Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization

Pierre Ocvirk, Dominique Aubert, Jenny G. Sorce, Paul R. Shapiro, Nicolas Deparis, Taha Dawoodbhoy, Joseph Lewis, Romain Teyssier, Gustavo Yepes, Stefan Gottlöber, et al.

To cite this version:

Pierre Ocvirk, Dominique Aubert, Jenny G. Sorce, Paul R. Shapiro, Nicolas Deparis, et al.. Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization. Monthly Notices of the Royal Astronomical Society, 2020, 496 (4), pp.4087-4107. 10.1093/mnras/staa1266 . hal-01960360

HAL Id: hal-01960360
https://hal.science/hal-01960360v1

Submitted on 21 Jun 2022

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **496,** 4087–4107 (2020)
Advance Access publication 2020 May 13

doi:10.1093/mnras/staa1266

Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization

Pierre Ocvirk,[1★] Dominique Aubert,[1] Jenny G. Sorce,[1,2,3] Paul R. Shapiro,[4] Nicolas Deparis,[1] Taha Dawoodbhoy,[4] Joseph Lewis,[1] Romain Teyssier,[5] Gustavo Yepes,[6,7] Stefan Gottlöber,[3] Kyungjin Ahn,[8] Ilian T. Iliev[9] and Yehuda Hoffman[10]

[1] *CNRS, Observatoire astronomique de Strasbourg, Université de Strasbourg, UMR 7550, F-67000 Strasbourg, France*
[2] *CNRS, Centre de Recherche Astrophysique de Lyon, Université Lyon 1, ENS de Lyon, UMR 5574, F-69230 Saint-Genis-Laval, France*
[3] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*
[4] *Department of Astronomy, University Texas, Austin, TX 78712-1083, USA*
[5] *Institute for Theoretical Physics, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[6] *Departamento de Física Teórica M-8, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[7] *Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*
[8] *Chosun University, 375 Seosuk-dong, Dong-gu, Gwangjiu 501-759, Korea*
[9] *Astronomy Center, Department of Physics & Astronomy, University of Sussex, Pevensey II Building, Falmer, Brighton BN1 9QH, UK*
[10] *Racah Institute of Physics, Hebrew University, Jerusalem 91904, Israel*

Accepted 2020 May 1. Received 2020 April 3; in original form 2019 October 30

ABSTRACT

Cosmic Dawn II (CoDa II) is a new, fully coupled radiation-hydrodynamics simulation of cosmic reionization and galaxy formation and their mutual impact, to redshift $z < 6$. With 4096^3 particles and cells in a 94 Mpc box, it is large enough to model global reionization and its feedback on galaxy formation while resolving all haloes above $10^8\,\mathrm{M}_\odot$. Using the same hybrid CPU–GPU code RAMSES–CUDATON as CoDa I in Ocvirk et al. (2016), CoDa II modified and re-calibrated the subgrid star formation algorithm, making reionization end earlier, at $z \gtrsim 6$, thereby better matching the observations of intergalactic Lyman α opacity from quasar spectra and electron-scattering optical depth from cosmic microwave background fluctuations. CoDa II predicts a UV continuum luminosity function in good agreement with observations of high-z galaxies, especially at $z = 6$. As in CoDa I, reionization feedback suppresses star formation in haloes below $\sim 2 \times 10^9\,\mathrm{M}_\odot$, though suppression here is less severe, a possible consequence of modifying the star formation algorithm. Suppression is environment dependent, occurring earlier (later) in overdense (underdense) regions, in response to their local reionization times. Using a constrained realization of lambda cold dark matter constructed from galaxy survey data to reproduce the large-scale structure and major objects of the present-day Local Universe, CoDa II serves to model both *global* and *local* reionization. In CoDa II, the Milky Way and M31 appear as individual islands of reionization, i.e. they were not reionized by the progenitor of the Virgo cluster, or by nearby groups, or by each other.

Key words: radiative transfer – methods: numerical – intergalactic medium – Local Group – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

The first billion years of the Universe is a key period for the evolution of the intergalactic medium (hereafter, ‘IGM’), and the formation of galaxies. As the first stars form, their UV photons propagate, carving ionized regions which ultimately percolate, marking the end of the epoch of reionization (hereafter, ‘EoR’). The EoR is considered as the next observational frontier, both for 21-cm radio telescopes such as LOFAR, MWA, SKA, NenuFAR, and infrared or near-infrared missions such as JWST. The recent tentative detection of a 21 cm from the dark ages by EDGES (Bowman et al. 2018) has stirred the community and prompted numerous teams to try to confirm the signal with other instruments.

Meanwhile, a new breed of numerical simulations of galaxy formation has appeared, which aims at accounting for the rich physics, in particular the impact of the ionizing radiation, necessary to describe the EoR.[1] Faithful modelling of reionization, though, requires us to be able to describe simultaneously the coupled multiscale problem of global reionization and individual galaxy formation, with gravity, hydrodynamics and radiative transfer. To represent the complex geometry of reionization in a statistically

---

★ E-mail: pierre.ocvirk@astro.unistra.fr

[1] See Dayal & Ferrara (2018) for a review of recent advances in numerical simulations of galaxy formation and reionization.

Published by Oxford University Press on behalf of the Royal Astronomical Society

meaningful way requires a comoving computational domain as large as $\sim$(100 Mpc)^3 e.g. Iliev et al. (2006, 2014). In order to follow the millions of galaxies in this volume over the full mass range of galactic haloes contributing to reionization,[2] while modelling the impact of reionization on these individual galaxies and the IGM, a physical resolution of a few kpc over the entire volume is needed.

The EoR is also important for its impact on low-mass galaxy formation. The baryonic gas content of low-mass galaxies is depressed by the hydrodynamical backreaction of their interstellar and surrounding intergalactic media, due to the photoheating that accompanies reionization by the rising ionizing UV background (Shapiro, Giroux & Babul 1994; Gnedin 2000; Hoeft et al. 2006). This, in turn, suppresses their star formation rates. The resulting 'fossil galaxies' may offer a credible solution to the missing satellite problem as formulated in the Local Group (see e.g. Bullock & Boylan-Kolchin 2017, and references, therein).

The faintest galaxies known, potentially susceptible to such radiative suppression of their gas and star formation, can only be detected nearby, in the Local Group. Therefore, to make direct comparisons possible between simulations and this observationally accessible sample, it is necessary to start from initial conditions preselected to produce the observed galaxies and large-scale structure of the Local Universe.

In the past few years, we have made significant progress in this direction, in our project of large-scale simulation of cosmic reionization and galaxy formation based upon fully coupled radiation hydrodynamics, the Cosmic Dawn ('CoDa') simulation project. The first of the CoDa simulations (hereafter, referred to as 'CoDa I'), described in Ocvirk et al. (2016) (hereafter, O16), with further results presented in Dawoodbhoy et al. (2018), was based upon the massively parallel, hybrid CPU–GPU code RAMSES–CUDATON and ran for 11 d on the Titan supercomputer at the Oak Ridge Leadership Computational Facility (OLCF), with 8092 GPUs and 8092 CPUs, utilizing 4096^3 particles and cells in a volume 94 comoving Mpc on a side. The second simulation, CoDa I-AMR, described in Aubert et al. (2018), was based upon the massively parallel, hybrid CPU–GPU code EMMA, with adaptive mesh refinement (AMR), which also ran on Titan, with 4096 GPUs and 32768 CPUs, utilizing 2048^3 particles, starting from a uniform grid of 2048^3 unrefined cells from which AMR increased the resolution locally, by up to a factor of 8, depending on the local overdensity, leading ultimately to 18 billion cells after refinement, in the same size box, starting from the same initial conditions (except coarsened to the initial, unrefined grid of CoDa I-AMR).

These CoDa simulations share four principal characteristics that make them special and a first of their kind, as follows. Their resolution is high enough to track the formation of the mass range of galactic haloes above $10^8\,\mathrm{M}_\odot$ believed to dominate reionization and the backreaction of reionization on their evolution as sources. Their volume is large enough to model the mean history of reionization and the hydrogen ionizing UV background, along with the inhomogeneity of reionization in time and space (Fig. 1 illustrates these aspects). They are based upon a constrained realization of the Gaussian random initial conditions for lambda cold dark matter (ΛCDM), derived from galaxy observations of the Local Universe, which are designed to reproduce the large-scale structure and familiar features observed in our Local Universe today when evolved to the present. By starting from this 'constrained realization' of initial conditions which reproduces a number of selected, familiar features of the Local Universe in a volume centred on the Local Group, such as the Milky Way, M31, and the Fornax and Virgo clusters, we are able to use these simulations to model both *global* and *local* reionization, simultaneously, including the impact of reionization on the formation of galaxies in the Local Group. Finally, by fully coupling the hydrodynamics to the radiative transfer of ionizing radiation, we move beyond simulations which adopt a pre-computed, uniform UV background, by computing the rise of the UV background, inhomogeneous in space and time, and its location-specific impact on galaxies and their emissivity, together, self-consistently.

In this paper, we present the Cosmic Dawn II ('CoDa II') simulation, which aims to improve upon our previous two simulations in several aspects. In CoDa I (O16), the end of global reionization was somewhat later than observed, at $z \lesssim 5$, a consequence of adopting a slightly lower efficiency for the subgrid star formation algorithm than necessary to end reionization by $z = 6$, despite our effort to calibrate this efficiency by a series of small-box simulations intended to predict the outcome for the large box. For CoDa II, the star formation algorithm parameters were recalibrated and the algorithm was also modified, to ensure that reionization ended earlier than CoDa I, by $z = 6$, in better agreement with the observations. While CoDa I-AMR (Aubert et al. 2018) was also successfully recalibrated to end reionization earlier than did CoDa I, both CoDa I and CoDa I-AMR started from the same generation of 'constrained realization' initial conditions developed by the *Constrained Local UniversE Simulations* ('CLUES') project which have since been updated and improved, including for instance more recent observational data. Those earlier initial conditions resulted in a Virgo replica less massive than observed, and were derived for a background universe with cosmological parameters for a WMAP5 cosmology (Hinshaw et al. 2009). The new initial conditions for CoDa II now feature a more realistic Virgo replica, and use a more recent set of cosmological parameters determined by the *Planck* collaboration (Planck Collaboration XVI 2014).

The main goal of this first CoDa II paper is to introduce the simulation and compare the results with current observational constraints on the global EoR and the high-redshift galaxies that caused it, such as the evolution of the cosmic neutral and ionized fractions, the mean intensity of the cosmic ionizing UV background, the cosmic star formation rate density, the Thomson-scattering optical depth through the intergalactic medium measured by cosmic microwave background fluctuation measurements, and the high-redshift galactic UV luminosity functions. In Section 2, we describe the RAMSES-CUDATON code on which the CoDa II simulation is based, and the simulation set-up, including the new 'constrained realization' initial conditions used here. Our results are presented in Section 3, including the global history of reionization compared with observational constraints, the star formation rates of galaxies and the net star formation rate density of the universe, and the UV continuum luminosity function of high-z galaxies during the EoR. We use the simulation to show how the star formation rate in low-mass haloes is affected by reionization and the environmental dependence of this process. Finally, we investigate the reionization history of the Local Group, and then finish with a short summary in Section 4.

---

[2] The so-called 'atomic-cooling haloes' (henceforth, 'ACHs'), those with virial temperatures above $\sim 10^4$ K, corresponding to halo masses above $\sim 10^8$ $\mathrm{M}_\odot$.

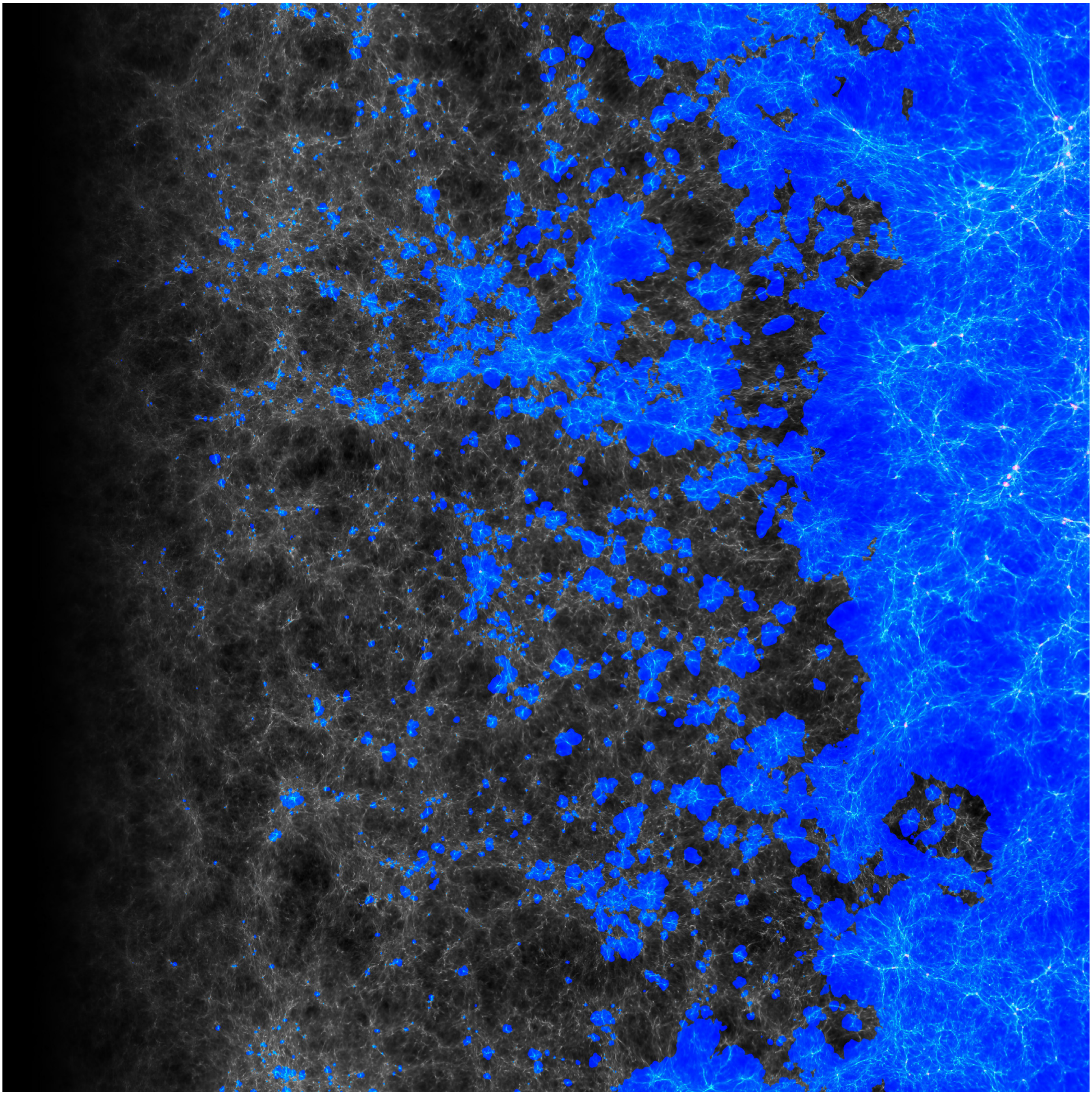

Figure 1. Illustration of the time history of reionization in the CoDa II simulation: a pseudo-light-cone image. A single, planar (x, y)-slice through the full (94 cMpc)^3 simulation cube was chosen for display, and a sequence of ∼1024 closely spaced time-slices of the 4096^2 grid cells in that plane was outputted, to sample the time-evolution from redshift $z = 150$ to $z = 5.8$. The image is made by the concatenation of vertical stripes, 4-cells wide, each stripe selected from another time-slice in the sequence, to show the time evolution as time increases to the right along the x-axis, from redshift $z = 150$ (left) to $z = 5.8$ (right). Blue regions are photoheated, while small, bright red regions (which can be seen more clearly by zooming the high-resolution figure) correspond to regions heated by supernovae feedback and accretion shocks. The green colour, on the other hand, denotes regions where ionization is ongoing and incomplete, and temperature has not yet risen to the ∼10^4 K typical of fully ionized regions. Brightness indicates the gas density contrast.

2 METHODOLOGY

The Cosmic Dawn simulation uses the fully coupled radiation hydrodynamics code RAMSES–CUDATON described in O16. This section outlines the basics of the code and the differences with CoDa I, as well as its deployment. The reader is referred to O16 for further detail. For quick reference, the parameters of the simulation are summarized in Table 1.

Table 1. Cosmic Dawn II simulation parameters summary.

