The impact of the first galaxies on cosmic dawn and reionization

Munoz, Julian B.; Qin, Yuxiang; Mesinger, Andrei; Murray, Steven G.; Greig, Bradley; Mason, Charlotte

Published in:
Monthly Notices of the Royal Astronomical Society

DOI:
10.1093/mnras/stac185

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Munoz, J. B., Qin, Y., Mesinger, A., Murray, S. G., Greig, B., & Mason, C. (2022). The impact of the first galaxies on cosmic dawn and reionization. Monthly Notices of the Royal Astronomical Society, 511(3), 3657-3681. https://doi.org/10.1093/mnras/stac185
The impact of the first galaxies on cosmic dawn and reionization

Julian B. Muñoz,1,2,3 Yuxiang Qin,2,3 Andrei Mesinger,4 Steven G. Murray,5 Bradley Greig2,3 and Charlotte Mason1,6,7

1Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
2School of Physics, University of Melbourne, Parkville, VIC 3010, Australia
3ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Parkville, VIC 3010, Australia
4Scuola Normale Superiore, Pisa, PI-56126, Italy
5School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85281, USA
6Cosmic Dawn Center (DAWN)
7Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200 København N, Denmark

Accepted 2022 January 14. Received 2022 January 13; in original form 2021 November 16

ABSTRACT
The formation of the first galaxies during cosmic dawn and reionization (at redshifts z = 5–30), triggered the last major phase transition of our universe, as hydrogen evolved from cold and neutral to hot and ionized. The 21-cm line of neutral hydrogen will soon allow us to map these cosmic milestones and study the galaxies that drove them. To aid in interpreting these observations, we upgrade the publicly available code 21cmFAST. We introduce a new, flexible parametrization of the radiative feedback from: an inhomogeneous, H2-dissociating (Lyman–Werner; LW) background; and dark matter – baryon relative velocities; which recovers results from recent, small-scale hydrodynamical simulations with both effects. We perform a large, ‘best-guess’ simulation as the 2021 installment of the Evolution of 21-cm Structure (EOS) project. This improves the previous release with a galaxy model that reproduces the observed UV luminosity functions (UVLFs), and by including a population of molecular-cooling galaxies. The resulting 21-cm global signal and power spectrum are significantly weaker, primarily due to a more rapid evolution of the star formation rate density required to match the UVLFs. Nevertheless, we forecast high signal-to-noise detections for both HERA and the SKA. We demonstrate how the stellar-to-halo mass relation of the unseen, first galaxies can be inferred from the 21-cm evolution. Finally, we show that the spatial modulation of X-ray heating due to relative velocities provides a unique acoustic signature that is detectable at z ≈ 10–15 in our fiducial model. Ours are the first public simulations with joint inhomogeneous LW and relative-velocity feedback across the entire cosmic dawn and reionization, and we make them available at this link https://scholar.harvard.edu/julianbmunoz/eos-21.

Key words: galaxies: high-redshift – intergalactic medium – cosmology: theory – dark ages, reionization, first stars – diffuse radiation.

1 INTRODUCTION
The epoch of reionization (EoR) and the cosmic dawn (CD) represent fundamental milestones in the history of our universe, and are rapidly becoming the next frontier in astrophysics. These two eras witnessed the last major phase change of our universe, as the intergalactic medium (IGM) evolved from being cold and neutral following recombination, to being hot and ionized due to the radiation emitted by the first galaxies. At present we only hold some pieces of this cosmic puzzle (e.g. Loeb & Furlanetto 2013; Mesinger 2016). Nevertheless, the next few years will see the advent of different observations targeting these cosmic eras (Mellera et al. 2013; Beardsley et al. 2016; DeBoer et al. 2017; Bowman et al. 2018; Greig et al. 2020b; Mertens et al. 2020), which will open a window to the astrophysics of the early universe.

While the stellar content of low-z galaxies is relatively well understood (see e.g. Wechsler & Tinker 2018 for a review), much less is known about the first galaxies that started the CD. Given the hierarchical paradigm of structure formation, we expect the first (Pop III) stars to form in small molecular-cooling galaxies (MCGs) at z ∼ 20–30 (Tegmark et al. 1997; Abel, Bryan & Norman 2002; Bromm & Larson 2004; Haiman & Bryan 2006; Trenti 2010), residing in haloes with virial temperatures Tvir ≲ 104 K (which corresponds to total halo masses Mh ≲ 108 M⊙, during the EoR and CD). Feedback eventually quenches star formation in these MCGs, and hierarchical evolution ushers in the era of heavier, atomically cooled galaxies (ACGs; Mh ⩾ 108 M⊙). With most ACGs forming out of pre-enriched MCG building blocks, it is likely that second-generation (Pop II) stars drove the bulk of cosmic reionization at z ∼ 5–10 (e.g. Greig, Mesinger & Bañados 2019; Mason et al. 2019a; Aghanim et al. 2020; Choudhury, Paranjape & Bosman 2021; Qin et al. 2021).

