



# The OBELISK simulation: Galaxies contribute more than AGN to HI reionization of protoclusters

Maxime Trebitsch, Yohan Dubois, Marta Volonteri, Hugo Pfister, Corentin Cadiou, Harley Katz, Joakim A Rosdahl, Taysun Kimm, Christophe Pichon, Ricarda Beckmann, et al.

## ► To cite this version:

Maxime Trebitsch, Yohan Dubois, Marta Volonteri, Hugo Pfister, Corentin Cadiou, et al.. The OBELISK simulation: Galaxies contribute more than AGN to HI reionization of protoclusters. Astronomy and Astrophysics - A&A, 2021, 653, pp.A154. 10.1051/0004-6361/202037698 . hal-03443317

**HAL Id: hal-03443317**

**<https://cnrs.hal.science/hal-03443317>**

Submitted on 23 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The OBELISK simulation: galaxies contribute more than AGN to H I reionization of protoclusters

Maxime Trebitsch<sup>1,2,3,\*</sup>, Yohan Dubois<sup>1</sup>, Marta Volonteri<sup>1</sup>, Hugo Pfister<sup>1,4,5</sup>, Corentin Cadiou<sup>1,6</sup>, Harley Katz<sup>7,\*\*</sup>, Joakim Rosdahl<sup>8</sup>, Taysun Kimm<sup>9</sup>, Christophe Pichon<sup>1,10</sup>, Ricarda S. Beckmann<sup>1</sup>, Julien Devriendt<sup>7,8</sup>, and Adrienne Slyz<sup>7</sup>

<sup>1</sup> Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France

<sup>2</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

<sup>3</sup> Zentrum für Astronomie der Universität Heidelberg, Institut für Theoretische Astrophysik, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany

<sup>4</sup> DARK Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark

<sup>5</sup> Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong

<sup>6</sup> Department of Physics and Astronomy, University College London, London WC1E 6BT, UK

<sup>7</sup> Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

<sup>8</sup> Univ Lyon, Univ Lyon 1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon, UMR5574, F-69230, Saint-Genis-Laval, France

<sup>9</sup> Department of Astronomy, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

<sup>10</sup> Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea

Received Month DD, YYYY; accepted Month DD, YYYY

## ABSTRACT

We present the OBELISK project, a cosmological radiation-hydrodynamics simulation following the assembly and reionization of a protocluster progenitor during the first two billions of years from the big bang, down to  $z = 3.5$ . The simulation resolves haloes down to the atomic cooling limit, and tracks the contribution of different sources of ionization: stars, active galactic nuclei, and collisions. The OBELISK project is designed specifically to study the coevolution of high redshift galaxies and quasars in an environment favouring black hole growth. In this paper, we establish the relative contribution of these two sources of radiation to reionization and their respective role in establishing and maintaining the high redshift ionizing background. Our volume is typical of an overdense region of the Universe and displays star formation rate and black hole accretion rate densities similar to high redshift protoclusters. We find that hydrogen reionization happens inside-out and is completed by  $z \sim 6$  in our overdensity, and is predominantly driven by galaxies, while accreting black holes only play a role at  $z \sim 4$ .

**Key words.** Methods: numerical – Galaxies: formation – Galaxies: high-redshift – intergalactic medium – quasars: supermassive black holes – dark ages, reionization, first stars

## 1. Introduction

Observations of galaxies across cosmic time indicate that the Universe was significantly more active during its infancy compared to today. The cosmic star formation rate density (SFRD) increases rapidly with cosmic time, reaching its peak around  $z \sim 2$  (*cosmic noon*). During the first three billion years from the big bang, galaxies convert their gas into stars very rapidly: the typical growth timescale of these early galaxies is 2 to 10 times shorter than in the present Universe (e.g. Madau & Dickinson 2014). As a result of this rapid growth, approximately the same fraction of the stellar mass observed today was formed before  $z > 2$  and after  $z \sim 0.7$ . As they quickly form stars, the first galaxies illuminate the initially neutral intergalactic medium (IGM). Accreting massive black holes (BHs) harboured in some of these galaxies also shine as they grow in mass. These sources of radiation start ionizing their environment from their formation at  $z \gtrsim 20$ , so that by  $z \sim 6$  (*cosmic dawn*), almost all of the

hydrogen in the Universe is (re)ionized (e.g. Fan et al. 2006a); a similar process happens for helium at a later time, with He II reionization being complete around  $z \sim 2.5 - 3.5$  (e.g. Shull et al. 2010; Worseck et al. 2016).

There has been tremendous progress in recent years in observing those sources of reionization: a large number of galaxies has now been found at  $z \geq 6$  (see e.g. the review of Stark 2016), with some candidates even reaching  $z \sim 10 - 11$  when the Universe was less than 500 Myr old (Oesch et al. 2016; Salmon et al. 2018; Lam et al. 2019; Bouwens et al. 2019). Yet, how the assembly of these early galaxies results in the large scale reionization of the Universe in less than one billion years is still largely an open problem. For instance, while we know that high redshift galaxies produce ionizing photons very efficiently, the fraction  $f_{\text{esc}}$  of these ionizing photons that manage to escape the interstellar (ISM) and circumgalactic medium (CGM), and hence contribute to reionization, is still largely unconstrained. Significant observational effort has been undertaken to measure  $f_{\text{esc}}$  both at high (e.g. Mostardi et al. 2015; Shapley et al. 2016; Grazian et al. 2017; Naidu et al. 2018; Vanzella et al. 2016, 2018; Steidel et al.

\* e-mail: trebitsch@mpia.de

\*\* Visitor

2018; Fletcher et al. 2019; Tanvir et al. 2019) and low redshift (Leitet et al. 2011, 2013; Borthakur et al. 2014; Leitherer et al. 2016; Izotov et al. 2016a,b, 2018a,b), yet we still seem to be only scratching the surface. The lack of constraints on  $f_{\text{esc}}$  limits our understanding of the epoch of reionization (EoR), as changing the value of  $f_{\text{esc}}$  by a factor of 2 drastically affects the timing of reionization (e.g. Madau 2017).

Moreover, the relative role of the different ionizing sources (star forming galaxies and active galactic nuclei: AGN) in reionizing the Universe is one of the most pressing question of the field (e.g. Madau et al. 1999; Haehnelt et al. 2001; Robertson et al. 2015; Haardt & Salvaterra 2015; Madau & Haardt 2015; Garaldi et al. 2019), and the need for sources beyond star forming galaxies depends strongly on the assumed value for  $f_{\text{esc}}$  (e.g. Yoshiura et al. 2017; Finkelstein et al. 2019; Naidu et al. 2019; Dayal et al. 2020, etc.). While bright quasars such as those found at  $z > 7$  by Mortlock et al. (2011); Bañados et al. (2018) are too rare to be dominant actors of reionization (e.g. Becker & Bolton 2013, but see Giallongo et al. (2019) for a different perspective), they have been suggested (e.g. Chardin et al. 2015) as a solution to explain the patchiness of the late stages of reionization (e.g. Becker et al. 2015). Revealing the detailed properties of these distant galaxies and black holes to build a consistent picture of the EoR is a key science project for next generation of observatories such as the *James Webb Space Telescope* or the *Square Kilometre Array*.

In the standard  $\Lambda$  Cold Dark Matter ( $\Lambda$ CDM) cosmological model, structures form hierarchically, meaning that low mass structures form first: this implies that at a given time, the most massive structures will on average be older. As a consequence, galaxies found in the most massive haloes have a mass assembly history even more skewed towards early times. For instance, models like those of Behroozi et al. (2013) predict that for galaxies found in today’s clusters, living in dark matter (DM) halos with virial mass of order  $M_{\text{vir}} \gtrsim 10^{14} M_{\odot}$ , most of the stellar mass was assembled before  $z \gtrsim 2$ .

Because of their accelerated evolution (Overzier 2016), protoclusters are unique laboratories to study the complex processes governing galaxy formation and evolution in the high redshift Universe, a crucial step towards understanding the assembly history of galaxies in our local Universe. Additionally, following the growth of a protocluster can prove extremely useful to study the accretion of gas onto galaxies. In the Birnboim & Dekel (2003); Dekel & Birnboim (2006) picture, accretion onto these haloes should transition from a “cold mode” to a “cold in hot” mode and finally reach the regime where cosmological filaments penetrating the halo are shock-heated and destroyed in the hot atmosphere; and this transition is expected to set the scene for the subsequent quenching of galaxies.

The most massive of these protoclusters are expected to harbour galaxies hosting some of the most massive black holes, with masses in excess of  $M_{\bullet} \gtrsim 10^9 M_{\odot}$  already at  $z \sim 6$ . The accretion onto these BHs results in feedback phenomena releasing copious amounts of energy into the surrounding environment (e.g. Fabian 2012), which has been suggested as a viable mechanism to explain the coevolution of supermassive BHs and their host galaxies (e.g. Silk & Rees 1998): understanding to which extent this AGN feedback operates is one of the main challenges in the field of galaxy formation today. Here again, protoclusters could be very useful probes of this phenomenon: compared to the field, these objects are found to be richer in AGN than the average environments (e.g. Casey 2016).

On the theoretical side, tremendous effort has been made in the recent years to improve cosmological simulations models.

Large simulation projects, such as HORIZON-AGN (Dubois et al. 2014a), EAGLE (Crain et al. 2015; Schaye et al. 2015), ILLUSTRIS (Vogelsberger et al. 2014) and its successor ILLUSTRIS-TNG (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018; Pillepich et al. 2018a; Springel et al. 2018), or MASSIVEBLACK-II (Khandai et al. 2015) and its high redshift successor BLUETIDES (Feng et al. 2016), have been designed to study the evolution of the galaxy population in a cosmological volume. These simulations have been very successful at reproducing and explaining for example the diversity of galaxies in the low redshift Universe. Motivated by this success, several teams have been developing a new generation of simulations that reach a much higher resolution while keeping the large scale cosmological environment, for instance ILLUSTRIS-TNG50 (Nelson et al. 2019; Pillepich et al. 2019), ROMULUS (Tremmel et al. 2017) or NEW-HORIZON (Dubois et al. in prep; Park et al. 2019).

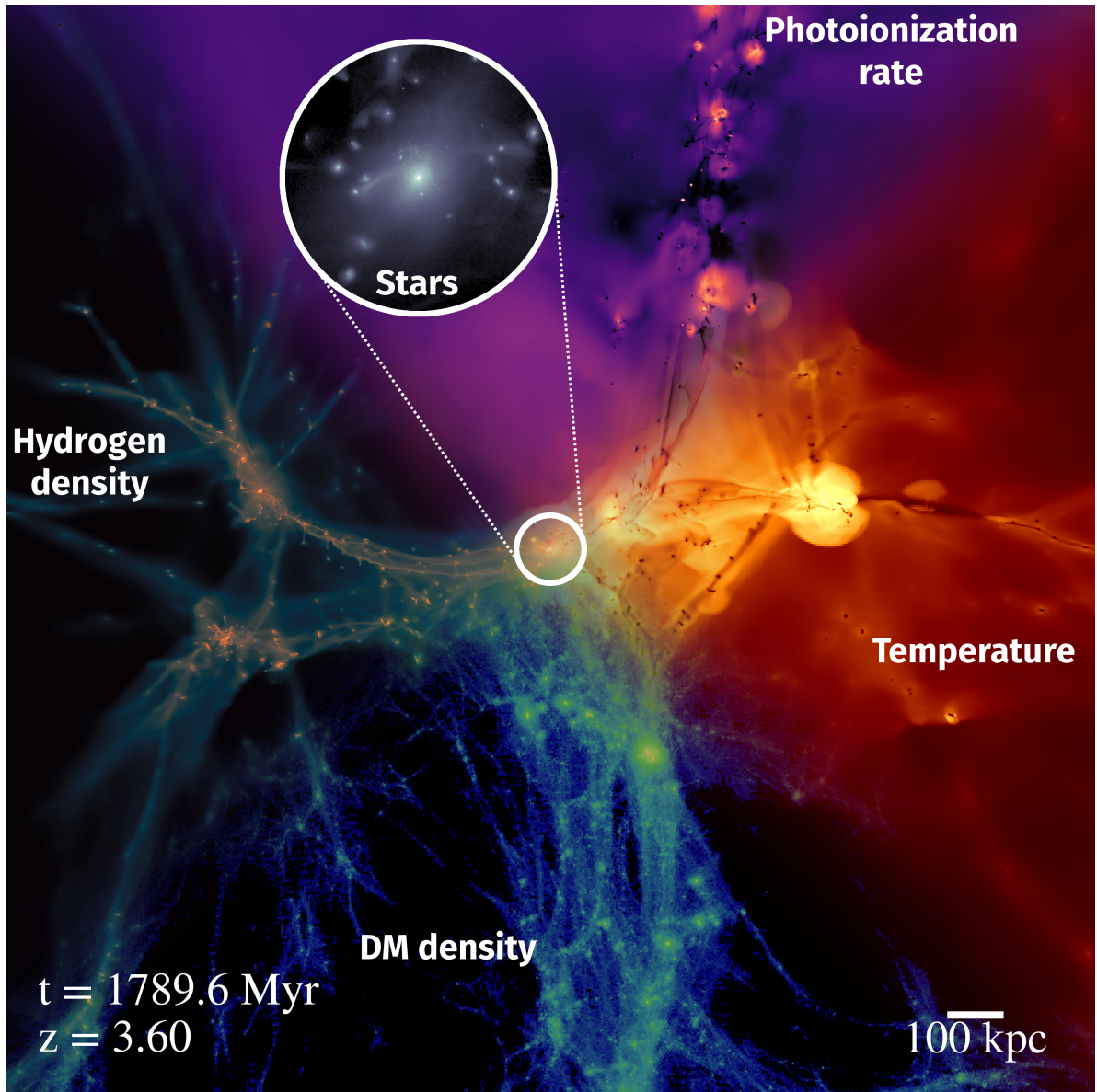
Because of the additional complexity introduced by the inclusion of radiation hydrodynamics and the resulting extra computational cost, the landscape of cosmological simulations attempting to model reionization self-consistently is more sparsely populated. On top of this, radiative transfer strongly couples the hierarchy of scales involved in galaxy formation: while reionization is a global process that needs to be modelled on scales larger than  $\gtrsim 200 h^{-1}$  comoving Mpc (cMpc) to sample cosmic variance (Iliev et al. 2014), the physical mechanisms affecting the source properties operate at the scale of the ISM (e.g. Kimm & Cen 2014; Paardekooper et al. 2015) with some studies even suggesting that simulations need to resolve molecular clouds to include all relevant processes (Howard et al. 2018; Kimm et al. 2019; Kim et al. 2019). At the same time, the inhomogeneous reionization can act as a feedback loop and affect galaxy formation, for instance by suppressing star formation in low mass haloes (e.g. Shapiro et al. 1994, 2004; Gnedin 2000; Katz et al. 2019c) or around rare bright sources.

Several projects such as the Croc project (Gnedin et al. 2014), the CoDA project (Ocvirk et al. 2016, 2018), the AURORA simulation (Pawlik et al. 2017) and the TECHNICAL-DAWN simulation (Finlator et al. 2018) have taken the approach of modelling a large volume and including the sources with sub-grid models. These have provided a very fruitful approach to study the evolution of the IGM and of the broad galaxy population at  $z \gtrsim 6$ . The other approach, taken for instance by the RENAISSANCE suite simulations (O’Shea et al. 2015) and the SPHINX project (Rosdahl et al. 2018), has been to carry out full volume simulations with a resolution comparable to zoom simulations ( $\lesssim 10$  pc), at the cost of strongly reducing the simulated volume. This has led to major advances in the modelling of the ISM of the galaxies responsible for reionizing the Universe. Because of the limited volume, these simulations can only model average<sup>1</sup> environments, dominated by low mass galaxies. Importantly, these simulations all focused on the study of galaxies *during* the EoR and are therefore terminated by  $z \sim 6$ .

In an ideal world, one would want to bridge the gap between “galaxy evolution”-oriented simulations (EAGLE, HORIZON-AGN, ILLUSTRIS, etc.) and the “reionization”-oriented simulations, in order to build a full picture of galaxy formation and evolution at high redshift, *during and after* reionization. In this work, we present our attempt at doing so through the OBELISK project: a simulation designed to study the high redshift Universe, following self-consistently the evolution of the largest protocluster of the HORIZON-AGN volume down to  $z \sim 3.5$ . This is an intermedi-

<sup>1</sup> An exception to this is the RENAISSANCE-Rare peak simulation, but it has only been run until  $z \sim 12$ .





**Fig. 1.** Snapshot of the central region of the OBELISK, illustrating the physics modelled in the simulation. The complex gas distribution is shown on the left, the corresponding dark matter skeleton on the bottom, the gas temperature on the right highlighting self-shielded filaments (dark brown) and hot feedback bubbles (in yellow), and the upper part shows the H I photoionization rate with the knots of the cosmic web lit up by bright sources. The inset zooms in on the stellar distribution around the central galaxy.

ate approach in many respects: while we do not simulate a large cosmological volume, we still capture the formation of a rare structure, unattainable in small boxes, and we retain a higher resolution than simulations like HORIZON-AGN while doing so. At the same time, the connection to the larger HORIZON-AGN simulation allows us to track the descendants of the OBELISK galaxies to connect the high and low redshift galaxy populations. The project focuses on the full high redshift evolution, going beyond the end of the EoR at  $z \sim 6$ , and is unique in that it follows the build up and maintenance of the ionizing background using both galaxies and AGN as sources.

This paper, the first of a series based on the exploitation of the OBELISK simulation, is structured as follow: we begin in Sect. 2 with a comprehensive description of our numerical methodology. In Sect. 3, we present the galaxy and black hole populations emerging from the OBELISK model, and in Sect. 4

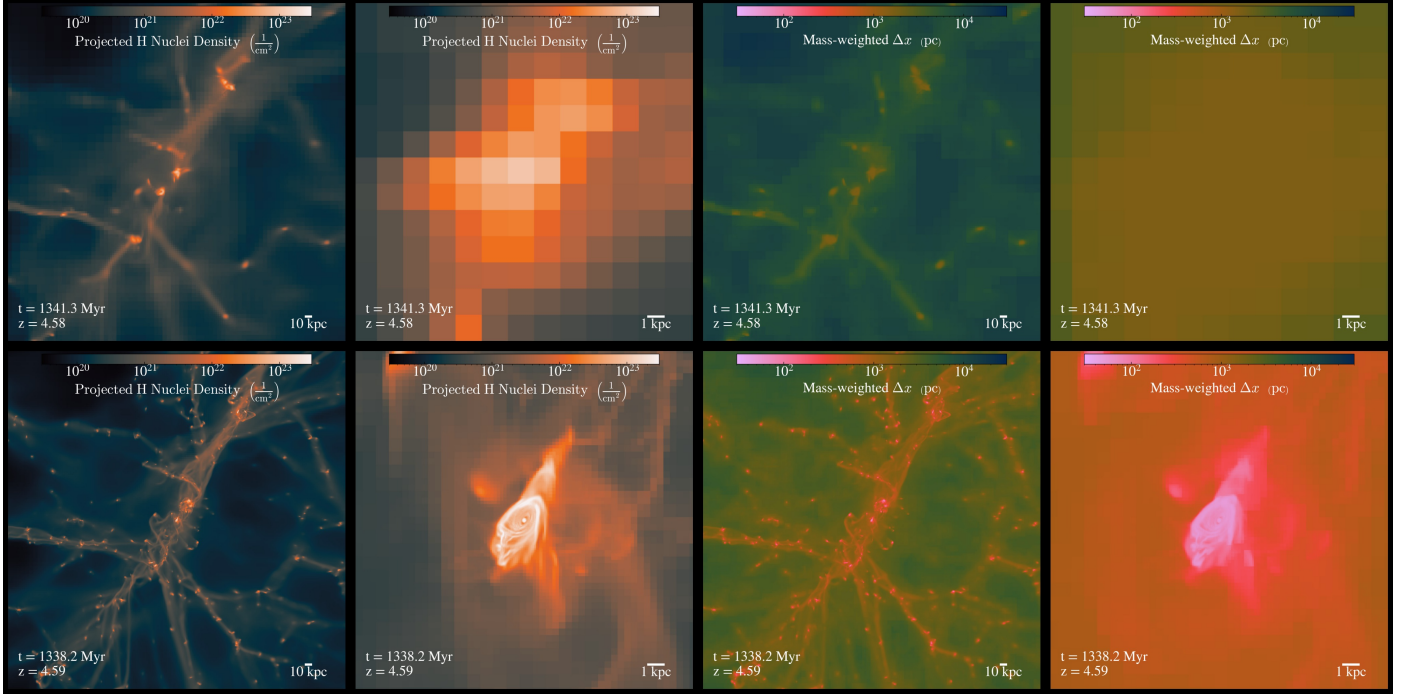
discuss the relative role of these populations in reionizing the volume. We finally present a summary of our results in Sect. 5

## 2. The OBELISK simulation

In this paper, we introduce the OBELISK simulation, a high-resolution ( $\Delta x \simeq 35$  pc), radiation-hydrodynamical simulation of a sub-volume of HORIZON-AGN<sup>2</sup> (Dubois et al. 2014a). We describe in this section the code and physical models that we use in the simulation: the initial conditions and volume selection (Sect. 2.1), the broad features of our radiation hydrodynamics simulation code (Sect. 2.2), and the physical models we employ for stars (Sect. 2.3), BHs (Sect. 2.4) and dust physics (Sect. 2.5). We describe our halo and galaxy identification strategy in Sect. 2.6.

<sup>2</sup> <https://www.horizon-simulation.org/>





**Fig. 2.** Illustration of the increased resolution between HORIZON-AGN (upper row) and OBELISK (lower row) for the projected gas density (first two columns) and typical cell size (last two columns). The first and third columns highlight the large-scale environment, where the resolution in the filaments goes from  $\Delta x_{\text{HORIZON-AGN}} \sim 4$  kpc to  $\Delta x_{\text{OBELISK}} \lesssim 500$  pc, and the second and fourth columns shows the improvement at the galaxy scale.

### 2.1. Initial conditions and volume selection

The initial conditions for the OBELISK simulation are chosen to follow the high redshift evolution of an overdense environment. For this purpose, we have chosen to re-simulate with a high resolution the region around the most massive halo at  $z \sim 2$  in the HORIZON-AGN simulation: selecting initial conditions based on this simulation allows our results to be compared and contrasted with the work that has already resulted from HORIZON-AGN.

The HORIZON-AGN simulation follows the evolution of a cosmological volume of side  $L_{\text{box}} = 100 h^{-1} \text{ cMpc}$  (and periodic boundary conditions), assuming a  $\Lambda$ CDM cosmology compatible with the 7-year data from the *Wilkinson Microwave Anisotropy Probe* (Komatsu et al. 2011): Hubble constant  $H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , total matter density  $\Omega_m = 0.272$ , dark energy density  $\Omega_\Lambda = 0.728$ , baryon density  $\Omega_b = 0.0455$ , amplitude of the matter power spectrum  $\sigma_8 = 0.81$ , and scalar spectral index  $n_s = 0.967$ . The initial conditions have been created with MPGRAFIC (Prunet et al. 2008; Prunet & Pichon 2013) with  $1024^3$  dark matter (DM) particles and as many gas cells (corresponding to a DM mass resolution of  $M_{\text{DM,LR}} = 8 \times 10^7 M_\odot$ ). In the original HORIZON-AGN run, the grid is then adaptively refined over the course of the simulation, maintaining a maximum spatial resolution of  $\Delta x = 1$  proper kpc at all redshifts down to  $z = 0$ . We improve on this resolution by a factor  $\sim 30$  in our OBELISK sub-volume, reaching a resolution of  $\Delta x = 35$  pc (see Sect. 2.2.1).