Cosmology (Planck14)	
Dark energy density Ω_Λ	0.693
Matter density $\Omega_{\rm m}$	0.307
Baryonic matter density $\Omega_{\rm b}$	0.048
Hubble constant $h = {\rm H}_0/(100\,{\rm km\,s^{-1}})$	0.677
Power spectrum	
Normalization σ_8	0.8288
Index n	0.963
Setup	
Number of nodes (GPUs, cores used)	16 384 (16 384, 65 536)
Grid size	4096^3
Comoving box size $L_{\rm box}$	94.44 Mpc (64 h^{-1}Mpc)
Grid cells per node	128 × 128 × 64
Comoving force resolution dx	23.06 kpc
Physical force resolution at $z = 6$	3.3 kpc
DM particle number $N_{\rm DM}$	4096^3
DM particle mass $M_{\rm DM}$	$4.07 \times 10^5\ {\rm M}_\odot$
Average cell gas mass	$0.75 \times 10^5\ {\rm M}_\odot$
Initial redshift $z_{\rm start}$	150
End redshift $z_{\rm end}$	5.8
Star formation	
Density threshold $\delta_\star$	$50\ \langle\rho_{\rm gas}\rangle$
Efficiency $\epsilon_\star$	0.02
Stellar particle birth mass $M_\star$	$11732\ {\rm M}_\odot$
Feedback	
Massive star lifetime $t_\star$	10 Myr
Supernova	
Mass fraction $\eta_{\rm SN}$	10 per cent
Energy $E_{\rm SN}$	10^{51} erg
Radiation	
Stellar ionizing emissivity	$4.32 \times 10^{46}\ {\rm ph\,s^{-1}\,M_\odot}^{-1}$
Stellar particle escape fraction $f_{{\rm esc},\star}$	0.42
Effective photon energy	20.28 eV
Effective H I cross-section σ_E	$2.493 \times 10^{-22}\ {\rm m}^2$
Speed of light c	$299\,792\,458\ {\rm m\,s^{-1}}$

2.1 RAMSES–CUDATON

Since, the code is already described in detail in O16, we here only give a brief summary of it and highlight the differences in implementation / parameter choices between CoDa I and CoDa II. RAMSES–CUDATON results from the coupling of RAMSES (Teyssier 2002) and ATON (Aubert & Teyssier 2008). RAMSES handles gravity, hydrodynamics, and star formation and its mechanical feedback, while ATON handles photon propagation, hydrogen-photon thermochemistry, and computes the cooling terms. Finally, CUDATON results from the porting of ATON to CUDA (Aubert & Teyssier 2010) to enable it to take advantage of hybrid architectures featuring GPGPU (general purpose graphics processing unit). Thanks to this unique feature, we were able to perform CoDa II using the full speed of light, therefore circumventing potential artefacts arising with reduced speed of light approaches, highlighted by Deparis et al. (2019) and Ocvirk et al. (2019), or arising with infinite speed of light approaches in many ray-tracing codes. However, with this optimization, we are only able to use RAMSES in unigrid mode, i.e. without adaptive mesh refinement (hereafter, 'AMR'). This is one of the shortcomings that the EMMA code (Aubert, Deparis & Ocvirk 2015) and the RAMSES-RT code (Rosdahl et al. 2013) attempt to address.

2.2 Differences with CoDa I

2.2.1 *Star formation and calibration*

As in O16, we consider star formation using a phenomenological approach. In each cell with gas density larger than a gas overdensity $\delta_\star = 50$, we spawn new star particles at a rate given by

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

where $t_{\rm ff} = \sqrt{\frac{3\pi}{32 G\rho}}$ is the free-fall time of the gaseous component and $\epsilon_\star = 0.02$ is the star formation efficiency, instead of 0.01 in O16. The higher star formation efficiency of CoDa II is intended to achieve the end of reionization by redshift 6, whereas CoDa I finished reionization by $z = 4.6$. In practice, we calibrate this subgrid star formation efficiency by running a suite of reionization simulations in much smaller boxes than our final production run, but at the same grid and particle resolution as the latter, and then test the best-choice efficiency parameter by another suite of simulations of increasing box size. We adjusted the value of $\epsilon_\star$ so as to achieve a good agreement with the evolution of the cosmic star formation rate density inferred from observations of high-redshift galaxies (discussed in Section 3.3).

Unlike the first Cosmic Dawn simulation reported in O16, we do not require the cell temperature to be lower than $T_\star = 2 \times 10^4$ K in order to form stars: all cells, no matter their temperature, are eligible to forming stars above the density threshold. This is to account for the fact that, at higher resolution, such ionized cells may still host cold, neutral regions, which may still form stars. We will see in Section 3 that this choice has an impact on the strength of the radiative suppression of star formation.

The star particle mass at birth depends on the cell gas density, but is always a multiple of a fixed elementary mass $M_\star^{\rm birth}$, chosen to be a small fraction of the baryonic mass resolution. In this framework, with the box size and resolution of CoDa (see Section 2.3), we have $M_\star^{\rm birth} = 11732\ {\rm M}_\odot$. This mass is small enough to sample adequately the star formation history of even low-mass galaxies, and still large enough to mitigate stochastic variations in the number of massive stars per star particle. It is larger than the elementary mass used in CoDa I, which had $M_\star^{\rm birth} = 3194\ {\rm M}_\odot$. This change reduces the number of star particles in the simulation, which in turn reduces their computational cost, in memory and processing.

2.2.2 *Ionizing source model*

The radiative transfer considered here is of H ionizing radiation released by stellar sources associated with the star particles described in the previous section, subject to absorption by H I bound-free opacity. As with our subgrid star formation algorithm, the rate of release of ionizing photons by star particles into the grid cells in which they are located must be parametrized, so as to represent the absorption taking place at the subpc scales of the interstellar molecular clouds where stars are formed inside each galaxy.

Hence, we must add to our subgrid star formation algorithm described above a parametrized ionizing photon efficiency (henceforth, 'IPE'), the number of ionizing photons released per unit stellar baryon per unit time, into the host grid cell of each star particle. This IPE is based on an assumed stellar IMF and, since star particles are assumed to form embedded in high-density molecular clouds which are unresolved by our grid cells, we also adopt a subgrid stellar-birthplace escape fraction. We define this IPE, therefore, as $\xi_{\rm IPE} \equiv f_{{\rm esc},\star}\xi_{\rm ph,IMF}$, where $f_{{\rm esc},\star}$ is the stellar-birthplace escape fraction and $\xi_{\mathrm{ph,IMF}}$ is the number of ionizing photons emitted per Myr per stellar baryon. The subgrid stellar-birthplace escape fraction $f_{\mathrm{esc},\star}$ should not be confused with the *galactic* escape fraction, which is not a subgrid quantity put in 'by hand' but is rather the self-consistent outcome of our radiative transfer module CUDATON between grid cells, for all the cells associated with a given galactic halo. There is no need for us to adopt a galactic escape fraction, therefore. Our RT calculation finds the actual time-varying escape fraction for each galactic halo, which differs from one galaxy to the next.

Each stellar particle is considered to radiate for one massive star lifetime $t_\star = 10$ Myr, after which the massive stars die (triggering a supernova explosion) and the particle becomes dark in the H-ionizing UV. We adopted an emissivity $\xi_{\mathrm{ph,IMF}} = 1140$ ionizing photons/Myr per stellar baryon. This is similar to the ionizing emissivity of a BPASS model (Eldridge et al. 2017) for a Z=0.001 binary population with Kroupa initial mass function (Kroupa 2001), as presented in Rosdahl et al. (2018), averaged over the first 10 Myr. Although we do not follow chemical enrichment in CoDa II, Pawlik et al. (2017) showed that the metallicity in gas at the average density of the ISM is indeed close to Z =0.001 during the epoch of reionization, which validates the choice of this metallicity. We used a mono-frequency treatment of the radiation with an effective frequency of 20.28 eV, as in Baek et al. (2009).

Finally, we calibrated the stellar-particle escape fraction $f_{\mathrm{esc},\star}$ by adjusting the value in our set of smaller box simulations as mentioned above, so as to obtain a reionization redshift close to $z = 6$, which led us to adopt a value of $\mathrm{f}_{\mathrm{esc},\star} = 0.42$, close to the value $\mathrm{f}_{\mathrm{esc},\star} = 0.5$ used in O16. We note that the *net* escape fraction from the cell in which the star particle is located is less than or equal to this subgrid value adopted for the unresolved stellar-birthplace, since there is additional absorption of ionizing photons inside the cell, by the bound-free opacity of the simulated gas in that cell, which depends upon that cell's neutral H fraction and gas density, as tracked self-consistently by RAMSES–UDATON. The overall *galactic* escape fraction will be even smaller, since galactic haloes typically involve multiple cells, and one cell can absorb the photons released in another. Our radiative transfer module CUDATON explicitly accounts for this.

As in Coda I, we neglect here the possible release of hard, ionizing UV radiation associated with a phase in the evolution of our simulated galaxies in which they host active galactic nuclei (hereafter, AGNs). Some AGN sources may arise during the epoch of reionization, e.g. in rare, massive protoclusters (Dubois et al. 2012). However, they are very rare and have long been thought to be minor contributors to the cosmic budget of ionizing photons responsible for the reionization of hydrogen (e.g. Shapiro & Giroux 1987; Giroux & Shapiro 1996; Haardt & Madau 2012; Haardt & Salvaterra 2015), although they could be important for explaining the line-of-sight variations of the properties of the Ly α forest just after reionization (Chardin et al. 2015). The question of the possible AGN contribution to reionization has recently been revisited following claims by Giallongo et al. (2015) that the observed AGN luminosity function at high redshift may have been underestimated (Madau & Haardt 2015). Further studies indicate that the AGN contribution to reionization must be subdominant or else would violate other observational constraints (e.g. Worseck et al. 2016; D'Aloisio et al. 2017; Oñorbe et al. 2017; Qin et al. 2017; Mitra, Choudhury & Ferrara 2018), while subsequent updates of the observed quasar luminosity function also conclude that the AGN contribution to reionization is very small (e.g. Parsa, Dunlop & McLure 2018; Kulkarni, Worseck & Hennawi 2019).

2.3 Simulation setup

2.3.1 Initial conditions

Our initial conditions are a constrained realization of the ΛCDM universe, intended to reproduce the observed features of our Local Universe in a box centred on the Milky Way if evolved to the present, while at the same time serving as a representative sample of the universe-at-large with which to model global reionization. We generated specifically the initial conditions for this CoDa II simulation as part of the CLUES (*Constrained Local UniversE Simulations*) project. These initial conditions are constrained by observational data on the positions and velocities of galaxies in the Local Universe described in Tully et al. (2013) (referred to as *Cosmicflows-2*) to result in a simulation that resembles the Local Universe within the framework of *Planck* cosmology ($\Omega_{\mathrm{m}} = 0.307$, $\Omega_\Lambda = 0.693$, $H_0 = 67.77$ km s^{-1} Mpc^{-1}, $\sigma_8 = 0.829$, Planck Collaboration XVI 2014), in a comoving box 94 (i.e. 64 h^{-1}) cMpc on a side. It contains 4096^3 dark matter particles on a cubic-lattice with the same number of cells, which are also the grid cells for the gas and radiation. The mass of each dark matter particle is then 4.07×10^5 M$_\odot$. The CoDa II initial conditions have the average universal density for this chosen cosmology.

Sorce et al. (2016a) describe in detail the steps of the method used to build these constrained initial conditions. The main steps are summarized as follows:

(i) We started with a set of data points, the radial peculiar velocities of observed galaxies, i.e. radial peculiar velocities v_r at discrete positions r at $z = 0$. Galaxies in the radial peculiar velocity catalogue were then grouped (e.g. Tully 2015a, b) to produce a data set that traces the coherent large-scale velocity field, with non-linear virial motions removed that would affect the reconstruction obtained with the linear method as shown by Sorce, Hoffman & Gottlöber (2017) and Sorce & Tempel (2017).

(ii) Biases inherent to any observational radial peculiar velocity catalogue were minimized (Sorce 2015).

(iii) The three-dimensional peculiar velocity field and associated cosmic displacement field were then reconstructed by applying the Wiener-filter (henceforth, WF) technique – a linear minimum variance estimator (Zaroubi et al. 1995; Zaroubi, Hoffman & Dekel 1999) – to the grouped and bias-minimized radial peculiar velocity constraints.

(iv) To account for the displacement of mass elements located at the positions of the galaxies in the catalogue at $z = 0$, away from the positions of their Lagrangian patch in the early universe, caused by the growth of structure over time, the positions of the constraints were relocated to the positions of their progenitors using the Reverse Zel'dovich Approximation (Doumler et al. 2013). Additionally, noisy radial peculiar velocities are replaced by their WF 3D reconstructions (Sorce et al. 2014).

(v) Linear density and velocity fields were then produced, constrained by the modified observational peculiar velocities combined with a random realization of ΛCDM to restore statistically the missing structures using the constrained realization technique (Hoffman & Ribak 1991, 1992).

(vi) These density and velocity fields, extrapolated to $z = 0$ according to linear theory, have to be rescaled to an actual starting redshift ($z_i = 150$ for the CoDa II simulation) to build its constrained initial conditions, and the resolution was increased (MUSIC; Hahn & Abel 2011). Increasing the resolution implies adding higher frequency modes to the initial conditions, which introduce random small-scale features. For this purpose, we created 200 random realizations of the initial conditions with these higher frequency modes added, and tried them out by dark matter-only N-body simulations [using the GADGET-2 code (Springel 2005)] evolved to $z = 0$, in order to select that with the best match to the observed features of the Local Universe at $z = 0$, including the Local Group[3] (according to its mass, the separation between the Milky Way and M31, their mass ratio, and very nearby environment with M33 and Centaurus A counterparts). The production of other prominent features of the Local Universe, like the Virgo cluster (with close to the right mass at close to the right position), was also used as a selection criterion.

The CoDa II initial conditions were thus chosen so as to reproduce familiar features of the observed Local Universe within a volume 94 cMpc on a side, centred on the LG, as closely as possible. Towards this end, all 200 of the dark matter-only N-body simulations from these 200 constrained realizations (as described above) were examined at $z = 0$, and, amongst these, we selected a reduced set which best reproduced the Local Group, then further selected amongst these to meet our additional requirements, as follows:

(a) A pair of galaxies is located at the centre (by construction) with masses, M_{200}, between 5.5×10^{11} and $2 \times 10^{12}\,h^{-1}\,M_{\odot}$.

(b) There is no other halo more massive than $5.5 \times 10^{11}\,h^{-1}\,M_{\odot}$ within a sphere of radius 2.5 h^{-1} Mpc.

(c) Their separation is smaller than 1.5 h^{-1}Mpc.

(d) Their mass ratio is smaller than 2.

(e) They are located between 10 h^{-1} and 14 h^{-1} Mpc away from the Virgo cluster replica.

(f) There are haloes that could stand for M33 and Centaurus A (by far the most restrictive criterion).

Left with a dozen simulations with halo pairs, we selected the pair (and, thus, the simulation) that satisfied both requirements, that of a small separation and a small mass ratio: at $z = 0$, the distance between the two haloes of the pair is 0.85 h^{-1} Mpc and their mass ratio is 1.2 (with masses of 1.55×10^{12} and $1.3 \times 10^{12}\,h^{-1}\,M_{\odot}$).

These N-body simulations were at lower mass-resolution than our final CoDa II simulation but at a resolution high enough (at 512^3 particles, with particle mass $1.7\times10^8\,h^{-1}\,M_{\odot}$) to ensure that the resulting Local Group candidates would be stable to increasing the resolution still further when the same initial conditions were resimulated in our CoDa II production run.[4] The resolution of the best initial conditions thereby selected was then increased still further, to 4096^3 particles and cells, adding additional modes at higher frequency, still with MUSIC. In order to provide a companion dark matter-only N-body simulation to $z = 0$ from these constrained-realization initial conditions, we also coarsened the final, high-resolution CoDa II initial conditions from 4096^3 down to 2048^3, to run a GADGET-2 N-body simulation to $z = 0$, called CoDa II-DM2048.[5] The CoDa II-DM2048 simulation was then used to make a final comparison of the simulated LG and other features of the Local Universe by our CoDa II initial conditions with the observed Local Universe at $z = 0$. A comparison of the observed Local Universe with the outcome of simulating the CoDa II initial conditions to $z = 0$ by CoDa II-DM2048, is shown in Fig. 2.

The advantage of these new initial conditions with respect to those used for the first generation of CoDa simulations, CoDa I and CoDa I-AMR, is a large-scale structure that matches the local one down to the linear threshold ($\sim 3\,h^{-1}$ Mpc) on larger distances (the entirety of the box if not for the periodic boundary conditions) and with more accurate positions (~ 3–$4\,h^{-1}$ Mpc). In particular, this simulation contains a Virgo cluster at the proper position and with a mass in better agreement with recent observational mass estimates (Sorce et al. 2016b). While the Virgo replica had a mass of $7\times10^{13}\,h^{-1}\,M_{\odot}$ in the CoDa I simulation, its mass is now $2.2\times10^{14}\;h^{-1}\,M_{\odot}$, i.e. the cluster is 3 times more massive.[6] These properties will be important for the planned follow-up projects.

