded regions represent the  $1\sigma$  Poisson uncertainties.<sup>10</sup> Estimates of the local black hole mass function are shown with points and grey shaded regions. We see that the discrepancy between various observations is substantial. This is a result of inconsistent correlations between black hole mass and observable quantities, such as the Sérsic indices, bulge velocity dispersion or luminosity and galaxy geometry, and from the intrinsic scatter of these adopted scaling relations. Extrapolations of these observed scaling relations have impacts on the black hole mass function at the high-mass end, while different treatments of the spiral galaxy bulge can significantly change the low-mass end (Shankar et al. 2009). In this work, the model is therefore calibrated against the black hole mass function between  $10^{7.5} M_{\odot}$  and  $10^9 M_{\odot}$ . In the bottom

left-hand panel of Fig. 1, we see that the mass function converges at lower redshifts above a black hole mass of  $10^6 M_{\odot}$  (shown as the vertical dotted line). The different mass resolutions of *Tiamat* and *Tiamat-125-HR* result in different merger rates, especially when approaching the resolution limit. At high redshift, because the growth of black hole is dominated by the merger triggered quasar mode, the number density of small black holes is relatively lower in the *Tiamat-125-HR* result (e.g. comparing the  $z = 8.0$  thick and thin lines). At low redshift,  $z \sim 0.6$ , the model is in agreement with the observational estimations.

The middle panel of Fig. 1 shows the relation between black hole mass and stellar mass (the Magorrian relation<sup>11</sup>, Magorrian et al. 1998). The 2D histogram indicates the distribution of galaxies<sup>12</sup> in logarithm from the fiducial model using the *Tiamat-125-HR* halo merger trees at  $z \sim 0.6$ , while the solid line represents the mean. The Magorrian relations at  $z = 2, 5$  and  $7$  from *Tiamat-125-HR* are also shown with dash-dotted, dashed and dash-dot-dotted lines, respectively. The right bottom subplot shows the  $z = 2, 5$  and  $7$  Magorrian relations of the *Tiamat* result

<sup>11</sup> A black hole mass – galaxy property scaling relation usually accounts for the bulge property. However, the majority of systems with black holes are expected to be bulge dominated. Therefore, using the total stellar mass as a proxy does not lead to a significant bias in Fig. 1. We also note that recent studies suggest that while the  $M_{\text{bh}}-M_{\star}$  relation is significantly biased (see Section 5), it is the bulge velocity dispersion that connects supermassive black holes and host galaxies. However, interpreting the scaling relation is beyond the scope of this work. We leave a detailed analysis of the black hole–galaxy scaling relation to future work when bulge properties are included (e.g. Tonini et al. 2016).

<sup>12</sup> We exclude the recently identified haloes in order to minimize the effect of black hole seeding. However, this does not have a significant impact in Fig 1.

with thick lines compared with the  $z \sim 0.6$  *Tiamat-125-HR* Magorrian relation. Observations from the local Universe are indicated with different symbols (Thornton et al. 2008; Jiang et al. 2011; Mathur et al. 2012; Jiang et al. 2013; Reines et al. 2013; Scott et al. 2013; Busch et al. 2014; Sanghvi et al. 2014; Yuan et al. 2014). We see that the model predicts a similar Magorrian relation at  $z \sim 0.6$  compared to the local observations and we find an increasing normalization towards lower redshifts in the mass range of  $10^{10} M_{\odot} < M_* < 10^{12} M_{\odot}$ .

The evolution of the Magorrian relation in our model is due to the black hole and stellar mass evolving with the underlying dark matter halo mass differently. In M16, we have shown that the median relation between stellar mass and virial mass does not evolve in our model and can be described by  $M_* \propto M_{\text{vir}}^{7/5}$  in the range of  $10^8 M_{\odot} < M_* < 10^{11} M_{\odot}$ , which is supported by a simple analytic model of supernova energy conversation (Wyithe & Loeb 2003). On the other hand, haloes with a given virial mass host less massive black holes at earlier times in our model (see the right-hand panel of Fig. 1). These result in an increasing normalization of the Magorrian relation towards lower redshifts.

From the right-hand panel of Fig. 1, we see that the  $M_{\text{bh}}-M_{\text{vir}}$  scaling relation does not get suppressed in massive haloes (at least to  $M_{\text{vir}} \sim 10^{14} M_{\odot}$  or  $M_{\text{bh}} \sim 10^9 M_{\odot}$ ). Both black holes and stars grow from the cold gas disc.<sup>13</sup> However, AGN feedback significantly suppresses the cooling flow in massive galaxies (Croton et al. 2006), preventing stellar mass from growing. On the other hand, black holes are able to continue accreting until there is enough energy in feedback to overcome the halo potential and unbind the gas (Booth & Schaye 2010). Because of these, the slope becomes steeper in the Magorrian relation at  $M_{\text{bh}} > 10^8 M_{\odot}$ .

### 3.2 Galaxy properties

Fig. 2 presents the galaxy stellar mass functions from our fiducial model for comparison with the available observational data (Pozzetti et al. 2007; Drory et al. 2009; Marchesini et al. 2009; González et al. 2011; Mortlock et al. 2011; Santini et al. 2012; Ilbert et al. 2013; Muzzin et al. 2013; Duncan et al. 2014; Tomczak et al. 2014; Grazian et al. 2015; Song et al. 2016; Huertas-Company et al. 2016; Stefanon et al. 2017; Davidzon et al. 2017) at redshifts 7 to  $\sim 0.6$ . The results calculated using the *Tiamat* and *Tiamat-125-HR* trees are shown with thick and thin lines, respectively. The shaded regions represent the  $1\sigma$  Poisson uncertainties. We see that the fiducial model is able to reproduce the observed galaxy stellar mass function across the redshift range of  $z = 7-0.6$ .

We also present two additional models in Fig. 2, M16 and M16BH. M16 adopts identical parameters as the redshift varying  $f_{\text{esc},*}$  model ( $f_{\text{esc},*,\text{M16}} = \min \left[ 0.04 \times \left( \frac{1+z}{6} \right)^{2.5}, 1.0 \right]$ , see Table 1) presented in M16. This model is able to reproduce the evolution of the stellar mass function at high redshift ( $z > 5$ ). The redshift dependence of  $f_{\text{esc},*,\text{M16}}$  was chosen to simultaneously reproduce the normalization and flat slope of the McQuinn, Peng Oh & Faucher-Giguère (2011) emissivity measurement at  $z \sim 5$  and the Planck 2015 optical depth measurement (Planck Collaboration XIII 2016, hereafter Planck15). Details of the M16 model at high redshift ( $z > 5$ ) can be found in M16. In this work, we extend the model<sup>14</sup>

<sup>13</sup> Radio mode accretion and merger-driven starburst are relatively less important to the growth.

<sup>14</sup> We note that the *Tiamat* halo merger trees have been improved (Poole et al. in preparation) since M16. However, the impact on galaxy formation is trivial.

to lower redshifts and find that, without AGN feedback regulating galaxy formation, the model fails to reproduce the observed stellar mass functions at  $z < 2$ , especially in larger mass ranges ( $M_* > 10^{11} M_{\odot}$ ). However, when AGN feedback is implemented, shown as the M16BH model, the model shows better agreement with observations (Croton et al. 2006). These two models also suggest that radio-mode feedback does not play a significant role in galaxy formation during the EoR, and because reionization is dominated by low-mass galaxies (Liu et al. 2016, see also Section 5), AGN feedback is expected to have no significant impact on reionization.<sup>15</sup>

With respect to M16, the fiducial model presented in this work employs a stronger star formation efficiency ( $\alpha_{\text{sf}}$ ) with maximized supernova feedback ( $\alpha_{\text{energy}}$  and  $\alpha_{\text{mass}}$ ) and more intense radio mode feedback ( $k_{\text{h}}$ ), in order to gain better agreement with the observed stellar mass function in the intermediate mass range ( $10^9 M_{\odot} < M_* < 10^{11} M_{\odot}$ ) at  $1 < z < 2$ .

### 3.3 Quasar luminosity function

In the previous two subsections, we presented the predicted properties of the central massive black holes and host galaxies, together with their correlations. These show that our model is able to produce the evolution of the galaxy stellar mass function over a large time-scale ( $z \sim 7-0.6$ ), as well as the observed black hole properties at low redshift. Before we start exploring the contribution of quasars to reionization, we discuss the quasar luminosity function in this section.

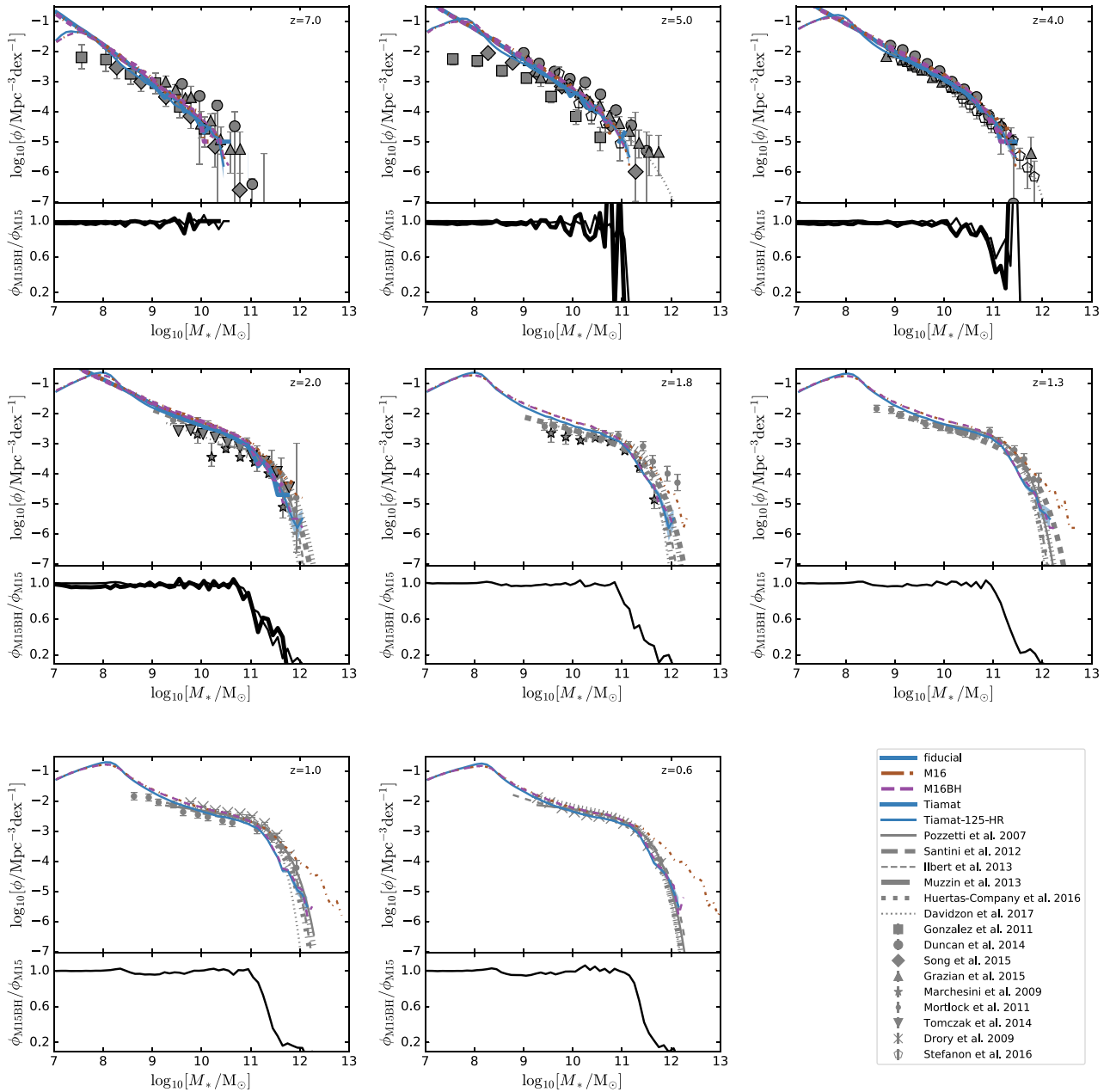
Semi-analytic models can predict quasar bolometric luminosity through the modelled central black hole mass. In order to compare with observations, three corrections are usually required: bolometric corrections, duty cycles and obscured fractions. In our model, we calculate the duty cycle self-consistently by assuming Eddington accretion of available gas and a random observation time between snapshots (see Section 2.2.2). Black holes that are inactive when observed are not considered in the census. We convert the total luminosity into a particular band using the bolometric correction of Hopkins et al. (2007) (see Section 2.3). Finally, we need to take the obscuration due to the presence of a dusty torus surrounding the AGN into account. In practice, each quasar is weighted by  $1 - \cos \frac{\theta}{2}$ , where  $\theta$  represents the opening angle of AGN radiation, during the calculation of the quasar luminosity function. We note that the dependence of  $\theta$  or the obscured fraction (the ratio of obscured to unobscured AGN) is very complicated. It has been suggested that  $\theta$  may depend on wavelength, redshift and luminosity (Elvis et al. 1994; Hopkins et al. 2007). For simplicity, we only consider constant  $\theta$ . The opening angle is a free parameter in our model, chosen to reproduce the quasar luminosity function amplitude. In this work,  $\theta = 80$  deg is adopted. This corresponds<sup>16</sup> to a fraction of visible objects,<sup>17</sup>  $f_{\text{obs}}$ , to be  $\sim 23.4$  per cent, in agreement with Hopkins et al. (2007), who suggests a luminosity-dependent observable fraction,  $f_{\text{obs}} = 0.26(L_{\text{bol}}/10^{12.4} L_{\odot})^{0.082}$  for the *B* band.