In order to achieve our target resolution at a reasonable computation cost, we only re-simulate a fraction of the initial HORIZON-AGN volume. To this end, we select the most massive halo (of virial mass is around  $M_{\text{vir}} \simeq 2.5 \times 10^{13} M_\odot$ ) in the simulation at  $z \sim 2$  as identified by the ADAPTAHOP halo finder (Aubert et al. 2004; Tweed et al. 2009). By  $z = 0$ , this halo remains the most massive cluster of HORIZON-AGN, with  $M_{\text{vir}} \sim 6.6 \times 10^{14} M_\odot$ . We first identify all the particles within  $4R_{\text{vir}}$  of the target halo at  $z = 1.97$ , i.e. in a sphere of radius

$2.51 h^{-1} \text{ cMpc}$  and track them back to the initial conditions. We compute the convex hull enclosing all these particles in the initial conditions and define this region as our high-resolution (HR) patch. We then create a re-sampled version of the HORIZON-AGN initial conditions with  $4096^3$  DM particles, corresponding to a mass resolution of  $M_{\text{DM,HR}} = 1.2 \times 10^6 M_\odot$ . Finally, we select all the HR particles belonging to the HR patch previously defined, and embed this patch in the larger HORIZON-AGN box, using successively lower and lower resolution regions as buffers, until we reach an effective resolution of  $256^3$  particles in the outer parts of the volume. We fill the full volume with a passive variable whose value is 1 within the HR patch and 0 outside, and use this as a refinement mask (see Sect. 2.2.1 for further details). Fig. 2 illustrates the gain in resolution between HORIZON-AGN (upper row) and OBELISK (lower row).

Finally, the OBELISK simulation improves upon its parent HORIZON-AGN in several important ways (beyond resolution) that we will describe in detail in Sect. 2.2, 2.3 and 2.4. We should note at this point that a similar methodology has been employed for the NEW-HORIZON simulation, which focuses on an average region of the Universe. Apart from the radiation-hydrodynamical evolution, our numerical methodology is kept as close as possible to the NEW-HORIZON simulation, so as to facilitate the comparison between the two simulations.

### 2.2. Radiation hydrodynamics with RAMSES-RT

The OBELISK simulation is run with RAMSES-RT (Rosdahl et al. 2013; Rosdahl & Teyssier 2015), a multi-group radiative transfer (RT) extension of the public, adaptive mesh refinement (AMR) code RAMSES<sup>3</sup> (Teyssier 2002). RAMSES follows the evolution of DM, gas, stars, and black holes, via gravity, hydrodynamics, radiative transfer, and non-equilibrium thermochemistry.

<sup>3</sup> <https://bitbucket.org/rteyssie/ramses/>

### 2.2.1. Hydrodynamics and gravity

The gas is evolved using an unsplit second-order MUSCL-Hancock scheme (van Leer 1979), based on the Harten-Lax-van Leer-Contact (HLLC) Riemann solver (Toro et al. 1994) to solve the Euler equations. A MinMod total variation diminishing scheme is used to reconstruct the inter-cell conservative variables from the cell-centred values. We assume an ideal gas equation of state with adiabatic index  $\gamma = 5/3$  to close the relation between internal energy and gas pressure. Gravity is modelled by projecting the collisionless particles (stars and DM) onto the AMR grid using cloud-in-cell interpolation and solving the Poisson equation using the multigrid particle-mesh method described in Guillet & Teyssier (2011) on coarse levels and conjugate gradient on fine levels with a transition at level  $\ell = 13$ .

In the HR region, the initial mesh is refined up to a spatial resolution of  $\Delta x \simeq 35 \text{ ckpc}$  (equivalent to  $4096^3$ , or an initial grid level  $\ell_{\text{min,HR}} = 12$ ), and a passive refinement scalar is set to a value of 1 within that region. We only allow refinement where the value of this refinement scalar exceeds 0.01, effectively ensuring that only the initial HR region is adaptively refined throughout its collapse. Within this region, we allow for ten extra levels of refinement, up to a maximal spatial resolution of  $\Delta x \simeq 35 \text{ pc}$  varying within a factor of two depending on the redshift. Our refinement criterion follows the standard RAMSES quasi-Lagrangian approach: a cell is selected for refinement if  $\rho_{\text{DM}} \Delta x^3 + (\Omega_{\text{DM}}/\Omega_b) \rho_{\text{gas}} \Delta x^3 + (\Omega_{\text{DM}}/\Omega_b) \rho_* \Delta x^3 > 8 M_{\text{DM,HR}}$ , where  $\rho_{\text{DM}}$ ,  $\rho_{\text{gas}}$  and  $\rho_*$  are the DM, gas and stellar densities in the cell, respectively. In a DM-only run, this would refine a cell as soon as it contains at least 8 high-resolution DM particles. Note that cells hosting sink particles and the associated clouds (see Sect. 2.4) are always maximally refined. In order to keep the *physical* resolution constant over the course of the simulation (the box has a constant *comoving* size), we only permit a new level of refinement when the expansion scale factor doubles (in our case,  $a_{\text{exp}} = 0.1$  and  $0.2$ ). While this is known to induce a small temporary increase of the star formation (e.g. Snaith et al. 2018), it ensures that the physical subgrid models that have been derived with a specific physical resolution in mind are always used at the appropriate scale.

While the baryonic mass (from gas, stars and black holes) is directly projected onto the maximally refined grid, we smooth the DM density field by depositing the mass of the DM particles on a coarser grid ( $\Delta x \simeq 540 \text{ pc}$ ). This level of smoothing corresponds to the maximal level of refinement triggered in an analogue DM-only simulation where no absolute maximal level of refinement was enforced. This ensures that the effective size of the DM particles correspond to their mass resolution.

The Courant-Friedrichs-Lewy stability condition is enforced using a Courant factor of 0.8, even though the duration of the timestep is predominantly set by the radiation solver (see Sect. 2.2.2).

### 2.2.2. Radiation

The details of the methods used for the injection, propagation and interaction of the radiation with hydrogen and helium are described in Rosdahl et al. (2013), and we therefore only summarize the main features here.

The RT module propagates the radiation emitted by massive stars and accreting black holes (see Sect. 2.3.3 and 2.4.6 for details on the source models) in three frequency intervals describing the H I, He I and He II photon fields (between 13.6–24.59 eV, 24.59–54.4 eV, and 54.4–1000 eV, respectively). The first two

moments of the equation of RT are solved on the AMR grid using an explicit first-order Godunov method with the M1 closure (Levermore 1984; Dubroca & Feugeas 1999) for the Eddington tensor. The radiation is then coupled to the hydrodynamical evolution of the gas through the non-equilibrium thermochemistry for hydrogen and helium and radiation pressure (Sect. 2.2.3).

As we use an explicit solver, we are subject to a Courant-like condition for the propagation of the radiation. This is an extremely stringent condition for the radiation: as the speed of light is much larger than any other velocity in the simulation, the RT timestep should in principle be extremely short. We mitigate this in two ways, following a similar approach to the SPHINX simulation (Rosdahl et al. 2018). First, we subcycle the RT timestep on each AMR level, with up to 500 RT steps for each hydro step, while preventing that photons cross level boundaries during the subcycling (see the discussion in Sect. 2.4 of Rosdahl et al. 2018). In addition to this, the traditional approach is to artificially reduce the speed of light by a constant factor  $f_c$  to prevent too short a timestep and too large a number of RT subcycles. This *reduced speed of light* approximation, initially proposed by Gnedin & Abel (2001) and used here following the implementation of Rosdahl et al. (2013), works well when studying the ISM and the CGM of individual galaxies, where the propagation of light is effectively limited by the propagation of ionization fronts. Contrary to single galaxy or small volume simulations performed with RAMSES-RT (e.g. Kimm et al. 2017; Trebitsch et al. 2017; Costa et al. 2019), the OBELISK simulation tracks the reionization process in a reasonably large volume of the Universe (typically  $\gtrsim 10^4 \text{ Mpc}^3$  comoving at  $z \sim 4$ ): because of this we can no longer fully employ this *reduced speed of light* approximation. We instead use the so-called *variable speed of light* approximation (VSLA, Katz et al. 2017), where the factor  $f_c$  is local. Ideally, one would wish to have a relatively low speed of light in the densest regions and a higher speed of light in voids. Following Katz et al. (2018) and the SPHINX simulation, we use  $f_c = 0.2$  for the coarsest cells in the HR region ( $\ell = 12$  or  $\Delta x \simeq 35 \text{ ckpc}$ ), as the reionization history should be fairly converged with this value (e.g. Deparis et al. 2019; Ocvirk et al. 2019). We then decrease  $f_c$  by a factor of two per level until  $f_c = 0.0125$  at all levels above  $\ell \geq 16$  ( $\Delta x \simeq 2 \text{ kpc}$ ). For the low resolution cells outside of the HR region, we decrease again the speed of light to a low value ( $f_c = 0.01$ ). While this creates some accumulation of radiation just outside of the region of interest, the effects on the HR region are negligible.

As one of the major endeavours of the OBELISK project is to study the relative contributions of massive stars and of AGN to the establishment and maintenance of the ionizing UV background, we use the photon tracer algorithm of Katz et al. (2018, 2019a) to keep track of the contribution of different sources to the local photon flux (and hence the strength of the UV background) and ionization of the gas. This allows us, for example, to track across cosmic time which sources contributes the most to H I-photoionization rate in the IGM, which sources are responsible for ionizing which environment, and to compute population-averaged escape fractions. This algorithm has already been applied in Katz et al. (2018) and Katz et al. (2019b) to study the contribution to reionization of stellar sources of different ages or mass, or residing in halos of different masses. In this work, we trace photons based on their original source: we follow explicitly the radiation produced by stellar populations and by accreting black holes. We refer the interested reader to the Sect. 2 of Katz et al. (2018) for details on the implementation of the algorithm.

### 2.2.3. Gas thermochemistry

RAMSES-RT features non-equilibrium thermochemistry by following the ionization state of hydrogen and helium: H, H<sup>+</sup>, He, He<sup>+</sup>, He<sup>++</sup>, as described in [Rosdahl et al. \(2013\)](#). The coupling with the radiation is performed via photoionization, collisional ionization, collisional excitation, recombination, bremsstrahlung, homogeneous cooling and heating off the cosmic microwave background, and di-electronic recombinations. The whole thermochemistry step is subcycled within every RT step, and uses the smooth injection approach from [Rosdahl et al. \(2013\)](#) to limit the amount of subcycling. We further assume the on-the-spot approximation: any ionizing photon emitted by recombination is assumed to be absorbed locally, and we thus ignore emission of ionizing radiation resulting from direct recombinations to the ground level.

We include a cooling contribution from metals using the standard approach in RAMSES. Above  $T \geq 10^4$  K, we use the cooling rates computed with CLOUDY<sup>4</sup> ([Ferland et al. 1998](#), version 6.02) assuming photoionization equilibrium with the redshift-dependent [Haardt & Madau \(1996\)](#) UV background. We stress that we do not use this UV background for the hydrogen and helium non-equilibrium photo-chemistry, but solely for computing the metal cooling rates. We also account for energy losses via metal line cooling below  $T \leq 10^4$  K, following the prescription of [Rosen & Bregman \(1995\)](#) based on [Dalgarno & McCray \(1972\)](#), and approximate the effect of the metallicity by scaling the metal cooling enhancement linearly (we still assume a Solar-like abundance pattern for simplicity). This allows the gas to cool down to a temperature floor of  $T_{\text{floor}} = 50$  K. We choose to use a density-independent temperature floor rather than a (density-dependent) pressure floor, usually used to prevent numerical fragmentation ([Truelove et al. 1998](#)), because our model for star formation (see the next section) is constructed to efficiently remove gas to form stars in regions with high density and low temperature (which would be susceptible to numerical fragmentation).

Because we do not include molecular hydrogen, we adopt a homogeneous initial metallicity floor of  $Z = 10^{-3}Z_{\odot}$  in the whole computational volume, to allow the gas to cool down below  $10^4$  K before the formation of the first stars (and the subsequent metal enrichment). Aside from this, the gas in the box is initially neutral and composed of  $X = 76\%$  hydrogen and  $X = 24\%$  helium by mass.

## 2.3. Stellar populations

We model stars as particles with mass  $m_{\star} \simeq 10^4 M_{\odot}$  representing a single stellar population, assuming a fully sampled [Kroupa \(2001\)](#) initial mass function (IMF) between 0.1 and  $100 M_{\odot}$ .

### 2.3.1. Star formation

We only consider cells to be star forming when the local number density  $n_{\text{gas}}$  (or equivalently the mass density  $\rho_{\text{gas}}$ ) exceeds a threshold  $n_{\text{SF}} = 5 \text{ H cm}^{-3}$  (chosen as the typical ISM density), and when the local turbulent Mach number defined as the ratio of the turbulent velocity to the sound speed exceeds  $\mathcal{M} \geq 2$ . The amount of gas converted into stars follows a [Schmidt \(1959\)](#) law:

$$\dot{\rho}_{\star} = \epsilon_{\star} \frac{\rho_{\text{gas}}}{t_{\text{ff}}}, \quad (1)$$

<sup>4</sup> <http://www.nublado.org/>

so that on average  $M_{\text{sf}} = \dot{\rho}_{\star} \Delta x^3 \Delta t = \epsilon_{\star} \rho_{\text{gas}} \Delta x^3 \Delta t / t_{\text{ff}}$  of gas is converted into star particles during one timestep  $\Delta t$ , where  $G$  is the gravitational constant,  $\epsilon_{\star}$  is the local star formation efficiency per free-fall time and  $t_{\text{ff}} = \sqrt{3\pi/32G\rho_{\text{gas}}}$  is the gas free-fall time.

One key difference between the method we use here, as already discussed e.g. in [Trebitsch et al. 2017](#) or [Kimm et al. 2017](#), and the traditional approach used e.g. in HORIZON-AGN is that the star formation efficiency  $\epsilon_{\star}$  is a local parameter, rather than a constant, and depends on the local gas density  $\rho_{\text{gas}}$ , sound speed  $c_s$ , and turbulent velocity  $\sigma_{\text{gas}}$ . Here, we approximate  $\sigma_{\text{gas}}$  by taking the velocity differences between the host cell and its immediate neighbours. The analytic expression for  $\epsilon_{\star}(\rho_{\text{gas}}, c_s, \sigma_{\text{gas}})$  follows the ‘multi-ff PN’ model of [Federrath & Klessen \(2012\)](#); [Padoan & Nordlund \(2011\)](#):

$$\epsilon_{\star} \propto e^{\frac{3}{8}\sigma_s^2} \left[ 1 + \text{erf} \left( \frac{\sigma_s^2 - s_{\text{crit}}}{\sqrt{2}\sigma_s^2} \right) \right], \quad (2)$$

for  $\mathcal{M} \geq 2$  and 0 otherwise, and where  $\sigma_s = \sigma_s(\sigma_{\text{gas}}, c_s)$  characterizes the turbulent density fluctuations,  $s_{\text{crit}} = \ln(\rho_{\text{gas,crit}}/\rho_{\text{gas}})$  is the critical density above which the gas will be accreted onto stars, and  $\rho_{\text{gas,crit}} \propto (\sigma_{\text{gas}}^2 + c_s^2)\sigma_{\text{gas}}^2/c_s^2$ . In practice, this means that  $\epsilon_{\star}$  increases with  $\sigma_{\text{gas}}$  and when the virial parameter decreases.

The actual number  $N$  of particles formed in one cell in one timestep  $\Delta t$  is drawn from a Poisson distribution  $P(N) = (\lambda^N/N!) \exp(-\lambda)$  of parameter  $\lambda = M_{\text{sf}}/m_{\star}$  (see [Rasera & Teyssier 2006](#), for details). Additionally, for numerical stability, we limit the number of star particles formed in one timestep so that the amount of gas removed from a cell is capped at 90% of the gas mass in that cell.

### 2.3.2. Supernova feedback

When a star particle reaches an age of  $t_{\text{SN}} = 5$  Myr, we assume that a mass fraction  $\eta_{\text{SN}} = 0.2$  of the initial stellar population explodes as SNe and returns mass and metals to its environment with a yield of 0.075. Each SN explosion releases an energy  $E_{\text{SN}} = 10^{51}$  erg assuming an average progenitor mass of  $m_{\text{SN}} \simeq 20 M_{\odot}$ , corresponding to an average of 1 SN explosion for  $100 M_{\odot}$  of stars. At our mass resolution, each star particle releases  $10^{53}$  erg per event, instantaneously.

The explosion itself is modelled following the mechanical feedback implementation of [Kimm & Cen \(2014\)](#); [Kimm et al. \(2015\)](#), which injects radial momentum according to the phase of the explosion (energy conserving or momentum conserving) that is resolved. We refer the reader to these works for the details of the algorithm implementation, and once again only recall the main features here. If we resolve the adiabatic expansion phase of the SN, we directly inject the kinetic energy and momentum corresponding to  $10^{51}$  erg to the cells neighbouring the SN site. If this energy conserving phase is not resolved, the SN explosion is only captured in its final snowplough phase; in this case we directly inject the terminal momentum  $p_{\text{snow}}$  in the neighbouring cells. We determine the phase of the SN explosion that we resolved on a cell-by-cell basis: for each of the cells neighbouring the SN site, we evaluate the mass ratio  $\chi_{\text{SN}} = dM_{\text{swept}}/dM_{\text{ej}}$  between the swept material (ejecta + swept ISM) and the ejecta, and compare it to a critical mass ratio  $\chi_{\text{SN,tr}}$ . For low values of  $\chi_{\text{SN}}$  (i.e. in the adiabatic phase), we inject

$$\Delta p_{\text{ad}} = f_{\text{geom}} \sqrt{2\chi_{\text{SN}} M_{\text{ej}} E_{\text{SN}}}, \quad (3)$$

where  $f_{\text{geom}}$  is a geometrical factor describing how the total energy and mass of the SN is split between the neighbouring cells,



and  $f_e = 1 - (\chi_{\text{SN}} - 1)/(3\chi_{\text{SN,tr}} - 1)$  ensures a smooth transition between the two modes of momentum injection. In the snowplough phase, we inject the terminal momentum (e.g. Thornton et al. 1998; Blondin et al. 1998; Kim & Ostriker 2015; Martizzi et al. 2015)

$$\Delta p_{\text{snow}} = 3 \times 10^5 \text{ M}_\odot \text{ km s}^{-1} f_{\text{geom}} E_{\text{SN},51}^{16/17} n_{\text{H}}^{-2/17} \tilde{Z}^{-0.14}, \quad (4)$$

where  $E_{\text{SN},51}$  is the total SN energy in units of  $10^{51}$  erg,  $n_{\text{H}}$  is the local hydrogen number density in units of  $\text{cm}^{-3}$ , and  $\tilde{Z}$  is the local metallicity in units of  $Z_\odot$  floored at  $0.01 Z_\odot$ .

We further assume that, the photo-ionization pre-processing of the ISM by young OB stars prior to the SN event augments the final radial momentum from a SN (Geen et al. 2015). While this should be taken into account by the radiative transfer in our simulation, Trebitsch et al. (2017) argued that a significant fraction of the ionizing radiation can be emitted in regions where the Strömgen radius,  $r_S$ , is not resolved in which case this momentum increase will often be missed. Thus, as in Trebitsch et al. (2018) and Rosdahl et al. (2018), we follow the subgrid model of Kimm et al. (2017) which adds this missing momentum when the Strömgen radius is locally not resolved, i.e. when  $r_S < \Delta x$ . We then increase the pre-factor in Eq. 4 to  $5 \times 10^5 \text{ M}_\odot \text{ km s}^{-1}$  following the results of Geen et al. (2015).

### 2.3.3. Stellar radiation

Using the SPHINX simulation, Rosdahl et al. (2018) have shown that the inclusion of binary stars has a strong effect on the timing of reionization, because binary interactions lead to an increased and more sustained production of ionizing radiation (e.g. Stanway et al. 2016; Göteborg et al. 2019, see also Topping & Shull 2015). We emit ionizing radiation from each star particle following the spectral energy distribution (SED) resulting from the Binary Population and Spectral Synthesis code v2.2.1 (BPASS, Eldridge et al. 2017; Stanway & Eldridge 2018, *Tuatara* version), which includes the effect of these binary interactions. In the *Tuatara* release, the binary fraction of stars depends on the initial stellar mass, with around 60% (6%) of low-mass (high-mass) stars being isolated. More specifically, we use the model `imf135_100` closest to a Kroupa (2001) IMF with slopes  $-1.30$  between  $0.1 - 0.5 \text{ M}_\odot$  and  $-2.35$  between  $0.5 - 100 \text{ M}_\odot$ . Each star particle is assigned a luminosity in each radiation bin as a function of its age and metallicity, scaled directly with the mass of the particle. As described in Rosdahl et al. (2013), the average energy of each radiation bin and the interaction cross-sections are updated every 5 coarse timesteps, so that they accurately represent the properties of the average source population.

While lower resolution simulations usually include a subgrid correction to account for the unresolved escape fraction and calibrated to reproduce reasonable reionization history (e.g. Gnedin et al. 2014; Ocvirk et al. 2016, 2018; Pawlik et al. 2017; Finlator et al. 2018), we follow here the approach taken by simulations reaching a spatial resolution better than a few tens of pc and use the luminosity of the star particle directly. We stress that we do not claim that  $\Delta x \simeq 35 \text{ pc}$  is sufficient to resolve in great detail the rich multiphase structure of the ISM; we acknowledge that the uncertainty on the ISM structure could affect the escape of ionizing radiation from birth clouds either way: unresolved ionized channels could lead to a higher escape fraction (e.g. Ma et al. 2015), while unresolved clumping could increase the amount of absorption in the ISM (e.g. Yoo et al. 2020). We note however that the tests performed in the SPHINX framework (appendix B of Rosdahl et al. 2018) suggest that a resolution of

$\Delta x \simeq 20 \text{ pc}$  yields galaxy properties similar to their fiducial run with  $\Delta x \simeq 10 \text{ pc}$ . Finally, we note that the absence of subgrid calibration of the escape fraction is consistent with the assumptions we made for the star formation model and for the black hole accretion (see Sect. 2.4): namely, that we broadly resolve the large-scale ISM density distribution and the formation of molecular clouds. Changing one ingredient (e.g. the unresolved escape of radiation) without changing the others would break that consistency.

## 2.4. BH model

We follow the same approach to model BHs, based on the implementation of Dubois et al. (2010, 2012, 2014c): BHs are modelled as sink particles, which are allowed to accrete gas and release energy, momentum and radiation into their environment. We shall now describe the various aspects of our BH model.

### 2.4.1. BH formation

We seed the sink particles representing our BHs with an initial mass  $M_{\bullet,0} = 3 \times 10^4 \text{ M}_\odot$  in cells where the following criteria are met:

- the gas density  $n_{\text{gas}}$  and stellar density  $n_\star$  must both locally exceed a threshold  $n_{\text{sink}} = 100 \text{ H cm}^{-3}$ ;
- the gas must be Jeans-unstable.

We impose an exclusion radius  $r_{\text{excl}} = 50 \text{ kpc}$  to avoid the formation of multiple BHs in the same galaxy. Each sink particle is then dressed with a swarm of ‘cloud’ particles, positioned on a regular grid lattice within a sphere of radius  $4\Delta x$  and equally spaced by  $\Delta x/2$ . These cloud particles provide a convenient way of probing and averaging the properties of the gas around the BH. At our resolution, this means we probe a sphere of radius  $4\Delta x \simeq 140 \text{ pc}$  around each BH with 2109 clouds. We set the initial velocity of the BH to that of its host cell, and assign it a spin  $a = 0$ .