Finally, the baryonic initial conditions at the initial redshift of $z = 150$ were generated assuming a uniform temperature equal to that of the CMB at that time, with identical gas and dark matter velocity fields. The initial value of the H ionized fraction was taken to be the homogeneous, post-recombination-era freeze-out value at $z = 150$, as computed following standard recipes such as in RECFAST (Seager et al. 1999).

2.3.2 *Code deployment*

Like CoDa I (O16) and CoDa I-AMR (Aubert et al. 2018), CoDa II was performed on the massively hybrid CPU–GPU supercomputer Titan, at OLCF. The code was deployed on 16 384 Titan nodes and 16 384 GPUs (1 GPU per node), with each GPU coupled to 4 cores, for a total of 65 536 cores. Each node hosted 4 MPI processes that each managed a volume of $64 \times 128 \times 128$ cells.

2.4 Post-processing: friends-of-friends halo catalogues

We used the massively parallel friends-of-friends (hereafter, FoF) halo finder of Roy, Bouillot & Rasera (2014) with a standard linking length of $b = 0.2$ to detect dark matter haloes in the CoDa II simulation. As shown in O16, they are reliably detected down to $\sim 10^8\,M_{\odot}$.

2.5 Online data publication

We plan to make a subset of the data and higher level products publicly available through the VizieR data base at CDS Strasbourg.[7] The latter link will be active only after acceptation of the paper. The CoDa II catalogues will also be made available through the cosmosim data base hosted by Leibniz Institut für Astrophysik Potsdam.[8] They will contain halo properties such as dark matter halo mass, position, stellar mass, the age of the oldest and youngest stellar particles in the halo, the star formation rate, and the AB magnitude at 1600 Angstrom.

---

[3] The Local Group belongs to the non-linear regime and is thus not directly constrained but is induced by the local environment (Carlesi et al. 2016).

[4] The resolution of each of these N-body simulations must indeed be large enough for their candidate Local Group to be affected only barely when resimulated with even higher resolution, when yet higher frequency random modes are introduced to the initial conditions, thereby adding even smaller scale features.

[5] It corresponds to the simulation ESMDPL_2048 from the MultiDark project. It is fully available at www.cosmosim.org

[6] Higher masses in prefect agreement with observational estimates for Virgo have been obtained in realizations of 500 h^{-1} Mpc, but with the fixed 4096^3 grid size used here, such a box size results in a spatial and mass resolution unsuitable for this study.

[7] http://vizier.u-strasbg.fr/viz-bin/VizieR?-source = VI/146

[8] https://www.cosmosim.org/cms/simulations/cosmic-dawn/

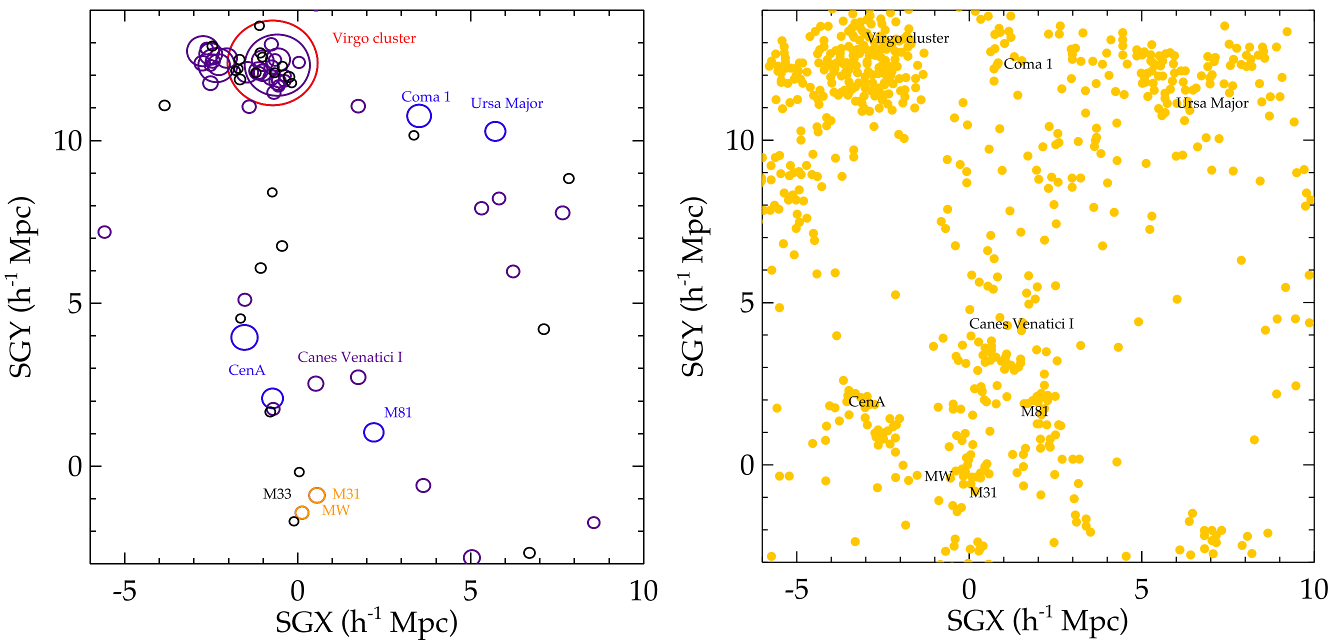

Figure 2. Comparison of the *N*-body haloes with masses $M > 2 \times 10^{11}$ M$_\odot$ in CoDa II-DM2048 at $z = 0$ in the Supergalactic Plane, in a subregion 16 h^{-1} Mpc on a side and 4 h^{-1} Mpc thick (left-hand panel) with the observed galaxies, groups, and clusters of the nearby Local Universe for the same region, a thin slice centred on the supergalactic equator, 4 h^{-1} Mpc thick (right panel), based upon the observed galaxies as plotted in fig. 4 of Courtois et al. (2013), taken from the Extragalactic Distance Database (http://edd.ifa.hawaii.edu). Circles of radius equal to the virial radii of the simulated haloes are shown on the left-hand panel. Features of the Local Universe and their analogues in the simulation are labelled. The correspondence between figure labels and Simbad (http://simbad.u-strasbg.fr/simbad/) names is given in Table 2.

Table 2. Correspondence table between names used in the article and Simbad names.

Name	Simbad name
Virgo cluster	Vir I
Coma I	Coma I group
Ursa Major	Ursa Major cluster
Canes Venatici I	CVn group
CenA	Cen A group
M81	M81 group

3 RESULTS

3.1 Global properties

In this section, we study the mean, global reionization history of the CoDa II simulation and compare it with available observational constraints. We demonstrate that the simulation is in agreement with most expectations from observations.

The most basic quantities to consider when gauging the success of a global EoR simulation are the evolution of the volume-weighted cosmic means of the neutral and ionized fractions, and the mean intensity of the ionizing radiation background. These are shown in Fig. 3, along with several observational constraints from Fan et al. (2006), Ouchi et al. (2010), and the *Planck* CMB Thomson-scattering optical depth τ (Planck Collaboration VI 2018).

Intergalactic H II regions grow in number and size over time until they overlap fully to finish reionizing the universe, before redshift 6. As shown in the plot of mean ionized hydrogen fraction versus redshift in Fig. 2, the ionized fraction was $x_{\rm H\,II} \sim 10^{-1}$, 0.5, 0.9, and 0.99 at $z = 8.5$, 7, 6.5, and 6.2, respectively, compatible with observationally inferred values by, e.g. Greig et al. (2017), Davies et al. (2018), Mason et al. (2018), and Hoag et al. (2019). At this point, the neutral fraction plot shows a characteristic, very steep decrease, down to $x_{\rm H\,I} \sim 10^{-4.6}$, after which the slope becomes slightly shallower. This transition marks the end of the EoR. According to this definition, reionization is complete in the simulation at redshift $z = 6.1$, in agreement with the brisk drop in neutral fraction between $z = 7$ and $z = 6$ indicated by the combined measurements of Ouchi et al. (2010) and Fan et al. (2006). The global timing of CoDa II is therefore in better agreement with these observations than was CoDa I. The latter ended reionization somewhat late, below redshift 5, although we found we could relabel the simulation redshifts of CoDa I by uniformly rescaling them by multiplying the simulation redshift by a factor of 1.3, in order to compare those results with observations or make observational predictions, with good agreement with these same observational constraints after this rescaling. Thanks to the new calibration and choice of parameters, our CoDa II simulation ended reionization earlier, by redshift 6, just as required by the observations, so we do not need to perform any such rescaling; we can compare its results here directly with observations and make direct predictions.

The post-reionization neutral hydrogen fraction in CoDa II is about 1 dex lower than observed. Such an offset is not uncommon and the literature shows that simulations of the EoR can exhibit a variety of such departures from the observed evolution of the ionized fraction measured from quasar lines of sight (e.g. Aubert & Teyssier 2010; Petkova & Springel 2011; Gnedin & Kaurov 2014; So et al. 2014; Zawada et al. 2014; Aubert et al. 2015; Bauer et al. 2015; Ocvirk et al. 2016; Aubert et al. 2018; Ocvirk et al. 2019; Wu et al. 2019). This issue is likely related to the evolution of the ionizing UV background in CoDa II, shown in Fig. 3. The surge between $z = 7$ and $z = 6$ correlates with the drop of the neutral fraction, and CoDa

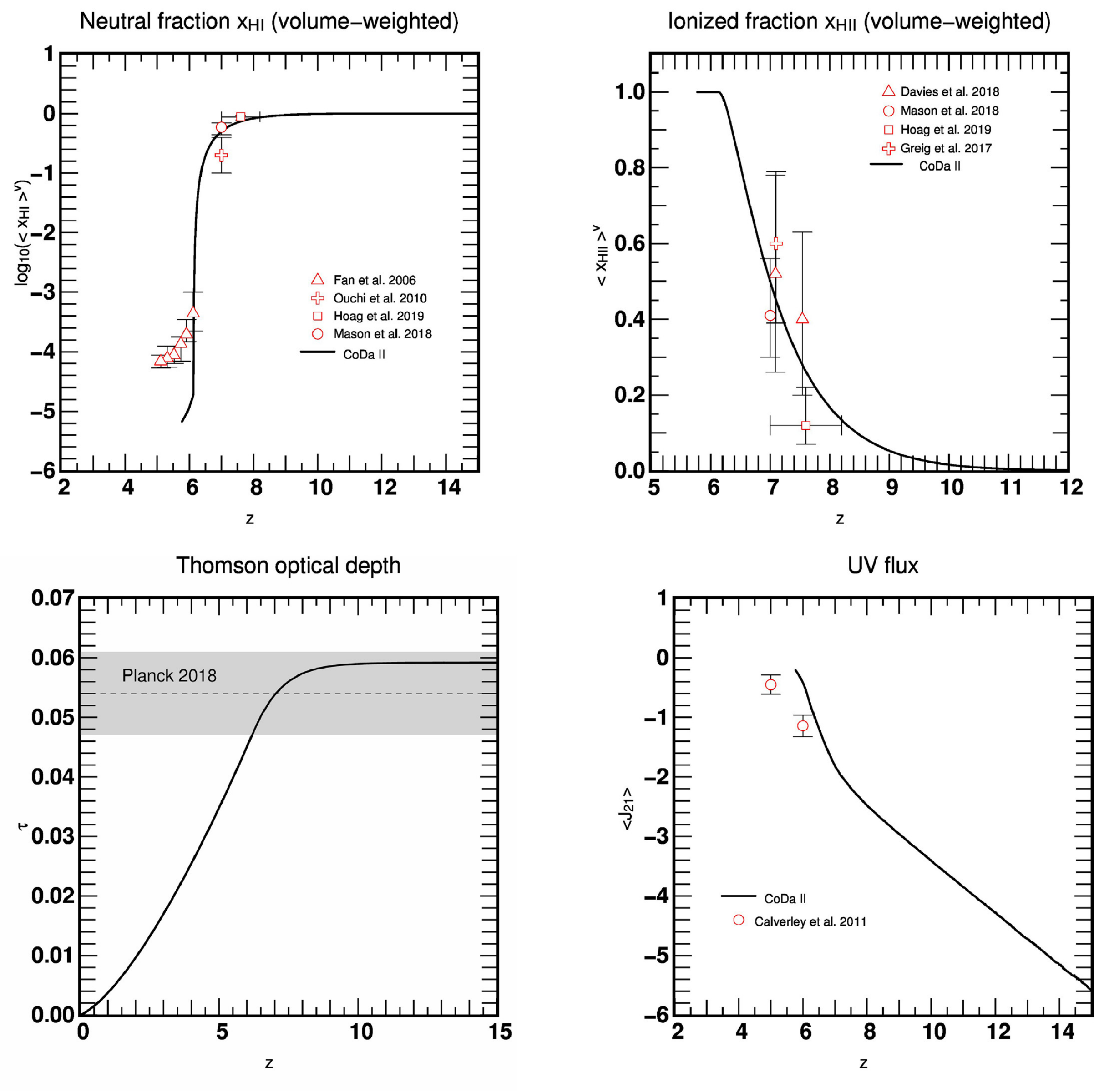

Figure 3. CoDa II simulation global reionization history, compared to observations: the evolution of the volume-weighted average of the neutral hydrogen fraction (top left) and ionized hydrogen fraction (top right), the Thomson scattering optical depth caused by free electrons in the IGM (integrated thru the simulated universe from $z = 0$ to each z, using the average ionized fraction in CoDa II at each z thru the end of the EOR and assuming the IGM was fully ionized thereafter), compared to the value integrated thru the entire post-recombination universe according to observations of CMB polarization anisotropy by Planck 2018 (bottom left), and the evolution of the average ionizing UV flux density (bottom right).

II overshoots the observations of Calverley et al. (2011), resulting in a lower-than-observed neutral fraction.

It is commonly accepted that resolving Lyman-limit systems is necessary to describe properly the population of absorbers in the IGM and predict the correct average neutral hydrogen fraction (Miralda-Escudé, Haehnelt & Rees 2000; Shukla et al. 2016). In particular, Rahmati & Schaye (2018) state a physical size for Lyman-limit systems (hereafter LLSs) of 1–10 kpc in their simulation. While this is compatible with the spatial resolution of CoDa II (∼3.3 physical kpc at $z = 6$ and less at higher redshifts), LLSs may be only marginally resolved and this could have an impact on our predicted neutral fraction after overlap. This is sufficient to resolve the scale of Jeans-mass filtering in the photoionized IGM after it is photoheated to $\sim 10^4$ K. However, according to Emberson, Thomas & Alvarez (2013) and Park et al. (2016), still smaller scale structure, on the scale of the Jeans mass in the cold, neutral IGM before it is reionized, must be resolved in order to capture fully this missing opacity of self-shielded regions, on the minihalo scale. A suite of higher resolution simulations with RAMSES–CUDATON are required to understand this issue, and are beyond the scope of this paper.

3.2 UV luminosity function

The UV luminosity function is the other important observable simulations of the EoR must match. To compute magnitudes, we first estimate the halo virial radius as R_{200}, the radius within which

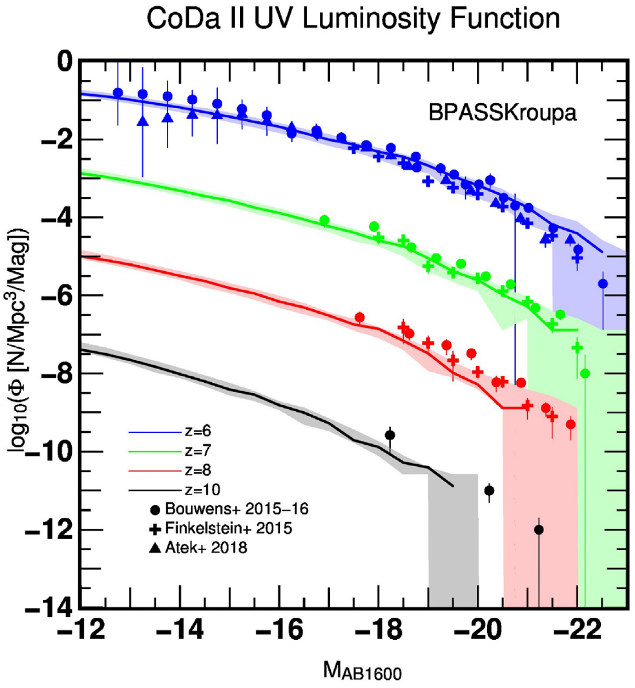

Figure 4. CoDa II UV luminosity functions, and comparison with observations. The full circles, crosses, and triangles with error bars are the observations from Bouwens et al. (2015, 2017), Finkelstein et al. (2015), and Atek et al. (2018), respectively, at $z = [6,7,8,10]$, from top to bottom, while the shaded area and the thick line show the envelope and the median of the LFs of 5 equal, independent, rectangular sub-volumes taken in the CoDa II simulation. For clarity, the LFs have been shifted downwards by 0, 2, 4, and 6 dex for redshifts $z = [6,7,8,10]$, respectively.

the average density of the dark matter is 200 times the cosmic mean dark matter density

$$R_{200} = \left(\frac{3\,M_{\rm FoF}}{4\pi \times 200\,\langle\rho_{\rm DM}\rangle}\right)^{1/3}, \tag{2}$$

where $M_{\rm FoF}$ is the halo mass given by FoF and $\langle\rho_{\rm DM}\rangle$ is the cosmic mean dark matter density. We then computed the $M_{\rm AB1600}$ magnitudes using our BPASS $Z = 0.001$ binary population model, assuming no dust extinction, for all the star particles associated with each individual halo, evolving the stellar populations inside each star particle over time, according to the birth-times of each particle.