We present the UV 1450 Å luminosity functions of quasars at redshift 6, 5, 4 and 3, and the *B*-band luminosity functions at  $z \sim 2$ ,

<sup>15</sup> Ignoring the impact on the ionizing photon escape fraction from AGN feedback.

<sup>16</sup> An opening angle of  $\theta = 80$  deg corresponds to a solid angle of  $\Omega = 2\pi(1 - \cos \frac{\theta}{2}) \sim 1.47$ . Considering a symmetric radiation from both sides of the accretion disc, the un-obscured fraction is  $f_{\text{obs}} \equiv \frac{2\Omega}{4\pi} \sim 0.234$ .

<sup>17</sup> The opening angle is used to illustrate the observable fraction, assuming an orientation model. We note that the observable fraction can also be interpreted by the line-of-sight absorption column density when the evolutionary model is assumed.

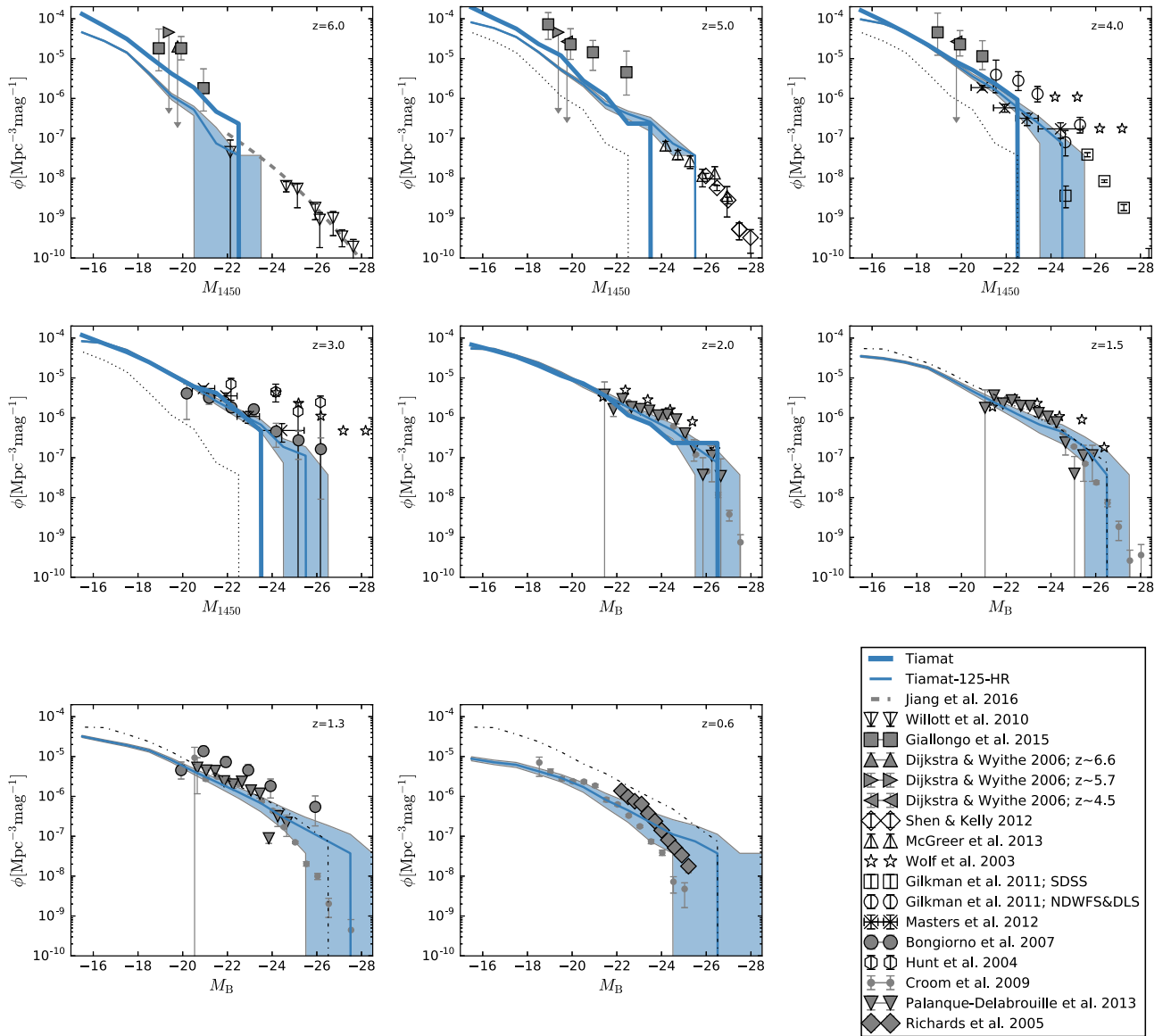


**Figure 2.** Galaxy stellar mass functions at  $z \sim 7$ –0.6 from the fiducial (—), M16 (—) and M16BH (—) models compared to the observational data (— Pozzetti et al. 2007,  $\times$  Drory et al. 2009,  $\star$  Marchesini et al. 2009,  $\bullet$  Mortlock et al. 2011,  $\blacksquare$  González et al. 2011,  $\blacksquare\blacksquare$  Santini et al. 2012,  $-\cdot-$  Ilbert et al. 2013,  $\blacksquare\blacksquare\blacksquare$  Muzzin et al. 2013,  $\blacktriangledown$  Tomczak et al. 2014,  $\bullet$  Duncan et al. 2014,  $\blacklozenge$  Song et al. 2016,  $\blacktriangle$  Grazian et al. 2015,  $\blacksquare\blacksquare\blacksquare\blacksquare$  Huertas-Company et al. 2016,  $\circ$  Stefanon et al. 2017,  $\dots$  Davidzon et al. 2017). The ratios of M16BH to the M16 result are shown in the bottom subpanels. The results calculated using the *Tiamat* and *Tiamat-125-HR* trees are shown with thick and thin lines, respectively. The shaded regions represent the  $1\sigma$  Poisson uncertainties for the *Tiamat-125-HR* result. [A colour version of this figure is available in the online version.]

1.5, 1.3 and 0.6 in Fig. 3, compared to the observational data from Wolf et al. (2003), Hunt et al. (2004), Richards et al. (2005), Dijkstra et al. (2006), Bongiorno et al. (2007), Croom et al. (2009), Willott et al. (2010), Glikman et al. (2011), Shen & Kelly (2012), Masters et al. (2012), McGreer et al. (2013), Palanque-Delabrouille et al. (2013), Giallongo et al. (2015) and Jiang et al. (2016). The results calculated using the *Tiamat* and *Tiamat-125-HR* trees are shown with thick and thin lines, respectively. The shaded regions represent the 95 per cent confidence intervals around the mean using 100 000 bootstrap re-samples (due to the random number  $t_{\text{obs}}$

in equation 22). Similarly to the black hole mass functions, the *Tiamat-125-HR* results show a lower number density at high redshift compared to the *Tiamat* results, and they converge at lower redshifts ( $z \lesssim 4$ ). We see that the model shows good agreement with observations across a large redshift range ( $z \sim 6$ –0.6). At high redshift, the model is consistent with the samples of bright quasars (Willott et al. 2010; Shen & Kelly 2012; McGreer et al. 2013; Jiang et al. 2016), while it predicts a lower number density of faint quasars compared to the Giallongo et al. (2015, hereafter G15) sample. On the other hand, the model produces a significant number of faint





**Figure 3.** UV 1450 Å luminosity functions of quasars at  $z \sim 6, 5, 4$  and 3, and  $B$ -band luminosity functions at  $z \sim 2.0, 1.5, 1.3$  and 0.6. The results using the *Tiamat* and *Tiamat-125-HR* halo merger trees are shown with blue thick and thin lines, respectively. The shaded regions represent the 95 per cent confidence intervals around the mean using 100 000 bootstrap re-samples for the *Tiamat-125-HR* result. In the panels showing UV luminosity functions at  $z \sim 5.0, 4.0$  and 3.0, the  $z = 6.0$  UV luminosity function of *Tiamat-125-HR* is indicated with black thin dotted lines for comparison. The  $z = 2.0$   $B$ -band luminosity function of *Tiamat-125-HR* is shown with black thin dash-dotted at  $z \sim 1.5, 1.3$  and 0.6. The observational data are shown with different symbols: Wolf et al. (2003) ( $\star$ ); Hunt et al. (2004) ( $\odot$ ); Richards et al. (2005) using the 2dF-SDSS LRG and QSO survey ( $\blacklozenge$ ); Dijkstra, Stuart & Wyithe (2006) using Ly  $\alpha$  constraints and providing upper limits at  $z \sim 4.5$  ( $\blacktriangleleft$ ),  $z \sim 5.7$  ( $\blacktriangleright$ ) and  $z \sim 6.6$  ( $\blacktriangle$ ); Bongiorno et al. (2007) using VVDS ( $\bullet$ ); Croom et al. (2009) using the 2dF-SDSS LRG and QSO survey ( $\circ$ ); Willott et al. (2010) using the CFHQS data ( $\nabla$ ); Glikman et al. (2011) at  $z \sim 4$  using the SDSS data ( $\square$ ) and using the NOAO Deep Wide-Field Survey and the Deep Lens Survey ( $\circ$ ); Shen & Kelly (2012) using SDSS DR7 data at  $z \sim 4.75$  ( $\diamond$ ); Masters et al. (2012) using COSMOS ( $\times$ ); McGreer et al. (2013) using SDSS, UKIDSS and MMT at  $z \sim 4.7$ –5.1 ( $\triangle$ ); Palanque-Delabrouille et al. (2013) with SDSS-III and MMT data; Giallongo et al. (2015) using *Chandra*, *HST*, *Spitzer* and various ground-based telescopes ( $\blacksquare$ ); Jiang et al. (2016) with SDSS ( $\blacksquare$ ). [A colour version of this figure is available in the online version.]

quasars. The turnover (not shown here) of the predicted UV luminosity function is around  $M_{1450} \sim -11$ , which is much fainter than the observed faintest quasars,  $M_{1450} \sim -18$  (G15). In addition, the observed population of bright quasars ( $M_{1450} < -23$ ) at high redshift is not present in our model due to our limited simulation volume. All of these factors may have an impact on the contribution of AGN to the reionization history, which is discussed in Section 4.