Our seed mass choice stems from physical as well as numerical considerations. The BH formation mechanism is highly uncertain, with different models predicting very different seed masses, from  $\sim 10 \text{ M}_\odot$  to  $\gtrsim 10^5 \text{ M}_\odot$  (e.g. Volonteri 2010). Numerically, having a BH seed less massive than the mass of star particles is not desirable: Pfister et al. (2019) suggests that otherwise, the BH trajectory becomes extremely sensitive to the fluctuation of the (stellar) gravitational potential. As our mass resolution for star particles is  $m_\star \simeq 10^4 \text{ M}_\odot$ , we choose a BH seed mass three times higher. The underlying physical picture would be that of a light seed BH that has already undergone some early growth or of a heavier seed, and is consistent with our choice of seeding BHs only in regions with a sufficiently high stellar density (thus mimicking pre-galactic centres).

### 2.4.2. BH accretion

Once a sink particle has formed, it grows through two channels: BH-BH mergers and gas accretion. We allow two BHs to merge when they are closer than  $4\Delta x$  from one another, and only if their relative velocity is lower than the escape velocity of the binary system they would form in vacuum. As we are far from resolving the gaseous accretion disc around BHs, we use the classical Bondi-Hoyle-Lyttleton (Bondi 1952) approach to compute the

accretion rate onto the BH:

$$\dot{M}_{\text{BHL}} = 4\pi G^2 M_\bullet^2 \frac{\bar{\rho}}{(\bar{c}_s^2 + \bar{v}_{\text{rel}}^2)^{3/2}}, \quad (5)$$

where  $M_\bullet$  is the BH mass,  $\bar{\rho}$  and  $\bar{c}_s$  are respectively the average gas density and sound speed, and  $\bar{v}_{\text{rel}}$  is the relative velocity between the BH and the surrounding gas. The bar notation denotes an averaging over the cloud particles using a Gaussian kernel  $w \propto \exp(-r^2/r_{\text{sink}}^2)$ , where  $r_{\text{sink}}$  is defined using the Bondi radius  $r_{\text{BHL}} = \frac{GM_\bullet}{c_s^2 + v_{\text{rel}}^2}$ :

$$r_{\text{sink}} = \begin{cases} \Delta x/4 & \text{if } r_{\text{BHL}} < \Delta x/4, \\ r_{\text{BHL}} & \text{if } \Delta x/4 \leq r_{\text{BHL}} < 2\Delta x, \\ 2\Delta x & \text{if } 2\Delta x \leq r_{\text{BHL}}. \end{cases} \quad (6)$$

We do not use any extra artificial boost for the gas accretion onto the BH. The accretion rate is capped at the Eddington rate:

$$\dot{M}_{\text{Edd}} = \frac{4\pi G M_\bullet m_p}{\epsilon_r \sigma_T c}, \quad (7)$$

where  $m_p$  is the proton mass,  $\sigma_T$  is the Thompson cross section,  $c$  is the speed of light and  $\epsilon_r$  is the radiative efficiency of the accretion flow onto the BH, which depends on the spin of the BH (see Sect. 2.4.4). Additionally, only a fraction  $1 - \epsilon_r$  of the mass accreted onto the accretion disc effectively reaches the BH: the rest is radiated away. At very low accretion rates, i.e. when  $\chi = \dot{M}_{\text{BHL}}/\dot{M}_{\text{Edd}}$  drops below  $\chi_{\text{crit}} = 0.01$ , the flow becomes radiatively inefficient, in which case we follow [Benson & Babul \(2009, Eq. 16\)](#) and reduce the radiative efficiency of the flow to  $\tilde{\epsilon}_r = \epsilon_r(\chi/\chi_{\text{crit}})$ . The final BH accretion rate is therefore:

$$\dot{M}_\bullet = (1 - \tilde{\epsilon}_r) \min(\dot{M}_{\text{BHL}}, \dot{M}_{\text{Edd}}). \quad (8)$$

Gas is then removed from cells within  $4\Delta x$  of the sink particle in a kernel-weighted fashion, and the accretion is capped to prevent the BH to remove more than 25% of the gas content of the cell in one timestep for numerical stability (i.e.  $w\dot{M}_\bullet\Delta t \lesssim 0.25\rho_{\text{gas}}\Delta x^3$ ).

### 2.4.3. BH dynamics

While the dynamics of the sink particles is computed at the most refined grid level (see Sect. 2.2.1), we lack the resolution to accurately capture the effects of dynamical friction that will affect the detailed BH dynamics. For instance, [Pfister et al. \(2017\)](#) showed with a resolution study that resolving the influence radius of a BH is crucial to get a chance to resolve the formation of BH binaries in the aftermath of a galaxy merger; this length-scale being of the order of 1 pc for a BH with mass  $M_\bullet \sim 10^6 M_\odot$  in a Milky-way like galaxy, and much lower for less massive BHs, it is well below the resolution that modern cosmological simulations like OBELISK can reach. While most of these simulations resort to moving the sink particle towards a local minimum of the potential (e.g. [Crain et al. 2015](#); [Pillepich et al. 2018b](#); [Davé et al. 2019](#)), we take here a different approach and use a sub-grid model to account for the effect of the unresolved dynamical friction (see also [Tremmel et al. 2015](#)).

We use the drag force implementation introduced in [Dubois et al. \(2013\)](#) to model the force exerted by the gaseous wake lagging behind the BH. The frictional force has an analytic expression given by [Ostriker \(1999\)](#), and is proportional to  $F_{\text{DF}} = \alpha f_{\text{gas}} 4\pi\rho(GM_\bullet/\bar{c}_s^2)$ , where  $\alpha$  is an artificial boost, with  $\alpha = (\rho/\rho_{\text{DF,th}})^2$  if  $\rho > \rho_{\text{th}}$  and 1 otherwise, and  $f_{\text{gas}}$  is a fudge factor

varying between 0 and 2 which depends on the BH Mach number  $\mathcal{M}_\bullet = \bar{v}_{\text{rel}}/\bar{c}_s$  ([Ostriker 1999](#); [Chapon et al. 2013](#)). In light of the results of [Beckmann et al. \(2018\)](#) who showed that this subgrid model begins to fail when the wake is resolved, we set  $F_{\text{DF}} = 0$  whenever the influence radius  $2GM_\bullet/\max(\bar{c}_s, \bar{v}_{\text{rel}})^2 > 0.2\Delta x$ . In this work, we have taken  $\rho_{\text{DF,th}} \approx 0.003$  to  $0.01 \text{ cm}^{-3}$ . We discuss the effect of this ad-hoc choice in Appendix A.

We also include a contribution to the dynamical friction caused by collisionless particles (stars and DM separately), analogous to what happens to the gas: a gravitational wake of stars and DM is created by the passage of a massive body (the BH) and will decelerate it ([Chandrasekhar 1943](#); [Binney & Tremaine 1987](#)). Our implementation, described in detail in [Pfister et al. \(2019\)](#), is somewhat similar to that of [Tremmel et al. \(2015\)](#) used in the ROMULUS simulation ([Tremmel et al. 2017](#)). We directly compute the contribution of the collisionless particles within a sphere of radius  $4\Delta x$ . The deceleration is parallel to the velocity  $\mathbf{v}_\bullet$  of the BH relative to the background (stars and DM) and has a magnitude  $a_{\text{DF}}$  is computed as follow:

$$a_{\text{DF}} = -4\pi \frac{G^2 M_\bullet}{v_\bullet^2} \left( \log \Lambda \int_0^{v_\bullet} 4\pi v_\bullet^2 f(v) dv + \int_{v_\bullet}^\infty 4\pi v^2 f(v) \left[ \log \left( \frac{v + v_\bullet}{v - v_\bullet} \right) - 2 \frac{v_\bullet}{v} \right] dv \right), \quad (9)$$

where  $v_\bullet$  is the norm of the relative velocity with respect to the mass-weighted velocity of the surrounding collisionless particles,  $\Lambda = 4\Delta x/r_{\text{def}}$  is the Coulomb logarithm with  $r_{\text{def}} = GM_\bullet/v_\bullet^2$ , and  $f$  is the distribution function defined with the velocities of the particles within the sphere of radius  $4\Delta x$ . As for the gas, we switch off the subgrid model when the influence radius is resolved by more than  $0.2\Delta x$

$$4\pi v^2 f(v) = \frac{3}{256\pi\Delta x^3} \sum_i m_i \delta(v_i - v). \quad (10)$$

We refer the interested reader to [Pfister et al. \(2019\)](#) for a more detailed discussion of the model.

### 2.4.4. BH spin evolution

We follow self-consistently the evolution of the spin magnitude and direction over the course of the simulation using the implementation presented in [Dubois et al. \(2014b,c\)](#), see also [Fiacconi et al. 2018](#); [Bustamante & Springel 2019](#) for similar implementations). Here again, we refer the interested reader to these works for an extensive discussion of the model and tests of its validity.

We evolve the magnitude of the spin following gas accretion through an expression derived in [Bardeen \(1970\)](#):

$$a^{n+1} = \frac{1}{3} \frac{r_{\text{isco}}^{1/2}}{M_{\text{ratio}}} \left[ 4 - \left( 3 \frac{r_{\text{isco}}}{M_{\text{ratio}}^2} - 2 \right)^{1/2} \right], \quad (11)$$

where  $M_{\text{ratio}} = M_{\bullet,n+1}/M_{\bullet,n}$  ( $M_{\bullet,n}$  being the mass of the BH at times  $t_n$ ), and

$$r_{\text{isco}} = R_{\text{isco}}/R_g = 3 + Z_2 \mp \text{sign}(a) \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)} \quad (12)$$

is the radius of the innermost stable circular orbit (ISCO) expressed in units of gravitational radius  $R_g$ .  $Z_1$  and  $Z_2$  are function of the spin magnitude  $a$ , given by

$$\begin{aligned} Z_1 &= 1 + (1 - a^2)^{1/3} \left[ (1 + a)^{1/3} + (1 - a)^{1/3} \right] \\ Z_2 &= \left( 3a^2 + Z_1^2 \right)^{1/2}. \end{aligned} \quad (13)$$

The  $\mp$  sign in Eq. 12 depends on whether the BH is co-rotating ( $a \geq 0$ ) or counter-rotating ( $a \leq 0$ ) with its accretion disc. For a co-rotating BH,  $1 \leq r_{\text{isco}} \leq 6$ , while  $6 \leq r_{\text{isco}} \leq 9$  for the counter-rotation case ( $r_{\text{isco}} = 6$  only for a non-spinning BH). The direction of the BH spin is evolved assuming that the angular momentum of the (unresolved) accretion disc aligns with that of the accreted gas. The potential mis-alignment between the BH spin and the accretion disc leads to the formation of a warped disc in the innermost regions of the accretion disc precessing about the spin axis because of the Lense-Thirring effect. This warped disc will eventually completely align or anti-align with the BH spin. The anti-alignment configuration occurs when the angle  $\theta$  between angular momenta of the disc  $\mathbf{J}_d$  and of the BH  $\mathbf{J}_\bullet$  fulfils  $\cos \theta < -0.5\|\mathbf{J}_d\|/\|\mathbf{J}_\bullet\|$  (King et al. 2005). We assume that the accretion disc is well described by the Shakura & Sunyaev (1973) thin disc, and define  $J_d$  as the value of the angular at the smallest radius between the warp radius and the self-gravity radius. We refer the reader to Sect. 3 of Dubois et al. (2014c) for the equations governing the details of this process, but we wish to stress here that we do not enforce the spin of the BH to always be aligned with the angular momentum of the accreted gas.

Our model assumes a thin disc solution: we only apply it at high accretion rates, when  $\chi \gtrsim 0.01$ . At lower accretion rate, when the accretion flow is radiatively inefficient, we modify our model following the results of the simulations of “magnetically choked” accretion flows from McKinney et al. (2012). In practice, we assume that each low accretion rate event fills an accretion disc, and over the course of the accretion event, the BH spin is evolved at a rate given by a “spin-up” parameter (or rather spin-down parameter, as in this regime, the absolute value of the spin magnitude tends to decrease systematically) given by a fourth-order polynomial fit of the results in Table 7 of McKinney et al. (2012), particularly their simulations AaaaN100 where aaa is the BH spin of each model. In any case, we limit the spin-up process to  $|a| \leq a_{\text{max}} = 0.998$  following Thorne (1974). Finally, when two BHs merge, we update the spin of the remnant using the fit of Rezzolla et al. (2008), according the measured pre-merger BH spins, orbital angular momentum, and mass ratio.

The spin evolution is not a purely passive quantity in OBELISK: the radiative efficiency  $\epsilon_r$  of the accretion flow is effectively set by the spin through  $r_{\text{isco}}$ :

$$\epsilon_r = 1 - \sqrt{1 - \frac{2}{3r_{\text{isco}}}}. \quad (14)$$

For a non-rotating BH, this leads to  $\epsilon_r \simeq 0.057$ , and the canonical  $\epsilon_r = 0.1$  corresponds to  $a \simeq 0.7$ .

#### 2.4.5. AGN feedback

We implement AGN feedback following accretion events using the dual mode approach of Dubois et al. (2012). At low Eddington ratio  $\chi < \chi_{\text{crit}}$ , the AGN is in “radio mode”, while it is in “quasar mode” when  $\lambda_{\text{Edd}} \geq 0.01$ .

For the quasar mode, each sink particle dumps an amount  $\dot{E}_{\text{AGN}}\Delta t$  of thermal energy over a timestep  $\Delta t$  in a sphere of radius  $\Delta x$  centred on the BH. The energy injection rate is calculated as

$$\dot{E}_{\text{AGN}} = \epsilon_f \epsilon_r \dot{M}_\bullet c^2, \quad (15)$$

where  $\epsilon_f = 0.15$  is the fraction of the bolometric luminosity that is transferred to the gas (Booth & Schaye 2009; Dubois et al. 2012), driving unresolved winds through Compton heating, and

UV and IR momentum coupling of radiation. This ad-hoc choice efficiency allows for the self-regulation of BH growth through their feedback. The actual input of (UV) photons from the AGN is also treated explicitly on resolved scales with the RAMSES-RT solver (see Sect. 2.4.6).

For the radio mode, we deposit momentum as a bipolar cylindrical outflow mimicking the propagation of a jet in the surrounding gas with a velocity  $u_j = 10^4 \text{ km s}^{-1}$ . The outflow profile corresponds to a cylinder of radius  $\Delta x$  and height  $2\Delta x$  weighted by a kernel function

$$\psi(r_{\text{cyl}}) = \frac{1}{2\pi\Delta x^2} \exp\left(-\frac{r_{\text{cyl}}^2}{\Delta x^2}\right), \quad (16)$$

with  $r_{\text{cyl}}$  the cylindrical radius. The outflow removes mass from the central cell and transport it to the cells enclosed by the jet at a rate  $\dot{M}_J$  with

$$\dot{M}_J(r_{\text{cyl}}) = \frac{\psi(r_{\text{cyl}})}{\Psi} \eta_j \dot{M}_\bullet, \quad (17)$$

where  $\Psi$  is the integral of  $\psi$  over the cylinder and  $\eta_j = 100$  is the mass-loading factor of the jet accounting for the interaction between the jet and the ISM at unresolved scales. The momentum is injected in a direction parallel to the BH spin (with opposite signs above and below the sink) with the norm of the momentum  $q(r_{\text{cyl}})$  given by

$$q_J(r_{\text{cyl}}) = \dot{M}_J(r_{\text{cyl}}) u_j. \quad (18)$$

We inject the corresponding kinetic energy into the gas:

$$\dot{E}_J(r_{\text{cyl}}) = \frac{q_J(r_{\text{cyl}})^2}{2\dot{M}_J(r_{\text{cyl}})} = \frac{\psi(r_{\text{cyl}})}{\Psi} \dot{E}_{\text{AGN}}. \quad (19)$$

Here, the injection rate  $\dot{E}_{\text{AGN}}$  is given by

$$\dot{E}_{\text{AGN}} = \epsilon_{\text{MCAF}} \dot{M}_\bullet c^2, \quad (20)$$

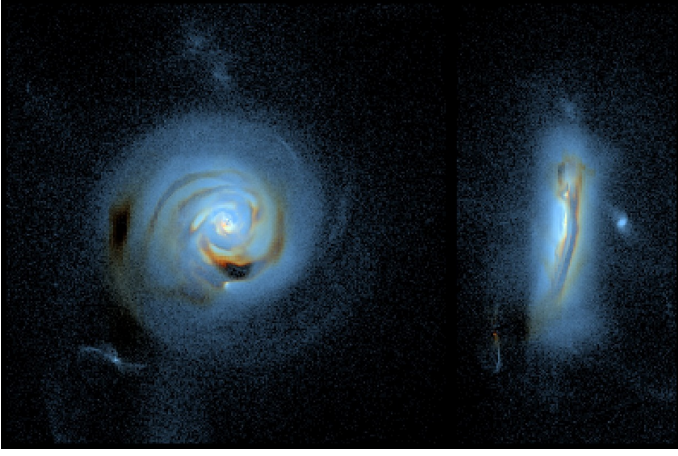
where  $\epsilon_{\text{MCAF}}$  is given by a fourth-order polynomial fit to the simulations of McKinney et al. (2012). More precisely, we use the same set of simulations as for the spin evolution (see Sect. 2.4.4), and sum the contributions of winds and jet based on Table 5:  $\epsilon_{\text{MCAF}} = \epsilon_j + \epsilon_{w,o}$  (using respectively  $\eta_j$  and  $\eta_{w,o}$  in their notations).

#### 2.4.6. BH radiation

In addition to the thermal and kinetic energy injection described in the previous section, we release ionizing energy radiation from the BHs to represent the contribution of AGN to the ionizing radiation field. For this, we apply the method presented in Trebitsch et al. (2019). We release radiation at each fine timestep, and the amount of radiation released in each frequency bin is given by the luminosity of the quasar in each band. In this section, we highlight the main aspects of our AGN SED model, and leave the details to Appendix B.

While the implementation of the photon injection is similar to the work of Bieri et al. (2017), it differs in the spectrum assumed for the AGN. Indeed, instead of a constant SED inspired by the averaged spectrum of Sazonov et al. (2004), we model the SED of the radiation produced by the accretion onto the BH as a multi-colour black-body spectrum corresponding to a Shakura & Sunyaev (1973) thin disc, and extend it at high energy with a power-law  $\alpha_{\text{UV}} = -1.5$ , consistent with the value derived by





**Fig. 3.** Mock (rest-frame) *ugr* image of a  $M_{\star} = 4 \times 10^{10} M_{\odot}$  galaxy at  $z = 6$ , accounting for the dust attenuation along the line of sight. The almost edge-on panel on the right shows a clear dust lane.

Lusso et al. (2015) for a sample of high redshift quasars. We then approximate the whole spectrum with a piecewise power-law for simplicity. We assume that  $f_{\text{IR}} = 30\%$  (consistent with the Sazonov et al. 2004 spectrum) of the bolometric luminosity of the disc is absorbed by dust and re-emitted as IR radiation (which will participate to the quasar mode feedback), that we do not model here, thus leaving a total luminosity  $L_{\text{rad}} = 0.7 \epsilon_r \dot{M}_{\bullet} c^2$  available for the radiation. To get the AGN luminosity in each frequency bin, we integrate the resulting SED in each frequency interval. Similarly, we compute the average photon energy in each bin, as well as the energy-weighted and photon-weighted interaction cross sections (see Rosdahl et al. 2013 for details on the role of these quantities).

As the disc profile is a function of  $M_{\bullet}$ ,  $\dot{M}_{\bullet}$  and  $a$ , the multi-colour black-body spectrum of the AGN will also depend on these parameters. To limit the computational cost, we pre-calculate all the quantities that depend on the AGN SED, and only interpolate between these values over the course of the simulation. Because the shape of the spectrum is weakly sensitive to the value of the spin parameter, we adopt the shape corresponding to  $a = 0$ . We have ensured that the adopted AGN SED yields an average spectrum similar to the AGN SED used in Volonteri et al. (2017) for a population of growing black holes at  $z \sim 6$ .

Finally, we note again that the thin disc solution assumes that the accretion flow is radiatively efficient: we therefore only release radiation when  $\chi \geq \chi_{\text{crit}}$ .

## 2.5. Dust model

We include in OBELISK a subgrid model for the evolution of dust treated as a separate constituent to that of metals locked in the gas phase. The details of the model will be described in a future work Dubois et al. (in prep), and we present here the main features. Our model assumes that dust grains are released in the ISM via SN explosions, grow in mass via accretion of gas-phase metals, and are destroyed by SN explosions and via thermal sputtering. Fig. 3 shows the resulting dust attenuation on a mock (rest-frame) *ugr* image of a  $z \sim 6$  galaxy with mass  $M_{\star} = 4 \times 10^{10} M_{\odot}$ .

Specifically, we consider that all dust grains belong to one single population and are perfectly coupled to the gas (no dust drift relative to gas), i.e. they can be described with only one scalar  $D$  describing the local dust mass fraction (this is similar to the approach taken by e.g. McKinnon et al. 2017 and Li et al.

2019). This scalar is passively advected like the metallicity  $Z$ , now representing the total metal mass fraction (in the gas phase as well as locked in the dust). We further neglect the size distribution of grains and assume that all grains have a unique size  $a_d$ . While the size distribution could in principle be followed in the model (e.g. McKinnon et al. 2018), this would come at a substantial memory overhead. In practice, we assume that the grains have an average size of  $a_g = 0.1 \mu\text{m}$  and density  $\mu_g = 2.4 \text{ g cm}^{-3}$ . Finally, we stress that the dust is purely passive in our simulation: while in reality, a fraction  $D/Z$  of the metals is locked into dust, we do not account for this when estimating the contribution of metal cooling.