The results are shown in Fig. 4, along with observational constraints. The latter are taken from Bouwens et al. (2015, 2017) and Finkelstein et al. (2015), which have been shown to be in broad agreement with a number of other studies including Oesch et al. (2013) and Bowler et al. (2015). We also show the $z = 6$ data of Atek et al. (2018), to highlight the current uncertainties of the very faint-end obtained using gravitational lensing in cluster fields. The luminosity functions (hereafter LF) have been shifted vertically for clarity. The shaded area shows the envelope of the LFs obtained for 5 non-overlapping, rectangular subvolumes of the CoDa II simulation. Each of these subvolumes spans ∼168 000 cMpc^3 ($\Delta x = {\rm L_{box}}/5$, $\Delta y = \Delta z = {\rm L_{box}}$, i.e. 1/5 of the full box volume), which is similar to the volume probed by CANDELS-DEEP at $z = 6$. The resulting envelope therefore illustrates the expected effect of cosmic variance at $M_{\rm AB1600} > -20$. The thick solid line shows the median of these 5 LFs. The CoDa II LF is in good agreement with the observations at all redshifts, although a small shift (i.e. in the sense of undershooting) can be seen for $z > 6$, which increases at higher redshifts. We note, however, that given the error bars and the spread of the observed LFs, our results remain compatible with the observations.

3.3 Cosmic star formation rate density

Observations of high-redshift galaxies have been used to infer the global star formation rate density (‘SFRD’) of the universe at early times, indirectly, by measuring the UV LF of galaxies and assuming an IMF for the stars responsible for the UV starlight that escapes from each galaxy at 1600 Å, subject to a possible correction for an unknown amount of attenuation inside each galaxy by interstellar dust. In fact, there is considerable uncertainty that results from the fact that the bulk of the UV luminosity per galaxy is released by the most massive stars in the galaxy’s IMF, as long as star formation continues, while the bulk of the mass in stars formed is generally assumed to be dominated by the unseen lower mass stars in the conventionally assumed IMFs. Furthermore, once star formation ceases or slows down significantly inside a galaxy, its stellar population will evolve passively and its UV luminosity decline, after its most massive stars have died a fiery death as SNe and no longer contribute their UV starlight. In a fully coupled radiation-hydrodynamical simulation of reionization and galaxy formation like CoDa II, in which the release of ionizing UV starlight and SN energy inside galaxies are very important in affecting the evolution both inside and between galaxies, it is only those massive stars that matter, and any inference about the total mass of stars formed along with those massive ones is similarly subject to an assumption about the shape of the IMF over the mass range which is well below the level of 10’s of $M_\odot$ which dominates the production of ionizing photons. Nevertheless, with these caveats, we can compare our predictions for the SFRD in CoDa II with the observational constraints by making the same underlying assumption about the IMF for both the simulated SFR and the inferred one from the observed galaxies, as follows.

We computed CoDa II’s cosmic SFRD as the global star formation rate of the whole CoDa II volume divided by the box volume. The results are shown in Fig. 5. As a first remark, we note that the cosmic SFRD in CoDa II, as in CoDa I (O16), increases at all times, unlike the simulation of So et al. (2014), which shows a decline at late times. In the latter, the authors suggested this could be due to the small box size they used. This may indeed be the case, since CoDa II is more than 95 times larger in volume. The fact that the global SFRD in CoDa II increases continuously over time right through the end of the EOR was noted previously for CoDa I in O16. As pointed out there, this conflicts with the prediction by Barkana & Loeb (2000) that the global SFR would suffer a sharp drop as the end of reionization approached, caused by a jump in the local IGM Jeans-mass filter scale inside H II regions as reionization overtook a significant fraction of the volume of the universe. While the SFR dependence on halo mass does reflect this suppression effect at lower mass, the global SFR integrated over halo mass in CoDa II, like in CoDa I, is apparently dominated by galaxies more massive than those whose SFR is suppressed by reionization feedback.

In order to compare this integrated SFRD of CoDa II with the observations of high-redshift galaxies, we must put the latter on a self-consistent footing with the assumptions of the CoDa II stellar emissivity and IMF. As described in Section 2.2.2, our adopted value of the emissivity in ionizing photons per stellar baryon, $\xi_{\rm ph,IMF} = 1140$, assigned to each star particle when they form, is consistent with the ionizing emissivity of a BPASS model for a $Z = 0.001$ binary population with Kroupa IMF, averaged over the first 10 Myr. In practice, this means that, to compare the mass in star particles formed per time in the simulation volume with that

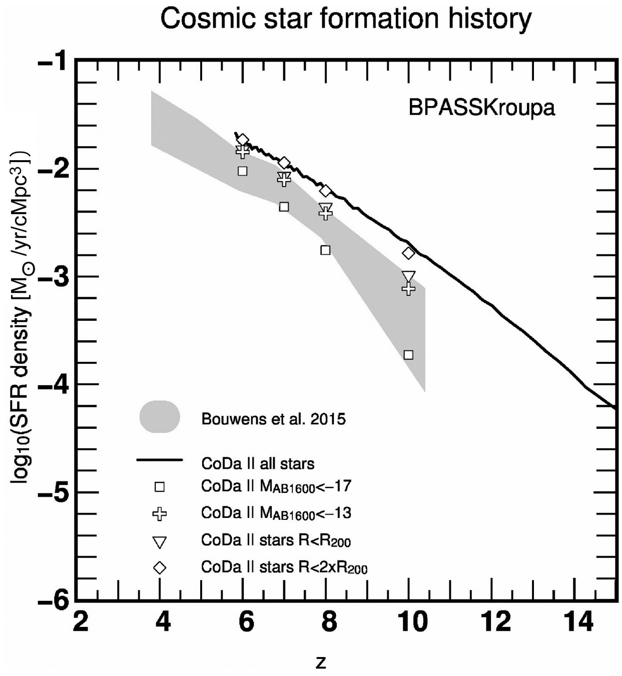

Figure 5. Evolution of the cosmic star formation rate density (SFRD) as simulated by CoDa II, compared with that inferred from observations: (1) Using all star particles in the simulation box (thick line), (2) accounting for haloes above realistic magnitude cuts (crosses and squares), and (3) using all star particles within the R_{200} and $2R_{200}$ spheres of haloes (triangles and diamonds). Observational result is taken from table 7 of Bouwens et al. (2015), but with the observed SFR decreased by 0.34 dex to account for the different $M_{\mathrm{AB}1600}$ – SFR conversion assumed here, resulting from our BPASS source model, rather than the Salpeter IMF adopted by Bouwens et al. (2015).

inferred from observations of the UV LF of high-redshift galaxies, those observations must be interpreted in terms of the same assumed IMF when inferring the observed SFRD and its cosmic evolution. Also shown in Fig. 5 are the observational constraints taken from Bouwens et al. (2015): the grey area shows the envelope bracketing the dust-corrected and dust-uncorrected SFRD's. We decreased this observed SFRD by 0.34 dex to account for the different $M_{\mathrm{AB}1600}$ – SFR conversions resulting from our assumption of the BPASS model, rather than the Salpeter IMF for single stars assumed by Bouwens et al. (2015). We computed this offset by applying the methodology of Madau, Pozzetti & Dickinson (1998) to our BPASS source model. This involves computing the 1600 Å luminosities of a population undergoing an exponentially decaying burst of star formation, resulting in a conversion factor:

$$L_{1600} = \mathrm{const}\,\frac{\mathrm{SFR}}{\mathrm{M}_\odot\mathrm{yr}^{-1}}\,\mathrm{erg\,s}^{-1}\,\mathrm{Hz}^{-1}, \qquad (3)$$

where const = 1.75×10^{28} with our source model, and const = 8×10^{27} in B15.

The CoDa II total SFRD (thick solid line) overshoots the observations by about 0.1 dex at $z = 6$. However, the cosmic SFRD inferred by B15 only includes galaxies with $M_{AB1600} < -17.7$, and does not account for fainter (undetected) objects. Therefore, to compare our CoDa II results to B15 in a fair way, we need to apply similar magnitude cuts to both. With such cuts (squares in Fig. 5), CoDa II falls within the observationally favoured SFRD at $z = 6$, and very close to the observations at higher redshifts.

At $z = 6$, the SFRD inferred from the CoDa II sample with $M_{\mathrm{AB}1600} < -17$ is 2.5 times smaller than the total SFRD in the box, and the offset grows with increasing redshift. However, not all that offset is due to the magnitude cuts: the triangles in Fig. 5 show the cosmic SFRD obtained when considering only those star particles which are located within a sphere of radius R_{200} centred on the centre of each halo, as opposed to considering all the star particles in the box. Even when considering all haloes, the R_{200}-bounded SFRD is always smaller than the full-box SFRD, showing that some of the star particles are located either beyond R_{200} for haloes detected by the FoF halo finder or else are not associated with haloes because their haloes are below our assumed lower halo mass limit of 10^8 $\mathrm{M}_\odot$ (246 dark matter particles). Haloes, in particular at the high redshifts studied here, are strongly triaxial, so that their particles and stars may extend beyond a sphere of radius R_{200}. This aspect is illustrated in more detail in Appendix C, where we show that stars beyond R_{200} are still associated with the most nearby dark matter halo. We further confirm this by considering all stars within $2R_{200}$ (diamonds in Fig. 5), yielding an SFRD which matches fhe total full box SFRD at redshifts below $z = 8$. Only at $z = 10$, some of the stars may be associated to small haloes below the assumed limit.

Finally, comparing the 'detectable' CoDa II galaxy population with the total SFRD within the R_{200} spheres of all haloes, we find that the CoDa II sample with $M_{\mathrm{AB}1600} < -17$ at $z = 6$ accounts for about 63 per cent of the halo $R < R_{200}$ SFRD, and only 18 per cent at $z = 10$. On the other hand, in the case of deeper detection limits such as in lensing cluster fields where lensing magnification makes it possible to detect fainter haloes with $M_{\mathrm{AB}1600} < -13$, these fractions increase to 94 per cent and 76 per cent, respectively.

3.4 The very faint end of the luminosity function

We can also use CoDa II to investigate the LF at magnitudes fainter than currently observed. We show in Fig. 6 the whole LF, going down to $M_{\mathrm{AB}1600} = -5$, for 4 redshifts. The progressive build-up of the LF with cosmic time is clearly apparent, both at the bright end, with more and more massive haloes hosting vigorous star formation, and at the faint end, made of passively evolving galaxies. The maximum of the LF is reached at $M_{\mathrm{AB}1600} \sim -11$ at all redshifts. At redshifts lower than 10, a plateau develops faintwards of −11, extending to fainter magnitudes as time goes by. At $z = 6$, the plateau extends down to $M_{\mathrm{AB}1600} = -6$. We show in Fig. 7 that this [−11, −6] magnitude range corresponds to masses of 10^8–10^9 $\mathrm{M}_\odot$, i.e. the mass range where SFR suppression by ionizing radiation takes place (see Section 3.5). Indeed, if haloes below $\sim 10^9$ $\mathrm{M}_\odot$ have their SFR's suppressed after they are overtaken by reionization and, thereafter, have only the luminosity associated with the passive evolution of their low-mass stars (after the high-mass stars finished their lives as SNe), we expect the number density of galaxies with magnitudes fainter than $M_{\mathrm{AB}1600} = -11$ to grow continually, as more and more volume is overtaken by reionization, as shown by our figure.

The full LFs for these 4 redshifts are given in the appendix in tabular form.

3.4.1 *Halo mass and UV luminosity*

The results of CoDa II show that the UV continuum luminosity of a galaxy is correlated with the mass of its galactic halo. The average luminosity of haloes of the same mass at a given redshift increases with halo mass, consistent with the correlation found for Coda I (O16). This is shown for CoDa II in the panels of Fig. 7,

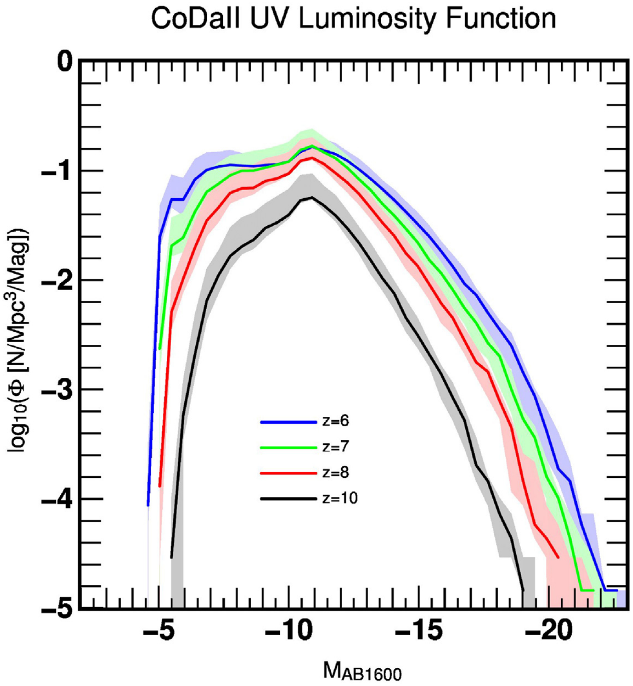

Figure 6. CoDa II UV luminosity functions, including the very faint end, at 4 epochs. The shaded area and the thick line show the envelope and the median of the LFs of 5 equal, independent, rectangular sub-volumes taken in the CoDa II simulation. The full LFs are given in a table in appendix.

which correspond to the galaxy populations of the $z = 6$ and $z = 10$ UV LF's in Fig. 4. Moreover, for each halo of a given mass, the luminosity can fluctuate significantly over time, which is reflected in the vertical spread of the distributions in Fig. 7. At $z = 6$, the vertical dispersion increases with decreasing halo mass. Moreover, the spread at low mass is larger at $z = 6$ than $z = 10$. The increase in dispersion is not driven by the brightest galaxies of the low-mass haloes. Indeed, the brightest galaxies residing in a haloes of mass a few times 10^8 $\mathrm{M}_\odot$ have $M_{\mathrm{AB1600}} \sim -12$, in both snapshots, and contain a couple of 1–2 Myr old elementary stellar particles in CoDa II. By contrast, the faintest galaxies of the $z = 10$ snapshot have $M_{\mathrm{AB1600}} \sim -5.5$, while the faintest galaxies of the $z = 6$ snapshot are as faint as $M_{\mathrm{AB1600}} \sim -5$. This is the signature of passively evolving galaxies: their star formation has stopped, and their stellar populations progressively fade, making them fainter as redshift goes down.

3.5 Impact of radiative feedback on galaxy formation

As described in Section 1, the CoDa suite of fully coupled radiation-hydrodynamics simulations was designed to address the long-standing question of how galaxies that released the UV H-ionizing radiation that reionized the universe were affected by the photoheating that accompanied photoionization during the EoR, acting back on their gas content and star formation rates, and on the progress of reionization, itself. In O16, we summarized the globally averaged effect of this feedback on the star formation rates of haloes of different mass at different redshifts in CoDa I, showing that the SFR in low-mass haloes below a few times $10^9\mathrm{M}_\odot$ was reduced towards the end of reionization, as more and more of the volume was reionized. In Dawoodbhoy et al. (2018), we studied these CoDa I results in more detail, going beyond the globally averaged results to identify the causal connection between the suppression of star formation in low-mass haloes and reionization by showing that SFR suppression followed the local time of arrival of reionization at the location of each galaxy. In that case, we showed, reionization was not limited to the late stage of reionization, as it appeared from the globally averaged results in O16, but was happening throughout the EoR, at different times in different places.

In Dawoodbhoy et al. (2018), we summarized the literature that had attempted to characterize this feedback effect and establish the mass range of haloes subject to suppression, by variety of approximations over the years, including simulations. We refer the reader to that summary and the references therein, for this background. In the context of 3D galaxy formation simulations, a number of authors have attempted to address this impact of radiative feedback on galaxy formation, by adopting a pre-determined, uniform UV background like that modelled by Haardt & Madau (1996, 2012), and more recently with fully coupled RHD simulations (Pawlik, Milosavljević & Bromm 2013; Jeon et al. 2014; Wise et al. 2014; Aubert, Deparis & Ocvirk 2015; Pawlik, Schaye & Dalla Vecchia 2015; Rosdahl et al. 2015; Ocvirk et al. 2016; Dawoodbhoy et al. 2018; Rosdahl et al. 2018; Katz et al. 2019; Wu et al. 2019), investigating a possible suppression or reduction of star formation in low-mass galaxies. Here, we will highlight some of the results of the CoDa II simulation on this issue and compare them with those in CoDa I.