### 3.4 The luminosity function of galaxies with AGN

The large number of faint AGN identified by G15 has prompted renewed discussion of the contribution of quasars to reionization (Madau & Haardt 2015; Mitra et al. 2015; Kulkarni et al. 2017). We therefore further discuss the low number density in the faint end predicted by the model compared with the G15 data.

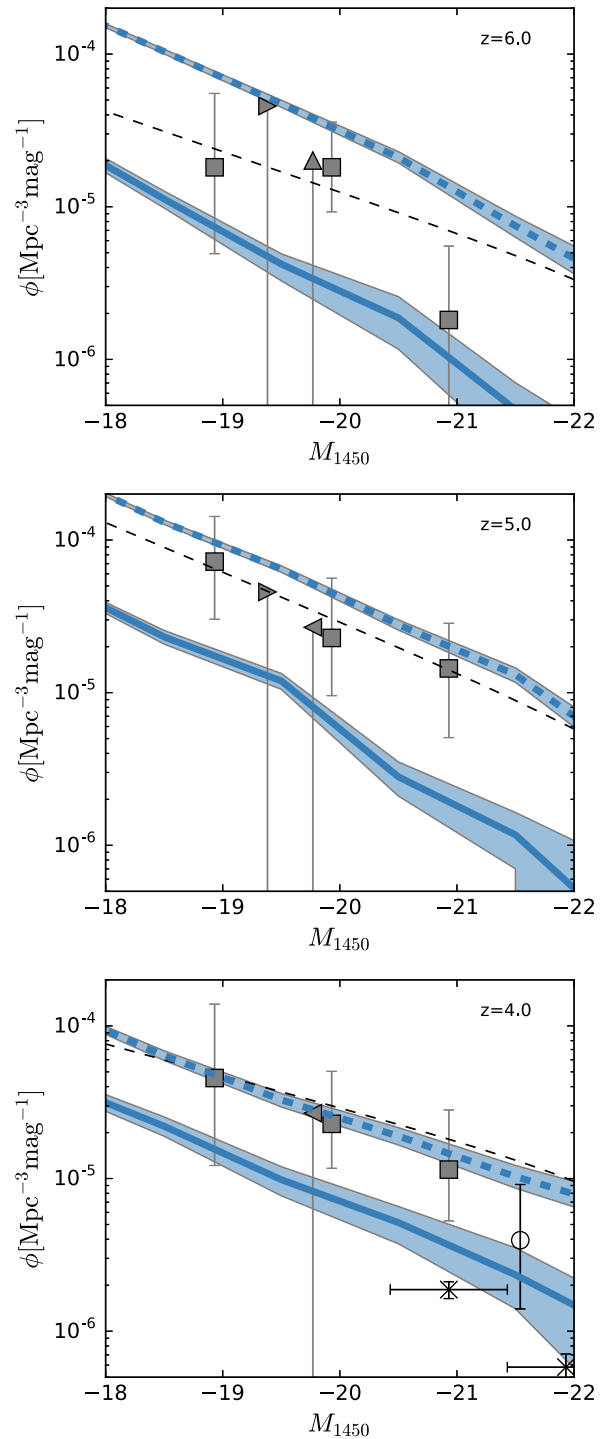
When constructing an observed AGN luminosity function, the sample is typically identified spectroscopically or via colour-colour

selections if spectra are not available, following which contamination from host galaxies is removed. For example, some observers model the surface brightness distributions of host galaxies and fit galaxies (e.g. Sérsic profiles) and point sources (using point-spread functions, Dunlop et al. 2003; Peng et al. 2006; Du et al. 2014; Martínez-Paredes et al. 2017) to the images.<sup>18</sup> Others examined the SED using a combination of AGN and galaxy emission with possible extinction of the AGN flux when spectroscopic data are available (Bongiorno et al. 2007; Croom et al. 2009; Masters et al. 2012; Mechtley et al. 2012; Lyu, Rieke & Alberts 2016). In most cases, ignoring the contribution from host galaxies does not make a significant difference to bright quasars (Hopkins et al. 2007). Therefore, some AGN samples do not exclude stellar light (Wolf et al. 2003, Hunt et al. 2004, Richards et al. 2005, Willott et al. 2010, Shen & Kelly 2012, McGreer et al. 2013, Palanque-Delabrouille et al. 2013; G15). However, this may not be the case for faint AGN. In particular, the G15 AGN sample is selected using X-ray activity and no AGN-galaxy separation is possible. Thus, the total UV luminosity may have a large fraction of stellar light, suggesting that the G15 sample may be potentially impacted by stellar light contamination. Despite the recent claim that only 12 of the 22 reported X-ray detections in G15 are high-redshift AGN (Parsa et al. 2017), this conjecture is supported by Ricci et al. (2017), who used X-ray observations as a proxy and derived the quasar UV luminosity function down to much fainter ranges. They showed that the luminosity function is in agreement with UV/optical observations (e.g. Croom et al. 2009; Glikman et al. 2010; Masters et al. 2012; Palanque-Delabrouille et al. 2013), and have much lower amplitudes than the G15 results, and that the high number density of faint AGN in G15 can be explained by the contribution from the luminosity of host galaxy with  $M_{1450} \sim -20$ .

In order to account for this, we calculate the galaxy UV luminosity by integrating model SEDs based on the modelled star formation history (Liu et al. 2016) and add the stellar light to mimic the observed total UV luminosity<sup>19</sup> (AGN+galaxy). We present the resulting luminosity function of faint objects ( $M_{1450} \sim -18$  to  $-22$ ) in Fig. 4 with dashed lines. The result is calculated using the *Tiamat* halo merger trees and is shown with shaded regions representing the 95 per cent confidence intervals around the mean using 100 000 bootstrap re-samples. The luminosity function calculated using only the AGN light (AGN only, as shown in Fig. 3) is shown as solid lines for comparison. We see that including stellar light can significantly increase the number density of faint AGN inferred at high redshift, by up to  $\sim 1$  order of magnitude. The fitting functions provided by G15 as shown with thin black dashed lines in Fig. 4 are more consistent with the AGN+galaxy luminosity function. If this is the case, the estimated emissivity at high redshift based on G15 is likely overestimated.

#### 4 REIONIZATION FROM QUASARS

Motivated by the G15 sample, which suggests a numerous population of faint AGN at  $z = 4-6$  (G15), Madau & Haardt (2015) extrapolated the emissivity calculated from G15 to higher redshifts and assessed a model of reionization, in which quasars are the dominant ionizing sources. They found that due to the high escape fraction of ionizing photons produced by those luminous objects, quasars are



**Figure 4.** UV 1450 Å luminosity functions of faint quasars at  $z \sim 6-4$  using the *Tiamat* halo merger trees. Solid lines represent the luminosity functions calculated through the light only from AGN (same as in Fig. 3), while dashed lines correspond to the calculation accounting for contributions from both AGN and their host galaxies. The shaded regions represent the 95 per cent confidence intervals around the mean using 100 000 bootstrap re-samples. Observations are shown with different symbols (see the caption of Fig. 3). In particular, the G15 sample is shown with squares and the fitting function provided by G15 is shown with thin black dashed lines. [A colour version of this figure is available in the online version.]

<sup>18</sup> Also with a constant to model the sky background.

<sup>19</sup> Dust attenuation is not considered because these relatively faint galaxies have little dust in our models (Liu et al. 2016).

able to ionize the neutral hydrogen by  $z \sim 5.7$  if the high emissivity of quasars derived at  $z \sim 5$  continues to higher redshifts. Later, Mitra et al. (2015) revisited the model with a revised extrapolation and also found that quasars have a significant role during the EoR. However, their analysis still prefers models with a non-zero escape fraction of  $\sim 12$  per cent from galaxies. It is worth noting that with different formalism, Manti et al. (2017) fit the observed quasar UV luminosity function at  $z = 0.5\text{--}6.5$ , including the G15 sample. They recalculated the emissivity by integrating the luminosity function and confirmed a large number of ionizing photons from quasars at high redshift using the Schechter luminosity function. However, the fitting result using a double power law presents a rapidly decreasing emissivity at  $z > 6$ , suggesting that the extrapolated high-redshift quasar emissivity is strongly dependent on the assumed shape of the quasar luminosity function. Taking advantage of the MERAXES semi-analytic model with 21cmFAST (Mesinger, Furlanetto & Cen 2011), we investigate the contribution of quasars to reionization, within a frame work that accounts for black hole growth and feedback on star formation.<sup>20</sup>

#### 4.1 Reionization model

The semi-numerical reionization code 21cmFAST (Mesinger et al. 2011) uses an excursion set formalism to identify H II bubbles in which the cumulative number of ionizing photons is more than the number of absorbing atoms:

$$N_* N_{\gamma,*} f_{\text{esc},*} + N_q N_{\gamma,q} f_{\text{obs}} f_{\text{esc},q} \geq (1 + \bar{N}_{\text{rec}}) N_{\text{H I}}, \quad (28)$$

where  $N_*$  and  $N_q$  are the numbers of baryons in stars and quasars<sup>21</sup>  $N_{\gamma,*} \sim 4000$  (Loeb & Barkana 2001) and  $N_{\gamma,q}$  (see the calculation in the Appendix) are the mean numbers of ionizing photons produced per baryon incorporated into the stellar or quasar components. The parameters  $f_{\text{esc},*}$  and  $f_{\text{esc},q}$  (see Table 1) are the escape fractions of ionizing photons produced by stars and quasars.  $f_{\text{obs}} \sim 0.234$  represents the observable fraction due to obscuration<sup>22</sup> (see Section 3.3).  $N_{\text{H I}}$  is the cumulative number of atoms being ionized and  $\bar{N}_{\text{rec}}$  is the mean number of recombinations per baryon. Inhomogeneous recombinations are ignored, which can have a large impact (Sobacchi & Mesinger 2014). In this work,  $\bar{N}_{\text{rec}}$  is set to be 0 as suggested by the high-redshift Ly  $\alpha$  forest in the IGM (Bolton & Haehnelt 2007; McQuinn et al. 2011). Expanding equation (28) gives

$$\frac{\xi_* m_* + \xi_q m_{\text{bh}}}{\frac{4}{3} \pi R^3 \Omega_m \rho_c(z)} \geq 1, \quad (29)$$

with

$$\xi_* = \frac{N_{\gamma,*} f_{\text{esc},*}}{f_b (1 - 0.75 Y_{\text{He}})} \quad \text{and} \quad \xi_q = \frac{N_{\gamma,q} f_{\text{obs}} f_{\text{esc},q}}{f_b (1 - 0.75 Y_{\text{He}})}, \quad (30)$$

where  $m_*$  is cumulative stellar mass that excludes the loss due to supernova<sup>23</sup> and  $R$  is the radius of the H II bubble.  $\xi_*$  and  $\xi_q$  are the H II ionizing efficiencies for stars and quasars,  $f_b = \frac{\Omega_b}{\Omega_m}$  and

<sup>20</sup> The *Tiamat-125-HR* halo merger trees cannot resolve small objects; therefore, we cannot consider ionizing photons from faint galaxies or quasars. We only show the result calculated using the *Tiamat* trees in the following sections.

<sup>21</sup> The mass of black hole seed is subtracted because they do not produce any ionizing photons. Reionization from the progenitor of black hole seeds will be considered in the future when Population III stars are implemented.

<sup>22</sup> One may also define  $f_{\text{esc},q}$  as the quasars escape fraction.

<sup>23</sup> The stellar mass recycled to the ISM through supernova also contributes ionizing photons in the H II bubble.

$Y_{\text{He}} = 0.24$  are the fraction of baryons in the Universe and the fraction of helium, and  $\rho_c(z)$  is the comoving critical density of the universe, respectively.