When supernovae release metals in the ISM, we assume that a fraction  $f_{d,\text{SN}} = 50\%$  of these metals are in the form of dust, and that the rest is in the gas phase. Because this parameter is very poorly constrained, we choose a value which falls in the middle of the range explored by Popping et al. (2017), who found that the value of  $f_{d,\text{SN}}$  mostly affects the dust content of low-mass, low-metallicity galaxies. The high-velocity shocks produced by the SN explosion will also partially destroy the dust already present near the SN site. The mass  $\Delta M_{\text{dest,SN}}$  of the dust destroyed in these events is related to the mass  $M_{s,100}$  of gas shocked at above  $100 \text{ km s}^{-1}$  via

$$\Delta M_{\text{dest,SN}} = 0.3 \frac{M_{s,100}}{M_{\text{gas}}} M_d, \quad (21)$$

where  $M_{\text{gas}}$  is the local gas mass and  $M_d$  is the local dust mass; meaning that 30% of the gas mass in the shocked gas is destroyed. We estimated the shocked gas mass by taking the Sedov solution in a homogeneous medium following McKee (1989):

$$M_{s,100} \simeq \frac{E_{\text{SN}}}{0.736 (100 \text{ km s}^{-1})^2} \simeq 6800 E_{\text{SN},51} M_{\odot}. \quad (22)$$

We model the dust destruction via thermal sputtering following Draine & Salpeter (1979), with a destruction timescale given by (Draine 2011, chapter 25):

$$t_{\text{sput}} \simeq 0.1 \left( \frac{a_g}{0.1 \mu\text{m}} \right) \left( \frac{n_{\text{gas}}}{\text{H cm}^{-3}} \right)^{-1} \left( 1 + \left( \frac{T}{10^6 \text{ K}} \right)^{-3} \right) \text{ Myr}. \quad (23)$$

In parallel, the dust content can grow in mass via accretion of metals from the gas phase. We estimate the competition between dust growth and destruction using an approach similar to that of Dwek (1998, albeit simplified) or Novak et al. (2012):

$$\dot{M}_{\text{growth}} = \left( 1 - \frac{M_d}{M_Z} \right) \frac{M_d}{t_{\text{growth}}} - \frac{M_d}{t_{\text{sput}}}, \quad (24)$$

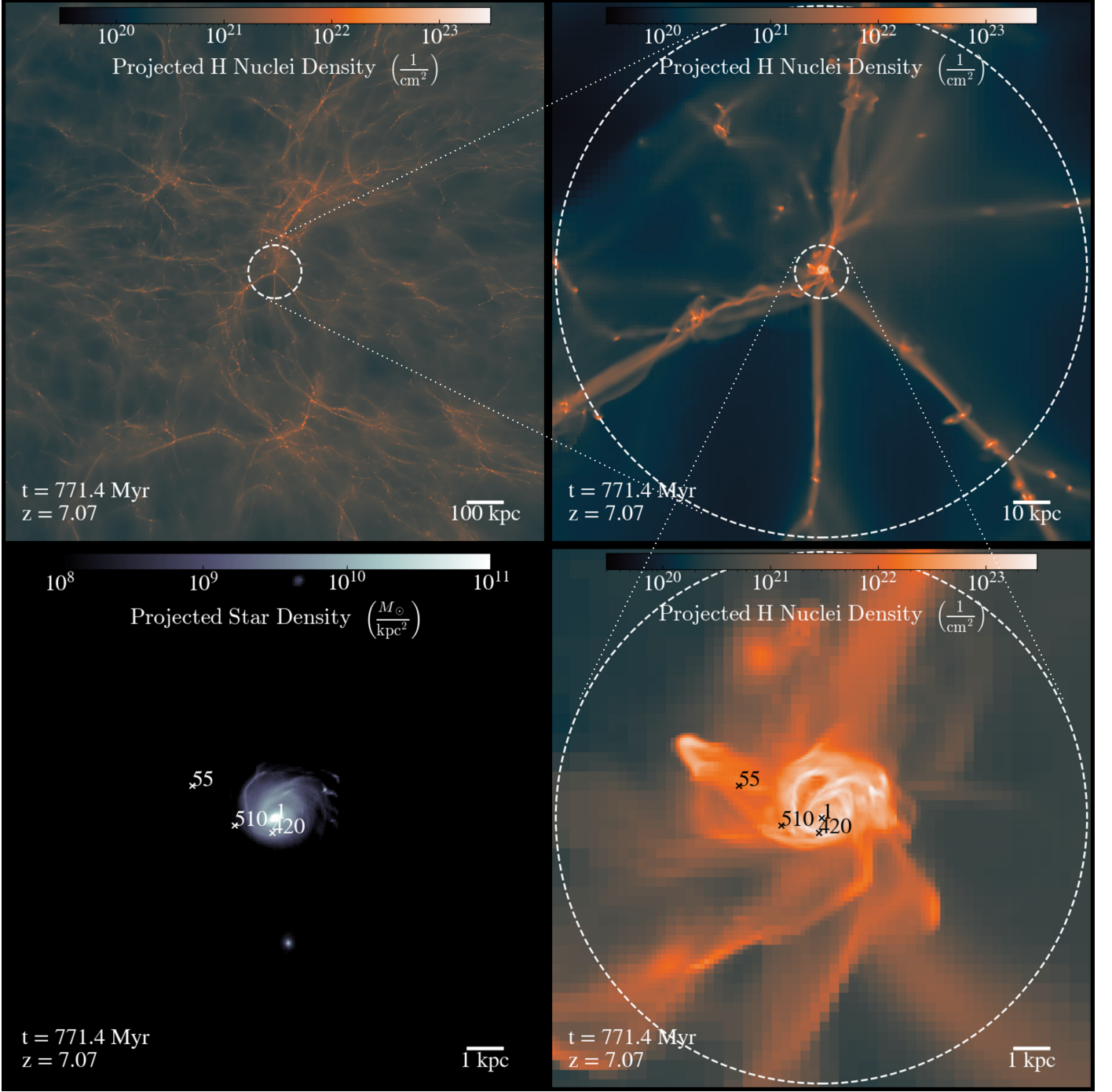
where  $M_Z$  the local metal mass in both the dust and gas phases, and  $t_{\text{growth}}$  is the dust growth timescale

$$t_{\text{growth}} = 100 \frac{1}{\alpha(T)} \left( \frac{a_g}{0.1 \mu\text{m}} \right)^{-2} \left( \frac{n_{\text{gas}}}{\text{H cm}^{-3}} \right)^{-1} \left( \frac{T}{20 \text{ K}} \right)^{-\frac{1}{2}} \text{ Myr}, \quad (25)$$

where  $\alpha(T)$  is the sticking coefficient of metals in the gas phase onto dust. The details of the impact of the choice of the sticking coefficient  $\alpha$  are discussed in Dubois et al. (in prep). Here, we use the results from laboratory experiments by Chaabouni et al. (2012):

$$\alpha(T) = 0.95 \frac{1 + \beta \frac{T}{T_0}}{\left( 1 + \frac{T}{T_0} \right)^{\beta}}, \quad (26)$$

with  $T_0 = 56 \text{ K}$  and  $\beta = 2.22$ . Overall, at temperatures below  $T \simeq 3 \times 10^4 \text{ K}$ , dust growth happens faster than destruction via thermal sputtering.

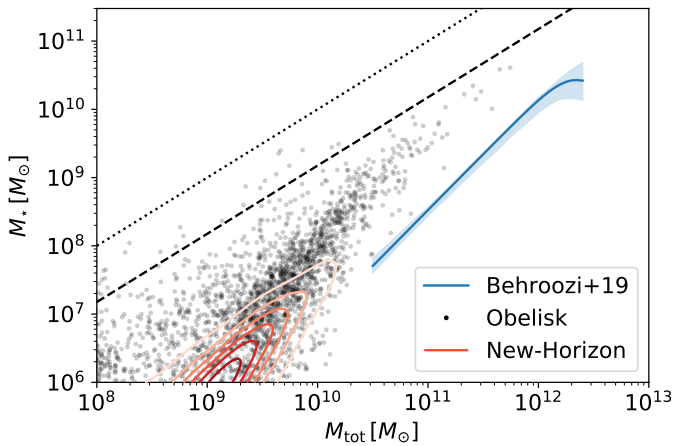


**Fig. 4.** Successive zooms on one of the most massive galaxies in the HR region at  $z \sim 7$ . From the top left clockwise, the first three panels show the hydrogen column density in a region of dimension 1.5 Mpc, 150 kpc and 15 kpc on a side, respectively. The bottom left panel shows the stellar density distribution in the same region as the bottom right panel. The numbered crosses mark the position of the corresponding BHs.

## 2.6. Halo and galaxy identification

We identify galaxies and haloes in each snapshot of the simulation using the ADAPTAHOP halo finder (Aubert et al. 2004; Tweed et al. 2009) with the most massive sub-maximum method (MSM) to separate between host haloes and substructures. In this framework, haloes and subhaloes are groups of particles located at maxima of the density field, and the MSM method requires that the most massive sub-structure is defined as the central object. Compared to previous works using ADAPTAHOP (e.g. Dubois et al. 2014a, for HORIZON-AGN), we have amended the halo finder to identify structures using all collisionless parti-

cles, both stars and dark matter. The DM halo is then identified to the DM component of the (sub-)structure, while the galaxy is defined as all stellar particles in the (sub-)structure. In the following, we will refer interchangeably to (sub-)structures and (sub-)haloes when discussing groups produced by our modified ADAPTAHOP. To be elected as a candidate halo, a structure has to exceed a threshold density  $\rho_t$ . Instead of using a fixed density (e.g. 200 times the average or critical density), we use the fit from Bryan & Norman (1998), yielding an density of roughly  $\rho_t \lesssim 178$  for the redshift interval studied here. We only consider structures with more than  $n_{\text{members}} \geq 100$  particles. No-



**Fig. 5.** Stellar-to-halo mass relation at  $z = 6$  in OBELISK (black dots), compared to the model of Behroozi et al. (2019, in blue), and to the NEW-HORIZON simulation Dubois et al. (in prep) (in red) at the same redshift but for an average environment. The dashed and dotted lines indicate the 1:1 relation and the universal baryon fraction, respectively. For haloes with  $M_{\text{vir}} > 10^{9.5} M_{\odot}$ , the stellar mass in OBELISK exceeds that of NEW-HORIZON, highlighting the role of the overdensity.

tably, we only consider in the analysis galaxies with more than 100 star particles. This yields 52 428 (69 235) host haloes, 41 244 (116 663) subhaloes and 12 549 (67 478) galaxies at  $z = 6.0$  ( $z = 3.53$ ), respectively.

Once a (sub-)halo has been identified, we fit a tri-axial ellipsoid to it, and we find the largest ellipsoid for which the virial theorem is verified. We use this ellipsoid to define the virial radius  $R_{\text{vir}}$ , and the virial mass  $M_{\text{vir}}$  is the mass enclosed in this ellipsoid. In addition to properties global to the total (DM + stellar) structure, we also measure several quantities separately for the stars and DM component, such as the half-mass radius  $R_{50}$ . For the galaxy, we also compute additional kinematical and morphological informations such as the projected effective radius  $R_{\text{eff}}$ , the star formation rate, or the mass-weighted age and metallicity. Note that a galaxy can correspond to a sub-structure and still lie outside of the virial radius of its parent structure. While we do not separate the populations of central and satellite galaxies in this work, it is worth emphasizing that not all galaxies in sub-structures will correspond to satellite galaxies.

### 3. Galaxies and BH populations in OBELISK

In this first paper based on the OBELISK simulation, we focus on one of the main goals of the project: establishing the respective role of galaxies and accreting black holes in reionizing the large-scale environment of a protocluster. In this section, we begin by presenting the global properties of both populations; their contribution to reionization will be discussed in the next section.

We illustrate in Fig. 4 the hierarchy of scales captured in the OBELISK simulation by zooming on one of the most massive galaxies at  $z \sim 7.1$ , with stellar mass  $M_{\star} = 2.3 \times 10^{10} M_{\odot}$ . Clockwise, the first three panels show the hydrogen column density  $N_{\text{H}}$  at large intergalactic scales ( $\sim 1.5$  Mpc), at the halo scale ( $\sim 150$  kpc) where the large scale filaments connect to the galaxy, and at galactic scales ( $\sim 15$  kpc) where the galactic disc is roughly face on. The bottom left panel shows the distribution of stars in the galaxy for the same projection as in the bottom right panel. Even for a galaxy as compact as this one (the effective radius is of order  $\lesssim 200$  pc and 90% of the stellar mass is contained within  $\sim 850$  pc from the centre), we start to capture

some of the structure of the disc (e.g. the spiral arms, and a hint of a bar structure). The numbered crosses in the two lower panels mark the position of each BH, independently of their mass. While there is a massive ( $M_{\bullet} \approx 1.3 \times 10^5 M_{\odot}$ ) BH exactly at the centre of the main galaxy, we can see three other black holes in the image: these are all lower mass BHs ( $M_{\bullet} \approx 3 - 4 \times 10^4 M_{\odot}$ ) and have been brought to the galaxy by previous mergers over the course of its assembly.

We first focus on the galaxies identified in the high-resolution region. Fig. 5 shows the stellar-to-halo mass relation at  $z = 6$  for OBELISK (black dots) compared to estimates from Behroozi et al. (2019, blue line). The red contours show the results from the NEW-HORIZON simulation Dubois et al. (in prep), which features the same subgrid models as OBELISK but focusing on an average environment. There is no overlap between the models from Behroozi et al. (2019) and NEW-HORIZON, but the extrapolation between the two seems reasonably consistent. By comparison, at high halo masses ( $M_{\text{vir}} > 10^{9.5} M_{\odot}$ ), the galaxies in OBELISK are more massive than in NEW-HORIZON; since both simulations share their subgrid models, this is indicative of the effect of the environment.

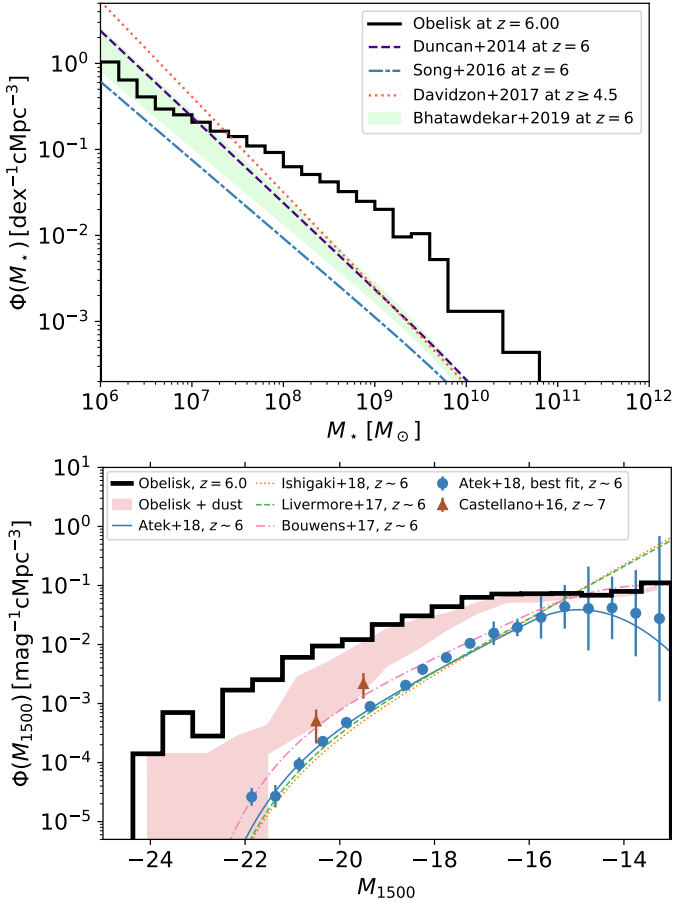
#### 3.1. Galaxy populations

Observations at lower redshift suggest that overdensities are not only comparatively richer in galaxies than the field, but also that the shape of the mass function may depend on the environment (e.g. Davidzon et al. 2016; Tomczak et al. 2017; Papovich et al. 2018). At a lower redshift, Shimakawa et al. (2018) compare the mass function of galaxies protoclusters to the results for field galaxies in COSMOS (Davidzon et al. 2017) at  $z \approx 2 - 2.5$  and found that at a fixed stellar mass, protoclusters have around 10 times more galaxies.

We present in the upper panel of Fig. 6 the galaxy stellar mass function at  $z \sim 6$ , as a thick solid line, and compare our results to average mass functions derived by Duncan et al. (2014, dashed purple line) and Song et al. (2016, dash-dotted blue line) based on CANDELS data, Davidzon et al. (2017, dotted red line) based on the COSMOS survey, and Bhatawdekar et al. (2019, green area) based on the Hubble Frontier Fields. As expected, the shape of the mass function is qualitatively different in our simulation, especially at the high-mass end. Around  $M_{\star} \sim 10^{9-9.25} M_{\odot}$ , we find a number density around  $\Phi \sim 10^{-2} \text{ dex}^{-1} h^3 \text{ cMpc}^{-3}$ . By comparison, Song et al. (2016) find  $\Phi \sim 7 \times 10^{-4} \text{ dex}^{-1} h^3 \text{ cMpc}^{-3}$  and Bhatawdekar et al. (2019) find  $\Phi \sim 10^{-3} \text{ dex}^{-1} h^3 \text{ cMpc}^{-3}$ , around an order of magnitude below OBELISK. At the very low-mass end ( $M_{\star} \lesssim 10^7 M_{\odot}$ ), the number density of galaxies in OBELISK recovers a reasonable agreement with observations (keeping in mind however that, in this mass regime, the assumed morphology of galaxies can change the estimation of the mass function derived by observations by  $\sim 0.5$  dex, Bhatawdekar et al. 2019).

The excess of massive galaxies compared to the fields can also be seen in the UV luminosity function (LF). We measure the UV luminosities of our galaxies by assigning a UV luminosity to each star particle in the simulation based on its age, mass and metallicity following the BPASS v2.2.1 SED (Eldridge et al. 2017; Stanway & Eldridge 2018), and then summing the luminosity of all star particles associated to a galaxy. We show the UV LF measured at  $z \sim 6$  in the OBELISK volume (thick solid line) in the lower panel of Fig. 6, compared to the luminosity functions derived from observations of  $z \sim 6$  lensed galaxies behind clusters by Bouwens et al. (2017); Livermore et al. (2017); Atek et al. (2018); Ishigaki et al. (2018). For clarity, we show the

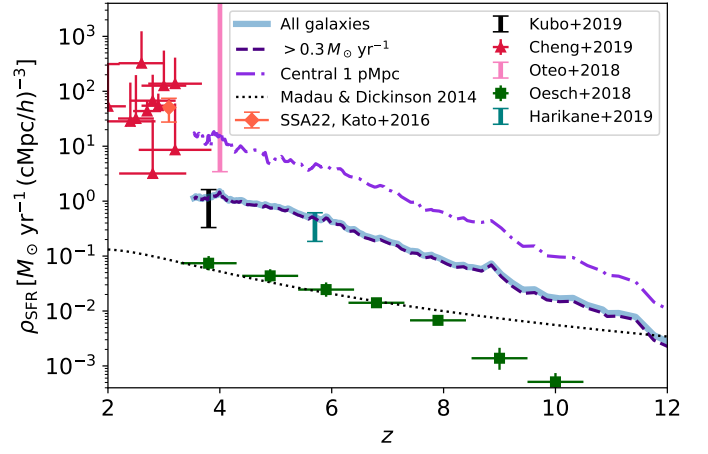




**Fig. 6.** *Top:* Galaxy stellar mass function in our high-resolution region at  $z = 6$  (thick black line) compared to the observational determination of Duncan et al. (2014, dashed purple line), Song et al. (2016, dash-dotted blue line), Davidzon et al. (2017, red dotted line) and Bhatawdekar et al. (2019, green area). *Bottom:* Corresponding intrinsic UV luminosity function (thick black line) and including an estimate for the expected dust attenuation based on the dust present in the simulation (red area). We compare our results to the UV luminosity functions for field galaxies found by Bowler et al. (2015); Bouwens et al. (2017); Livermore et al. (2017); Atek et al. (2018); Ishigaki et al. (2018) and the BDF field from Castellano et al. (2016). Details of the legend are given in the text.

best fit LF resulting from these works and only the actual data points from Atek et al. (2018). We show the data from Bowler et al. (2015) at the bright end, and the determination of the LF by Castellano et al. (2016) for the marginally overdense *Bremer Deep Field* (BDF, Lehnert & Bremer 2003).

Not only do we expect more galaxies in our simulation (due to the fact that we probe a more biased regions), but we also expect galaxies to be, on average, more evolved at a given redshift compared to the field (e.g. Overzier 2016). In a  $\Lambda$ CDM universe, it is expected that the densest structures collapse first, and, due to the so-called halo assembly bias (e.g. Sheth & Tormen 2004; Harker et al. 2006; Borzyszkowski et al. 2017; Musso et al. 2018), that galaxies in denser regions are older and more massive compared to the regions of average density. This is exactly the case here: at  $z = 6$ , galaxies in OBELISK reach masses in excess of  $10^{9-10} M_{\odot}$  and UV magnitude brighter than  $M_{UV} \lesssim -22$ , and are therefore likely to be significantly enriched in dust. To get an estimate of the impact of the UV extinction due to dust, we cast 192 rays from the centre of each galaxy and integrate the dust column density out to the virial radius of the host halo. We



**Fig. 7.** Cosmic SFRD in the OBELISK volume for all galaxies (in light blue), only for galaxies forming stars faster than  $0.3 M_{\odot} \text{ yr}^{-1}$  (dashed dark purple line), and for all galaxies within 1 Mpc of the most massive galaxy (dash-dotted light purple line). The simulation should be compared to the protocluster observations of Kato et al. (2016, green diamond), Kubo et al. (2019, black error bar), Cheng et al. (2019, red triangles) and Harikane et al. (2019, teal error bar). The model of Madau & Dickinson (2014, dotted black line) and the observations of Oesch et al. (2018, green squares) are only here to guide the eye.

then convert this column density to an optical depth using the dust extinction law fits from Gnedin et al. (2008) for the Large Magellanic Cloud. This parametrization assumes that the dust extinction is proportional to the neutral hydrogen column density (their Eq. 5): we rescale the neutral column density by the dust-to-gas ratio measured in our simulation along each line of sight to measure the dust optical depth. We repeat the experiment by casting rays from outside of the half-mass radius of each galaxy, to account for stars outside of the centre of the galaxy that might be less attenuated. This results in the red shaded area: the brighter galaxies are more affected than their faint counterparts. Yet, at all magnitudes, the OBELISK volume contains more galaxies than the field. Using simulations based on constrained initial conditions, Yajima et al. (2015) have found a similar increase in the number density of galaxies compared to their simulation of an average patch of the Universe.

We then compare the evolution of the total SFR density  $\rho_{SFR}$  measured in the OBELISK volume to observations both in the field and in high redshift protoclusters. In Fig. 7, we show the total  $\rho_{SFR}$  as a thick, light blue line. We additionally compute the SFR density for only the galaxies with a SFR higher than  $\dot{M}_{\star} > 0.3 M_{\odot} \text{ yr}^{-1}$  (corresponding to  $M_{UV} \approx -17$ , dashed dark purple line) and for galaxies within 1 physical Mpc from the most massive halo (dash-dotted light purple line) as a proxy for the central region of our protocluster. For comparison, we show the best fit cosmic SFR density from Madau & Dickinson (2014, dotted black line) and the observational constraints of Oesch et al. (2018) as green squares including previous determinations of the cosmic SFR density from Oesch et al. (2013, 2014); Bouwens et al. (2016). The other data points are a very heterogeneous compilation of high redshift protoclusters and overdensities taken from Kato et al. (2016); Harikane et al. (2019); Kubo et al. (2019) and the compilation of Cheng et al. (2019). The enhancement of the SFR density measured in OBELISK compared to the field is in broad agreement with these observations. For instance, Harikane et al. (2019) found  $\rho_{SFR} \approx 0.32 M_{\odot} \text{ yr}^{-1} \text{ cMpc}^{-3}$  at  $z \sim 5.7$  for a sample including both Ly $\alpha$  emitters (LAEs) and sub-millimetre galaxies, assuming an overdensity radius of

10 cMpc. This is very comparable to what we find at  $z \sim 6$ . At  $z \sim 4$ , Oteo et al. (2018) detected an overdensity that could correspond to a protocluster core, with a measured SFR around  $\sim 6500 M_{\odot} \text{ yr}^{-1}$  in a projected area of  $260 \times 310 \text{ kpc}^2$  (physical). The total structure, which might extend over more than  $2.3 \text{ Mpc}^2$  would have a total SFR of  $14\,400 M_{\odot} \text{ yr}^{-1}$ . We bracket these two values in Fig. 7 as the pink error bar.

Similarly, Kubo et al. (2019) quote an average total SFR of  $\sim 2.1 \times 10^3 M_{\odot} \text{ yr}^{-1}$  for their  $z \sim 3.8$  candidate protoclusters. The upper and lower limits of the error bar in Fig. 7 indicate the resulting SFR density assuming that all the star formation happens in the central (physical) Mpc or in a larger 8 arcmin region, equivalent to 3.4 Mpc (physical) in diameter. Finally, at a slightly lower redshift, Kato et al. (2016) report a total SFR of  $\sim 4.7 \times 10^3 M_{\odot} \text{ yr}^{-1}$  for the concentration of dusty star-forming galaxies in the SSA22 field, corresponding to a SFR density of  $\sim 50 M_{\odot} \text{ yr}^{-1} \text{ cMpc}^{-3}$  assuming that the protocluster size is around 1 Mpc (physical). While OBELISK has not reached  $z \approx 3.1$ , it seems that extrapolating the trend of the central SFR density would lead to a reasonable (qualitative) agreement with the SSA22 value.