We computed the instantaneous SFR of CoDa II haloes as the stellar mass formed within a sphere of radius R_{200} centred on the dark matter halo centre of mass, during the last 10 Myr, divided by a duration of 10 Myr. Fig. 8 shows the instantaneous SFR that results, as a function of the instantaneous mass of the dark matter halo, for several redshifts.

There is a general trend for more massive haloes at all epochs to form more stars, as seen in Ocvirk, Pichon & Teyssier (2008) and O16. Here, we find a trend: SFR $\propto M^{\alpha}$ for $M > 10^{10}$ $\mathrm{M}_\odot$ with a slope $\alpha \sim 5/3$, just as in CoDa I. This slope is compatible with the values found in the literature i.e. $1 < \alpha < 2.5$, in numerical and semi-analytical studies, such as Tescari et al. (2009), Hasegawa & Semelin (2013), Yang et al. (2013), Gong et al. (2014), Aubert et al. (2015), and Ocvirk et al. (2016). Moreover, Fig. 8 shows a smooth decrease in SFR for the low-mass haloes, around $\sim 2 \times 10^9$ $\mathrm{M}_\odot$, relative to this slope at higher mass. Below this mass, the SFR–mass relation is slightly steeper, and gets even steeper at lower redshifts, as reionization advances. This steepening at late times during the EoR is a signature of suppression by reionization feedback, also seen in the globally averaged results from CoDa I. However, the strength of the suppression is stronger in CoDa I than in CoDa II.

In order to understand the origin of this evolution in the SFR–mass relation, in particular with respect to the relative importance of supernova feedback and radiative feedback, we performed two additional control simulations, with the same parameters as CoDa II, but in a smaller box, 8 h^{-1}cMpc on a side, isolating the feedback physics:

(i) *SN, RT*: this simulation has the exact same physics as CoDa II, i.e. SNe and radiation.

(ii) *SN, no RT*: this simulation was run with SNe but without radiation (and therefore does not undergo global reionization).

Both simulations ran to redshift $z = 5$. Fig. 9 shows the star formation history of haloes present in the final snapshot at $z = 5$, binned according to their final masses at this redshift. Again, the trend of increasing SFR with halo mass is clear. Also, expectedly, since $z = 5$ haloes were less massive at, say, $z = 15$, their SFR is lower at higher redshift. For all mass bins, the SFR of the run with radiation is smaller than without radiation. While the effect is very small

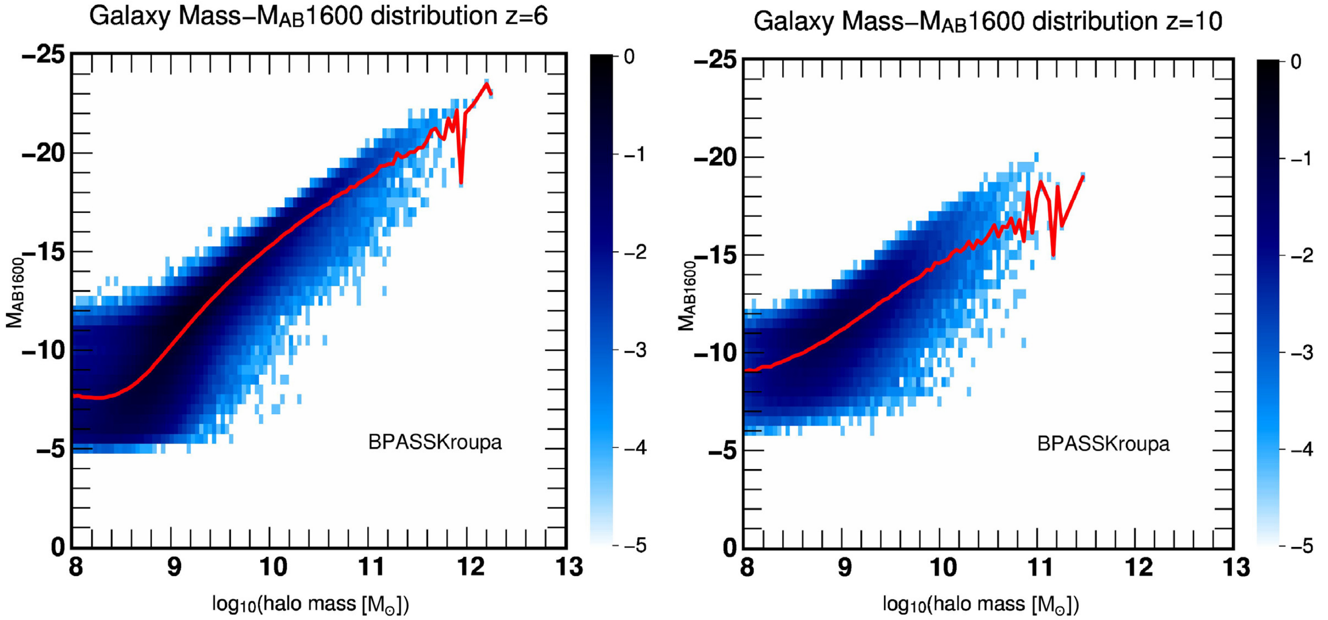

Figure 7. CoDa II galaxy mass – magnitude distributions at $z = 6$ and $z = 10$. The colour indicates the galaxy comoving number density in N/Mpc^3/Mag/log(M$_\odot$). The red line indicates the average magnitude for each mass bin.

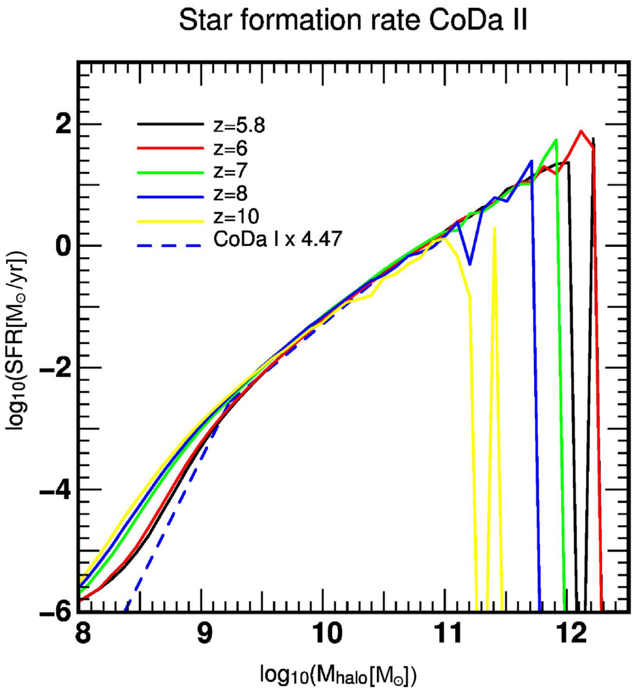

Figure 8. Instantaneous star formation rate per halo as a function of instantaneous halo mass, for various redshifts. The instantaneous SFR is computed as the stellar mass formed within an R_{200} radius sphere centred on the dark matter halo centre of mass, during the last 10 Myr, divided by a duration of 10 Myr. The solid lines show CoDa II data while the dashed line shows the post-overlap CoDa I data, multiplied by a factor to make it match the high-mass regime end, for easier comparison.

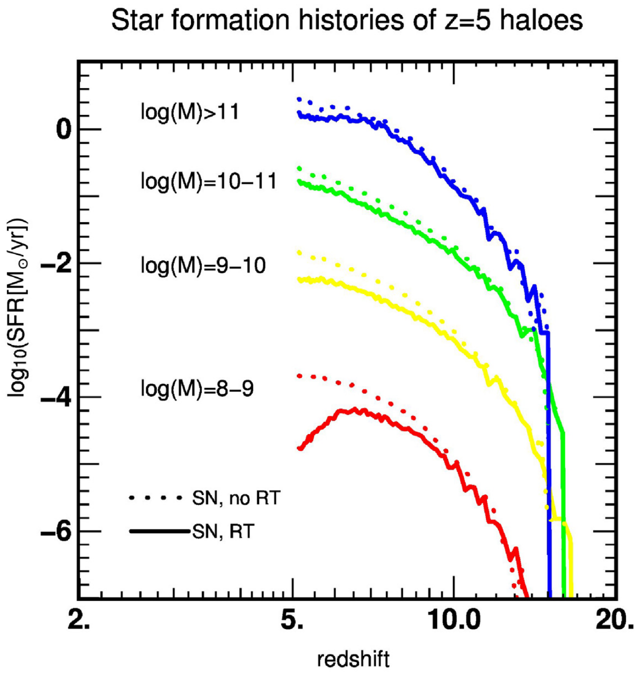

Figure 9. Star formation histories of 4 halo mass bins for two simulations in a test box 8 h^{-1}cMpc on a side. *Solid line*: with full fiducial physics (supernovae and radiative transfer), *dotted line*: with supernovae only, no radiation. The mass bins correspond to the haloes final mass (i.e. halo mass measured in the final simulation time-step, at $z \sim 5$.).

in the highest mass bin, it gets stronger the lower the mass. The lowest mass bin, in particular, shows a striking behaviour: the SFR increases with time, and then starts to decrease around $z = 6.5$, i.e. during overlap and shortly before the end of reionization. This is markedly different from the run without radiation, in which the low-mass haloes have a monotonically increasing SFR. This demonstrates that the UV ionizing radiation suffusing the Universe during the EoR is indeed the cause of the withering SFR of the low-mass haloes in CoDa II.

This gradual suppression reflects a radiation-driven reduction in the baryonic fraction of low-mass haloes (Dawoodbhoy et al. 2018; Sullivan, Iliev & Dixon 2018). In contrast, more massive haloes are

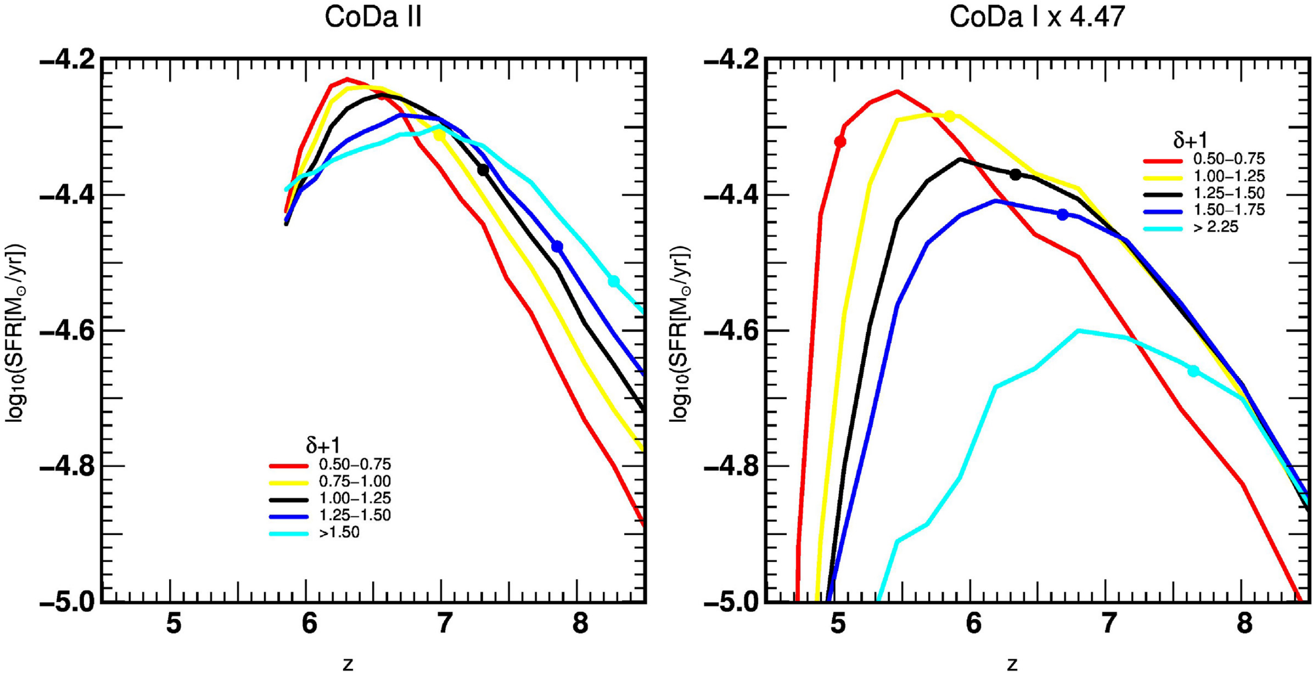

Figure 10. Average star formation histories of low-mass haloes ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$) in CoDa II (left) and CoDa I (right), for haloes present in that mass bin just after global reionization ended ($z = 5.8$ for CoDa II and $z = 4.44$ for CoDa I) for different environments, quantified by their overdensity $\delta + 1 = \rho/\bar{\rho}$ computed on 4 h^{-1}cMpc sub-boxes. The filled circle on each curve shows the average reionization redshift (when ionized fraction was 0.5) of each overdensity bin. Note that the overdensity bins we used are slightly different between the two simulations, because the lower reionization redshift of CoDa I results in a higher maximum overdensity.

able to retain their gas and keep on forming stars at a very similar rate pre- and post-overlap. We note that the reduction in the SFR of low-mass haloes due to global photoheating during reionization is less severe than in CoDa I. Indeed, the sub-grid model for star formation in CoDa I included an additional criterion, allowing stars to form only in cells colder than $T = 20\,000$ K, thereby making a smaller amount of cells in the simulation eligible for star formation after local reionization occurs. This more stringent requirement led to a more severe suppression of star formation in low-mass haloes (dashed curve in Fig. 8). This requirement was relaxed in CoDa II, where star formation is purely based on a gas overdensity threshold and an efficiency parameter (see Table 1). It is interesting to note that suppression nevertheless still clearly happens in Coda II at low masses, although in a slightly less severe form.

We performed a resolution study, presented in Appendix A, where we show that increasing the spatial resolution of our setup by up to a factor 4 in space and 64 in mass (although in a smaller test box), results in similar star formation suppression in the low-mass halo range (10^8–10^9 M$_\odot$) after reionization. A quick comparison with the literature, e.g. Pawlik et al. (2015) and Wu et al. (2019), shows reasonable agreement. Therefore, star formation suppression in CoDa II is not likely to be an artefact due to numerical resolution.

3.6 Environment and the timing of SFR suppression in low-mass galaxies

Reionization is a patchy process, as shows Fig. 1. Therefore in the context of the radiative feedback demonstrated in Section 3.5, we may expect different environments (galaxy cluster or group progenitor, galaxy neighbourhood, voids), to affect low-mass galaxies' star formation rates with different strengths and at different times. The CoDa II simulation is well suited to such investigations thanks to the variety of environments it contains and its self-consistent modelling of the inhomogeneous progress of reionization. Here, we wish to illustrate this aspect, while leaving a more complete, detailed study in the spirit of Dawoodbhoy et al. (2018) for a later paper.

We computed the overdensity in which CoDa II galaxies reside at the end of the simulation, as $\delta + 1 = \rho/\bar{\rho}$ averaged in sub-boxes of 4h^{-1}cMpc on a side. This is the typical scale of a galaxy group progenitor, and therefore this quantifies the environment at an intermediate, meta-galactic scale. We divide our galaxy population into 5 overdensity bins and compute the average star formation histories of CoDa II low-mass haloes ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$ at $z = 5.8$. The results for CoDa II are shown in the left-hand panel of Fig. 10 for the 5 overdensity bins. Each curve shows the same characteristic parabolic shape seen for low-mass haloes in Fig. 9, with a timing and amplitude clearly correlated with the local overdensity.

In particular, haloes in dense environments have a more intense early ($z = 8$) star formation history. This is likely caused by the inherently faster formation of their dark matter haloes and therefore larger mass accretion rates, as compared to underdense regions, as shown in Maulbetsch et al. (2007).

The filled circle overplotted on each curve shows the average redshift at which the cells in that overdensity bin reached the point of ionized fraction 0.5.

In the highest density environments, the star formation in low-mass haloes starts early, culminates around $z = 7$, and then decreases.

In underdense regions, the same scenario unfolds, with a delay: star formation starts later, reionization happens later, and therefore the average star formation rate peaks and decreases later as well.

Strikingly, at $z \sim 6.2$, low-mass haloes in underdense regions form stars at a higher rate than their counterpart in overdense regions. However, this episode is short-lived, and as the epoch of reionization comes to an end, the average star formation rates of low-mass haloes seem to converge to a similar value independent of their environment.