When the local volume around a galaxy is ionized, the UV background provides an extra heating mechanism, which modifies the baryonic fraction of the host halo (see  $\chi_r$  in equation 1). Following Sobacchi & Mesinger (2013), when the virial mass,  $M_{\text{vir}}$  is smaller than a filtering mass, which can be calculated through

$$M_{\text{filt}} = 2.8 \times 10^9 M_{\odot} J_{21}^{0.17} \left( \frac{1+z}{10} \right)^{-2.1} \left[ 1 - \left( \frac{1+z}{1+z_{\text{ion}}} \right)^2 \right]^{2.5}, \quad (31)$$

the suppression of gas becomes significant ( $\chi_r \equiv 2^{-M_{\text{filt}}/M_{\text{vir}}}$ ). Here,  $z_{\text{ion}}$  is the redshift when the local volume is first ionized, which is determined by the criteria given by equation (28).  $J_{21}$  represents the intensity of the local UV background. This can be calculated through

$$J_{21} = \frac{3(1+z)^2}{8\pi^2 R^3 m_p} \lambda_{\text{mfp}} h (\alpha_* f_{\text{esc},*} N_{\gamma,*} \dot{m}_* + \alpha_q f_{\text{obs}} f_{\text{esc},q} N_{\gamma,q} \dot{m}_q), \quad (32)$$

where  $h$  is the Planck constant.  $\lambda_{\text{mfp}}$  is the mean-free path of ionizing photons, which is approximated by the H II bubble radius,  $R$ . The parameters  $\alpha_* = 5.0$  (Loeb & Barkana 2001) and  $\alpha_q = 1.57$  (G15) are the spectral indexes for a stellar-driven and a quasar-driven spectra in the UV band.  $\dot{m}_*$  and  $\dot{m}_q$  are the growth rates of stellar mass and black hole mass, respectively.

#### 4.2 Reionization history

Fig. 5 shows the instantaneous ionizing emissivity averaged over the entire *Tiamat* simulation volume from different models as a function of redshift:

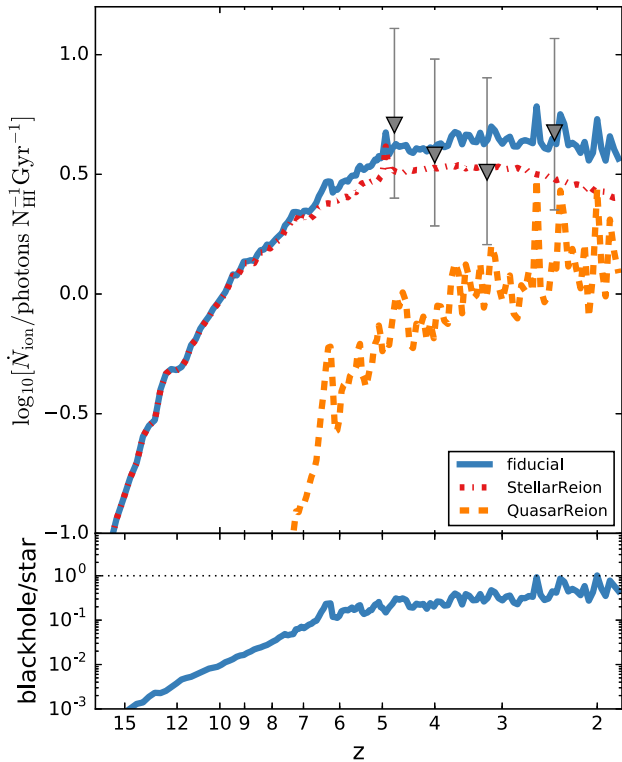
$$\dot{N}_{\text{ion},*} = \frac{N_{\gamma,*} f_{\text{esc},*} \dot{m}_{*,\text{tot}}}{f_b (1 - 0.75 Y_{\text{He}}) \rho_c V_{\text{tot}}}, \quad (33)$$

and

$$\dot{N}_{\text{ion},q} = \frac{N_{\gamma,q} f_{\text{obs}} f_{\text{esc},q} \dot{m}_{\text{bh,tot}}}{f_b (1 - 0.75 Y_{\text{He}}) \rho_c V_{\text{tot}}}, \quad (34)$$

where  $\dot{N}_{\text{ion},*}$ ,  $\dot{N}_{\text{ion},q}$ ,  $\dot{m}_{*,\text{tot}}$  and  $\dot{m}_{\text{bh,tot}}$  are the ionizing emissivities for the stellar component and quasars, and the total growth rates of stellar mass excluding the loss from supernovae and black holes in the model, respectively. From the Ly  $\alpha$  opacity, several measurements of the total emissivity (AGN and stars) at high redshift have been estimated (Bolton & Haehnelt 2007; McQuinn et al. 2011; Becker & Bolton 2013). There are relatively large uncertainties in these measurements. In this work, we compare our models using the most recent data<sup>24</sup> from Becker & Bolton (2013, hereafter BB13). The top panel of Fig. 5 shows that the fiducial model agrees with the BB13 data and the bottom shows the ratio between ionizing photons from black holes and stars in the fiducial model, suggesting that during the EoR, quasars are subdominant in our model. The evolutions of the mass-weighted global neutral hydrogen fraction and the integrated Thomson scattering optical depth  $\tau_e$  (see M16 for the calculation of  $\tau_e$ ) are shown in Fig. 6. It shows that the fiducial model has a reasonable reionization history, with the mean global neutral hydrogen fraction decreasing from 90 per cent at  $z \sim 10$  to 0 by  $z \sim 6$  and a Thomson scattering optical depth in agreement with

<sup>24</sup> The systematic error due to recombination radiation is ignored.

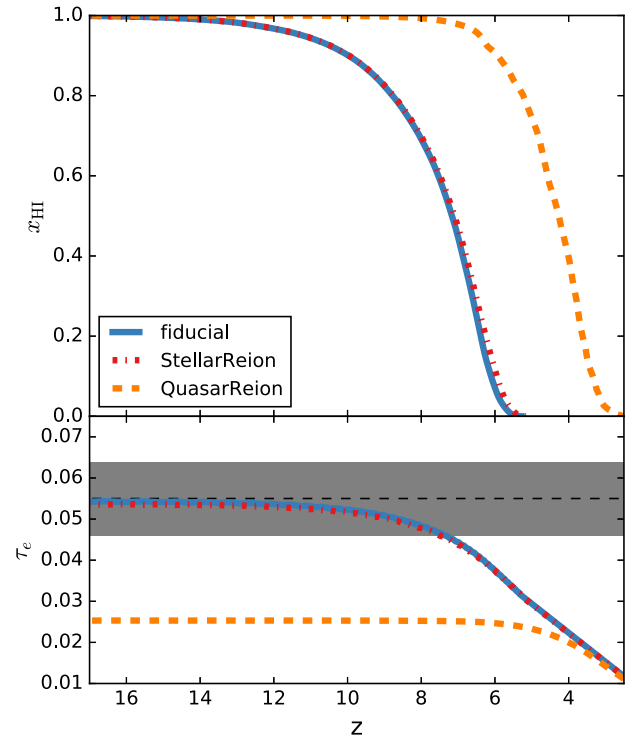


**Figure 5.** *Top panel:* the evolution of the total ionizing emissivity in units of number of ionizing photons per Hydrogen atom per Gyr for the fiducial model (—), the StellarReion model (⋯) in which stars are the only ionizing photon source, and the QuasarReion model (---) where quasars are the only reionization contributor. The total emissivity from BB13 are indicated with grey triangles ( $\nabla$ ). *Bottom panel:* the ratio of emissivities from black holes to the stellar component in the fiducial model. [A colour version of this figure is available in the online version.]

the latest Planck limits (Planck Collaboration XLVI 2016, hereafter Planck16).

There are two additional models in the top panel of Fig. 5 as well as in Fig. 6: StellarReion and QuasarReion. The ionizing source in the StellarReion model is only stars, with  $f_{\text{esc},*} = \min\left[0.06 \times \left(\frac{1+z}{6}\right)^{0.5}, 1.0\right]$  and  $f_{\text{esc},q} = 0$ , while quasars are the only reionization contributor in the QuasarReion model, with  $f_{\text{esc},q} = 1$  and  $f_{\text{esc},*} = 0$  (see Table 1). Note that only changing the feedback from reionization has little impact on the stellar mass function (see M16), the quasar luminosity function or the Magorrian relation. By preventing gas infall, reionization only affects less massive objects in our model.

In Fig. 5, we see that the emissivity of quasars grows rapidly in the QuasarReion model (also in the fiducial model) by a factor of 10 from  $z \sim 7$ –5. However, if quasars are the only reionization contributor, even with  $f_{\text{esc},q} = 1$  the number of ionizing photons cannot reach the BB13 data. Moreover, due to the deficiency in the photon budget at high redshift, quasars can only start reionization at  $z \sim 6$  resulting in an end at  $z \sim 3$ . Together with the predicted optical depth, our model rules out the quasar-only reionization scenario. We note that one may recalibrate the model with a more efficient black hole growth rate at high redshift (e.g. by incorporating a redshift dependence in the equation 19, see Bonoli et al. 2009), in order to match the G15 luminosity function and the estimated emissivity. Mitra et al. (2015) also suggest that if G15 emissivity is correct, quasar-only reionization is possible and it results in a small value of  $\tau_e$  due to the rapid evolution of the Lyman-limit



**Figure 6.** *Top panel:* the evolution of the mass weighted global neutral hydrogen fraction for the fiducial model (—), the StellarReion model (⋯) in which stars are the only ionizing photon source, and the QuasarReion model where quasars are the only reionization contributor (---). *Bottom panel:* the Thomson scattering optical depth as a function of redshift. The dashed line with shaded regions indicates the Planck 2016 measurement with  $1\sigma$  uncertainties (Planck16). [A colour version of this figure is available in the online version.]

systems. However, simultaneously matching the model with the G15 faint AGN luminosity function and the other observations of bright systems at high redshift is difficult. For example, comparing to observations at  $z \sim 6$ , our models produce a flatter luminosity function, which is more consistent with the bright quasar sample. This suggests that a mass-dependent black hole growth efficiency (e.g.  $k_c$ , see Section 2.2.2) would be required, in order to steepen the luminosity function and produce more small quasars. In the following sections, we explore the relative contribution of quasars to reionization based only on the presented black hole growth model.

Comparing the neutral hydrogen fraction and the optical depth between the fiducial and StellarReion models also suggests that quasars do not have a significant role in reionization in this model. Their contribution helps finish reionization earlier by  $\Delta z \lesssim 0.1$  and decreases the optical depth by less than 10 per cent.

## 5 DISCUSSION

Our models are calibrated against the black hole – galaxy scaling relation and quasar luminosity function, in order to reproduce a realistic AGN catalogue for the study of the contribution of quasars to hydrogen reionization. However, it has recently been suggested that the black hole sample used to derive scaling relations is likely different from the entire population, leading to a selection bias<sup>25</sup> (Bernardi et al. 2007). For instance, using Monte Carlo simulations

<sup>25</sup> If it is not a selection bias, it is possible that not every galaxy hosts a central massive black hole.

Shankar et al. (2016) recovered the intrinsic scaling relation assuming the selection bias comes from unresolved black holes (e.g. for galaxies with a given velocity dispersion,  $\sigma_*$ , it is more difficult to resolve smaller black holes; Batcheldor et al. 2010) and showed that such a bias can lead to factors of  $\geq 3$  and  $\sim 50$ – $100$  higher normalizations of the  $M_{\text{bh}}-\sigma_*$  and  $M_{\text{bh}}-M_*$  relations, respectively. We note that a Magorrian relation with a smaller normalization can be achieved in our model using a smaller black hole growth efficiency (i.e.  $k_c$  in equation 19). In this case, the reconstructed black hole population becomes less massive, more efficient AGN feedback (e.g.  $k_{\text{h}}$  in equation 14) and a larger fraction of observable AGN in the UV band are required to simultaneously reproduce the observed stellar mass function and quasar luminosity function. With the model calibrated against the quasar luminosity function, the black hole–galaxy scaling relation is coupled with the observable AGN fraction – a lower Magorrian relation requires a larger fraction of observable AGN. We note that the total emissivity of quasars is integrated from the luminosity function. Therefore, scaling relations do not have a significant impact on reionization in this work. Noting the difficulty of observationally determining the fraction of obscured AGN and the large uncertainties in the black hole–galaxy scaling relation, in this section we use a range of models which predict similar Magorrian relations as shown in Fig. 1 and explore the contribution of quasars to reionization.