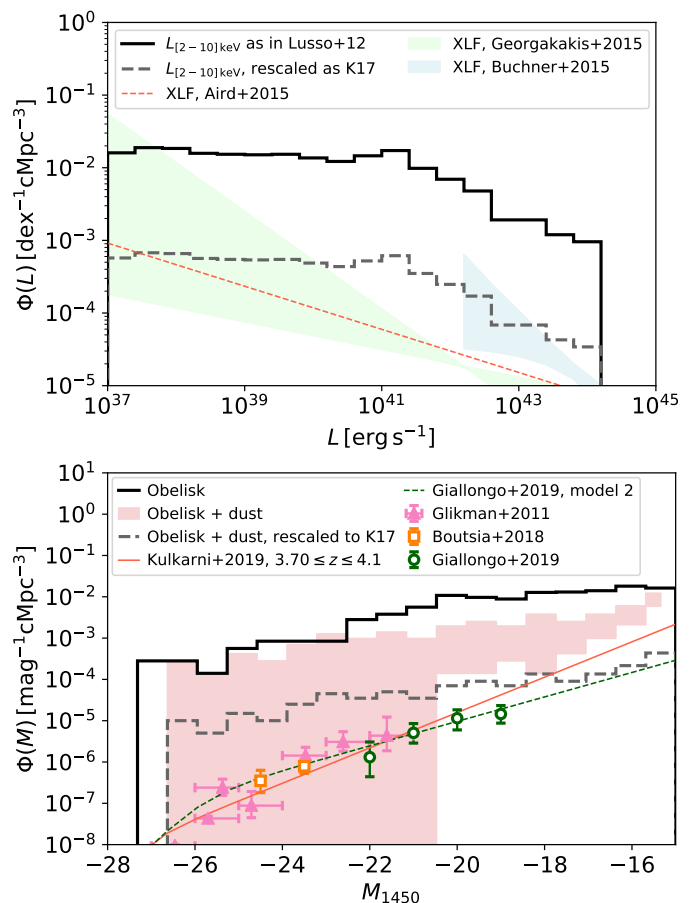
### 3.2. BH populations

Moving on to the BH population of OBELISK, we show in the upper panel of Fig. 8 the AGN luminosity function in the simulation, with the bolometric luminosity function as a solid black line and the hard X-ray luminosity function (XLF) as a dashed red line. We estimate the X-ray luminosity in the  $[2 - 10] \text{ keV}$  band using the bolometric correction from Lusso et al. (2012).

As is the case for the galaxy UV luminosity function, most observational estimates focus on the field, and there are no high redshift samples in overdense regions to compare our results to; the only estimate of the AGN luminosity function in a protocluster (Krishnan et al. 2017) is at  $z = 1.62$ . We discuss first the comparison with the field luminosity function, and then discuss the effect of the overdensity on the comparison.

As expected, our estimated XLF is significantly above the observed XLF in the field (e.g. Ueda et al. 2014; Aird et al. 2015; Buchner et al. 2015; Georgakakis et al. 2015; Miyaji et al. 2015; Vito et al. 2016, 2018). We only show a sample of these observational determinations in the figure for readability. For instance, Buchner et al. (2015) infer  $\Phi(L_X = 10^{43} \text{ erg s}^{-1}) \approx 4 \times 10^{-5} - 2 \times 10^{-4} \text{ dex}^{-1} h^3 \text{ cMpc}^{-3}$  at  $4 \leq z \leq 7$ , while we find a value between 5 and 25 times higher. In a protocluster at  $z = 1.62$  Krishnan et al. (2017) find that the XLF is higher than in the field by a factor  $\sim 28$  at  $L_X = 10^{43} - 10^{44}$ , compatible with our result. If we rescale our estimated XLF by dividing the AGN number density by a constant factor 28 (dashed grey line), we find a reasonable agreement between the simulation and the observed XLF from Buchner et al. (2015).

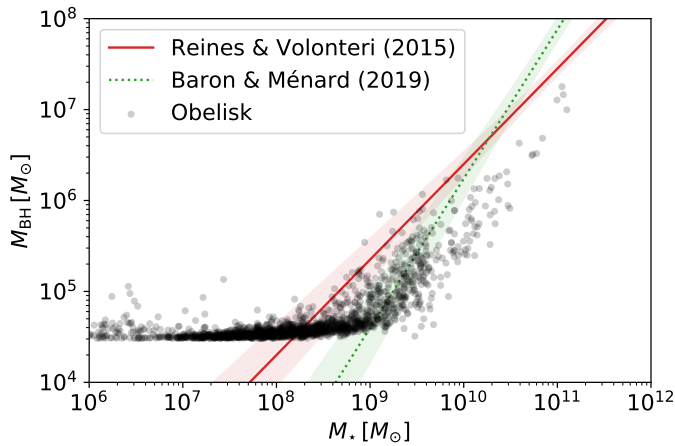
In the lower panel of Fig. 8, we repeat the same exercise for the UV luminosity of the AGN population in OBELISK, compared to a sample of observations from Glikman et al. (2011), Boutsia et al. (2018), and Giallongo et al. (2019), and to the fits by Giallongo et al. (2019) and Kulkarni et al. (2019). Here, the AGN overdensity is even more pronounced: this is in part because we have included all accreting BHs from the simulation, and a significant fraction of these are thought to be obscured (e.g. Vito et al. 2018). Obscuration affects somewhat the XLF, via a correction for Compton-thick AGN, although this is mitigated in high-redshift observations because redshift shifts the restframe band to higher energies, and strongly the UVLF. Trebitsch et al. (2019) using a zoom simulation of a high-redshift galaxy show



**Fig. 8.** *Top:* AGN X-ray luminosity functions at  $z \sim 4$  from the simulation (solid black line) and rescaled to the field (dashed grey line) using the AGN excess found by Krishnan et al. (2017) at  $z \approx 1.6$ . We compare our bolometric LF to the fit of Hopkins et al. (2007, blue line), our X-ray LF to the results of Buchner et al. (2015, red area), Aird et al. (2015, thin dashed line), and Georgakakis et al. (2015, green area). *Bottom:* AGN UV LF at  $z \sim 4$  without taking any obscuration into account (solid black line) and using the dust present in the simulation (red area, see text for details), compared to the data of Glikman et al. (2011, pink triangles), Boutsia et al. (2018, orange squares) and Giallongo et al. (2019, green circles) and to UV LF fits from Giallongo et al. (2019, dashed green line) and Kulkarni et al. (2019, red line). We rescale again our AGN UV LF with a lower limit of the dust attenuation (dashed grey line) following the excess found by Krishnan et al. (2017).

that UV observations capture only about 3% of the accreting AGN. We apply a correction for obscuration, similar to what we have done in Fig. 6 to allow for a fairer comparison, which however still does not account for the OBELISK region being an overdensity, which as noted above is an additional factor  $\sim 28$  in X-ray. The red area on the plot shows the effect of dust attenuation on the UVLF, based on the dust column density measured within 10 kpc from each BH, with the upper limit taking only into account dust beyond 100 pc of each BH (mimicking dust attenuation coming from the transfer through the ISM and CGM). We further rescale this upper limit on the UVLF using the factor  $\sim 28$  from the X-ray (dashed grey line).

While there is, to the best of our knowledge, no observational constraint on the AGN luminosity function in protoclusters at  $z > 1.6$ , there has been evidence for enhanced AGN activity in dense environments at  $z \sim 2 - 3$  (e.g. Lehmer et al. 2009; Digby-North et al. 2010). An interesting comparison is in SSA22 at  $z \sim 3$ , where there have been studies on both the galaxy



**Fig. 9.** BH mass vs stellar mass at  $z = 4$  in OBELISK (black dots) compared to the observational constraints from Reines & Volonteri (2015) and Baron & Ménard (2019). The plateau at low  $M_*$  corresponds to the mass of the BH seeds.

and AGN populations. Lehmer et al. (2009) find that the AGN fraction in SSA22 is increased by a factor 6 with respect to the field, which means that AGN are enhanced more than galaxies in protoclusters. Similarly, Digby-North et al. (2010) and Krishnan et al. (2017) find that the enhancement in AGN is higher than expected simply taking into account the higher number of galaxies, i.e., a higher fraction of galaxies in protoclusters exhibit AGN activity. This has been suggested to be a byproduct of protoclusters hosting more massive galaxies (see e.g. Yang et al. 2018, for a discussion) rather than of enhanced interactions or more sustained BH growth.

Finally, in Fig. 9 we compare the BH mass versus the galaxy stellar mass at  $z = 4$  in OBELISK with the local scaling relations from Reines & Volonteri (2015) and Baron & Ménard (2019). Consistent with previous work (e.g. Dubois et al. 2015; Bower et al. 2017; Habouzit et al. 2017; Weinberger et al. 2017), we find that BH growth is inefficient in galaxies with masses below  $M_* \lesssim 10^{9.5} M_\odot$ , as indicated by the plateau around the BH seed mass. In the high mass regime, BHs grow rapidly and start to follow scaling relations similar to those observed at  $z \sim 0$ . We leave the detailed study of the black hole population and its growth history for a future work.

In the context of the  $z \gtrsim 6$  Universe, we note that we do not find in OBELISK any BH with mass of order  $M_\bullet \gtrsim 10^9 M_\odot$ , the expected mass range of BHs powering the brightest quasars observed in the reionization era (e.g. Mazzucchelli et al. 2017b; Bañados et al. 2018; Reed et al. 2019, for recent results). This is entirely expected because these are rare objects with a number density of order  $\sim 10^{-9} \text{ cMpc}^{-3}$  (Wang et al. 2019) and can therefore only be modelled in simulations representing a much larger volume than HORIZON-AGN, not even BLUETIDES simulates a large enough volume (Di Matteo et al. 2017). By construction, as the OBELISK region was selected as the most massive region in the HORIZON-AGN box at  $z \sim 2$ , it corresponds to a number density of order  $\sim 10^{-6} h^3 \text{ cMpc}^{-3}$ : we do not expect our volume to contain any bright  $z \gtrsim 6$  quasar. Regardless, we note that OBELISK contains a large number of AGN that can in principle act as sources of reionization, as we will explore in Sect. 4.2.

## 4. Reionization of the OBELISK universe

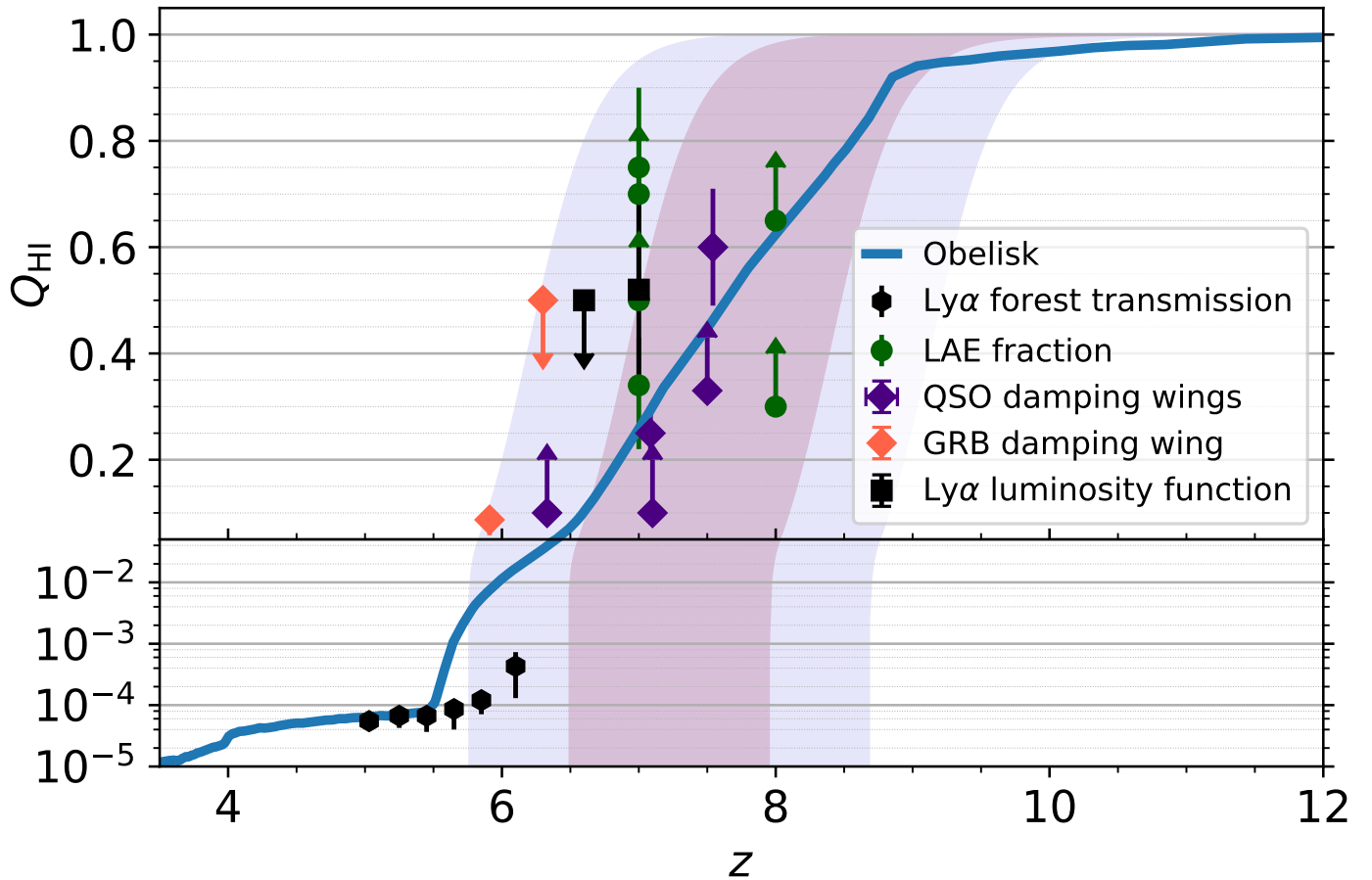
Now that we have established the main properties of the populations of sources (and in particular that the distribution of sources is in line with the expectations for protocluster environments), we can focus on their respective role in the reionization history of the OBELISK simulation.

### 4.1. Hydrogen reionization history

The global volume-weighted neutral fraction of hydrogen  $Q_{\text{H I}}$  in the high-resolution volume is presented as the thick blue line in Fig. 10. The reionization process starts when the first stars are born, and by  $z_{50} \simeq 7.63$ , half of the volume is reionized. We identify the redshifts at which 1%, 10%, 50%, 90% and 99% of the volume is ionized as  $z_{01} = 11.13$ ,  $z_{10} = 8.68$ ,  $z_{50} = 7.63$ ,  $z_{90} = 6.58$ , and  $z_{99} = 5.92$  respectively; corresponding to a reionization duration  $\Delta z = z_{99} - z_{10} = 2.8$  ( $\Delta t \simeq 385 \text{ Myr}$ ), broadly consistent with the estimates of Robertson et al. (2015). The dark and light shaded areas in Fig. 10 correspond to the  $1\sigma$  and  $2\sigma$  constraints on the redshift of reionization from the cosmic microwave background measurements of the Planck mission (Planck Collaboration et al. 2018), with a reionization midpoint  $z_{\text{re}} = 7.67 \pm 0.73$ . We also compare the OBELISK reionization history to a selection of observational constraints: black hexagons correspond to the measurements of the Lyman- $\alpha$  forest transmission (Ly $\alpha$  forest, Fan et al. 2006b), the green circles show constraints on the IGM opacity from the fraction of Lyman- $\alpha$  emitters in Lyman-break galaxy samples (Schenker et al. 2014; Ono et al. 2012; Pentericci et al. 2014; Robertson et al. 2013; Tilvi et al. 2014), the purple diamonds show measurements from quasar damping wings by Mortlock et al. (2011); Schroeder et al. (2013); Bañados et al. (2018); Ľurovčíková et al. (2019), the red diamonds show similar measurements on gamma-ray bursts (GRB, Totani et al. 2006, 2016), and the black squares from Ouchi et al. (2010); Ota et al. (2008) represent constraints derived from the evolution of the Lyman- $\alpha$  luminosity function. Some of these data points come from the compilations of Bouwens et al. (2015). Overall, we find that the simulation agrees with most observations in terms of reionization history, despite the fact that we focus on an overdense region. Interestingly, the simulation manages to capture residual neutral fraction after reionization is complete at  $z < 6$  similar to what is observed. We discuss this point further in Appendix C.

We illustrate the reionization process on the scale of our high resolution region in Fig. 11. The four rows represent four different snapshots of the simulation, at  $z = 11.13, 8.68, 7.63, 5.92$ , corresponding to a ionized volume fraction of 1%, 10%, 50%, and 99%. Each panel is a projection in a  $20 \times 20 \times 2 h^{-3} \text{ cMpc}^3$ . The first column shows both the gas density and the ionization state of the gas. Brighter regions on the maps are denser, and colourful regions are ionized while grey regions are still neutral. We see that as ionized bubbles grow and expand, the matter collapsed in haloes and voids appear on large scales. While the bottom row corresponds to a 99% ionized volume, we see some still neutral regions remaining on the map: this illustrates well the contours of the high-resolution region. The second column shows the temperature distribution: ionized regions reach  $T \gtrsim 1 - 2 \times 10^4 \text{ K}$ , and we can identify hot bubbles around the knots of the cosmic web that are created by feedback from AGN and supernovae. The third column presents the ionizing flux in units of  $J_{21} = 10^{-21} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1} \text{ cm}^{-2}$ . We can again identify the ionized regions at early time (before overlap) and the contours of the high-resolution region in the last panel. Interest-





**Fig. 10.** Evolution of the volume-filling fraction of neutral gas in the OBELISK volume. The shaded area indicate the cosmic microwave background constraints from *Planck*, and the data points show various observational constraints (see text for details). While reionization starts at  $z > 12$ , the volume is only significantly ionized ( $Q_{\text{HI}} < 90\%$ ) at  $z_{10} \sim 8.68$  and reionization finishes around  $z_{99} \sim 5.92$ , with a midpoint around  $z_{50} \sim 7.53$ .

ingly, even when the region is completely ionized, we note that the ionizing flux varies by more than two orders of magnitude.

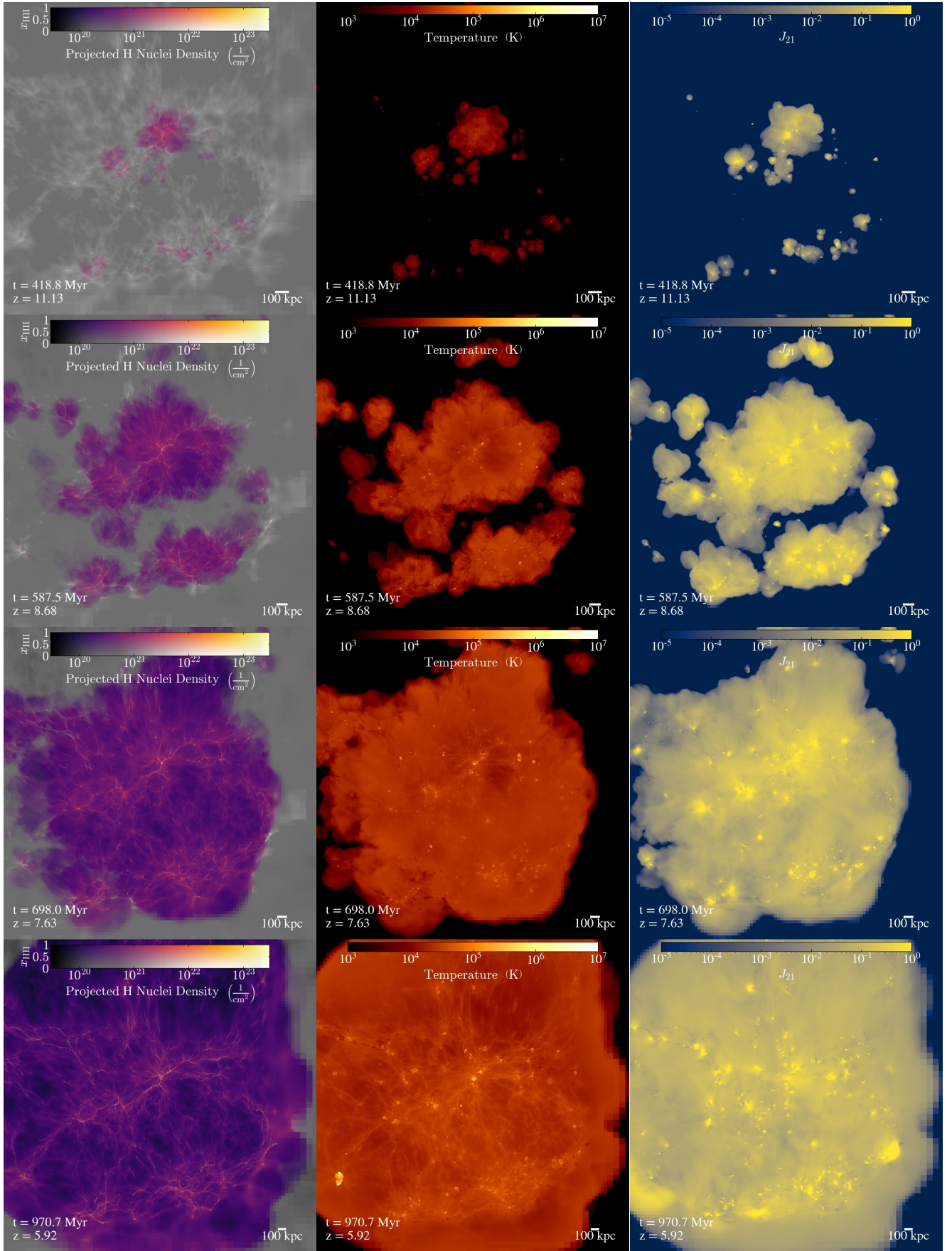
This good agreement with observational constraints is not necessarily expected: while overdense regions such as the one we are modelling here are rich in ionizing sources (as shown in the previous section), there is also more gas to reionize compared to a field environment. Earlier studies have found conflicting results: for instance, the simulations of Ciardi et al. (2003) have suggested that protoclusters could actually be completely reionized later than average environments, while the beginning of the reionization process happens earlier. Using a semi-analytical approach, Kulkarni & Choudhury (2011) found that on the contrary, overdense regions are reionized earlier in their model. More recent very large scale simulations by e.g. Iliev et al. (2006, 2014) point toward the opposite direction: they found a positive correlation between the reionization redshift and the overdensity of the region. We should note however that reionization simulations have tremendously progressed these past few years: for instance, OBELISK has a mass resolution 100 times better than the Ciardi et al. (2003) simulation, and evolves the radiation field directly coupled with the hydrodynamical evolution. Nevertheless, this suggests that the complex balance between an increased number of both sources and sinks of ionizing photons needs to be studied in more detail. In OBELISK, we find that even though our reionization mid-point is fully consistent with constraints on the average reionization redshift, the end of reionization is a bit delayed:  $z_{99} = 5.92$ . This can be seen again in Fig. 10: the sim-

ulation overshoots the data points from Fan et al. (2006b). It is entirely possible that this comes from the details of the simulation setup, but this is still suggestive that the very end of reionization happens later in OBELISK than in an average environment. By comparison, with a similar simulation setup<sup>5</sup> (albeit with higher spatial resolution), Rosdahl et al. (2018) finds that the SPHINX volume is 99% reionized at  $z \sim 7$ . This is reminiscent of the results of e.g. Aubert et al. (2018), who showed using the CoDa I-AMR simulation that in environments like the Local Group, reionization of progenitors of the most massive haloes starts earlier but last significantly longer.