These results for CoDa II are consistent with the results found for CoDa I by Dawoodbhoy et al. (2018), both for the correlation of the local reionization times (or redshifts) with local overdensity and the suppression of SFR in low-mass haloes following their local reionization time. We note that in Dawoodbhoy et al. (2018), the definition used for reionization time (or redshift) was the 90 per cent-ionized time, which is somewhat lower redshift than the 50 per cent-ionized time used here. Hence, the delay found here for CoDa II between the half-ionized redshift and the redshift at which the SFR peaks and then declines for a given overdensity bin, due to suppression, is larger here due to this choice of half-ionized reionization redshift.

To make a direct comparison here between CoDa II and CoDa I, we show in the right-hand panel of Fig. 10 the same analysis, performed this time on CoDa I data, for the same range of halo masses ($2.5 \times 10^8\,\mathrm{M}_\odot < M < 7.5 \times 10^8\,\mathrm{M}_\odot$) just after the end of global reionization, i.e. $z = 4.4$. The SFRs for CoDa I have been multiplied by 4.47 as in Fig. 8 to make them comparable in amplitude to CoDa II results. Because of the lower redshift, the maximum overdensity of the simulation is higher. The behaviour of low-mass galaxies share several similarities between the two simulations: in the mass range considered, the SFRs of all overdensity bins rise, culminate, and decay, correlated with the average reionization times of their bin. Also, the earlier suppression taking place in overdense regions is confirmed in CoDa II, although in a less severe form, as was already noted in Section 3.5. We note however that we are unable to predict what would happen if CoDa II had run to lower redshift. It could be that at $z = 5$, CoDa II low-mass haloes would end up with very low SFRs, comparable to those of CoDa I. Further simulations will be needed to address this aspect conclusively.

3.7 Mapping the inhomogeneous timing of reionization in the Local Universe: ‘islands in the stream’

The effects of reionization arrive at different times in different places, not only globally, but within the Local Universe, as well. The CoDa II simulation, with its constrained initial conditions, was designed to capture this aspect of the reionization of the Local Universe. To illustrate it, we computed the reionization redshift of CoDa II gas cells as the redshift at which the ionized fraction of the cells reaches 50 per cent for the first time. This was done using the level 10 ionized fraction of the simulation, yielding a reionization redshift field of 1024^3 cells (4096^3 is level 12). Using the level 10 instead of level 12 reduces the memory requirements of this process by 64, but does not impact our results. A reionization map is then created by a slice through this reionization redshift field 1 h^{-1}cMpc thick, shown in Fig. 11. The progenitors of several major Local Universe objects analogues were traced back from $z = 0$ to $z = 6$ using the CoDa II-DM2048 simulation. They are labelled on Fig. 11 to allow us to gauge their reionization history and radiative influence region during the epoch of reionization in the Local Universe.

The map shows local maxima of z_r, which are ‘islands’ of early reionization surrounded by a ‘sea’ in which reionization happened at a later time. Such regions are internally reionized, in an inside-out pattern.

As can be seen in Fig. 11, the MW and M31 are examples of such islands, not only surrounded by the sea, but separated from each other by a channel, so they appear not to have been ionized from outside or to have significantly influenced each other during their reionization. For comparison, An example of a reionization map in a scenario where M31’s progenitor partially or completely reionizes the MW progenitor is shown in fig. 4 of Ocvirk et al. (2013). However, the latter scenario is disfavoured because it requires an extreme feedback model, as discussed in Ocvirk et al. (2013).

Investigating further out, we see that the Virgo cluster’s progenitor, is also separated from the LG progenitors by later-to-reionize regions. It is therefore not responsible for the reionization of the LG in the CoDa II simulation. A similar result was obtained by the CoDa I-AMR simulation with the EMMA code, as described in Aubert et al. (2018), starting from CoDa I initial conditions, which were an earlier version of CLUES constrained realization initial conditions of the Local Universe.

Finally, the CenA and M81 groups’ progenitors also produced their own self-reionized islands. Despite their proximity, they do not appear to have affected the reionization of the Local Group, or each other, as the respective islands of CenA, M81, MW, and M31 are well-separated from each other.

4 CONCLUSIONS

CoDa II is a state-of-the-art, fully coupled radiation-hydrodynamics simulation of reionization and galaxy formation, in a volume large enough to model global reionization and with a uniform spatial resolution high enough to resolve the formation of all the galactic haloes in that volume above $\sim 10^8\,\mathrm{M}_\odot$. Based upon the massively parallel, hybrid CPU–GPU code RAMSES–CUDATON, CoDa II utilizes 4096^3 dark matter particles on a uniform cubic lattice of 4096^3 grid cells, which is also used for the gas and radiation field, in a comoving box 94 Mpc on a side, centered on the Local Group in a constrained realization of the Local Universe. CoDa II is a successor to our first such simulation, CoDa I, described in O16 and Dawoodbhoy et al. (2018), but with new initial conditions, updated parameters for the background universe, and a recalibration and modification of our subgrid model for star formation, chosen to ensure that CoDa II finished reionizing the universe before $z \sim 6$. For both CoDa I and II, our initial conditions were based upon a constrained realization of the Gaussian random density fluctuations for the ΛCDM model, chosen by ‘reverse engineering’ the structure of the present-day Local Universe, to yield initial conditions in the linear regime that, when evolved over cosmic time, will reproduce that observed structure. This makes CoDa I and II both suitable for modelling both global and local reionization and its impact on galaxy formation, including the progenitors of the Local Group. The galaxy data and background cosmology on which the constrained realization for CoDa II is based were updated and improved relative to those used for CoDa I.

CoDa II was performed on the hybrid CPU–GPU supercomputer Titan at OLCF, using 16 384 nodes and 16 384 GPUs (one GPU per node, 4 cores per GPU), for a total of 65 536 cores. This is 2 times the node and GPU count of CoDa I, and an 8-fold increase in CPU count. This is the first time a GPU-accelerated, fully coupled radiation-hydrodynamics galaxy formation code has been used on such a scale.

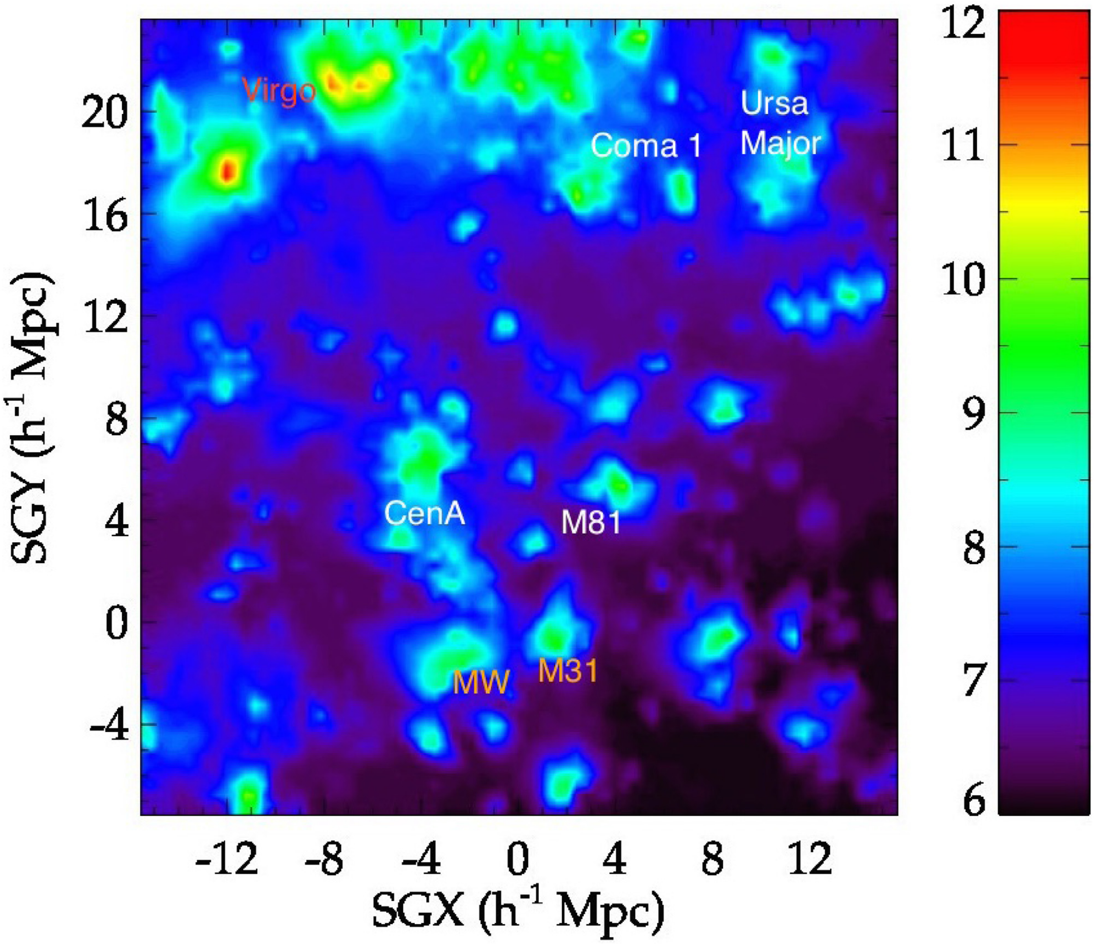

Figure 11. Reionization map of the Local Universe: Islands in the stream. The redshift at which each of the 1024^3 grid cells first reached an ionization of 50 per cent is displayed in a planar-slice 1 h^{-1} cMpc thick through the CoDa II simulation volume, labelled with the names of progenitors of familiar objects in the Local Universe today at their respective locations in the map.

The simulation accurately describes the properties of the gas and its interaction with ionizing radiation, self-consistently, including the growth of typical butterflyshaped ionized regions around the first stars and first galaxies, accompanied by photoheating and the subsequent dynamical back-reaction of the gas in response to pressure-gradients. This back-reaction includes the progressive smoothing of small-scale gas structures and the resistance of ionized intergalactic gas to its gravitational capture by galactic haloes too small to produce potential wells deep enough to overcome these pressure forces.

The simulation is in broad agreement with several observational constraints on the EoR, such as the Thomson scattering optical depth integrated through the IGM, as inferred from measurements of fluctuations in the cosmic microwave background by the *Planck* satellite, and the redshift by which reionization ended, as estimated from measurements of the Lyman α opacity of neutral H atoms in the absorption spectra of high-redshift quasars, due to the intervening IGM. While the timing of the evolution of the ionizing flux density is reasonable in CoDa II, its rapid rise at the end of reionization overshoots the observed value inferred from observations of quasar absorption spectra at $z = 6$ by a factor ~10. This in turn causes the neutral fraction in the post-reionization IGM to be too small compared to observations of the Lyman α forest at this redshift, by a similar factor. This may be a consequence of the existence of additional small-scale structure in the real universe, responsible for additional H I bound-free opacity that limits the mean-free-path of ionizing photons when reionization ends, that was missed by the simulation because it was unresolved.

We compute the UV continuum luminosity functions for our simulated galaxy population, and find them to be in very good agreement with high-redshift galaxy observations, in particular at $z = 6$. At higher redshifts, a small offset appears, increasing to ~1 mag at $z = 10$. We apply magnitude cuts to mimic observational detection limits, and find that the cosmic star formation rate associated with haloes with $M_{\rm AB1600} < -17$ only captures 63 per cent (18 per cent) of the total SFR in haloes at $z = 6$ ($z = 10$). Integrating down to fainter magnitudes, $M_{\rm AB1600} < -13$, as attempted in deep cluster fields, captures a much larger fraction of the total SFR in haloes, 94 per cent (76 per cent) at $z = 6$ ($z = 10$). This suggests that calibrating a numerical simulation of galaxy formation by comparing its total star formation rate density to observations without applying the magnitude cut may lead to non-optimal calibration, and in particular too small a simulated star formation rate.

The average star formation rate of individual galaxies increases with their dark matter halo mass, yielding a typical SFR $= M_{\rm halo}^{\alpha}$

relation, when the SFR is globally averaged, with a slope $\alpha = 5/3$. At low-halo masses (below $\sim 2 \times 10^9$ M$_\odot$), the slope becomes steeper, as galaxies become less efficient at forming stars. As time goes by and reionization progresses, the average slope for these low-mass haloes becomes even steeper, reflecting the suppression of the SFR in low-mass haloes caused by reionization feedback.

We show unambiguously that this behaviour at low-halo mass is the result of the spreading, rising ionizing UV background, by comparing our results with a simulation without radiative transfer. We note that this radiative reduction of the SFR appears to be somewhat less severe than in CoDa I due to different sub-grid physics, but it is still robustly measured.

In contrast, the gas core of high-mass haloes is dense enough to remain cool and/or cool down fast enough to keep forming stars, even if in bursts.

Furthermore, we show that environment has a strong impact on the star formation histories of low-mass galaxies: galaxies in overdense regions stop forming stars earlier than galaxies in underdense regions, supporting the analysis of CoDa I by Dawoodbhoy et al. (2018).

The CoDa II UV luminosity function peaks at $M_{\mathrm{AB1600}} = -11$ at all the studied redshifts, and develops into a plateau at fainter magnitudes, populated by radiation-suppressed, fading galaxies. We provide the full CoDa II UV luminosity functions for 4 epochs in tabular form, across 8 decades in luminosity, from $M_{\mathrm{AB1600}} = -3$ to $M_{\mathrm{AB1600}} = -24$, in the hope that these predictions will be useful for comparison with other simulations, but also to observers analysing high-redshift galaxy observations or designing new deep surveys. We analyse CoDa II's Local Universe analogue region, and find that the Milky Way and its neighbour M31 appear as individual islands in the reionization redshift map. This means that both galaxies reionized in isolation, i.e. they were not reionized by the progenitor of the Virgo cluster, or by nearby groups, or by each other.

ACKNOWLEDGEMENTS

This study was performed in the context of several French ANR (Agence Nationale de la Recherche) projects. PO acknowledges support from the French ANR funded project ORAGE (ANR-14-CE33-0016). ND and DA acknowledge funding from the French ANR for project ANR-12-JS05-0001 (EMMA). The CoDa II simulation was performed at Oak Ridge National Laboratory/Oak Ridge Leadership Computing Facility on the Titan supercomputer (INCITE 2016 award AST031). Processing was performed on the Eos and Rhea clusters. Resolution study simulations were performed on Piz Daint at the Swiss National Supercomputing Center (PRACE Tier 0 award, project id pr37). The authors would like to acknowledge the High Performance Computing center of the University of Strasbourg for supporting this work by providing scientific support and access to computing resources. Part of the computing resources were funded by the Equipex EquipMeso project (Programme Investissements d'Avenir) and the CPER Alsacalcul/Big Data.

The CoDaII-DM2048 / ESMDPL_2048 simulation has been performed at Leibniz Rechenzentrum Munich within the project pr74no.

ITI was supported by the Science and Technology Facilities Council [grant number ST/L000652/1]. JS acknowledges support from the 'l'Oréal-UNESCO Pour les femmes et la Science' and the 'Centre National d'Études Spatiales (CNES)' postdoctoral fellowship programs. YH has been partially supported by the Israel Science Foundation (1013/12). KA was supported by National Research Foundation grants NRF-2016R1D1A1B04935414 and NRF-2016R1A5A1013277

GY also acknowledges support from Ministerio de Economia Industria y Competitividad and Fondo Europeo de Desarrollo Regional (MINECO–FEDER) under research grants AYA2012-31101, AYA2015-63810-P, and AYA2015-63810-P, as well as Ministerio de Ciencia, Innovacion y Universidades (MICIU/FEDER) under grant PGC2018-094975-C21.

PRS was supported in part by U.S. National Science Foundation (NSF) grant AST-1009799, National Aeronautics and Space Administration (NASA) grant NNX11AE09G, NASA/Jet Propulsion Laboratory grant RSA Nos. 1492788 and 1515294, and supercomputer resources from NSF XSEDE grant TG-AST090005 and the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. PO thanks Y. Dubois, F. Roy, and Y. Rasera for their precious help dealing with SN feedback in RAMSES and various hacks in pFoF. This work made use of v2.1 of the Binary Population and Spectral Synthesis (BPASS) models as last described in Eldridge et al. (2017).