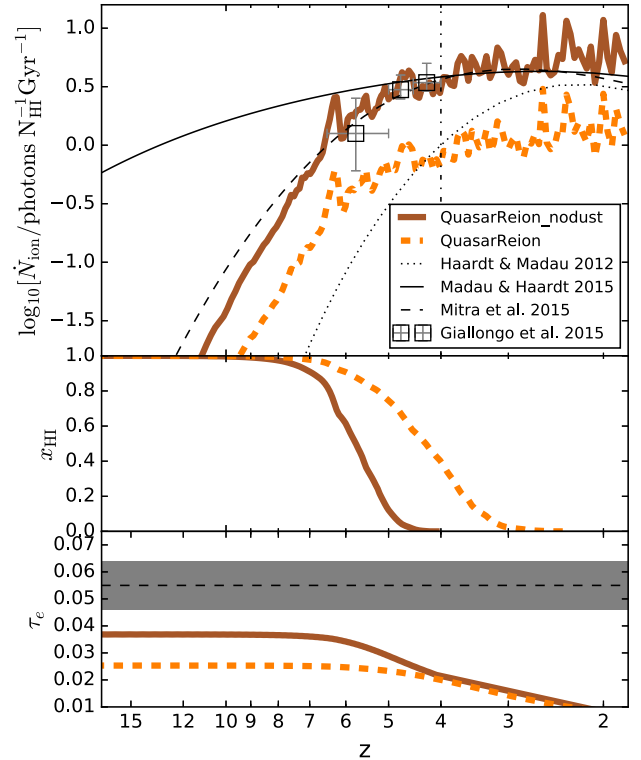
### 5.1 A larger opening angle

We have shown that with an opening angle of 80 deg, the model is able to reproduce the observed quasar luminosity function from  $z \sim 6$ – $0.6$  (see Fig. 3). However, at high redshift, the model predicts significant stellar contribution to UV flux in the G15 sample (see Fig. 4), and consequently less ionizing photons from AGN. Based on this, we find that quasars do not have a significant role during EoR.

In the top panel of Fig. 7, we show the estimated total emissivity from G15 with the model proposed by Haardt & Madau (2012); Madau & Haardt (2015) and Mitra et al. (2015). The modelled emissivity from QuasarReion is shown for comparison. Our quasar reionization-only model (QuasarReion) predicts lower emissivities compared to these two estimations, with only a third of the G15 value at  $z \sim 6$ . Consequently, in disagreement with Madau & Haardt (2015) and Mitra et al. (2015), we conclude that quasars cannot be the dominate sources during reionization. We could increase the emissivity by excluding the obscuration from dust (setting  $\theta = 180$ , shown as QuasarReion\_nodust in Fig. 7), which gives a closer emissivity compared to the G15 estimation and in agreement with the model proposed by Mitra et al. (2015). In this model, quasars have a more significant role during the EoR and can reionize the IGM alone by  $z \sim 4.5$ . However, this model overestimates the number density of bright- and low-redshift AGN, leading to an incorrect evolution of the quasar luminosity function. A lower fraction of observable AGN,  $f_{\text{obs}}$ , towards brighter luminosities and lower redshifts is required to solve this conflict. However, observations suggest the opposite trend in optical, infrared and X-ray bands (Hopkins et al. 2007) and more constraints are required to clearly establish a  $f_{\text{obs}}-z$  relation.

### 5.2 Escape fraction of ionizing photons from quasars

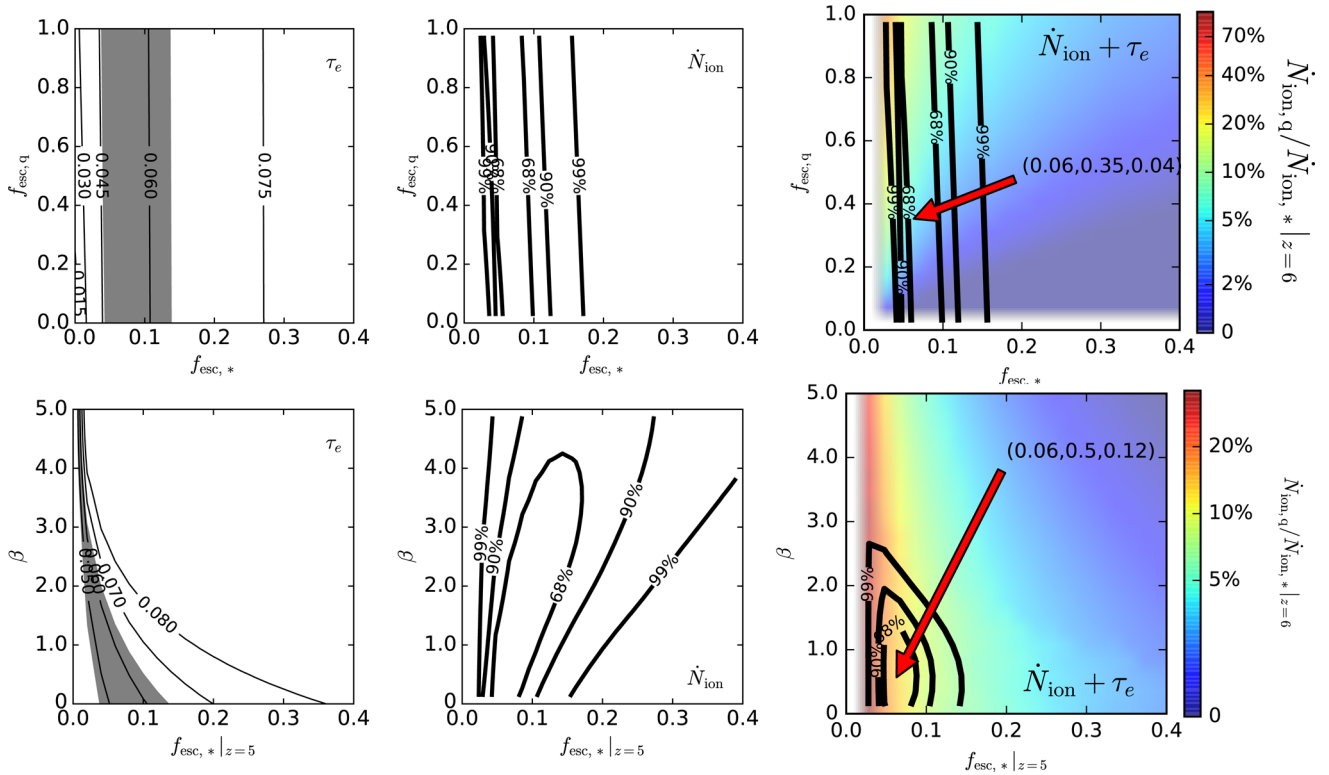
In this section, we explore possible combinations of stars and quasars that could result in an overall photon budget at  $z > 5$  consistent with the observed optical depth and ionizing flux at



**Figure 7.** *Top panel:* the evolution of ionizing emissivity for the QuasarReion (orange squares) and QuasarReion\_nodust (brown line) models. The estimation from G15 is shown as squares, while models from Haardt & Madau (2012); Madau & Haardt (2015) and Mitra et al. (2015) are indicated using black dotted, solid and dash-dotted lines, respectively. The vertical dash-dotted line represents the redshift when the model starts underestimating the quasar emissivity. *Middle panel:* the evolution of the mass-weighted global neutral hydrogen fraction. *Bottom panel:* the Thomson scattering optical depth as a function of redshift. The dashed line with shaded region indicates the Planck 2016 measurement with  $1\sigma$  uncertainties (Planck16). [A colour version of this figure is available in the online version.]

$z \sim 2$ – $5$ . In the Section 4, as well as in M16 where only galaxies are considered, we demonstrated the requirement of an evolving escape fraction for stars to explain the observed emissivity at  $z \sim 5$ . Noting this requirement, in this section, we assume constant escape fractions both for simplicity and to ease interpretation.

Motivated by the recent claim that the escape fraction of low-luminosity AGN is possibly less than unity at high redshift (Micheva et al. 2017), we run MERAXES with different combinations of  $f_{\text{esc},*}$  and  $f_{\text{esc},q}$ , without any changes to the other parameters. In the top panel of Fig. 8, the left panel shows the Thomson scattering optical depth,  $\tau_e$ . For comparison, shaded regions are shown corresponding to the best fit and  $1\sigma$  range of the Planck16 measurements. Based on the BB13 data at  $z \sim 2$ – $5$  and the Planck16 measurement, the top right-hand two panels show the 68, 90 and 99 per cent confidence contours on each parameter of the best fit via the standard minimum- $\chi^2$  technique. The 2D histogram shows the distribution of the ratios of quasar emissivity to stellar emissivity at  $z \sim 6$ . We see that a lower escape fraction of ionizing photons from stars,  $f_{\text{esc},*}$ , requires a higher contribution from quasars, in order to reach the observational constraint. This also returns a higher ratio of quasar emissivity to stellar emissivity. However, because there is not a significant number of quasars at high redshift, changing  $f_{\text{esc},q}$  has little impact to the optical depth. Through the best-fitting contours, we see that if the escape fraction of ionizing photons from stars is only a few per cent



**Figure 8.** *Left panel:* the Thomson scattering optical depth  $\tau_e$ . The measurements with  $1\sigma$  uncertainties by [Planck16](#) are shown in the shaded region. *Middle panel:* the 68, 90 and 99 per cent confidence limits on each parameter of the best fit via the minimum- $\chi^2$  technique based on the estimated total emissivities and their errors from [BB13](#) at  $z \sim 2-5$ . *Right panel:* the 68, 90 and 99 per cent confidence limits based on [BB13](#) and [Planck16](#). The 2D histogram shows the distribution of the ratios of quasar emissivity to stellar emissivity at  $z \sim 6$ . The position of the minimum  $\chi^2$  value and the corresponding value of the 2D histogram are indicated with arrows. *Top panel:* the results as functions of the escape fractions of stars,  $f_{\text{esc},*}$  and quasars,  $f_{\text{esc},q}$ . *Bottom panel:* the results as functions of  $f_{\text{esc},*|z=5}$  and  $\beta$ , which are the normalization and scaling of the redshift-dependent stellar escape fraction,  $f_{\text{esc},*}(z) = \min[f_{\text{esc},*|z=5} \times \left(\frac{1+z}{6}\right)^\beta, 1]$ . In these models, the quasar escape fraction is assumed to be  $f_{\text{esc},q} = 1$ . [A colour version of this figure is available in the online version.]

(<5 per cent; [Ciardullo et al. 2014](#); [Matthee et al. 2016b](#)), the model requires  $f_{\text{esc},q} \sim 1.0$ .