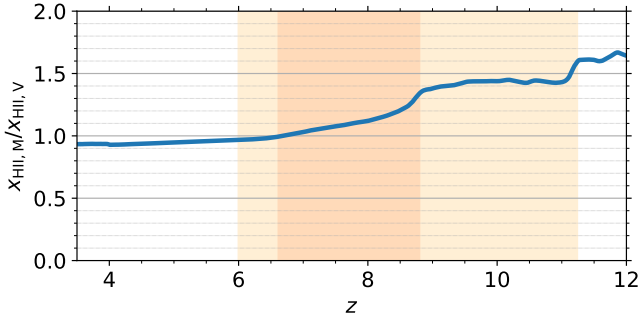
To explore this picture in which reionization happens first inside-out and then outside-in, we follow the approach of Iliev et al. (2006); Bauer et al. (2015) and show in Fig. 12 the ratio of the ionized mass fraction to the ionized volume fraction,  $x_{\text{H II},\text{M}}/x_{\text{H II},\text{V}}$ . This ratio is directly related to the (over)density of ionized gas  $\delta_{\text{H II}}$  via<sup>6</sup>  $x_{\text{H II},\text{M}}/x_{\text{H II},\text{V}} = \rho_{\text{H II}}/\bar{\rho} = 1 + \delta_{\text{H II}}$ , with  $\rho_{\text{H II}}$  the density of ionized hydrogen in the volume and  $\bar{\rho}$  the universal cosmic hydrogen density. Here, the light (dark) shaded region marks the  $z_{01}-z_{99}$  ( $z_{10}-z_{90}$ ) redshift intervals. We see that at early times, the ratio is above unity, indicating that  $\delta_{\text{H II}} > 0$ : early ion-

<sup>5</sup> This also comes from the fact that Rosdahl et al. (2018) used the v2.0 of the BPASS library. Using BPASS v2.1.0, they find a significantly delayed reionization Rosdahl et al. (in prep).

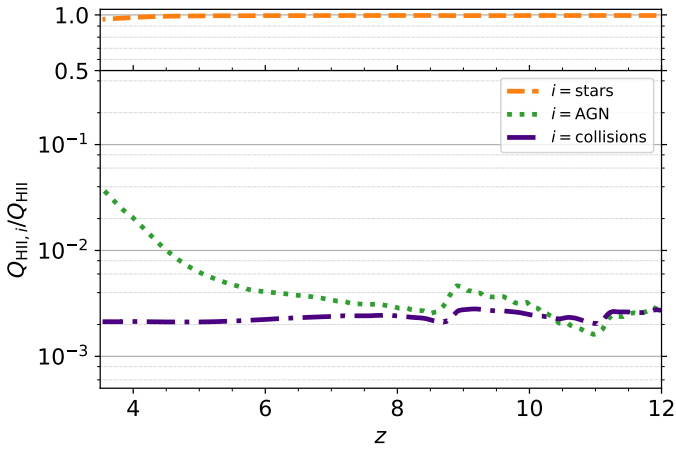
<sup>6</sup> If we note  $M_{\text{box}}$  and  $V_{\text{box}}$  the mass and volume of the box, we have:  $x_{\text{H II},\text{M}}/x_{\text{H II},\text{V}} = x_{\text{H II},\text{M}}M_{\text{box}}/V_{\text{box}} \times V_{\text{box}}/(x_{\text{H II},\text{V}}M_{\text{box}}) = (x_{\text{H II},\text{M}}M_{\text{box}})/(x_{\text{H II},\text{V}}V_{\text{box}}) \times V_{\text{box}}/M_{\text{box}} = M_{\text{H II}}/V_{\text{H II}} \times V_{\text{box}}/M_{\text{box}} = \rho_{\text{H II}}/\bar{\rho}$ .



**Fig. 11.** Illustration of the reionization of the volume. The four rows are four different snapshots at  $z = 11.13, 8.68, 7.63, 5.92$ , corresponding to a ionized volume fraction of the OBELISK universe of 1%, 10%, 50%, and 99%. The left column shows the gas density (brighter is denser) with the coloured vs. greyscale regions indicating ionized vs. neutral gas. The central column shows the gas temperature, with cold neutral gas in black, warm photoionized gas in dark orange, and hot shocked gas in yellow. The ionized regions can be mapped in the right panel to the local ionizing flux in units of  $J_{21} = 10^{-21} \text{ erg s}^{-1} \text{ Hz}^{-1} \text{ sr}^{-1} \text{ cm}^{-2}$ . All three columns are mass-weighted projections.



**Fig. 12.** Ratio of the mass-weighted ionized fraction  $x_{\text{H II, M}}$  to the volume-weighted ionized fraction  $x_{\text{H II, V}}$ , showing how reionization happens inside-out: overdense regions get ionized first by their central sources ( $x_{\text{H II, M}} > x_{\text{H II, V}}$ ), followed by voids. The final ratio is just below 1 because the gas in the densest regions (i.e. haloes and filaments) can recombine, so that  $x_{\text{H II, M}} \lesssim x_{\text{H II, V}}$  after reionization is complete. The light (dark) shaded region marks the  $z_{01} - z_{99}$  ( $z_{10} - z_{90}$ ) redshift intervals.

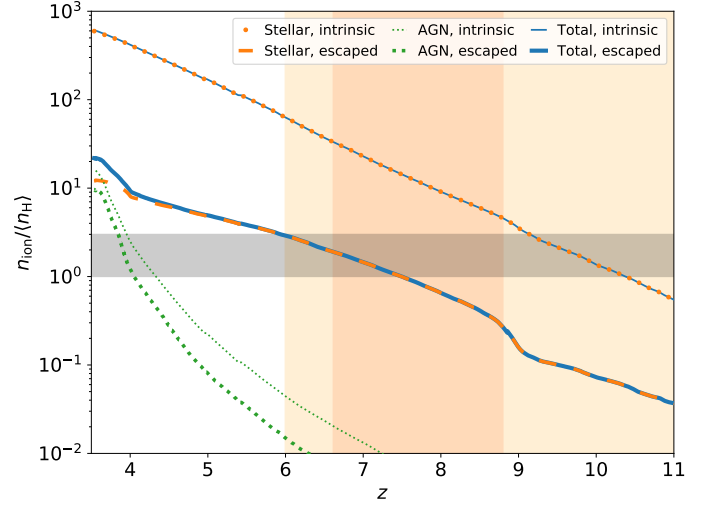


**Fig. 13.** Contributions of various sources to the volume-weighted ionization: at all times, most of the ionized volume is ionized by stellar populations, while AGN only start to be relevant at  $z \lesssim 4$ .

ized regions typically correspond to overdensities. At later times, this ratio decreases, and by the time the OBELISK universe is 90% ionized, ionized regions are typically at the average density. Interestingly,  $x_{\text{H II, M}} / x_{\text{H II, V}}$  seems to converge to a value right below unity: this is because the collapsed regions (e.g. haloes) are included in this analysis, and they represent the densest regions where gas can recombine efficiently.

#### 4.2. Relative contribution of galaxies and black holes

Using the photon tracer method of Katz et al. (2018), we can now to relate the source populations described in Sect. 3 to the reionization history presented in the previous section. To this end, we measure what fractions of the ionized volume have been carved out by stellar photons, AGN photons, and collisional ionizations. These fractions are shown in Fig. 13. The orange dashed line corresponds to the contribution of photons of stellar origin, and dominates overwhelmingly the contributions of photons produced by AGN (green dotted line) and collisional ionization (purple dash-dotted line). In the rest of this section, we will use the same colour and line-style convention for all figures, unless stated otherwise. This means that the OBELISK volume is predominantly reionized by stellar populations. We note however that at  $z \lesssim 5$ , the contribution to photo-ionization by AGN becomes



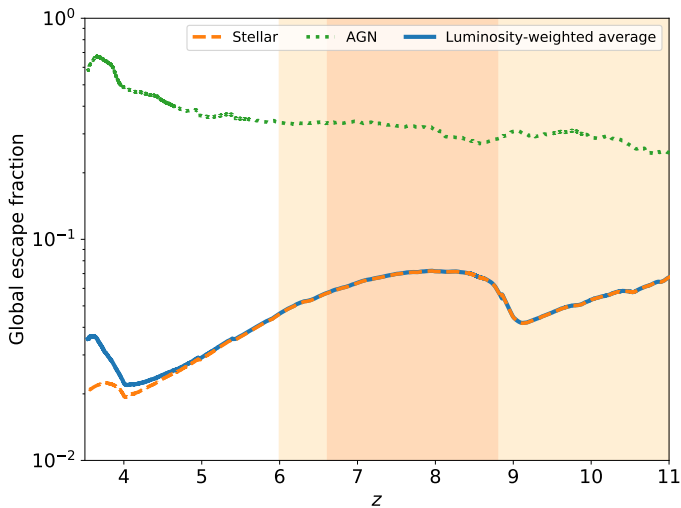
**Fig. 14.** Cumulative ratio of the number of produced (thin lines) and escaped (thick lines) ionizing photon to the average hydrogen density. The solid blue, dashed orange and dotted green lines correspond to the total contribution and that of stellar populations and AGN, respectively. For better readability, we show the intrinsic stellar contribution with dots instead of a line. The background shaded area indicates the  $z_{01} - z_{99}$  ( $z_{10} - z_{90}$ ) redshift intervals. The grey area highlights a photon-to-baryon ratio between 1 and 3, necessary to reionize the Universe. Stellar populations alone provide the necessary number of photons by  $z \sim 6$ .

more and more important, consistent with the picture that AGN maintain the ionization state of the Universe post-reionization (e.g. Haardt & Madau 2012; Becker & Bolton 2013; Faucher-Giguère 2020). Overall, collisional ionizations are completely irrelevant: this is not unexpected, as they predominantly occur in the vicinity of galaxies and haloes, which are already not very volume-filling and already largely photo-ionized.

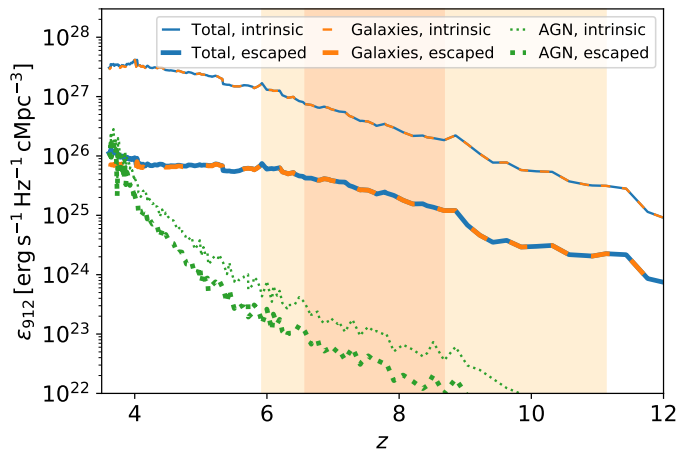
We explore this further in Fig. 14, which shows the cumulative photon-to-baryon ratio for different sources. The intrinsic ratio, i.e. the total number of photons produced divided by the total hydrogen number density, is depicted with thin lines, while the ratio of photons in the IGM over total hydrogen number density is shown with thick lines. Here, we define photons in the IGM (*escaped* photons) as photons in cells where the gas density is below 180 times the average density at that redshift. The light (dark) shaded region marks again the  $z_{01} - z_{99}$  ( $z_{10} - z_{90}$ ) redshift intervals, and the grey horizontal region highlights a photon-to-baryon ratio between 1 and 3, corresponding to the typical photon budget required to reionize the Universe. The most striking feature of this figure is that until  $z \lesssim 4$ , AGN are completely irrelevant to the photon budget of the Universe, despite the fact that we are studying a region that is particularly rich in AGN. Stellar populations alone account for most of the photons, both intrinsically and after transfer through the ISM. Around our reionization midpoint, they have contributed around 1 photon per atom, and by  $z_{99}$ , their contribution reaches the critical value of around 3 photons per hydrogen atom. By comparison, AGN reach this value only at  $z \sim 4$ . We note however that, analogous to Fig. 13, the AGN contribution increases quickly at  $z \lesssim 4$ , but still represents a small fraction of the number of photons in the IGM.

A second aspect of Fig. 13 is that on average, the fraction of photons of stellar origin escaping into the IGM is lower than for the photons of AGN origin. Indeed, the ratio of the thick to thin lines correspond to the population-averaged escape fractions,  $\langle f_{\text{esc}} \rangle$ . We show this global escape fraction in Fig. 15 for the different source populations. We choose to leave the analysis of the





**Fig. 15.** Global escape fraction for different sources with the same colour scheme as in Fig. 14, defined as the total escaped luminosity of a population divided by its intrinsic luminosity. The luminosity-weighted average escape fraction is very close to that of stellar populations at  $z \gtrsim 4$ , highlighting that they largely dominate the photon budget.



**Fig. 16.** Ionizing emissivity, intrinsic (thin lines) and after transfer in the ISM (thick lines), keeping again the same colour coding as in Fig. 14.

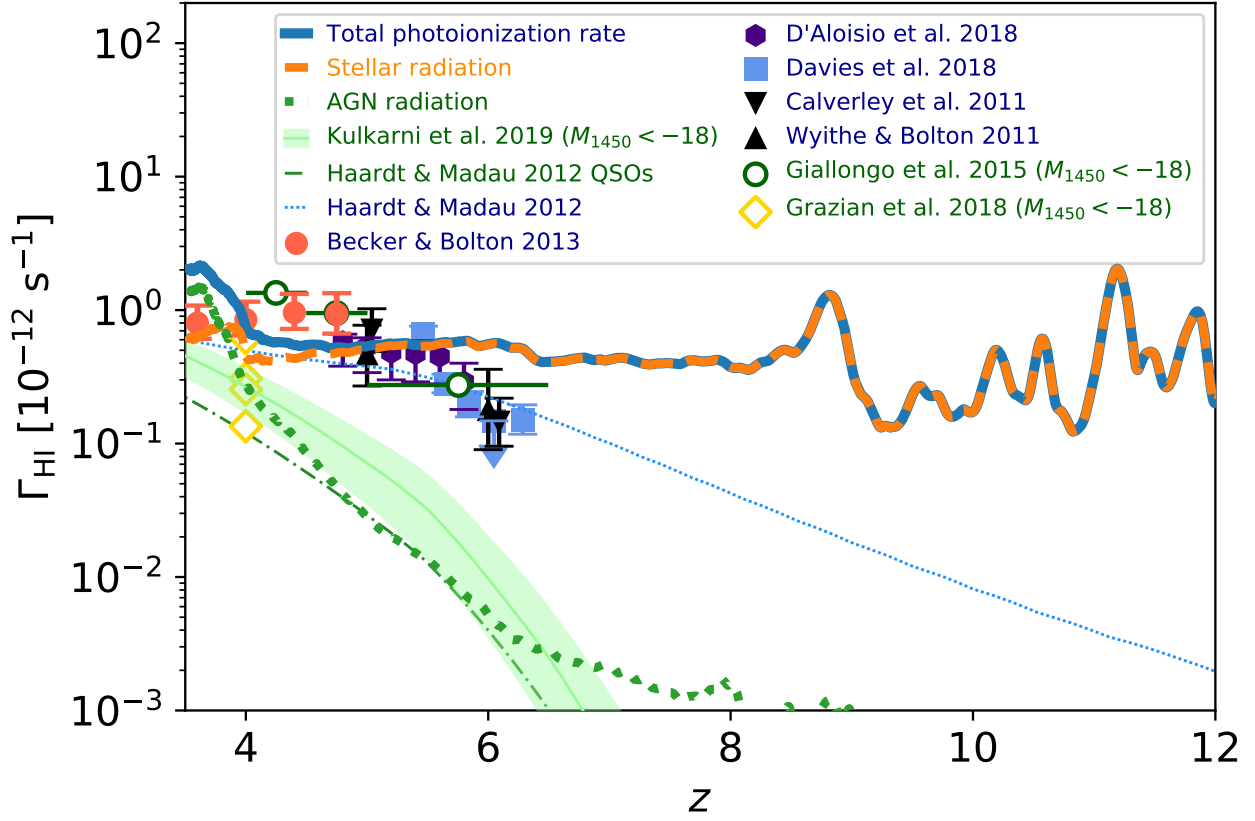
variation of  $f_{\text{esc}}$  on an object-by-object basis to a forthcoming work, but we can note two main results from this. Overall, AGN have a very high (population-averaged) escape fraction, on the order of  $\langle f_{\text{esc}}^{\text{AGN}} \rangle \sim 20 - 50\%$ . This is in good agreement with the estimates of Cristiani et al. (2016) at lower redshift, although this does not correspond to the escape fraction of individual AGN in our simulation. We defer a detailed analysis of  $f_{\text{esc}}^{\text{AGN}}$  to a further study. By comparison, stellar sources exhibit  $\langle f_{\text{esc}}^{\text{star}} \rangle \lesssim 10\%$ , consistent with the values found by other high-resolution reionization simulations (e.g. Kimm & Cen 2014; Ma et al. 2015; Xu et al. 2016; Rosdahl et al. 2018; Yoo et al. 2020). Here again, we stress that this is a value averaged over the whole population, and that the individual  $f_{\text{esc}}^{\text{star}}$  can vary by order of magnitudes between objects and even for a given galaxy over the course of its evolution (e.g. Kimm & Cen 2014; Wise et al. 2014; Paardekooper et al. 2015; Trebitsch et al. 2017). Nevertheless, the population-averaged  $\langle f_{\text{esc}}^{\text{star}} \rangle$  (or  $\langle f_{\text{esc}}^{\text{AGN}} \rangle$ , for that matter) is a useful figure to compare to global reionization models.

Combining our estimate for the source populations and their respective escape fractions, we can now directly estimate the

emissivity of both types of sources, as this is the quantity that typically enters in reionization models. We measure this by summing over all sources (stellar populations or accreting black holes) their ionizing luminosity at  $912 \text{ \AA}$  using their respective SED (see Sect. 2.3.3 and 2.4.6), divided by the total volume of the high-resolution region: this results in the intrinsic emissivity  $\epsilon_{912}^{\text{int}}$  shown in Fig. 16 as thin lines with the same colour-coding as previously. We then combine this with our global escape fractions to get the actual emissivity  $\epsilon_{912} = \langle f_{\text{esc}} \rangle \epsilon_{912}^{\text{int}}$ , shown as thick lines in Fig. 16. We observe the same pattern as previously: galaxies dominate the ionizing photon production, even after transfer through the ISM. The total emissivity  $\epsilon_{912}$  is higher by a factor of  $\sim 10$  compared to models of e.g. Haardt & Madau (2012) or Faucher-Giguère (2020): this results directly from the fact that our source density is significantly higher than in average environments (see Sect. 3).

Finally, we plot in Fig. 17 the  $\text{H I}$  photoionization rate  $\Gamma_{\text{H I}}$  in ionized gas as a function of cosmic time. The thick lines correspond to the simulation: once again the total  $\Gamma_{\text{H I}}$  from all sources is the solid blue line, the contribution  $\Gamma_{\text{H I}}^{\text{star}}$  from stars is illustrated as the dashed orange line, and the contribution from AGN  $\Gamma_{\text{H I}}^{\text{AGN}}$  is shown as the dotted green line. We include the model of Haardt & Madau (2012) as a thin dotted blue line, and the contribution from quasars as a thin dash-dotted green line. The green shaded area corresponds to the determination of the contributions of quasars to  $\Gamma_{\text{H I}}^{\text{AGN}}$  by Kulkarni et al. (2019) integrating the quasar UV LF down to  $M_{1450} = -18$ . We also plot various  $\text{H I}$  photoionization rates measurements taken from the compilation of Kulkarni et al. (2019): Calverley et al. (2011) as downward-pointing black triangles, Wyithe & Bolton (2011) as upward pointing black triangles, Becker & Bolton (2013) as red circles, D’Aloisio et al. (2018) as purple hexagons, and Davies et al. (2018) as blue squares. We also show the value of  $\Gamma_{\text{H I}}^{\text{AGN}}$  derived by Kulkarni et al. 2019 based on the Giallongo et al. (2015) AGN luminosity function as empty green circles. The four empty yellow diamonds correspond to the estimates of  $\Gamma_{\text{H I}}^{\text{AGN}}$  by Grazian et al. (2018) based on the luminosity functions of Giallongo et al. (2015); Gliksman et al. (2011); Parsa et al. (2018); Akiyama et al. (2018), from top to bottom. These estimates give a measure of the effect of the uncertainty on the faint end of the AGN UV luminosity function on the determination of the contributions of quasars to the  $\text{H I}$  photoionization rate. For most of the reionization era, the simulated  $\text{H I}$  photoionization rate remains around  $\Gamma_{\text{H I}} \simeq 5 \times 10^{-12} \text{ s}^{-1}$  and is dominated by the contribution of stellar populations. The initial large fluctuations at  $z \gtrsim 9$  correspond to the very early stages of reionization: at this epoch, only a small fraction of the volume is ionized (see e.g. Fig. 10), so the value of  $\Gamma_{\text{H I}}$  will be very sensitive to the stochasticity of both star formation and the subsequent escape of ionizing radiation. Perhaps more interesting, it appears that AGN start to represent a significant contribution to the  $\text{H I}$  photoionization rate after  $z \lesssim 4$ . This is consistent with estimates from e.g. Kulkarni et al. (2019) for the whole AGN population<sup>7</sup> (i.e. not only overdense regions), and suggests that even in region where the growth of AGN is favoured, they are not important contributors to the reionization of their large scale environment.

<sup>7</sup> We should note that the determination of the faint end of the AGN UV luminosity function has been heavily discussed in the literature recently (e.g. Vito et al. 2016; Parsa et al. 2018, but see also Giallongo et al. 2019 for an extended discussion).

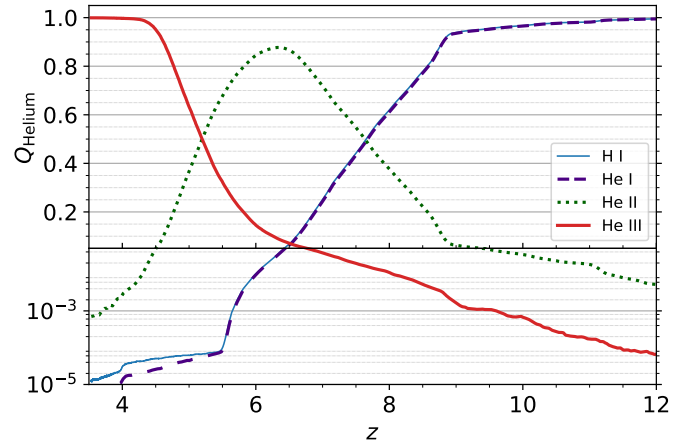


**Fig. 17.** Evolution of the H I photoionization rate in ionized gas (thick blue line) and the contributions of stellar populations (thick orange dashed line) and AGN (thick green dotted line), compared to the models of [Haardt & Madau \(2012\)](#) (thin dotted blue and dash-dotted green lines) as well as the constraints from the homogenized sample of [Kulkarni et al. \(2019\)](#) extending the AGN UV luminosity function down to  $M_{1450} = -18$  and other measurements (see text for details). Blue, orange and green labels correspond to the total photoionization rate, the stellar, and the AGN contribution, respectively. At  $z \gtrsim 4$ , and in particular during the whole epoch of reionization, the H I photoionization rate is completely dominated by the contribution of stellar populations.

#### 4.3. Helium reionization

Because OBELISK includes a model for the production of far-UV radiation by AGN, we can follow the ionization state of helium through cosmic time. Indeed, while helium is singly ionized by the same sources that ionize hydrogen, the second ionization of helium can only happen when through photoionization by photons with energies above 54.4 eV, which are almost entirely produced by AGN. We show the OBELISK helium reionization history in Fig. 18, but leave a more detailed discussion on the properties of He II-reionization sources to a future paper. The volume fractions of neutral, singly- and doubly-ionized helium are shown as a purple dashed line, a green dotted line, and a solid red line, respectively. As expected, the He I volume fraction follows a trend very similar to that of H I reionization. The double reionization of helium starts before He I single reionization is complete, and finishes fairly early ( $z \gtrsim 4$ ) compared the predicted  $z \sim 3$  He II reionization redshift (e.g. [Haardt & Madau 2012](#)).

Albeit perhaps surprising, the fact that helium is doubly reionized earlier than for the average universe still results from our choice to model an overdensity. In the case of H I reionization, the dominant sources are (faint) galaxies, which are not strongly clustered. In particular, the typical size of an H II region around galaxies (before overlap) is small compared to the size of our high-resolution region. The situation is very different for He II reionization, for which the sources are very clustered. The models of [McQuinn et al. \(2009\)](#) or [Dixon et al. \(2014\)](#) suggest that He III bubbles around bright sources can extend beyond



**Fig. 18.** Helium ionization history in the OBELISK volume. The evolution of the neutral fraction of helium,  $Q_{\text{He I}}$ , follows that of the neutral hydrogen  $Q_{\text{H I}}$ . He II reionization is complete by  $z \gtrsim 4$ : our region correspond to one of the growing He III bubbles around massive hosts.