REFERENCES

Atek H., Richard J., Kneib J.-P., Schaerer D., 2018, MNRAS, 479, 5184
Aubert D., Teyssier R., 2008, MNRAS, 387, 295
Aubert D., Teyssier R., 2010, ApJ, 724, 244
Aubert D., Deparis N., Ocvirk P., 2015, MNRAS, 454, 1012
Aubert D. et al., 2018, ApJ, 856, L22
Baek S., Di Matteo P., Semelin B., Combes F., Revaz Y., 2009, A&A, 495, 389
Barkana R., Loeb A., 2000, ApJ, 539, 20
Bauer A., Springel V., Vogelsberger M., Genel S., Torrey P., Sijacki D., Nelson D., Hernquist L., 2015, MNRAS, 453, 3593
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, ApJ, 843, 129
Bowler R. A. A. et al., 2015, MNRAS, 452, 1817
Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J., Mahesh N., 2018, Nature, 555, 67
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
Calverley A. P., Becker G. D., Haehnelt M. G., Bolton J. S., 2011, MNRAS, 412, 2543
Carlesi E. et al., 2016, MNRAS, 458, 900
Chardin J., Haehnelt M. G., Aubert D., Puchwein E., 2015, MNRAS, 453, 2943
Courtois H. M., Pomarède D., Tully R. B., Hoffman Y., Courtois D., 2013, AJ, 146, 69
D'Aloisio A., Upton Sanderbeck P. R., McQuinn M., Trac H., Shapiro P. R., 2017, MNRAS, 468, 4691
Davies F. B. et al., 2018, ApJ, 864, 142
Dawoodbhoy T. et al., 2018, MNRAS, 480, 1740
Dayal P., Ferrara A., 2018, Phys. Rep., 780, 1
Deparis N., Aubert D., Ocvirk P., Chardin J., Lewis J., 2019, A&A, 622, A142
Doumler T., Hoffman Y., Courtois H., Gottlöber S., 2013, MNRAS, 430, 888
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, MNRAS, 420, 2662
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, PASA, 34, e058
Emberson J. D., Thomas R. M., Alvarez M. A., 2013, ApJ, 763, 146
Fan X. et al., 2006, AJ, 132, 117
Finkelstein S. L., Ryan R. E., Jr, Papovich C., Dickinson M., Song M., Somerville R. S., Ferguson H. C., 2015, ApJ, 810, 71
Giallongo E. et al., 2015, A&A, 578, A83
Giroux M. L., Shapiro P. R., 1996, ApJS, 102, 191
Gnedin N. Y., 2000, ApJ, 542, 535
Gnedin N. Y., Kaurov A. A., 2014, ApJ, 793, 30
Gong Y., Silva M., Cooray A., Santos M. G., 2014, ApJ, 785, 72
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
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
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Hasegawa K., Semelin B., 2013, MNRAS, 428, 154
Hinshaw G., Weiland J. L., Hill R. S., Odegard N., 2009, ApJS, 180, 225
Hoag A. et al., 2019, ApJ, 878, 12
Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Hoffman Y., Ribak E., 1992, ApJ, 384, 448
Iliev I. T., Mellema G., Pen U.-L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625
Iliev I. T., Mellema G., Ahn K., Shapiro P. R., Mao Y., Pen U.-L., 2014, MNRAS, 439, 725
Jeon M., Pawlik A. H., Bromm V., Milosavljević M., 2014, MNRAS, 444, 3288
Katz H. et al., 2020, MNRAS, 494, 2200
Kroupa P., 2001, MNRAS, 322, 231
Kulkarni G., Worseck G., Hennawi J. F., 2019, MNRAS, 488, 1035
Madau P., Haardt F., 2015, ApJ, 813, L8
Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
Maulbetsch C., Avila-Reese V., Colín P., Gottlöber S., Khalatyan A., Steinmetz M., 2007, ApJ, 654, 53
Miralda-Escudé J., Haehnelt M., Rees M. J., 2000, ApJ, 530, 1
Mitra S., Choudhury T. R., Ferrara A., 2018, MNRAS, 473, 1416
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Ocvirk P., Aubert D., Chardin J., Knebe A., Libeskind N., Gottlöber S., Yepes G., Hoffman Y., 2013, ApJ, 777, 51
Ocvirk P. et al., 2016, MNRAS, 463, 1462 (O16)
Ocvirk P., Aubert D., Chardin J., Deparis N., Lewis J., 2019, A&A, 626, A77
Oesch P. A. et al., 2013, ApJ, 773, 75
Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63
Ouchi M. et al., 2010, ApJ, 723, 869
Park H., Shapiro P. R., Choi J.-h., Yoshida N., Hirano S., Ahn K., 2016, ApJ, 831, 86
Parsa S., Dunlop J. S., McLure R. J., 2018, MNRAS, 474, 2904
Pawlik A. H., Milosavljević M., Bromm V., 2013, ApJ, 767, 59
Pawlik A. H., Schaye J., Dalla Vecchia C., 2015, MNRAS, 451, 1586
Pawlik A. H., Rahmati A., Schaye J., Jeon M., Dalla Vecchia C., 2017, MNRAS, 466, 960
Petkova M., Springel V., 2011, MNRAS, 412, 935
Planck Collaboration XVI, 2014, A&A, 571, A16
Planck Collaboration VI, 2018, preprint (arXiv:1807.06209)
Qin Y. et al., 2017, MNRAS, 472, 2009
Rahmati A., Schaye J., 2018, MNRAS, 478, 5123
Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188
Rosdahl J., Schaye J., Teyssier R., Agertz O., 2015, MNRAS, 451, 34
Rosdahl J. et al., 2018, MNRAS, 479, 994
Roy F., Bouillot V. R., Rasera Y., 2014, A&A, 564, A13
Seager S., Sasselov D. D., Scott D., 1999, ApJ, 523, L1
Shapiro P. R., Giroux M. L., 1987, ApJ, 321, L107
Shapiro P. R., Giroux M. L., Babul A., 1994, ApJ, 427, 25
Shukla H., Mellema G., Iliev I. T., Shapiro P. R., 2016, MNRAS, 458, 135
So G. C., Norman M. L., Reynolds D. R., Wise J. H., 2014, ApJ, 789, 149
Sorce J. G., 2015, MNRAS, 450, 2644
Sorce J. G., Tempel E., 2017, MNRAS, 469, 2859
Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014, MNRAS, 437, 3586
Sorce J. G. et al., 2016a, MNRAS, 455, 2078
Sorce J. G., Gottlöber S., Hoffman Y., Yepes G., 2016b, MNRAS, 460, 2015
Sorce J. G., Hoffman Y., Gottlöber S., 2017, MNRAS, 468, 1812
Springel V., 2005, MNRAS, 364, 1105
Sullivan D., Iliev I. T., Dixon K. L., 2018, MNRAS, 473, 38
Tescari E., Viel M., Tornatore L., Borgani S., 2009, MNRAS, 397, 411
Teyssier R., 2002, A&A, 385, 337
Tully R. B., 2015a, AJ, 149, 54
Tully R. B., 2015b, AJ, 149, 171
Tully R. B. et al., 2013, AJ, 146, 86
Wise J. H., Demchenko V. G., Halicek M. T., Norman M. L., Turk M. J., Abel T., Smith B. D., 2014, MNRAS, 442, 2560
Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, ApJ, 825, 144
Wu X., Kannan R., Marinacci F., Vogelsberger M., Hernquist L., 2019, MNRAS, 488, 419
Yang X., Mo H. J., van den Bosch F. C., Bonaca A., Li S., Lu Y., Lu Y., Lu Z., 2013, ApJ, 770, 115
Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, ApJ, 449, 446
Zaroubi S., Hoffman Y., Dekel A., 1999, ApJ, 520, 413
Zawada K., Semelin B., Vonlanthen P., Baek S., Revaz Y., 2014, MNRAS, 439, 1615

APPENDIX A: RESOLUTION STUDY ON SUPPRESSION OF STAR FORMATION

We conducted a resolution study to determine the impact of spatial resolution on the suppression of star formation in low-mass haloes due to radiative feedback. To this end we performed a series of three simulations of a test box 4 h^{-1}cMpc on a side, with the full CoDa II physics, using exactly the same parameters as given in Table 1, except for grid cell size and dark matter particle mass, which therefore impact spatial resolution and mass resolution. We explored spatial (mass) resolutions 2 (8) times and 4 (64) times higher than the fiducial CoDa II setup. These simulations are described in Table A1. The average star formation histories of $z = 5$ haloes were then computed for three different bins of dark matter halo mass and are shown in the left-hand panel of Fig. A1. The global shapes of the star formation histories for each mass bin are similar to those of Fig. 9 in the full physics case ('SN, RT'). The highest mass bin of Fig. 9 is missing in this resolution study because it was performed in a 4 h^{-1}cMpc instead of 8 h^{-1}cMpc box, to keep the computational cost of the study in check. This is not a problem, however, since it is in low-mass haloes that suppression is strongest and matters the most. That is the focus of our resolution study.

The left-hand panel of Fig. A1 shows that for the two most massive bins, i.e. between 10^9 and 10^{11} $M_\odot$, where the suppression is almost non-existent, increasing spatial resolution by a factor of 2 yields an increase of star formation rate of a factor ~2. A similar increase is also seen at early times in the 10^{8-9} $M_\odot$ halo mass bin, where star formation rises earlier in the high-resolution runs. However, in all 3 runs, the low-mass haloes experience a star formation plateau at a similar level of $\sim 10^{-4}$ $M_\odot$ yr^{-1}. Reionization ends for all three simulations at redshifts in the range $z = 6$–7 (shown by the vertical solid, dashed, and dotted lines), and the post-reionization epochs are marked, as in Fig. A1, by a decline of SFR in the low-mass haloes. The decrease is seen at all resolutions, and the timing and slope of the suppression are similar across all three runs.

To account for the slight differences in reionization redshifts between the three runs, we shifted the runs *in time* to align their reionization histories, so that they have the same reionization redshift $z = 6$. The resulting average halo star formation histories and J_{21} are shown in Fig. A2.

With these small time shifts applied to each, the curves of J_{21} lie virtually on top of each other, especially after ~550 Myr, demonstrating that the 3 runs have very similar levels of irradiation at each stage of their respective reionization histories.

Table A1. Summary table of the simulations of our resolution study, giving the number of particles, the dark matter particle mass, and the spatial resolution in comoving and physical kpc at $z = 5$. All simulations are performed with full CoDa II physics.

Simulation	$N_{\rm DM}$	$M_{\rm DM}$ ($\mathrm{M}_\odot$)	dx (ckpc)	$\mathrm{dx}_{\rm phys}$ (kpc, $z = 5$)	Comment
L04N0256	256^3	4.07×10^5	23.06	3.84	Fiducial
L04N0512	512^3	6.21×10^4	11.53	1.92	$\mathrm{dx}_{\rm fiducial}/2$
L04N1024	1024^3	7.77×10^3	5.76	0.96	$\mathrm{dx}_{\rm fiducial}/4$

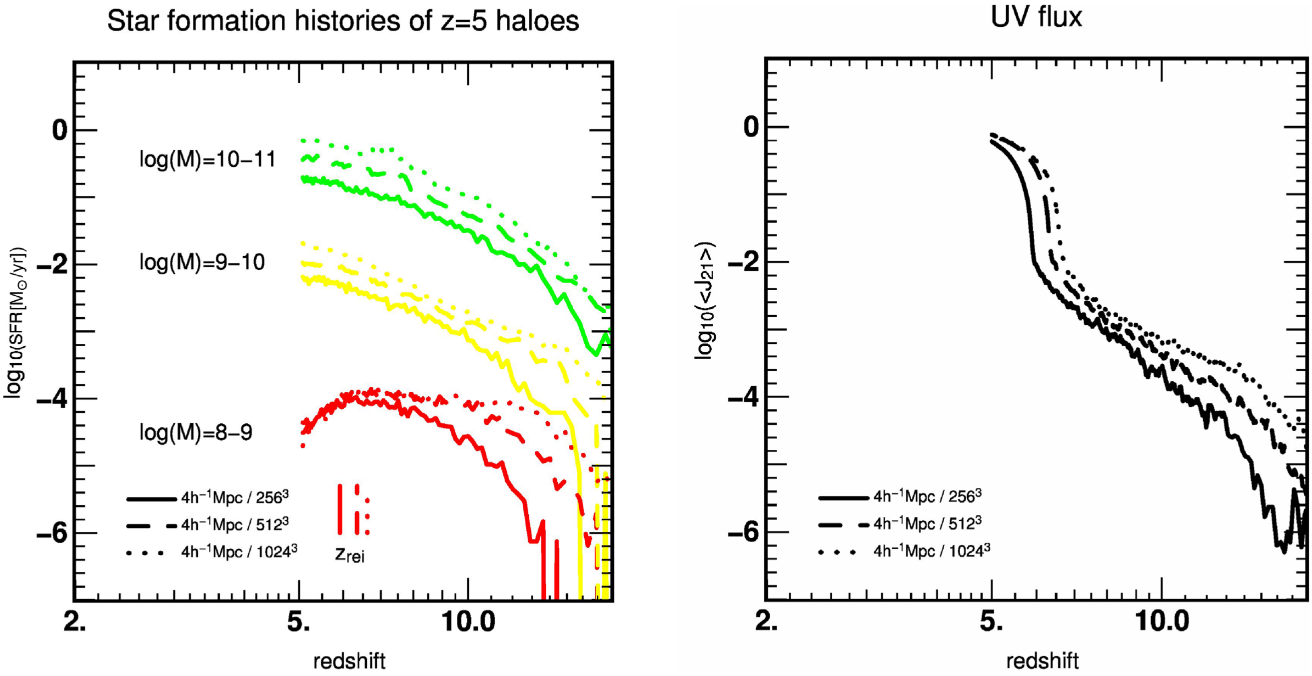

Figure A1. Results of the resolution study simulations with full CoDaII physics listed in Table A1. The test box is 4 h^{-1}cMpc on a side, at different spatial resolutions. *Solid line*: fiducial CoDa II resolution (256^3), *dashed line*: 512^3, *dotted line*: 1024^3. *Left:* Average star formation histories for 3 halo mass bins. The vertical lines above the '$z_{\rm rei}$' label indicate the reionization redshift of each simulation. The mass bins correspond to the haloes final mass (i.e. halo mass measured in the final simulation time-step, at $z = 5$. Right: Volume-averaged ionizing UV flux J_{21} for these simulations.

We can now take a closer look at the SFRs of the low-mass bin, as shown in the zoomed left-hand panel of Fig. A2. The slightly earlier onset and rise of SFR at higher resolutions is confirmed, as is the quasi-plateau at an SFR = 10^{-4} $\mathrm{M}_\odot$ yr^{-1}, from 600 Myr to $z = 6$, shared by all 3 resolution runs. During this plateau, the SFRs of all 3 runs are very consistent, within 0.1 dex of each other. The same sharp decline of SFR takes place at $z = 6$ for all 3 runs.

We already explored the impact of numerical resolution on star formation suppression in CoDa I in Ocvirk et al. (2016), and found little impact when increasing resolution by a factor 2. The new results of this appendix confirm that, despite the revised subgrid model for star formation for CoDaII, which results in a milder suppression than CoDa I, suppression is still seen in 10^{8-9} $\mathrm{M}_\odot$ haloes, even at a resolution 4 times higher.

Some other efforts have also been made to address the question of radiative suppression of star formation in low-mass haloes based upon cosmological hydrodynamical simulations with RT. Although their methodology and metrics for measuring suppression may differ from ours, there is reasonable agreement with at least some of them. For instance, Wu et al. (2019) is particularly valuable because of the use of a completely different code, AREPO-RT. They found that photoheating reduces the SFR by 50 per cent in haloes below $10^{8.4}$ $\mathrm{M}_\odot$ at $z = 6$ and $10^{8.8}$ $\mathrm{M}_\odot$ at $z = 5$. At both redshifts, their suppression mass is within the 10^{8-9} $\mathrm{M}_\odot$ mass bin where suppression can be seen to take place in Fig. A1.