### 5.3 An evolving escape fraction

Although the escape fraction of ionizing photons depends on the local environment, many theoretical and observational works suggest an evolving or mass-dependent escape fraction with a decreasing average value at lower redshifts or in more massive galaxies ([Kuhlen & Faucher-Giguère 2012](#); [Haardt & Madau 2012](#); [Paardekooper et al. 2013](#); [Kimm & Cen 2014](#); [Wise et al. 2014](#); [Bauer et al. 2015](#); [Price et al. 2016](#)). Therefore, as discussed in [M16](#), in order to simultaneously match the normalization and flat slope of the observed ionizing emissivity at  $z \lesssim 6$ , and the Planck  $\tau_e$  measurements,<sup>26</sup> a redshift-dependent escape fraction for galaxies was proposed:

$$f_{\text{esc},*}(z) = \min \left[ f_{\text{esc},*|z=5} \times \left( \frac{1+z}{6} \right)^\beta, 1 \right]. \quad (35)$$

We have shown two models with evolving escape fractions (fiducial and StellarReion) in Section 4. In this section, we further explore

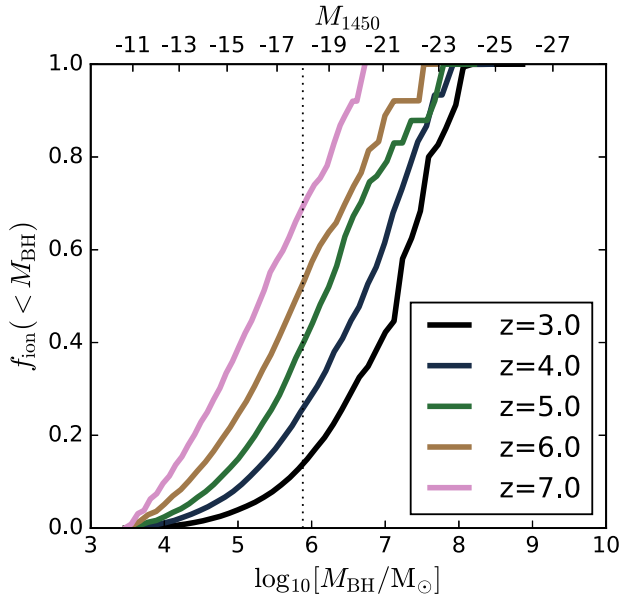
<sup>26</sup> [M16](#) constrained the model using the observed emissivities from [McQuinn et al. \(2011\)](#), which are lower than the [BB13](#) data, and the [Planck15](#) result, which has a larger optical depth than the [Planck16](#) data.

the possible evolution of the escape fraction by running the semi-analytic model with different combinations of  $f_{\text{esc},*|z=5}$  and  $\beta$ . Note that all of the ionizing photons from quasars are included in this section ( $f_{\text{esc},q} = 1$ ) in order to investigate the evolution of the stellar escape fraction with the contribution of quasars to reionization.

The bottom panels of [Fig. 8](#) show the optical depth and the best-fitting confidence limits as functions of the normalization of  $f_{\text{esc},*}(z)$ ,  $f_{\text{esc},*|z=5}$  and the scaling,  $\beta$ , when  $f_{\text{esc},q} = 1$ . We see that because a larger scaling suppresses the escape fraction at lower redshifts, which results in a lower emissivity at  $z \lesssim 5$ , a larger normalization is required. In addition, a larger  $\beta$  gives less ionizing photons at  $z \sim 8$ , which slows the process of reionization and consequently increases the optical depth. When  $f_{\text{esc},*|z=5}$  reaches 0, the model becomes quasar dominated, returning a low  $\tau_e$  of around a half of the [Planck16](#) measurement (see [Fig. 6](#)). We see that when including the contribution from quasars, the model prefers a combination of  $f_{\text{esc},*|z=5} \sim 6$  per cent and  $\beta \sim 0.5$ . This corresponds to  $f_{\text{esc},*} \sim 6.5$  per cent at  $z = 6$  with a ratio between the emissivities of quasars and stars of  $\sim 0.12$ .

### 5.4 Faintest and brightest quasar contributors

In addition to the possibility that quasars do not have a very high escape fraction ([Barkana & Loeb 2001](#)), we note that the very faintest quasars predicted in the model are not observed. The recent detection by [G15](#) only reaches to  $M_{1450} \sim -18$ , while our fiducial models predict a significant population of faint quasars down to



**Figure 9.** The cumulative fraction of ionizing photons as a function of black hole mass (or the corresponding UV magnitude  $M_{1450}$  during the Eddington state as shown in the top axis) at  $z \sim 7-3$  in the *Tiamat* volume. The vertical dotted line represents the faintest AGN detected by G15. [A colour version of this figure is available in the online version.]

$M_{1450} \sim -11$ . Whether those undetected quasars are able to contribute a significant amount of ionizing photons is still unknown. For instance, they might be buried in the dust with a large obscuration fraction (Hopkins et al. 2007). The critical mass above which quasars can contribute ionizing photons, coupled with the previously discussed escape fraction, represents limiting cases of a mass-dependent escape fraction for quasars. In the top panel of Fig. 9, we present the cumulative fraction of ionizing photons as a function of black hole mass (or the corresponding UV magnitude  $M_{1450}$  during the Eddington state as shown in the top axis) assuming  $f_{\text{esc,q}} = 1$  from the fiducial model using the *Tiamat* simulation. We see that quasars fainter than  $M_{1450} = -18$  contribute approximately 80 per cent of the total emissivity at  $z \sim 7$  with a decreasing contribution towards lower redshifts (10 per cent at  $z \sim 3$ , the end of the EoR in the QuasarReion model). This suggests that the number of fainter quasars becomes relatively smaller at later times, which can also be observed from the slope of the predicted quasar luminosity function becoming flatter from  $z \sim 7-3$  (see Fig. 3). However, at redshifts higher than  $z \sim 6$ , the total emissivity from quasars is low. For example, the total emissivity at  $z \sim 7$  is five times lower than  $z \sim 5$ , suggesting that faint quasars below current observational limits provide only a small contribution to reionization.

In addition, the AGN light curve adopted in this work, which assumes that black holes are either accreting with  $\epsilon = 1$  or stay quiescent ( $\epsilon = 0$ ) depending on the amount of accretion mass (see Section 2.2.2), has been shown to underestimate the number density of faint AGN at low redshift (Bonoli et al. 2009). For instance, allowing  $\epsilon$  to decrease progressively when the accretion disc has been mostly consumed predicts more faint AGN with  $M_B \sim -16$  by a factor of 2 at  $z \sim 0.1$ . However, the impact becomes insignificant at brighter ranges and higher redshifts. Due to the small contribution of ionizing photons from faint quasars, AGN light curves are therefore not expected to have a significant impact on our conclusions regarding reionization.

On the other hand, due to the limited simulation volume the brightest quasars at high redshift ( $z \geq 4$ ) in our model only reach

$M_{1450} \sim -23$ , above which the contribution of ionizing photons is not considered. However, the G15 emissivity accounts for bright quasars up to  $M_{1450} = -28$ , 100 times brighter than the brightest quasar in our model. In order to estimate the emissivity of the missing bright quasars, we integrate the fitting functions provided by G15 with a magnitude interval of  $-28 < M_{1450} < -23$ . We find that the total emissivity at high redshift increases by less than 1 per cent with the inclusion of the ionizing photons from these quasars. Therefore, the conclusion that quasars do not have a significant role during the EoR is not affected by the volume size. However, with the flattening luminosity function at lower redshifts ( $z < 4$ , indicated with the vertical dash-dotted line in Fig. 7), bright quasars become more important and their contribution to reionization is not ignorable.<sup>27</sup> This results in a lower emissivity of quasars in our model compared to the Haardt & Madau (2012) model at low redshift (see Fig. 7). We note that including these objects will bring forward reionization in the QuasarReion model, but have no impact to the fiducial model.

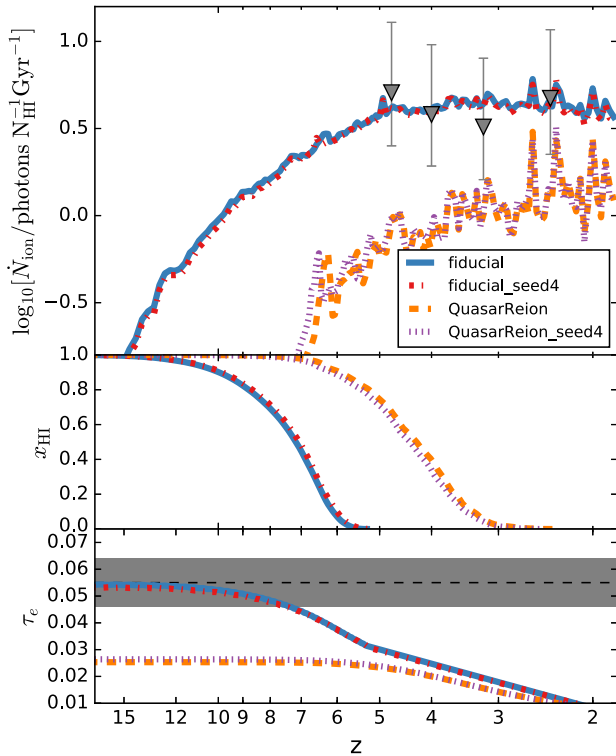
### 5.5 Black hole seed mass

Our choice of black hole seed mass,  $1000 h^{-1} M_\odot$ , lies between the light seed ( $\sim 10^2-10^3 M_\odot$ ) from a remnant Pop III star and the massive seed ( $\sim 10^3-10^5 M_\odot$ ) from the direct collapse of a gas cloud at early times (Greene 2012). The massive seed mass is frequently used to initialize massive haloes ( $\gtrsim 10^6-10^{12} M_\odot$ ) in hydrodynamic simulations (Springel, Di Matteo & Hernquist 2005; Vogelsberger et al. 2014; Schaye et al. 2014; Feng et al. 2016), while the  $\sim 10^3 M_\odot$  seeds are also often adopted in semi-analytic models (Somerville et al. 2008; Bonoli et al. 2009). We note that this seed mass assumption only affects the black hole mass at early times and in the least massive galaxies. The main conclusions of this work are not significantly affected by this assumption. For example, with exactly the same adopted parameters (see Table 1) but 10 times larger seed mass, the properties such as the black hole mass function, the Magorrian relation and the UV luminosity function are changed by less than 5 per cent in massive galaxies ( $M_* > 10^9 M_\odot$ ). On the other hand, the model predicts a significant number of less massive black holes with masses  $\sim 10^5 M_\odot$ , which is more than  $\sim 1$  order of magnitude larger than the Shankar et al. (2009) sample. However, this has negligible impact on the total instantaneous emissivity, and consequently, the reionization history does not change significantly. Fig. 10 presents the evolution of emissivity, neutral hydrogen fraction and optical depth for the fiducial and QuasarReion models with larger black hole seed masses of  $10^4 h^{-1} M_\odot$ . Compared to the original models, we see that the quasar emissivity increases with a larger seed mass, while the stellar emissivity decreases due to the stronger feedback from black holes. However, the changes are negligible, resulting in a small perturbation to the reionization history and optical depth.

## 6 CONCLUSIONS

We have updated the MERAXES semi-analytic model (M16) with a detailed prescription of black hole evolution as part of the Dark-ages Reionization And Galaxy formation Observables from

<sup>27</sup> We fit the predicted quasar luminosity function at  $z = 2$  using a single power law at  $M_{1450} > -24$  and estimate the number of missing photons with a magnitude interval of  $-24 < M_{1450} < -22.5$ . We find the emissivity of quasars can be increased by a factor of 2 with the inclusion of the missing bright quasars.



**Figure 10.** *Top panel:* the evolution of ionizing emissivity for the fiducial models with black hole seed masses of  $10^3 h^{-1} M_{\odot}$  (—) and  $10^4 h^{-1} M_{\odot}$  (---), and for the QuasarReion models with  $10^3 h^{-1} M_{\odot}$  (· · ·) and  $10^4 h^{-1} M_{\odot}$  (- · - ·) seed masses. The total emissivity from BB13 are indicated with grey triangles. *Middle panel:* the evolution of the mass-weighted global neutral hydrogen fraction. *Bottom panel:* the Thomson scattering optical depth as a function of redshift. The dotted and dashed lines with shaded regions indicate the Planck16 measurement (Planck16). [A colour version of this figure is available in the online version.]