$R_{\text{He III}} \gtrsim 35$  Mpc even at  $z \sim 6$  (when helium is on average not doubly-ionized). This scale is larger than the size of our high-resolution: in other words, we are probing the expansion of He III bubbles around sources in an epoch where the majority of the Universe is still not affected by these bubbles.

## 5. Summary and Conclusions

We have introduced the OBELISK project: a fully coupled radiation-hydrodynamical cosmological simulation that follows the assembly of a massive protocluster during the first few billions of years of its history. This simulation combines the power of modern cosmological codes to simulate a large overdensity at high resolution with the ability to capture self-consistently the evolution of the intergalactic UV background from sources (i.e. galaxies and black holes) to sinks (i.e. neutral gas in the intergalactic medium). While modelling the assembly of a protocluster, the OBELISK simulation resolves haloes down to the atomic cooling limit, therefore capturing the bulk of the potential sources of ionizing photons in this volume. We have presented in some detail the improvements to the subgrid physical models we have used with respect to the previous generation of simulations, e.g. HORIZON-AGN.

In this paper, we concentrated on describing the global properties of galaxies and black holes in our simulation, and their contribution to reionization: indeed, OBELISK is unique in that it follows the radiation produced by both types of sources, allowing us to study directly their relative role in setting the ionization state of the Universe. We focussed on an overdense region in order to probe the contribution of both types of sources in an environment where black holes are expected to make the largest contribution. Our main results are as follows:

1. Stellar populations overwhelmingly dominate over the AGN as sources of reionization, and provide enough ionizing photons to complete reionization alone. At  $z \gtrsim 6$ , in spite of the relatively higher escape fraction, AGN are responsible for less than 1% of the total H I photoionization rate, and represent a similarly low fraction of the total ionizing emissivity.
2. Both star formation and black hole accretion are strongly enhanced in OBELISK compared to an average environment, in good agreement with extrapolations from the protocluster population observed at  $z \gtrsim 2$ .
3. The delicate balance between the larger number of sources and the higher gas density leads to a reionization history close to that of an average environment ( $z_{\text{reion}} \sim 6$ ). The reionization proceeds first inside-out, with the most massive galaxies reionizing their close environment first, before the ionization fronts propagate to the voids.
4. In our protocluster environment, the global escape fractions from stars and AGN during reionization are  $f_{\text{esc}}^* \approx 3 - 8\%$  and  $f_{\text{esc}}^{\text{AGN}} \approx 30 - 40\%$ , respectively.
5. In high densities environments, helium double-reionization happens early, predominantly because of the large density of He II-ionizing AGN sources compared to the field.

The broad picture emerging from this first analysis of the OBELISK simulation is in agreement with the traditional reionization picture, in which AGN are sub-dominant in establishing the ionizing UV background, but contribute to maintaining the Universe reionized in the post-overlap era (e.g. Madau et al. 1999; Haehnelt et al. 2001). Our paper shows that this result holds even in dense regions, where the AGN contribution is expected to be enhanced, but was here found to be still mostly irrelevant at  $z \gtrsim 6$ .

The results presented here are intended as an introduction to the simulation, upon which further analysis will build. For example, we will study if the harder ionizing spectrum of AGN, which can penetrate denser gas, means that AGN play a more important role than stars in ionizing intergalactic filaments, despite the fact that they ionize only a small fraction of the volume (Fig. 13).

In the same spirit, we will benefit from the comparison with a twin simulation that has been run without radiative transfer to study the impact of the inhomogeneous reionization on the suppression of low-mass galaxies, in particular around quasars (see e.g. Mazzucchelli et al. 2017a, for an observational perspective). On the galaxy formation side, further work is needed in conjunction with simulations of field environments (e.g. SPHINX or NEW-HORIZON) in order to understand how the overdensity of the OBELISK volume affects the concurrent growth of galaxies and black holes, and if this type of environment favours AGN activity beyond the global increase in the number density of massive haloes. This type of analysis will obviously greatly benefit from the fact that OBELISK is a sub-volume of the HORIZON-AGN simulation: we will be able to connect our high redshift results to the  $z = 0$  Universe. Finally, we will address more thoroughly the comparison of our simulated galaxies to the observed populations at high redshift, in preparation for surveys with the upcoming JWST, as well as with ground based instrument such as MUSE on the VLT or ALMA. This will of course benefit from the inclusion of a model for dust, allowing us for instance to weigh the contributions of obscured vs UV-bright star formation.

**Acknowledgements.** We would like to thank Rebekka Bieri, Jérémy Blaizot, Sam Geen, Mélanie Habouzit, Sandrine Lescaudron and Pierre Ocvirk for helpful comments and stimulating discussions on this work, and over the inception of the OBELISK project in general. This work made use of v2.2.1 of the Binary Population and Spectral Synthesis (BPASS) models as described in Eldridge et al. (2017) and Stanway & Eldridge (2018). We also used the models produced by Kulkarni (2019). MT and MV acknowledge funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 614199, project 'BLACK'). MT is supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2181/1 - 390900948 (the Heidelberg STRUCTURES Cluster of Excellence). HP is indebted to the Danish National Research Foundation (DNRF132) and the Hong Kong government (GRF grant HKU27305119) for support. TK was supported in part by the Yonsei University Future-leading Research Initiative (RMS2-2019-22-0216) and in part by the National Research Foundation of Korea (No. 2017R1A5A1070354 and No. 2018R1C1B5036146). CC has received partial funding from the Institut Lagrange Paris (ILP) and is supported by the European Union's Horizon 2020 research and innovation programme (under grant agreement No. 818085 GM-Galaxies). This work has made use of the Horizon Cluster hosted by Institut d'Astrophysique de Paris; we thank Stéphane Rouberol for running smoothly this cluster for us. We acknowledge PRACE for awarding us access to Joliot Curie at GENCI@CEA, France, which was used to run most of the simulations presented in this work. Additionally, this work was granted access to the HPC resources of CINES under allocations A0040406955 and A0040407637 made by GENCI. This work has made extensive use of the Yr analysis package Turk et al. (2011) and NASA's Astrophysics Data System, as well as the MATPLOTLIB Hunter (2007), NUMPY/SCIPY Jones et al. (2001) and IPYTHON Perez & Granger (2007) packages.

## References

- Aird, J., Coil, A. L., Georgakakis, A., et al. 2015, MNRAS, 451, 1892  
Akiyama, M., He, W., Ikeda, H., et al. 2018, PASJ, 70, S34  
Atek, H., Richard, J., Kneib, J.-P., & Schaerer, D. 2018, MNRAS, 479, 5184  
Aubert, D., Deparis, N., Ocvirk, P., et al. 2018, ApJ, 856, L22  
Aubert, D., Pichon, C., & Colombi, S. 2004, MNRAS, 352, 376  
Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018, Nature, 553, 473  
Bardeen, J. M. 1970, Nature, 226, 64  
Baron, D. & Ménard, B. 2019, MNRAS, 487, 3404  
Bauer, A., Springel, V., Vogelsberger, M., et al. 2015, MNRAS, 453, 3593  
Becker, G. D. & Bolton, J. S. 2013, MNRAS, 436, 1023  
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, MNRAS, 447, 3402  
Beckmann, R. S., Slyz, A., & Devriendt, J. 2018, MNRAS, 478, 995  
Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143  
Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57  
Benson, A. J. & Babul, A. 2009, MNRAS, 397, 1302  
Bhatawdekar, R., Conselice, C. J., Margalef-Bentabol, B., & Duncan, K. 2019, MNRAS, 486, 3805  
Bieri, R., Dubois, Y., Rosdahl, J., et al. 2017, MNRAS, 464, 1854



- Binney, J. & Tremaine, S. 1987, *Galactic dynamics*, Princeton series in astrophysics (Princeton University Press)
- Birnboim, Y. & Dekel, A. 2003, *MNRAS*, 345, 349
- Blondin, J. M., Wright, E. B., Borkowski, K. J., & Reynolds, S. P. 1998, *ApJ*, 500, 342
- Bondi, H. 1952, *MNRAS*, 112, 195
- Booth, C. M. & Schaye, J. 2009, *MNRAS*, 398, 53
- Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. 2014, *Science*, 346, 216
- Borzyszkowski, M., Porciani, C., Romano-Díaz, E., & Garaldi, E. 2017, *MNRAS*, 469, 594
- Boutsia, K., Grazian, A., Giallongo, E., Fiore, F., & Civano, F. 2018, *ApJ*, 869, 20
- Bouwens, R. J., Aravena, M., Decarli, R., et al. 2016, *ApJ*, 833, 72
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, *ApJ*, 811, 140
- Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2017, *ApJ*, 843, 129
- Bouwens, R. J., Stefanon, M., Oesch, P. A., et al. 2019, *ApJ*, 880, 25
- Bower, R. G., Schaye, J., Frenk, C. S., et al. 2017, *MNRAS*, 465, 32
- Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2015, *MNRAS*, 452, 1817
- Bryan, G. L. & Norman, M. L. 1998, *ApJ*, 495, 80
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2015, *ApJ*, 802, 89
- Bustamante, S. & Springel, V. 2019, *MNRAS*, 2455
- Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, *MNRAS*, 412, 2543
- Casey, C. M. 2016, *ApJ*, 824, 36
- Castellano, M., Dayal, P., Pentericci, L., et al. 2016, *ApJ*, 818, L3
- Chaabouni, H., Bergeron, H., Baouche, S., et al. 2012, *A&A*, 538, A128
- Chandrasekhar, S. 1943, *ApJ*, 97, 255
- Chapon, D., Mayer, L., & Teyssier, R. 2013, *MNRAS*, 429, 3114
- Chardin, J., Haehnelt, M. G., Aubert, D., & Puchwein, E. 2015, *MNRAS*, 453, 2943
- Cheng, T., Clements, D. L., Greenslade, J., et al. 2019, *MNRAS*, 490, 3840
- Ciardi, B., Stoeck, F., & White, S. D. M. 2003, *MNRAS*, 343, 1101
- Costa, T., Rosdahl, J., & Kimm, T. 2019, *MNRAS*, 489, 5181
- Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, *MNRAS*, 450, 1937
- Cristiani, S., Serrano, L. M., Fontanot, F., Vanzella, E., & Monaco, P. 2016, *MNRAS*, 462, 2478
- Dalgarno, A. & McCray, R. A. 1972, *ARA&A*, 10, 375
- D'Aloisio, A., McQuinn, M., Davies, F. B., & Furlanetto, S. R. 2018, *MNRAS*, 473, 560
- Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, *MNRAS*, 486, 2827
- Davidzon, I., Cucciati, O., Bolzonella, M., et al. 2016, *A&A*, 586, A23
- Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, 605, A70
- Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018, *ApJ*, 864, 142
- Dayal, P., Volonteri, M., Choudhury, T. R., et al. 2020, *arXiv e-prints*, arXiv:2001.06021
- Dekel, A. & Birnboim, Y. 2006, *MNRAS*, 368, 2
- Deparis, N., Aubert, D., Ocvirk, P., Chardin, J., & Lewis, J. 2019, *A&A*, 622, A142
- Di Matteo, T., Croft, R. A. C., Feng, Y., Waters, D., & Wilkins, S. 2017, *MNRAS*, 467, 4243
- Digby-North, J. A., Nandra, K., Laird, E. S., et al. 2010, *MNRAS*, 407, 846
- Dixon, K. L., Furlanetto, S. R., & Mesinger, A. 2014, *MNRAS*, 440, 987
- Draine, B. T. 2011, *Physics of the Interstellar and Intergalactic Medium*, Princeton Series in Astrophysics (Princeton University Press)
- Draine, B. T. & Salpeter, E. E. 1979, *ApJ*, 231, 77
- Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2010, *MNRAS*, 409, 985
- Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2012, *MNRAS*, 420, 2662
- Dubois, Y., Pichon, C., Devriendt, J., et al. 2013, *MNRAS*, 428, 2885
- Dubois, Y., Pichon, C., Welker, C., et al. 2014a, *MNRAS*, 444, 1453
- Dubois, Y., Volonteri, M., & Silk, J. 2014b, *MNRAS*, 440, 1590
- Dubois, Y., Volonteri, M., Silk, J., Devriendt, J., & Slyz, A. 2014c, *MNRAS*, 440, 2333
- Dubois, Y., Volonteri, M., Silk, J., et al. 2015, *MNRAS*, 452, 1502
- Dubroca, B. & Feugeas, J. 1999, *Academie des Sciences Paris Comptes Rendus Serie Sciences Mathematiques*, 329, 915
- Duncan, K., Conselice, C. J., Mortlock, A., et al. 2014, *MNRAS*, 444, 2960
- Dwek, E. 1998, *ApJ*, 501, 643
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, *PASA*, 34, e058
- Fabian, A. C. 2012, *ARA&A*, 50, 455
- Fan, X., Carilli, C. L., & Keating, B. 2006a, *ARA&A*, 44, 415
- Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, *AJ*, 132, 117
- Faucher-Giguère, C.-A. 2020, *MNRAS*, 259
- Federrath, C. & Klessen, R. S. 2012, *ApJ*, 761, 156
- Feng, Y., Di-Matteo, T., Croft, R. A., et al. 2016, *MNRAS*, 455, 2778
- Ferland, G. J., Korista, K. T., Verner, D. A., et al. 1998, *PASP*, 110, 761
- Fiacconi, D., Sijacki, D., & Pringle, J. E. 2018, *MNRAS*, 477, 3807
- Finkelstein, S. L., D'Aloisio, A., Paardekoooper, J.-P., et al. 2019, *ApJ*, 879, 36
- Finlator, K., Keating, L., Oppenheimer, B. D., Davé, R., & Zackrisson, E. 2018, *MNRAS*, 480, 2628
- Fletcher, T. J., Tang, M., Robertson, B. E., et al. 2019, *ApJ*, 878, 87
- Garaldi, E., Compostella, M., & Porciani, C. 2019, *MNRAS*, 483, 5301
- Geen, S., Rosdahl, J., Blaizot, J., Devriendt, J., & Slyz, A. 2015, *MNRAS*, 448, 3248
- Georgakakis, A., Aird, J., Buchner, J., et al. 2015, *MNRAS*, 453, 1946
- Giallongo, E., Grazian, A., Fiore, F., et al. 2015, *A&A*, 578, A83
- Giallongo, E., Grazian, A., Fiore, F., et al. 2019, *ApJ*, 884, 19
- Glikman, E., Djorgovski, S. G., Stern, D., et al. 2011, *ApJ*, 728, L26
- Gnedin, N. Y. 2000, *ApJ*, 542, 535
- Gnedin, N. Y. & Abel, T. 2001, *New A*, 6, 437
- Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, *ApJ*, 672, 765
- Gnedin, O. Y., Ostriker, J. P., & Tremaine, S. 2014, *ApJ*, 785, 71
- Götberg, Y., de Mink, S. E., Groh, J. H., Leitherer, C., & Norman, C. 2019, *A&A*, 629, A134
- Grazian, A., Giallongo, E., Boutsia, K., et al. 2018, *A&A*, 613, A44
- Grazian, A., Giallongo, E., Paris, D., et al. 2017, *A&A*, 602, A18
- Guillet, T. & Teyssier, R. 2011, *Journal of Computational Physics*, 230, 4756
- Haardt, F. & Madau, P. 1996, *ApJ*, 461, 20
- Haardt, F. & Madau, P. 2012, *ApJ*, 746, 125
- Haardt, F. & Salvaterra, R. 2015, *A&A*, 575, L16
- Habouzit, M., Volonteri, M., & Dubois, Y. 2017, *MNRAS*, 468, 3935
- Haehnelt, M. G., Madau, P., Kudritzki, R., & Haardt, F. 2001, *ApJ*, 549, L151
- Harikane, Y., Ouchi, M., Ono, Y., et al. 2019, *ApJ*, 883, 142
- Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, *MNRAS*, 367, 1039
- Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, *ApJ*, 654, 731
- Howard, C. S., Pudritz, R. E., Harris, W. E., & Klessen, R. S. 2018, *MNRAS*, 475, 3121
- Hunter, J. D. 2007, *Computing in Science and Engineering*, 9, 90
- Iliev, I. T., Mellema, G., Ahn, K., et al. 2014, *MNRAS*, 439, 725
- Iliev, I. T., Mellema, G., Pen, U. L., et al. 2006, *MNRAS*, 369, 1625
- Ishigaki, M., Kawamata, R., Ouchi, M., et al. 2018, *ApJ*, 854, 73
- Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016a, *Nature*, 529, 178
- Izotov, Y. I., Schaerer, D., Thuan, T. X., et al. 2016b, *MNRAS*, 461, 3683
- Izotov, Y. I., Schaerer, D., Worseck, G., et al. 2018a, *MNRAS*, 474, 4514
- Izotov, Y. I., Worseck, G., Schaerer, D., et al. 2018b, *MNRAS*, 478, 4851
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, *SciPy: Open source scientific tools for Python*
- Kato, Y., Matsuda, Y., Smail, I., et al. 2016, *MNRAS*, 460, 3861
- Katz, H., Kimm, T., Haehnelt, M., et al. 2018, *MNRAS*, 478, 4986
- Katz, H., Kimm, T., Haehnelt, M. G., et al. 2019a, *MNRAS*, 483, 1029
- Katz, H., Kimm, T., Sijacki, D., & Haehnelt, M. G. 2017, *MNRAS*, 468, 4831
- Katz, H., Laporte, N., Ellis, R. S., Devriendt, J., & Slyz, A. 2019b, *MNRAS*, 484, 4054
- Katz, H., Ramsay, M., Rosdahl, J., et al. 2019c, *arXiv e-prints*, arXiv:1905.11414
- Khandai, N., Di Matteo, T., Croft, R., et al. 2015, *MNRAS*, 450, 1349
- Kim, C.-G. & Ostriker, E. C. 2015, *ApJ*, 802, 99
- Kim, J.-h., Wise, J. H., Abel, T., et al. 2019, *ApJ*, 887, 120
- Kimm, T., Blaizot, J., Garel, T., et al. 2019, *MNRAS*, 486, 2215
- Kimm, T. & Cen, R. 2014, *ApJ*, 788, 121
- Kimm, T., Cen, R., Devriendt, J., Dubois, Y., & Slyz, A. 2015, *MNRAS*, 451, 2900
- Kimm, T., Katz, H., Haehnelt, M., et al. 2017, *MNRAS*, 466, 4826
- King, A. R., Lubow, S. H., Ogilvie, G. I., & Pringle, J. E. 2005, *MNRAS*, 363, 49
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, 192, 18
- Krishnan, C., Hatch, N. A., Almaini, O., et al. 2017, *MNRAS*, 470, 2170
- Kroupa, P. 2001, *MNRAS*, 322, 231
- Kubo, M., Tashikawa, J., Kashikawa, N., et al. 2019, *ApJ*, 887, 214
- Kulkarni, G. 2019, *QLF: Luminosity function analysis code*
- Kulkarni, G. & Choudhury, T. R. 2011, *MNRAS*, 412, 2781
- Kulkarni, G., Worseck, G., & Hennawi, J. F. 2019, *MNRAS*, 488, 1035
- Lam, D., Bouwens, R. J., Coe, D., et al. 2019, *arXiv e-prints*, arXiv:1903.08177
- Laor, A. & Netzer, H. 1989, *MNRAS*, 238, 897
- Lehmer, B. D., Alexander, D. M., Geach, J. E., et al. 2009, *ApJ*, 691, 687
- Lehnert, M. D. & Bremer, M. 2003, *ApJ*, 593, 630
- Leitet, E., Bergvall, N., Hayes, M., Linné, S., & Zackrisson, E. 2013, *A&A*, 553, A106
- Leitet, E., Bergvall, N., Piskunov, N., & Andersson, B. G. 2011, *A&A*, 532, A107
- Leitherer, C., Hernandez, S., Lee, J. C., & Oey, M. S. 2016, *ApJ*, 823, 64
- Levermore, C. D. 1984, *J. Quant. Spectr. Rad. Transf.*, 31, 149
- Li, Q., Narayanan, D., & Davé, R. 2019, *MNRAS*, 490, 1425
- Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, *ApJ*, 835, 113
- Lusso, E., Comastri, A., Simmonds, B. D., et al. 2012, *MNRAS*, 425, 623
- Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, *MNRAS*, 449, 4204
- Ma, X., Kasen, D., Hopkins, P. F., et al. 2015, *MNRAS*, 453, 960
- Madau, P. 2017, *ApJ*, 851, 50