A similar agreement is found with Pawlik et al. (2015), also using a completely different methodology (smoothed particles hydrodynamics and a form of adaptive ray-tracing). They found a break at the low-mass end of the SFR-halo mass relation, which evolves with redshift. Their simulations reionize earlier than CoDa II, at $z \sim 8$. At this redshift, they find a break in the low-mass end of the SFR-halo mass relation, at 5.10^9 $\mathrm{M}_\odot$, characteristic of suppression of star formation by photoheating and supernova feedback. Altough this mass increases with decreasing redshift, the overlap between

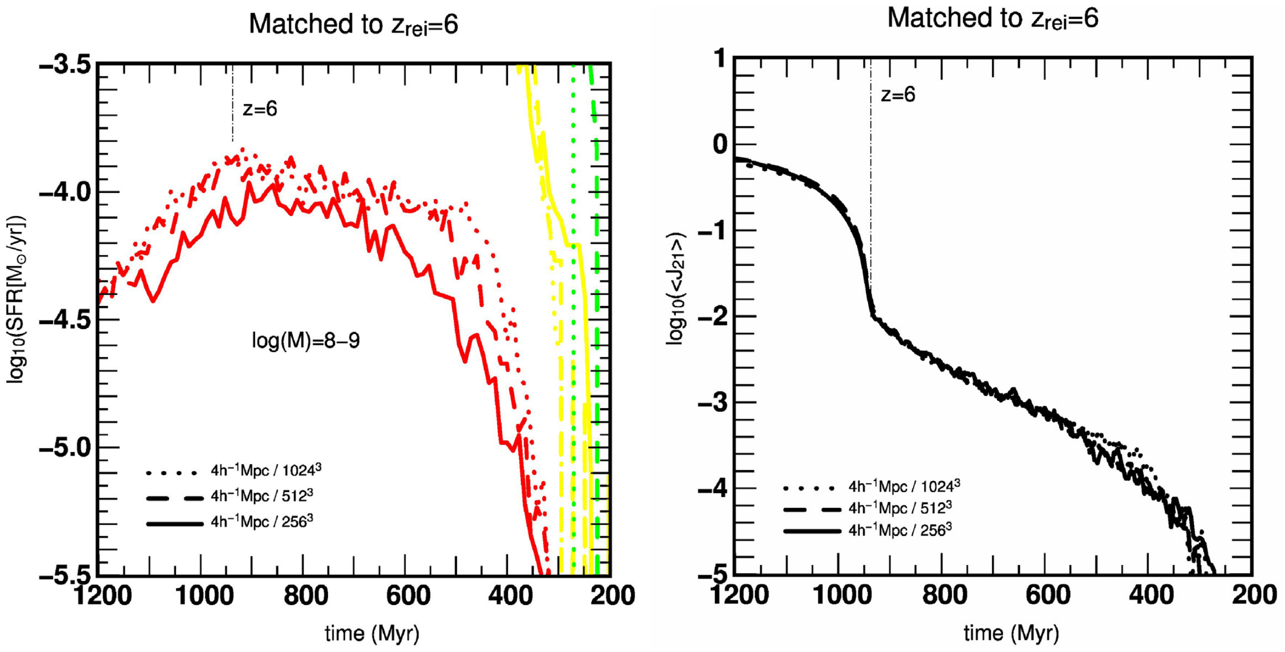

Figure A2. Same as Fig. A1, but the simulations have been shifted in time to align them, so that all 3 runs have $z_{\rm rei} = 6$. This results in a shift of −17, 66, and 112 Myr for the 256^3, 512^3, and 1024^3 run, respectively. The *x*-axis now shows time since big bang, with $z = 6$ marked as a vertical dot–dashed line for reference. The left-hand panel is zoomed on the low-mass bin.

the halo mass range where suppression takes place in CoDa II and Pawlik et al. (2015) is encouraging.

The rough agreement between these examples from the literature and our own results, along with the resolution study presented here, suggests that the suppression in star formation seen in CoDa II is not an artefact due to numerical resolution.

APPENDIX B: LUMINOSITY FUNCTIONS

We give the CoDa II luminosity functions at 4 epochs in Table B1, in the hope that they may be of use to the community and in the design, analysis, and interpretation of future surveys.

Table B1. CoDa II UV luminosity functions at 4 different epochs. Each group of 4 columns corresponds to a different redshift. Within each group, the 4 columns give the base-10 logarithm of (1): the median LF, (2): the average LF, (3): the minimum LF, (4): the maximum LF. The LF statistics have been determined using 5 independent sub-boxes each spanning 1/5 of the total CoDa II volume.

M_{AB1600}	$\log_{10}(\Phi[{\rm Mag}^{-1}{\rm Mpc}^{-3}]$ median (1), average (2), min (3), max (4))															
	$z = 6$				$z = 7$				$z = 8$				$z = 10$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
−24.0																
−23.5		−5.58		−4.88												
−23.0		−5.58		−4.88		−5.58		−4.88								
−22.5	−4.88	−4.80		−4.58		−5.28		−4.88								
−22.0	−4.41	−4.41		−4.11	−4.88	−4.98		−4.58		−5.28		−4.88				
−21.5	−4.18	−4.17	−4.41	−3.93	−4.88	−4.74		−4.41		−5.28		−4.58				
−21.0	−3.74	−3.80	−4.18	−3.65	−4.28	−4.26	−4.58	−4.04	−4.88	−4.74		−4.41				
−20.5	−3.44	−3.47	−3.74	−3.28	−3.98	−4.04	−4.88	−3.80	−4.88	−4.50	−4.88	−4.18				
−20.0	−3.18	−3.21	−3.50	−3.06	−3.60	−3.62	−3.88	−3.50	−4.28	−4.17	−4.41	−3.98		−5.28		−4.58
−19.5	−2.95	−3.00	−3.36	−2.87	−3.38	−3.38	−3.65	−3.25	−3.98	−3.89	−4.11	−3.71	−4.88	−4.88		−4.58
−19.0	−2.67	−2.70	−2.86	−2.55	−3.04	−3.07	−3.38	−2.90	−3.49	−3.50	−3.93	−3.35	−4.41	−4.44	−4.58	−4.41
−18.5	−2.45	−2.50	−2.73	−2.39	−2.74	−2.77	−3.00	−2.67	−3.16	−3.17	−3.39	−3.06	−4.28	−4.24	−4.58	−4.11
−18.0	−2.32	−2.33	−2.51	−2.22	−2.59	−2.61	−2.73	−2.49	−2.85	−2.89	−3.08	−2.81	−3.88	−3.90	−4.11	−3.74
−17.5	−2.13	−2.15	−2.31	−2.07	−2.38	−2.41	−2.60	−2.30	−2.74	−2.73	−2.89	−2.60	−3.71	−3.60	−3.77	−3.44
−17.0	−2.01	−1.98	−2.16	−1.87	−2.23	−2.22	−2.37	−2.11	−2.51	−2.51	−2.64	−2.43	−3.27	−3.25	−3.47	−3.15
−16.5	−1.84	−1.83	−1.98	−1.73	−2.05	−2.05	−2.22	−1.95	−2.31	−2.32	−2.48	−2.20	−3.00	−3.02	−3.30	−2.87
−16.0	−1.68	−1.68	−1.84	−1.59	−1.88	−1.88	−2.03	−1.78	−2.16	−2.13	−2.31	−2.01	−2.81	−2.81	−2.93	−2.72
−15.5	−1.55	−1.53	−1.69	−1.44	−1.74	−1.73	−1.89	−1.62	−1.95	−1.95	−2.11	−1.84	−2.54	−2.57	−2.70	−2.45
−15.0	−1.42	−1.41	−1.55	−1.32	−1.57	−1.57	−1.71	−1.48	−1.81	−1.79	−1.93	−1.69	−2.41	−2.37	−2.56	−2.24

Table B1 – *continued*

M_{AB1600}	$\log_{10}(\Phi[\mathrm{Mag}^{-1}\mathrm{Mpc}^{-3}]$ median (1), average (2), min (3), max (4))															
	$z = 6$				$z = 7$				$z = 8$				$z = 10$			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
−14.5	−1.30	−1.30	−1.41	−1.20	−1.45	−1.43	−1.57	−1.34	−1.63	−1.62	−1.79	−1.53	−2.20	−2.16	−2.32	−2.05
−14.0	−1.18	−1.18	−1.30	−1.10	−1.32	−1.31	−1.43	−1.22	−1.50	−1.47	−1.59	−1.38	−2.01	−1.96	−2.11	−1.86
−13.5	−1.08	−1.08	−1.19	−0.99	−1.19	−1.17	−1.29	−1.08	−1.35	−1.34	−1.45	−1.24	−1.82	−1.80	−1.97	−1.68
−13.0	−0.98	−0.98	−1.08	−0.89	−1.07	−1.06	−1.17	−0.97	−1.21	−1.19	−1.31	−1.10	−1.65	−1.62	−1.75	−1.51
−12.5	−0.90	−0.89	−0.98	−0.81	−0.96	−0.94	−1.05	−0.84	−1.09	−1.06	−1.18	−0.96	−1.50	−1.46	−1.58	−1.34
−12.0	−0.84	−0.82	−0.88	−0.74	−0.87	−0.84	−0.93	−0.75	−1.00	−0.95	−1.06	−0.84	−1.39	−1.33	−1.43	−1.20
−11.5	−0.81	−0.78	−0.83	−0.70	−0.81	−0.78	−0.86	−0.67	−0.92	−0.87	−0.96	−0.75	−1.31	−1.24	−1.33	−1.11
−11.0	−0.79	−0.76	−0.80	−0.69	−0.77	−0.72	−0.79	−0.61	−0.88	−0.81	−0.89	−0.68	−1.24	−1.16	−1.26	−1.02
−10.5	−0.88	−0.85	−0.89	−0.77	−0.87	−0.81	−0.88	−0.70	−0.97	−0.90	−0.97	−0.77	−1.32	−1.25	−1.34	−1.09
−10.0	−0.95	−0.93	−0.97	−0.85	−0.95	−0.90	−0.97	−0.78	−1.07	−1.01	−1.10	−0.87	−1.48	−1.41	−1.53	−1.24
−9.5	−0.93	−0.92	−0.97	−0.84	−0.96	−0.91	−1.00	−0.79	−1.10	−1.03	−1.13	−0.88	−1.51	−1.45	−1.58	−1.28
−9.0	−0.97	−0.96	−1.02	−0.86	−1.01	−0.97	−1.07	−0.83	−1.15	−1.10	−1.21	−0.94	−1.63	−1.56	−1.70	−1.38
−8.5	−0.95	−0.94	−1.01	−0.83	−1.00	−0.97	−1.08	−0.83	−1.17	−1.13	−1.26	−0.97	−1.70	−1.63	−1.78	−1.44
−8.0	−0.95	−0.93	−1.01	−0.82	−1.05	−1.02	−1.14	−0.86	−1.22	−1.18	−1.32	−1.02	−1.79	−1.74	−1.92	−1.54
−7.5	−0.97	−0.96	−1.08	−0.83	−1.14	−1.11	−1.25	−0.94	−1.35	−1.32	−1.48	−1.14	−2.01	−1.97	−2.17	−1.76
−7.0	−1.04	−1.03	−1.16	−0.88	−1.26	−1.24	−1.39	−1.07	−1.55	−1.52	−1.71	−1.32	−2.32	−2.28	−2.55	−2.03
−6.5	−1.14	−1.12	−1.28	−0.95	−1.45	−1.42	−1.60	−1.23	−1.80	−1.76	−1.96	−1.55	−2.90	−2.81	−3.05	−2.51
−6.0	−1.16	−1.12	−1.26	−0.95	−1.56	−1.51	−1.66	−1.32	−2.09	−2.03	−2.21	−1.81	−3.84	−3.74	−4.18	−3.47
−5.5	−1.56	−1.49	−1.65	−1.28	−2.50	−2.43	−2.65	−2.18	−3.60	−3.60	−3.84	−3.39				
−5.0	−3.23	−3.08	−3.45	−2.76												
−4.5																
−4.0																
−3.5																
−3.0																

APPENDIX C: STARS BEYOND R_{200}

Following our finding in Section 3.3 that a fraction of the stars are found beyond R_{200}, we provide a few maps to illustrate this aspect and gain insight into the relative distributions of stars and dark matter haloes in CoDa II.

The left column of Fig. C1 shows a gas density map overplotted with the R_{200} spheres of dark matter haloes more massive than $M > 10^8$ $\mathrm{M}_\odot$ (black circles), along with stars, within and beyond R_{200}, for 3 typical star-forming regions of the simulation. The slice is 30 cells thick, i.e. about 470 h^{-1}ckpc. The fraction of stars beyond R_{200} increases from top to bottom, and is given in the title of each panel. In most cases, star particles are contained within the R_{200} of their parent haloes (blue circles), and the latter are well aligned along gas filaments, forming strings of haloes. Stars beyond R_{200} are mostly located between pairs of neighbour haloes within these strings. This is confirmed in the bottom plot, which contains a higher density region and more massive haloes, and where the fraction of stars beyond R_{200} reaches 21 per cent, comparable to what we found globally in Section 3.3 at $z = 6$. Here, we show that the fraction of stars beyond R_{200} is primarily the result of approximating FoF halo shapes by a sphere. The right column of Fig. C1 shows the particles of the FoF dark matter haloes of each region. They are clearly not spherical, always triaxial, often elongated in the direction of the filament, and often feature bridge-like structures linking them. Stars beyond R_{200} fall along these aspherical features. At higher redshifts, halo shapes become even more aspherical, accentuating this effect and increasing the number of stars found beyond R_{200}. If one considers more complex halo shapes (using convex hulls for instance) one would expect that the census of stars within haloes would be close to complete. Still a small fraction of stars might be formed in small haloes below the assumed minimum halo mass.

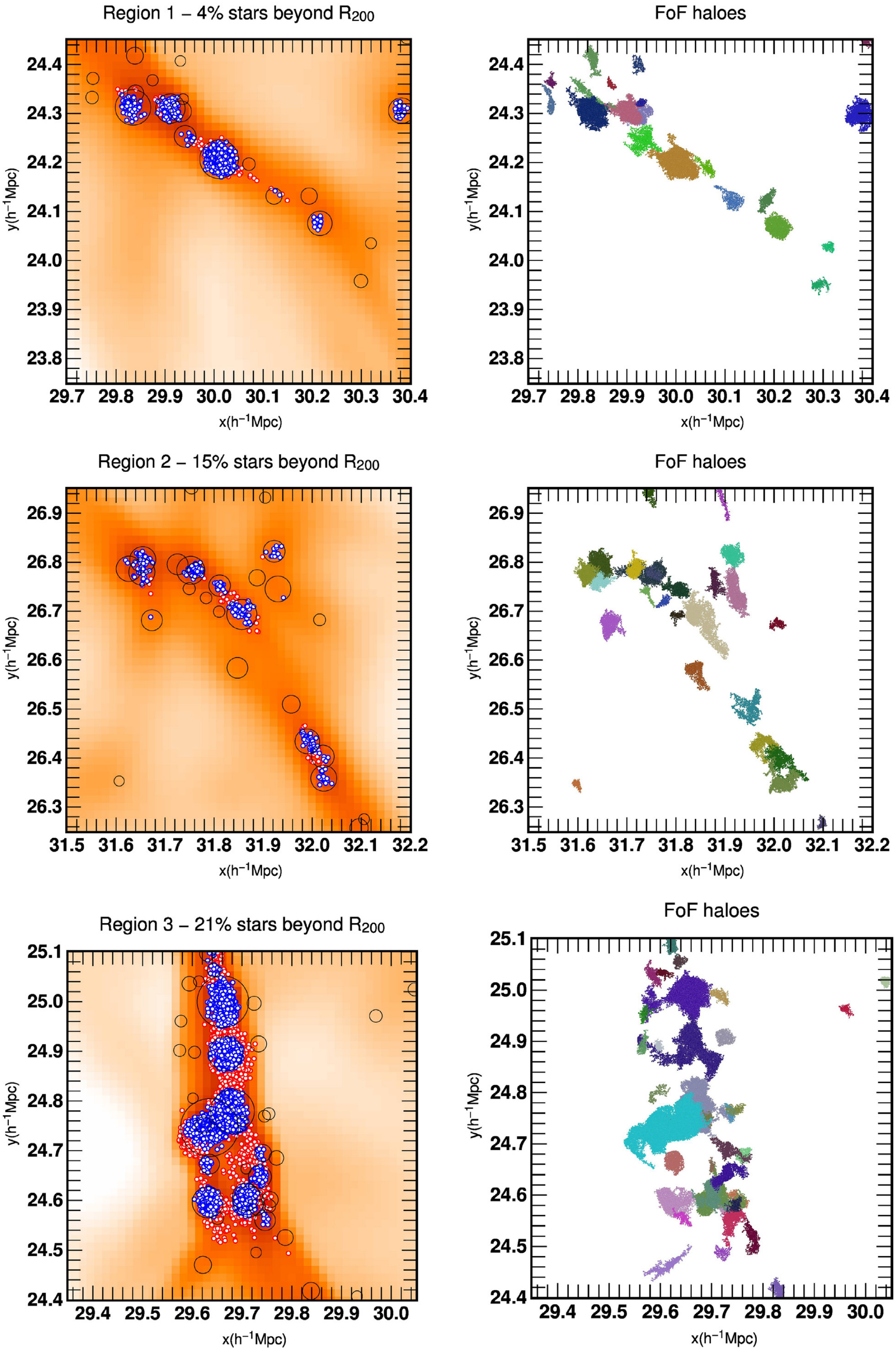

Figure C1. *Left:* Maps of the gas density for three different example regions in CoDa II, overplotted with the R_{200} spheres of dark matter haloes more massive than $M > 10^8$ M$_\odot$ (black circles). Also shown are the stellar particles, in blue if the particle is inside of the R_{200} sphere of a dark matter halo and red otherwise. The percentage of stars beyond R_{200} is indicated in the title of each panel. The most massive halo in the top, middle, and bottom panel is 5.5×10^9, 2.6×10^9, and 1.8×10^{10} M$_\odot$, respectively. *Right:* Maps of the dark matter particles of the FoF haloes of the corresponding left-hand panel. A colour is chosen randomly for each halo.