Numerical Simulations (DRAGONS) project to study the role of AGN in reionization and galaxy formation at high redshift. The model is calibrated against the observed stellar mass function ( $z \sim 7-0.6$ ), black hole mass function ( $z \lesssim 0.5$ ), quasar luminosity function ( $z \sim 6-0.6$ ), ionizing emissivity ( $z \sim 5-2$ ) and the Thomson scattering optical depth. The model is in agreement with the observed Magorrian relation at low redshift ( $z < 0.5$ ) and predicts a decreasing black hole mass towards higher redshifts at a fixed stellar mass. An opening angle of 80 deg, which corresponds to an un-obscured fraction of  $\sim 23.4$  per cent, allows the model to reproduce the observed quasar luminosity function across a large redshift range ( $z \sim 6-0.6$ ).

Our model suggests that the radiation observed from recently discovered faint AGN at high redshift G15 may include a significant fraction of UV flux from stars. Previous direct estimates of quasar contributions to reionization based on these observations (Madau & Haardt 2015; Mitra et al. 2015) therefore result in an overestimate of the emissivity of quasars by a factor of 3 at  $z \sim 6$ .

When we include the contribution of AGN to reionization, we find that quasars do not dominate the ionizing photon budget at  $z > 6$ . In a quasar-only reionization model, where the escape fractions of ionizing photons are 1 and 0 for quasars and stars, respectively, we find that reionization happens very late,  $z \sim 3$ , with a Thomson scattering optical depth of only half of the Planck16 measurement (Planck16). However, at low redshift, quasars are able to provide

a large number of ionizing photons. With quasars contributing all of their ionizing photons ( $f_{\text{esc,q}} = 1$ ), our model prefers a redshift-dependent escape fraction for stars, having the form of  $f_{\text{esc,*}}(z) = \min\left[0.06 \times \left(\frac{1+z}{6}\right)^{0.5}, 1\right]$ . This corresponds to quasars contributing 10 per cent of the total ionizing photons at  $z \sim 6$ .

## ACKNOWLEDGEMENTS

We would like to thank the anonymous referees for providing helpful suggestions that improves the paper substantially. This research was supported by the Victorian Life Sciences Computation Initiative (VLSCI), grant no. UOM0005, on its Peak Computing Facility hosted at the University of Melbourne, an initiative of the Victorian Government, Australia. Part of this work was performed on the gSTAR national facility at Swinburne University of Technology. gSTAR is funded by Swinburne and the Australian Governments Education Investment Fund. This research programme is funded by the Australian Research Council through the ARC Laureate Fellowship FL110100072 awarded to JSBW. This work was supported by the Flagship Allocation Scheme of the NCI National Facility at the ANU, generous allocations of time through the iVEC Partner Share and Australian Supercomputer Time Allocation Committee. AM acknowledges support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant no. 638809 – AIDA).

## REFERENCES

- Barkana R., Loeb A., 2001, *Phys. Rep.*, 349, 125  
 Batcheldor D., 2010, *ApJ*, 711, L108  
 Bauer A., Springel V., Vogelsberger M., Genel S., Torrey P., Sijacki D., Nelson D., Hernquist L., 2015, *MNRAS*, 453, 3594  
 Becker G. D., Bolton J. S., 2013, *MNRAS*, 436, 1023 (BB13)  
 Bernardi M., Sheth R. K., Tundo E., Hyde J. B., 2007, *ApJ*, 660, 267  
 Blanton M. R., Roweis S., 2007, *AJ*, 133, 734  
 Bolton J. S., Haehnelt M. G., 2007, *MNRAS*, 382, 325  
 Bongiorno A. et al., 2007, *A&A*, 472, 443  
 Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, *MNRAS*, 396, 423  
 Booth C. M., Schaye J., 2010, *MNRAS*, 405, L1  
 Bouwens R. J., Illingworth G. D., Oesch P. A., Caruana J., Holwerda B., Smit R., Wilkins S., 2015, *ApJ*, 811, 140  
 Busch G. et al., 2014, *A&A*, 561, A140  
 Capelo P. R., Volonteri M., Dotti M., Bellovary J. M., Mayer L., Governato F., 2015, *MNRAS*, 447, 2123  
 Chabrier G., 2003, *PASP*, 115, 763  
 Ciardullo R. et al., 2014, *ApJ*, 796, 64  
 Comerford J. M., Greene J. E., 2014, *ApJ*, 789, 112  
 Comerford J. M., Schlus K., Greene J. E., Cool R. J., 2013, *ApJ*, 777, 64  
 Comerford J. M., Pooley D., Barrows R. S., Greene J. E., Zakamska N. L., Madejski G. M., Cooper M. C., 2015, *ApJ*, 806, 219  
 Cristiani S., Serrano L. M., Fontanot F., Vanzella E., Monaco P., 2016, *MNRAS*, 462, 2478  
 Croom S. M. et al., 2009, *MNRAS*, 399, 1755  
 Croton D. J. et al., 2006, *MNRAS*, 365, 11  
 Croton D. J. et al., 2016, *ApJS*, 222, 22  
 Davidzon I. et al., 2017, preprint (arXiv:1701.02734)  
 Davis B. L. et al., 2014, *ApJ*, 789, 124  
 Dijkstra M., Stuart J., Wyithe B., 2006, *MNRAS*, 372, 1575  
 Drory N. et al., 2009, *ApJ*, 707, 1595  
 Du P. et al., 2014, *ApJ*, 782, 45  
 Duffy A. R., Wyithe J. S. B., Mutch S. J., Poole G. B., 2014, *MNRAS*, 443, 3435  
 Duncan K. et al., 2014, *MNRAS*, 444, 2960



- Dunlop J. S., McLure R. J., Kukuła M. J., Baum S. A., O’Dea C. P., Hughes D. H., 2003, *MNRAS*, 340, 1095
- Elvis M. et al., 1994, *ApJS*, 95, 1
- Faisst A. L., 2016, *ApJ*, 829, 99
- Feng Y., Di-Matteo T., Croft R. A., Bird S., Battaglia N., Wilkins S., 2016, *MNRAS*, 455, 2778
- Fiore F. et al., 2012, *A&A*, 537, A16
- Fontanot F., Cristiani S., Pfrommer C., Cupani G., Vanzella E., 2014, *MNRAS*, 438, 2097
- Giallongo E. et al., 2015, *A&A*, 578, A83 (G15)
- Glikman E., Bogosavljević M., Djorgovski S. G., Stern D., Dey A., Jannuzi B. T., Mahabal A., 2010, *ApJ*, 710, 1498
- Glikman E., Djorgovski S. G., Stern D., Dey A., Jannuzi B. T., Lee K.-S., 2011, *ApJ*, 728, L26
- Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, *ApJ*, 672, 765
- González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, *ApJ*, 735, L34
- Graham A. W., Driver S. P., Allen P. D., Liske J., 2007, *MNRAS*, 378, 198
- Grazian A. et al., 2015, *A&A*, 575, A96
- Greene J. E., 2012, *Nature Commun.*, 3, 1304
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Hassan S., Davé R., Finlator K., Santos M. G., 2016, *MNRAS*, 457, 1550
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005a, *ApJ*, 630, 705
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005b, *ApJ*, 630, 716
- Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005c, *ApJ*, 632, 81
- Hopkins P. F., Hernquist L., Martini P., Cox T. J., Robertson B., Di Matteo T., Springel V., 2005d, *ApJ*, 625, L71
- Hopkins P. F., Richards G. T., Hernquist L., 2007, *ApJ*, 654, 731
- Huertas-Company M. et al., 2016, *MNRAS*, 462, 4495
- Hunt M. P., Steidel C. C., Adelberger K. L., Shapley A. E., 2004, *ApJ*, 605, 625
- Ilbert O. et al., 2013, *A&A*, 556, A55
- Jiang N., Ho L. C., Dong X.-B., Yang H., Wang J., 2013, *ApJ*, 770, 3
- Jiang L. et al., 2016, *ApJ*, 833, 222
- Jiang Y.-F., Greene J. E., Ho L. C., 2011, *ApJ*, 737, L45
- Kauffmann G., 1996, *MNRAS*, 281, 475
- Kennicutt R. C., Jr, 1998, *ApJ*, 498, 541
- Khairé V., Srianand R., Choudhury T. R., Gaikwad P., 2016, *MNRAS*, 457, 4051
- Kimm T., Cen R., 2014, *ApJ*, 788, 121
- Kroupa P., 2001, *MNRAS*, 322, 231
- Kuhlen M., Faucher-Giguère C.-A., 2012, *MNRAS*, 423, 862
- Kulkarni G., Choudhury T. R., Puchwein E., Haehnelt M. G., 2017, *MNRAS*, 469, 4283
- Liu C., Mutch S. J., Angel P. W., Duffy A. R., Geil P. M., Poole G. B., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, 462, 235
- Loeb A., Barkana R., 2001, *Annu. Rev. Astron. Astrophys.*, 39, 19
- Lyu J., Rieke G. H., Alberts S., 2016, *ApJ*, 816, 85
- Ma X., Hopkins P. F., Kasen D., Quataert E., Faucher-Giguère C.-A., Kereš D., Murray N., Strom A., 2016, *MNRAS*, 459, 3614
- Madau P., Haardt F., 2015, *ApJ*, 813, L8
- Magorrian J. et al., 1998, *AJ*, 115, 2285
- Manti S., Gallerani S., Ferrara A., Greig B., Feruglio C., 2017, *MNRAS*, 466, 1160
- Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, *ApJ*, 701, 1765
- Martínez-Paredes M. et al., 2017, *MNRAS*, 468, 2
- Masters D. et al., 2012, *ApJ*, 755, 169
- Mathur S., Fields D., Peterson B. M., Grupe D., 2012, *ApJ*, 754, 146
- Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, *MNRAS*, 458, 449
- Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2016b, *MNRAS*, 000, 1
- McGreer I. D. et al., 2013, *ApJ*, 768, 105
- McQuinn M., Peng Oh S., Faucher-Giguère C.-A., 2011, *ApJ*, 743, 82
- Mechtley M. et al., 2012, *ApJ*, 756, L38
- Mesinger A., Furlanetto S., Cen R., 2011, *MNRAS*, 411, 955
- Mesinger A., Greig B., Sobacchi E., 2016, *MNRAS*, 459, 2342
- Micheva G., Iwata I., Inoue A. K., 2017, *MNRAS*, 465, 302
- Mitra S., Roy Choudhury T., Ferrara A., 2015, *MNRAS*, 454, L76
- Mortlock A., Conselice C. J., Bluck A. F. L., Bauer A. E., Grützbauch R., Buitrago F., Ownsworth J., 2011, *MNRAS*, 413, 2845
- Müller-Sánchez F., Comerford J. M., Nevin R., Barrows R. S., Cooper M. C., Greene J. E., 2015, *ApJ*, 813, 103
- Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016, *MNRAS*, 462, 250 (M16)
- Mutlu-Pakdil B., Seigar M. S., Davis B. L., 2016, *ApJ*, 830, 117
- Muzzin A. et al., 2013, *ApJ*, 777, 18
- Paardekooper J.-P., Khochfar S., Dalla C. V., 2013, *MNRAS*, 429, L94
- Palanque-Delabrouille N. et al., 2013, *A&A*, 551, A29
- Parsa S., Dunlop J. S., McLure R. J., 2017, preprint ([arXiv:1704.07750](https://arxiv.org/abs/1704.07750))
- Peng C. Y., Impey C. D., Rix H.-W., Kochanek C. S., Keeton C. R., Falco E. E., Lehár J., McLeod B. A., 2006, *ApJ*, 649, 616
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Planck Collaboration XLVI, 2016, *A&A*, 596, A107 (Planck16)
- Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, *MNRAS*, 459, 3025
- Portinari L., Chiosi C., Bressan A., 1997, *A&A*, 334, 38
- Pozzetti L. et al., 2007, *A&A*, 474, 443
- Price L. C., Trac H., Cen R., 2016, preprint ([arXiv:1605.03970](https://arxiv.org/abs/1605.03970))
- Reines A. E., Greene J. E., Geha M., 2013, *ApJ*, 775, 116
- Ricci F., Marchesi S., Shankar F., La Franca F., Civano F., 2017, *MNRAS*, 465, 1915
- Richards G. T. et al., 2005, *MNRAS*, 360, 839
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sanghvi J., Kotilainen J. K., Falomo R., Decarli R., Karhunen K., Uslenghi M., 2014, *MNRAS*, 445, 1261
- Santini P. et al., 2012, *A&A*, 538, A33
- Schaye J. et al., 2014, *MNRAS*, 446, 521
- Schirber M., Bullock J. S., 2003, *ApJ*, 584, 110
- Scott N., Graham A. W., Schombert J., 2013, *ApJ*, 768, 76
- Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, *ApJ*, 690, 20
- Shankar F. et al., 2016, *MNRAS*, 460, 3119
- Sharma M., Theuns T., Frenk C., Bower R., Crain R., Schaller M., Schaye J., 2016, *MNRAS*, 458, L94
- Shen Y., Kelly B. C., 2012, *ApJ*, 746, 169
- Shields G. A., Rosario D. J., Junkkarinen V., Chapman S. C., Bonning E. W., Chiba T., 2012, *ApJ*, 744, 151
- Sobacchi E., Mesinger A., 2013, *MNRAS*, 432, L1
- Sobacchi E., Mesinger A., 2014, *MNRAS*, 440, 1662
- Somerville R. S., Primack J. R., Faber S. M., 2001, *MNRAS*, 320, 504
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, *MNRAS*, 391, 481
- Song M. et al., 2016, *ApJ*, 825, 5
- Springel V., Di Matteo T., Hernquist L., 2005, *MNRAS*, 361, 776
- Stefanon M., Bouwens R. J., Labbé I., Muzzin A., Marchesini D., Oesch P., Gonzalez V., 2017, *ApJ*, 843, 36
- Steinborn L. K., Dolag K., Comerford J. M., Hirschmann M., Remus R.-S., Teklu A. F., 2016, *MNRAS*, 458, 1013
- Sun G., Furlanetto S. R., 2016, *MNRAS*, 460, 417
- Sutherland R. S., Dopita M. A., 1993, *ApJS*, 88, 253
- Thornton C. E., Barth A. J., Ho L. C., Rutledge R. E., Greene J. E., 2008, *ApJ*, 686, 892
- Tomczak A. R. et al., 2014, *ApJ*, 783, 85
- Tonini C., Mutch S. J., Croton D. J., Wyithe J. S. B., 2016, *MNRAS*, 459, 4109
- Vika M., Driver S. P., Graham A. W., Liske J., 2009, *MNRAS*, 400, 1451
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Volonteri M., Gnedin N. Y., 2009, *ApJ*, 703, 2113
- Volonteri M., Capelo P. R., Netzer H., Bellovary J., Dotti M., Governato F., 2015a, *MNRAS*, 452, L6