- Madau, P. & Dickinson, M. 2014, *ARA&A*, 52, 415
- Madau, P. & Haardt, F. 2015, *ApJ*, 813, L8
- Madau, P., Haardt, F., & Rees, M. J. 1999, *ApJ*, 514, 648
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, 480, 5113
- Martizzi, D., Faucher-Giguère, C.-A., & Quataert, E. 2015, *MNRAS*, 450, 504
- Mazzucchelli, C., Bañados, E., Decarli, R., et al. 2017a, *ApJ*, 834, 83
- Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017b, *ApJ*, 849, 91
- McKee, C. 1989, in *IAU Symposium*, Vol. 135, *Interstellar Dust*, ed. L. J. Allamandola & A. G. G. M. Tielens, 431
- McKinnon, J. C., Tchekhovskoy, A., & Blandford, R. D. 2012, *MNRAS*, 423, 3083
- McKinnon, R., Torrey, P., Vogelsberger, M., Hayward, C. C., & Marinacci, F. 2017, *MNRAS*, 468, 1505
- McKinnon, R., Vogelsberger, M., Torrey, P., Marinacci, F., & Kannan, R. 2018, *MNRAS*, 478, 2851
- McQuinn, M., Lidz, A., Zaldarriaga, M., et al. 2009, *ApJ*, 694, 842
- Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, *ApJ*, 804, 104
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Nature*, 474, 616
- Mostardi, R. E., Shapley, A. E., Steidel, C. C., et al. 2015, *ApJ*, 810, 107
- Musso, M., Cadiou, C., Pichon, C., et al. 2018, *MNRAS*, 476, 4877
- Naidu, R. P., Forrest, B., Oesch, P. A., Tran, K.-V. H., & Holden, B. P. 2018, *MNRAS*, 478, 791
- Naidu, R. P., Tacchella, S., Mason, C. A., et al. 2019, *arXiv e-prints*, arXiv:1907.13130
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, 477, 1206
- Nelson, D., Pillepich, A., Springel, V., et al. 2019, *MNRAS*, 490, 3234
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, 475, 624
- Novak, G. S., Ostriker, J. P., & Ciotti, L. 2012, *MNRAS*, 427, 2734
- Novikov, I. D. & Thorne, K. S. 1973, in *Black Holes (Les Astres Occlus)*, 343–450
- Ocvirk, P., Aubert, D., Chardin, J., Deparis, N., & Lewis, J. 2019, *A&A*, 626, A77
- Ocvirk, P., Aubert, D., Sorce, J. G., et al. 2018, *arXiv e-prints*, arXiv:1811.11192
- Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, *MNRAS*, 463, 1462
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2013, *ApJ*, 773, 75
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, *ApJ*, 786, 108
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, *ApJ*, 855, 105
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, 819, 129
- Ono, Y., Ouchi, M., Mobasher, B., et al. 2012, *ApJ*, 744, 83
- O’Shea, B. W., Wise, J. H., Xu, H., & Norman, M. L. 2015, *ApJ*, 807, L12
- Ostriker, E. C. 1999, *ApJ*, 513, 252
- Ota, K., Iye, M., Kashikawa, N., et al. 2008, *ApJ*, 677, 12
- Oteo, I., Ivison, R. J., Dunne, L., et al. 2018, *ApJ*, 856, 72
- Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2010, *ApJ*, 723, 869
- Overzier, R. A. 2016, *A&A Rev.*, 24, 14
- Paardekooper, J.-P., Khochfar, S., & Dalla Vecchia, C. 2015, *MNRAS*, 451, 2544
- Padoan, P. & Nordlund, Å. 2011, *ApJ*, 730, 40
- Page, D. N. & Thorne, K. S. 1974, *ApJ*, 191, 499
- Papovich, C., Kawinwanichakij, L., Quadri, R. F., et al. 2018, *ApJ*, 854, 30
- Park, M.-J., Yi, S. K., Dubois, Y., et al. 2019, *ApJ*, 883, 25
- Parsa, S., Dunlop, J. S., & McLure, R. J. 2018, *MNRAS*, 474, 2904
- Pawlik, A. H., Rahmati, A., Schaye, J., Jeon, M., & Dalla Vecchia, C. 2017, *MNRAS*, 466, 960
- Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, *ApJ*, 793, 113
- Perez, F. & Granger, B. E. 2007, *Computing in Science Engineering*, 9
- Pfister, H., Lupi, A., Capelo, P. R., et al. 2017, *MNRAS*, 471, 3646
- Pfister, H., Volonteri, M., Dubois, Y., Dotti, M., & Colpi, M. 2019, *MNRAS*, 486, 101
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, *MNRAS*, 475, 648
- Pillepich, A., Nelson, D., Springel, V., et al. 2019, *MNRAS*, 490, 3196
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, *MNRAS*, 473, 4077
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, *arXiv e-prints*, arXiv:1807.06209
- Popping, G., Somerville, R. S., & Galametz, M. 2017, *MNRAS*, 471, 3152
- Prunet, S. & Pichon, C. 2013, *MPgratic: A parallel MPI version of Gratic-1*
- Prunet, S., Pichon, C., Aubert, D., et al. 2008, *ApJS*, 178, 179
- Rasera, Y. & Teyssier, R. 2006, *A&A*, 445, 1
- Reed, S. L., Banerji, M., Becker, G. D., et al. 2019, *MNRAS*, 487, 1874
- Reines, A. E. & Volonteri, M. 2015, *ApJ*, 813, 82
- Rezzolla, L., Barausse, E., Dorband, E. N., et al. 2008, *Phys. Rev. D*, 78, 044002
- Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, *ApJ*, 802, L19
- Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, *ApJ*, 768, 71
- Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., & Teyssier, R. 2013, *MNRAS*, 436, 2188
- Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, *MNRAS*, 479, 994
- Rosdahl, J. & Teyssier, R. 2015, *MNRAS*, 449, 4380
- Rosen, A. & Bregman, J. N. 1995, *ApJ*, 440, 634
- Salmon, B., Coe, D., Bradley, L., et al. 2018, *ApJ*, 864, L22
- Sazonov, S. Y., Ostriker, J. P., & Sunyaev, R. A. 2004, *MNRAS*, 347, 144
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521
- Schenker, M. A., Ellis, R. S., Konidaris, N. P., & Stark, D. P. 2014, *ApJ*, 795, 20
- Schmidt, M. 1959, *ApJ*, 129, 243
- Schroeder, J., Mesinger, A., & Haiman, Z. 2013, *MNRAS*, 428, 3058
- Shakura, N. I. & Sunyaev, R. A. 1973, *A&A*, 24, 337
- Shapiro, P. R., Giroux, M. L., & Babul, A. 1994, *ApJ*, 427, 25
- Shapiro, P. R., Iliiev, I. T., & Raga, A. C. 2004, *MNRAS*, 348, 753
- Shapley, A. E., Steidel, C. C., Strom, A. L., et al. 2016, *ApJ*, 826, L24
- Sheth, R. K. & Tormen, G. 2004, *MNRAS*, 350, 1385
- Shimakawa, R., Koyama, Y., Röttgering, H. J. A., et al. 2018, *MNRAS*, 481, 5630
- Shull, J. M., France, K., Danforth, C. W., Smith, B., & Tumlinson, J. 2010, *ApJ*, 722, 1312
- Silk, J. & Rees, M. J. 1998, *A&A*, 331, L1
- Snaith, O. N., Park, C., Kim, J., & Rosdahl, J. 2018, *MNRAS*, 477, 983
- Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, *ApJ*, 825, 5
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, 475, 676
- Stanway, E. R. & Eldridge, J. J. 2018, *MNRAS*, 479, 75
- Stanway, E. R., Eldridge, J. J., & Becker, G. D. 2016, *MNRAS*, 456, 485
- Stark, D. P. 2016, *ARA&A*, 54, 761
- Steidel, C. C., Bogosavljević, M., Shapley, A. E., et al. 2018, *ApJ*, 869, 123
- Tanvir, N. R., Fynbo, J. P. U., de Ugarte Postigo, A., et al. 2019, *MNRAS*, 483, 5380
- Teyssier, R. 2002, *A&A*, 385, 337
- Thorne, K. S. 1974, *ApJ*, 191, 507
- Thornton, K., Gaudlitz, M., Janka, H. T., & Steinmetz, M. 1998, *ApJ*, 500, 95
- Tilvi, V., Papovich, C., Finkelstein, S. L., et al. 2014, *ApJ*, 794, 5
- Tomczak, A. R., Lemaux, B. C., Lubin, L. M., et al. 2017, *MNRAS*, 472, 3512
- Topping, M. W. & Shull, J. M. 2015, *ApJ*, 800, 97
- Toro, E. F., Spruce, M., & Speares, W. 1994, *Shock Waves*, 4, 25
- Totani, T., Aoki, K., Hattori, T., & Kawai, N. 2016, *PASJ*, 68, 15
- Totani, T., Kawai, N., Kosugi, G., et al. 2006, *PASJ*, 58, 485
- Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., & Slyz, A. 2017, *MNRAS*, 470, 224
- Trebitsch, M., Volonteri, M., & Dubois, Y. 2019, *MNRAS*, 487, 819
- Trebitsch, M., Volonteri, M., Dubois, Y., & Madau, P. 2018, *MNRAS*, 478, 5607
- Tremmel, M., Governato, F., Volonteri, M., & Quinn, T. R. 2015, *MNRAS*, 451, 1868
- Tremmel, M., Karcher, M., Governato, F., et al. 2017, *MNRAS*, 470, 1121
- Truelove, J. K., Klein, R. I., McKee, C. F., et al. 1998, *ApJ*, 495, 821
- Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, *ApJS*, 192, 9
- Tweed, D., Devriendt, J., Blaizot, J., Colombi, S., & Slyz, A. 2009, *A&A*, 506, 647
- Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, *ApJ*, 786, 104
- van Leer, B. 1979, *Journal of Computational Physics*, 32, 101
- Vanzella, E., de Barros, S., Vasei, K., et al. 2016, *ApJ*, 825, 41
- Vanzella, E., Nonino, M., Cupani, G., et al. 2018, *MNRAS*, 476, L15
- Đurovčiková, D., Katz, H., Bosman, S. E. I., et al. 2019, *arXiv e-prints*, arXiv:1912.01050
- Vito, F., Brandt, W. N., Yang, G., et al. 2018, *MNRAS*, 473, 2378
- Vito, F., Gilli, R., Vignali, C., et al. 2016, *MNRAS*, 463, 348
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, 444, 1518
- Volonteri, M. 2010, *A&A Rev.*, 18, 279
- Volonteri, M., Reines, A. E., Atek, H., Stark, D. P., & Trebitsch, M. 2017, *ApJ*, 849, 155
- Wang, F., Yang, J., Fan, X., et al. 2019, *ApJ*, 884, 30
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, *MNRAS*, 465, 3291
- Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, *MNRAS*, 442, 2560
- Worseck, G., Prochaska, J. X., Hennawi, J. F., & McQuinn, M. 2016, *ApJ*, 825, 144
- Wyithe, J. S. B. & Bolton, J. S. 2011, *MNRAS*, 412, 1926
- Xu, H., Wise, J. H., Norman, M. L., Ahn, K., & O’Shea, B. W. 2016, *ApJ*, 833, 84
- Yajima, H., Shlosman, I., Romano-Díaz, E., & Nagamine, K. 2015, *MNRAS*, 451, 418
- Yang, G., Brandt, W. N., Darvish, B., et al. 2018, *MNRAS*, 480, 1022
- Yoo, T., Kimm, T., & Rosdahl, J. 2020, *arXiv e-prints*, arXiv:2001.05508
- Yoshiura, S., Hasegawa, K., Ichiki, K., et al. 2017, *MNRAS*, 471, 3713

## Appendix A: Effect of the low threshold for the dynamical friction

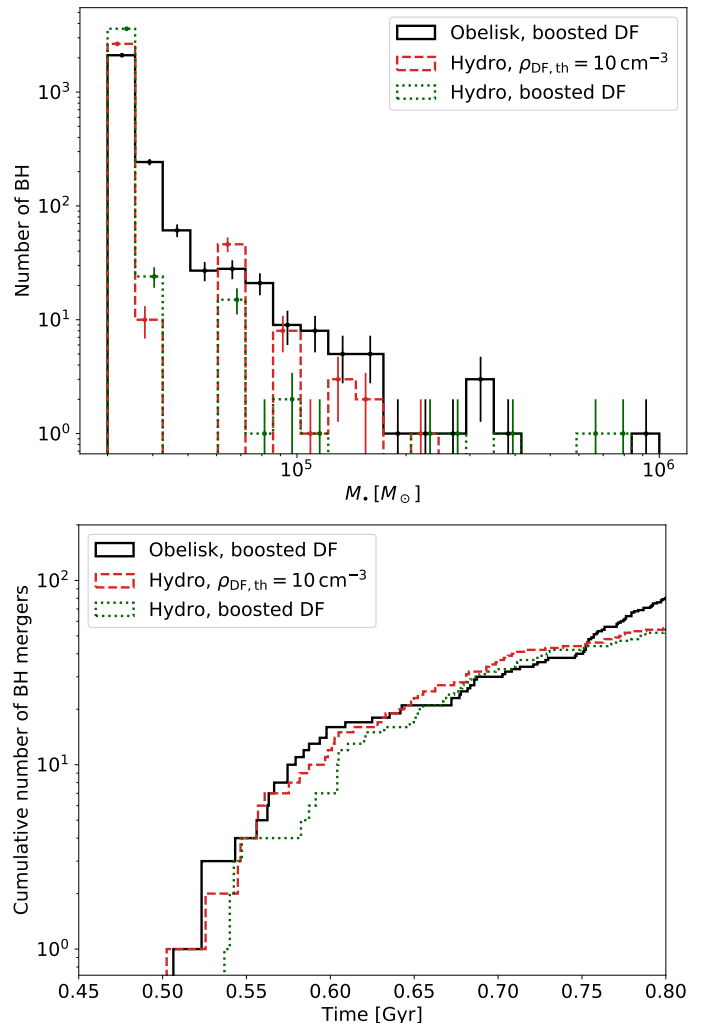
Let us investigate the effect of our very low threshold for including the dynamical friction,  $\rho_{\text{DF,th}}$ . We first recall that in order to follow BH dynamics correctly, simulations are in principle required to resolve the influence radius of a the BH. Because we cannot do so in a fully cosmological context, we have to rely to a subgrid model to account for unresolved dynamical friction from the gas onto the BH. This model will tend to align the velocity of the BH to that of the gas. In order to account for the unresolved structure of the gas in the close vicinity of the BH, we have chosen to boost the frictional force by an ad-hoc factor  $\alpha = (\rho/\rho_{\text{DF,th}})^2$  if  $\rho > \rho_{\text{th}}$ , the choice of the threshold  $\rho_{\text{DF,th}}$  controls the strength of the frictional force being arbitrary.

Here, we have used a very low value for this threshold: the main effect is to increase the dynamical friction significantly, effectively sticking the BH to the gas cloud it is embedded in. We stress that this is at most comparable to the effect of artificially redirecting the BH towards the centre of the cloud. Nevertheless, we want to quantify the effect of this on the BH population in OBELISK. To do so, we make use of a twin simulation that includes the exact same physical model as OBELISK except for the radiation hydrodynamics and that will be presented in another paper (Cadiou et al. (in prep)). Because it is significantly cheaper, we have run two versions of this simulation, varying the value of the density threshold from  $\rho_{\text{DF,th}} = 0.01 \text{ cm}^{-3}$  to  $\rho_{\text{DF,th}} = 10 \text{ cm}^{-3}$ , similar to the value used in the NEW-HORIZON simulation.

We show the BH mass function at  $z \sim 6.9$  for the fiducial OBELISK simulation and for both hydro runs in the upper panel of Fig. A.1. The hydro run using the same value of  $\rho_{\text{DF,th}}$  as OBELISK is show as a green dotted line and the run with the higher value corresponds to the red dashed line. The error bars indicate the  $\sqrt{N}$  Poisson error in each mass bin. Comparing the two hydro runs, it seems that the run with stronger dynamical friction is marginally depleted in BHs with masses around  $5 \times 10^4 - 8 \times 10^4 M_\odot$ , and that the most massive BH are somewhat more massive. The RHD run, with strong dynamical friction, behaves differently: there are overall more BHs at all masses, except just around the seed mass. This suggests that compared to other processes, the details of the dynamical friction from gas have a relatively minor effect, once it is strong enough to maintain the BH in the centre of the galaxy. The lower panel shows the cumulative number of BH-BH mergers measured in these three runs, with the same legend. Here, we see that in the hydro run with weaker dynamical friction experiences slightly more mergers than the other hydro run: this suggests that the main effect of our stronger dynamical friction is to slightly increase the gas density around the BH, enhancing a bit the accretion. However, we see here as well that changing the strength of the dynamical friction does not affect how BHs merge more than the inclusion of radiation hydrodynamics. To summarize, while our choice to strongly boost the dynamical friction seems artificial, it leads to a BH growth history that is similar to that from simulations with a weaker prescription for the BH dynamics.

## Appendix B: AGN SED model

Let us present more explicitly the model for the AGN radiation already described in Sect. 2.4.6, in order to facilitate the reproducibility of our numerical experiment. We highlight that our endeavour here is not to provide an AGN spectrum comparable with observations on a one-to-one basis, but rather to derive



**Fig. A.1.** Distribution of BH masses (*top*) and cumulative number of BH mergers (*bottom*) for the fiducial RHD run (in black) and for the two twin non-RHD runs with  $\rho_{\text{DF,th}} = 10 \text{ cm}^{-3}$  (red dashed line) and with the value described in Sect. 2.4.3 (green dotted line) at  $z \sim 6.9$ .

a spectral shape that will broadly capture how the ionizing luminosity of the AGN varies for different accreting black holes. Specifically, we focus on the ionizing UV bands that we consider in OBELISK. In particular, we do not model the infrared emission from the dust, and only assume that a fraction  $f_{\text{IR}} = 30\%$  of the total AGN luminosity is absorbed by dust and re-emitted in the IR. This is close to the value suggested by the average Sazonov et al. (2004) spectrum. We assume that each BH for which the accretion rate exceeds  $\dot{M}_{\text{BHL}} \geq \chi_{\text{crit}} \dot{M}_{\text{Edd}} = 0.01 \dot{M}_{\text{Edd}}$  is surrounded by an optically thick, geometrically thin  $\alpha$ -disc<sup>8</sup>, as described by Shakura & Sunyaev (1973), Novikov & Thorne (1973) and Page & Thorne (1974).

The emission from a column of gas located a radius  $R$  in the disc can be described, under the assumption of local thermodynamical equilibrium between the gas and the radiation, as that of a blackbody of temperature  $T_{\text{BB}}(R)$ , such that the energy flux crossing the surface of the disc  $\mathcal{F}(R)$  can be equated to  $\sigma_{\text{SB}} T_{\text{bb}}(R)^4$ . This energy flux can be computed analytically for

<sup>8</sup> Strictly speaking, this solution is only valid if the disc luminosity stays below  $0.3 L_{\text{Edd}}$  to ensure that the disc stays well described by a thin disc. We still choose to use the thin disc solution up to  $L = L_{\text{Edd}}$ .



an  $\alpha$ -disc as

$$\mathcal{F}(R) = \frac{3GM_\bullet\dot{M}_\bullet}{8\pi R^3} f\left(\frac{R}{R_g}\right) \quad (\text{B.1})$$

where  $f(r)$  is specified by the disc profile. For the [Shakura & Sunyaev \(1973\)](#) solution<sup>9</sup>, we have

$$f(r) = f(R/R_g) = 1 - \sqrt{r_{\text{isco}}/r}. \quad (\text{B.2})$$

We derive the blackbody temperature of a ring at radius  $R$  as

$$\begin{aligned} T_{\text{BB}}(R) &= \left( \frac{3GM_\bullet\dot{M}_\bullet}{8\pi R^3} f(r) \right)^{1/4} \\ &= \left( \frac{3GM_\bullet\dot{M}_\bullet}{8\pi R_{\text{isco}}^3} \right)^{1/4} \left( \frac{R}{R_{\text{isco}}} \right)^{-3/4} \left( 1 - \sqrt{\frac{R_{\text{isco}}}{R}} \right)^{1/4}. \end{aligned} \quad (\text{B.3})$$

The ring at a radius  $R$  will then emit radiation with a Planckian spectrum:

$$B_\nu(T_{\text{BB}}(R)) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T_{\text{BB}}(R)}} - 1}, \quad (\text{B.4})$$

with  $h$  and  $k_B$  the Planck and Boltzmann constants, respectively.

The total spectrum of the disc can then readily be computed by integrating Eq. B.4 between the ISCO radius and the outer edge of the disc. We take this outer edge to be the self-gravitating radius  $r_{\text{sg}}$  of the disc, i.e. the radius at which the gravity of the central BH does not dominate anymore. [Laor & Netzer \(1989\)](#) give for a radiation-dominated thin disc:

$$r_{\text{sg}} \simeq 2150\alpha^{2/9} \left( \frac{M_\bullet}{10^9 M_\odot} \right)^{-2/9} \dot{m}^{4/9}, \quad (\text{B.5})$$

where  $\alpha \sim 0.1$  is the disc viscosity and  $\dot{m} = \frac{\epsilon_r(0)}{\epsilon_r(a_\star)} \frac{L}{L_{\text{Edd}}}$  is the reduced mass accretion rate of a BH with spin  $a_\star$  and luminosity  $L$  normalized to that of a non-spinning BH. We note that for a moderately luminous BH, the assumption that the radiation pressure dominates over the gas pressure can break down at a radius smaller than the self-gravitating radius. However, only the outermost regions of the disc, for which the blackbody temperature are the lowest, could be affected. These regions contribute very little to the ionizing UV radiation, and our results are consequently not strongly affected by this.

The total luminosity of the disc at a frequency  $\nu$  thus reads

$$L_\nu = 2\pi \int_{r_{\text{isco}}}^{r_{\text{sg}}} \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T_{\text{BB}}(R)}} - 1} 2\pi R dR. \quad (\text{B.6})$$

The spectral shape resulting from Eq. B.6 can be approximated at low energy (in the Rayleigh-Jeans regime) by  $L_\nu \propto \nu^2$ , and by  $L_\nu \propto \nu^{1/3}$  at intermediate energies, corresponding to the part of the disc where the temperature profile follows  $T(R) \propto R^{-3/4}$ . At high frequencies (corresponding to the innermost regions of the disc), however, the spectrum is exponentially cut off. Rather than trying to find a physically-motivated approximate model for the high energy part, we choose to replace the exponential cut-off by a power law,  $L_\nu \propto \nu^{-\alpha_{\text{UV}}}$ , with  $\alpha_{\text{UV}} = 1.5$ , in broad agreement with the value  $\alpha_{\text{UV}} = -1.7 \pm 0.61$  derived by [Lusso et al. \(2015\)](#) for a sample of high redshift quasars. The fraction of the total luminosity in a given frequency interval

$[\nu_{\text{min}}; \nu_{\text{max}}]$  can then be easily obtained by integrating Eq. B.6 over the interval. For instance, the fraction of the luminosity in the first UV band considered in OBELISK is obtained by

$$f_{\text{UV},1} = \int_{13.6 \text{ eV}/h}^{24.59 \text{ eV}/h} L_\nu d\nu. \quad (\text{B.7})$$

This leads to a spectrum that depends on  $M_\bullet$ ,  $\dot{M}_\bullet$ , and  $a_\star$  that can be in principle readily used in RAMSES-RT in the same manner as the stellar population models. For this latter case, it is necessary to integrate the spectrum on-the-fly, as the average energy and the interaction cross-sections in each radiation bin can vary (up to a factor of a few for the cross-sections) as a function of the age and the metallicity of the stellar population (see e.g. [Rosdahl et al. 2013](#), appendix B).

In this work, we have opted for a slightly different approach: we have tabulated the values of  $f_{\text{UV},i}$  for each bin  $i = 1, 2, 3$  and only interpolate between the tabulated values over the course of the simulation. In principle, this requires to interpolate in a three-dimensional table (for the BH mass, accretion rate and spin). However, the model described here is very weakly dependent on the BH spin  $a_\star$ . Effectively, we checked that the variation of  $f_{\text{UV},i}$  as  $a_\star$  varies is much weaker than when changing the  $M_\bullet$  or  $\dot{M}_\bullet$ . To limit the cost of our model, we drop the spin dependence and assume  $a_\star \sim 0.7$  (corresponding to a radiative efficiency  $\epsilon_r \simeq 0.1$ ) when computing the spectrum, leading to a simpler two-dimensional interpolation. Similarly, it appears that the average energy and cross-sections in each frequency interval depend very weakly not only on the spin, but also on  $M_\bullet$  and  $\dot{M}_\bullet$ . We therefore choose to fix their values to that of a non-spinning BH with mass  $M_\bullet = 10^7 M_\odot$  accreting at 10% of its Eddington luminosity.

## Appendix C: Residual neutral fraction at $z < 6$

The fact that our residual neutral fraction at  $z < 6$  is consistent with the observations might seem at first in contradiction with the results of [Ocvirk et al. \(2019\)](#). Indeed, with our type of implementation of the “reduced speed of light” approximation, [Ocvirk et al. \(2019\)](#) found that the residual neutral fraction after reionization is complete scales inversely with the adopted speed of light reduction factor  $f_c$ . This is a consequence of the fact that the photoionization equilibrium of strongly ionized gas can be approximated by  $x_{\text{H I}} \sim \frac{\alpha_B \rho_H}{\rho_\gamma \sigma f_c c}$  where  $\alpha_B$  is the case-B recombination coefficient,  $\rho_H$  the total hydrogen density, and  $\rho_\gamma$  the ionizing photon density.

In our case, there are however two mitigating factors to this problem. First, the use of the VSLA has been shown by [Katz et al. \(2018, Appendix B\)](#) to yield converged results for reduction factors of order  $f_{c,\text{max}} \sim 0.1 - 0.4$ , and we use  $f_{c,\text{max}} = 0.2$ . A second difference is that our cosmological setup is very different compared to [Ocvirk et al. \(2019\)](#): because we zoom within the HORIZON-AGN volume, radiation is allowed to leak out of the high-resolution region. The argument in [Ocvirk et al. \(2019\)](#) is that, in the post-overlap universe, the volume is roughly at photoionization equilibrium and all emitted photons contribute to this equilibrium. Here, the situation is different: photons can either be absorbed within the high-resolution region or escape the volume. This will shift the photoionization equilibrium towards a higher neutral fraction. This also means that the dependency on  $f_c$  is reduced: a higher (lower) reduced speed of light would lead to more (fewer) photons escaping the volume by the same redshift, shifting the equilibrium towards a higher (lower) neutral fraction, counterbalancing somewhat the effect found by [Ocvirk et al. \(2019\)](#).

<sup>9</sup> The complete solution for a Kerr spacetime is given in [Page & Thorne \(1974\)](#), but we have checked that using the [Shakura & Sunyaev \(1973\)](#) profile does not change significantly the resulting spectrum.