Volonteri M., Capelo P. R., Netzer H., Bellovary J., Dotti M., Governato F., 2015b, MNRAS, 449, 1470  
 Weigel A. K., Schawinski K., Treister E., Urry C. M., Koss M., Trakhtenbrot B., 2015, MNRAS, 448, 3167  
 Willott C. J. et al., 2010, AJ, 139, 906  
 Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560  
 Wolf C., Wisotzki L., Borch A., Dye S., Kleinheinrich M., Meisenheimer K., 2003, A&A, 408, 499  
 Wyithe J. S. B., Loeb A., 2003, ApJ, 595, 614  
 Yuan W., Zhou H., Dou L., Dong X.-B., Fan X., Wang T.-G., 2014, ApJ, 782, 55

## APPENDIX A: CALCULATING THE MEAN NUMBER OF IONIZING PHOTONS PRODUCED PER BLACK HOLE

During one time step, for a black hole with a given initial mass of  $M_{\text{BH}}$ , its bolometric luminosity at Eddington rate can be calculated through the right hand of equation (16). Since the bolometric correction (Hopkins et al. 2007) adopted in this work is dependent on the bolometric luminosity (see equation 25), the UV magnitude,  $M_{1450}$  of the quasar changes during its accretion, so does the emissivity. In our model, because the accretion mass is always smaller than the black hole mass ( $\Delta M_{\text{BH}} < M_{\text{BH}}$ , see Fig. A1), for the sake of calculation speed, we estimate the mean number of ionizing photons produced per black hole,  $N_{\gamma,q}$  with the bolometric luminosity at the beginning of accretion. We calculate  $N_{\gamma,q}$  as follows:

(i) We calculate the UV magnitude,  $M_{1450}$  using equations (22)–(27).

(ii) We calculate the UV flux with  $M_{1450}$  in units of  $\text{erg s}^{-1}\text{Hz}^{-1}$  through

$$F_{1450} = 10^{(M_{1450}-48.6)/-2.5} \times 4\pi \left( \frac{10 \text{ pc}}{1 \text{ cm}} \right)^2. \quad (\text{A1})$$

(iii) We calculate the flux at Lyman limit following G15

$$F_{912} = F_{1450} \left( \frac{1200}{1450} \right)^{\alpha_{q,\text{optical}}} \left( \frac{912}{1200} \right)^{\alpha_q}, \quad (\text{A2})$$

where  $\alpha_{q,\text{optical}} = 0.44$  and  $\alpha_q = 1.57$  correspond to a double power-law AGN SED.

(iv) We calculate the instantaneous emissivity by

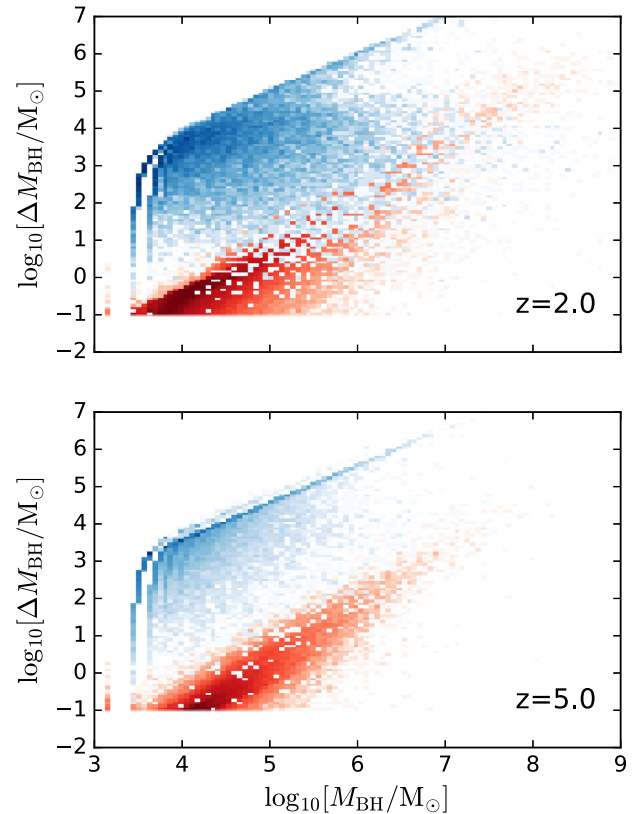
$$\dot{N}_{\text{ion}} \equiv \int_{\nu_{912}}^{\infty} F_{912} \left( \frac{\nu}{\nu_{912}} \right)^{-\alpha_q} \frac{d\nu}{h\nu} = \frac{F_{912}}{h\alpha_q}. \quad (\text{A3})$$

(v) The duration of accreting mass,  $\Delta M_{\text{BH}}$  can be calculated through

$$t_{\text{acc}} = \ln \left( \frac{\Delta M_{\text{BH}}}{M_{\text{BH}}} + 1 \right) \times \frac{\eta t_{\text{Edd}}}{\epsilon}. \quad (\text{A4})$$

Therefore, the total number of ionizing photons emitted is  $\dot{N}_{\text{ion}} t_{\text{acc}}$  and the mean number of ionizing photons produced per black hole is

$$N_{\gamma,q} = \frac{\int_0^{t_{\text{acc}}} \dot{N}_{\text{ion}}(t) dt}{(1-\eta) \Delta M_{\text{BH}}/m_p} \approx \frac{\dot{N}_{\text{ion}}|_{t=t_{\text{acc}}/2} t_{\text{acc}}}{(1-\eta) \Delta M_{\text{BH}}/m_p}, \quad (\text{A5})$$

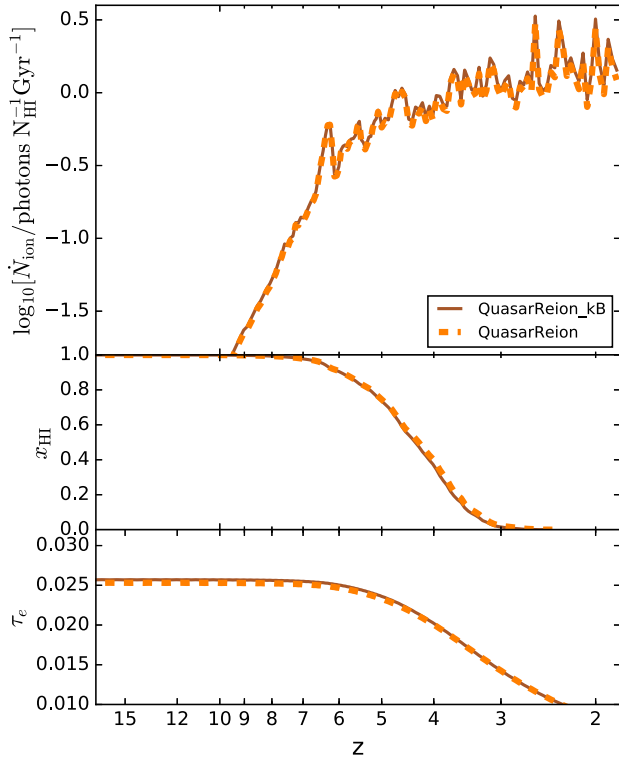


**Figure A1.** Accretion mass ( $\Delta M_{\text{BH}}$ ) versus black hole mass ( $M_{\text{BH}}$ ) at  $z = 2$  and 5 in the *Tiamat* result. The accretion masses from the radio and quasar modes are indicated red and blue, respectively. [A colour version of this figure is available in the online version.]

where the last step adopts the instantaneous emissivity at the middle of accretion for the sake of computational speed.

We note that during the accretion, with an exponential increase of black hole mass, the AGN bolometric luminosity,  $L_{\text{bol}}$  increases exponentially (equation 22). Since the *B*-band bolometric correction,  $k_B$ , decreases with increasing luminosity following a double power law (equation 25),  $\dot{N}_{\text{ion}}$  is a convex function of time. Therefore, the approximation in equation (A5) underestimates the number of ionizing photons produced by black hole. In order to test whether this has a significant impact on our conclusion, we rerun the QuasarReion model assuming a constant bolometric correction with  $k_B(t) \approx k_B|_{t=t_{\text{acc}}}$ . Eliminating the complex dependence of time from  $k_B$ ,  $N_{\gamma,q}$  can be analytically calculated by integrating the AGN light curve. However, we note that since  $k_B(t) \leq k_B|_{t=t_{\text{acc}}}$ , this approximation overestimates  $N_{\gamma,q}$ .

Fig. A2 presents the evolution of emissivity, neutral hydrogen fraction and optical depth for different QuasarReion models assuming constant  $\dot{N}_{\text{ion}}$  (QuasarReion) and  $k_B$  (QuasarReion\_kB), respectively. Since the time interval between two snapshots is much smaller than the Eddington accretion time-scale ( $t_{\text{Edd}} \sim 450 \text{ Myr}$ ), the black hole mass increment is still within the linear regime and therefore we see that the impact from the calculation of  $N_{\gamma,q}$  is not significant.



**Figure A2.** *Top panel:* the evolution of ionizing emissivity for models assuming constant  $\dot{N}_{\text{ion}}$  (QuasarReion,  $\square$ ) and  $k_B$  (QuasarReion\_kB,  $\text{—}$ ). *Middle panel:* the evolution of the mass-weighted global neutral hydrogen fraction. *Bottom panel:* the Thomson scattering optical depth as a function of redshift. [A colour version of this figure is available in the online version.]

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