Current data, from ultraviolet (UV) luminosity functions (UVLFs; Bouwens et al. 2014, 2015, 2021; Atek et al. 2015, 2018; Livermore, Finkelstein & Lotz 2017; Ishigaki et al. 2018; Oesch et al. 2018), the Lyman α forest (Becker et al. 2015; Bosman et al. 2018; Qin et al. 2021b), and the evolution of the volume-averaged hydrogen neutral fraction (ξHI; Mason et al. 2019b; Aghanim et al. 2020), provide some constraints (Park et al. 2019; Naidu et al. 2020) on
the halo–galaxy connection for ACGs at $z \lesssim 10$. However, very little is known about MCGs or the $z > 10$ universe. The brightest galaxies in this regime will be reached by the James Webb Space Telescope (JWST) and the Nancy Grace Roman Space Telescope (Roman). Unfortunately, the bulk of ACGs and MCGs are too faint to be seen with these telescopes, and must be studied indirectly with 21-cm observatories. These include both global-signal efforts like the Experiment to Detect the Global EoR Signature (EDGES), the Shaped Antenna measurement of the background Radio Spectrum (SARAS), the Large-aperture Experiment to Detect the Dark Ages (LEDA), as well as interferometers like the Low-Frequency Array (LOFAR), the Murchison Widefield Array (MWA), the Hydrogen Epoch of Reionization Array (HERA) and the Square Kilometre Array (SKA). Therefore, it is of paramount importance to develop flexible and robust models of the astrophysics of the CD and the EoR to compare against these upcoming data. In this work we build upon the public 21cmFAST code to include MCGs with all of the relevant sources of feedback.

Stellar formation in MCGs is easily disrupted by different processes, given their shallow potential wells. In addition to feedback from photoheating and supernovae (Draine & Bertoldi 1996; Barkana & Loeb 1999; Wise & Abel 2008; Sobacchi & Mesinger 2013, which also affects stellar formation in ACGs), MCGs suffer from two distinct sources of feedback. The first is driven by Lyman–Werner (LW) radiation, composed of photons in the 11.2–13.6 eV band, which efficiently photodissociate molecular hydrogen ($H_2$), hampering gas cooling, and subsequent star formation in MCGs (Machacek, Bryan & Abel 2001; Johnson, Greif & Bromm 2007; O’Shea & Norman 2008; Safranek-Shrader et al. 2012; Vishal et al. 2014; Schauer et al. 2017; Skinner & Wise 2020). The second are driven by the dark matter (DM)-baryon streaming velocities ($v_{	ext{DB}}$), which impede gas from efficiently accreting and cooling on to DM haloes, thus slowing the formation of the first stars (Tielkakhovich & Hirata 2010; Greif et al. 2011; O’Leary & McQuinn 2012; Naoz, Yoshida & Gnedin 2013; Hirano et al. 2018; Schauer et al. 2019). Until recently it was not clear how these two feedback channels interacted. A clearer picture has emerged from the small-scale hydrodynamical simulations of Pop III star formation in Schauer et al. (2021) and Kulkarni, Vishal & Bryan (2021). In this picture, the two sources of feedback are additive, enhancing the minimum mass that an MCG ought to have to be able to form stars. We synthesize the results from these simulations into a flexible fitting formula that jointly includes both LW and $v_{	ext{DB}}$ feedback. Along with a calibration for this formula, we assume MCGs host Pop III stars with a simple stellar-to-halo mass relation (SHMR), distinct from that of ACGs hosting Pop II stars.

We include these results into the seminumerical simulation code 21cmFAST\(^1\) (Mesinger & Furlanetto 2007; Mesinger, Furlanetto & Cen 2011; Murray et al. 2020). This builds upon the implementation of the streaming velocities in 21cmFAST\(^2\) by Muñoz (2019a,b, though in that work it was assumed MCGs and ACGs shared the same SHMR and that LW feedback was isotropic), and the first implementation of MCGs with inhomogeneous LW feedback in 21cmFAST by Qin et al. (2020a, 2021a, though there relative velocities were not considered and the old LW prescription from Machacek et al. 2001 was used). This code is now able to self-consistently evolve the anisotropic LW background, the streaming velocities, as well as the X-ray, ionizing, and non-ionizing UV backgrounds, necessary to predict the 21-cm signal.

We run a large simulation suite (1.5 Gpc comoving on a side, with 1000\(^3\) cells), as the 2021 installment of the Evolution Of 21-cm Structure (EoS) project, updating those in Mesinger, Greig & Sobacchi (2016). This represents our state of knowledge about cosmic dawn and reionization. We dub this EoS simulation AllGalaxies, and we make its most relevant light-cones public.\(^2\) This simulation goes beyond the previous FaintGalaxies model, both by including Pop III-hosting MCGs as well as including a SHMR that fits current UVLFs. As a consequence, we predict a slower evolution of the 21-cm signal, and 21-cm fluctuations that are nearly an order of magnitude smaller (see also Mirocha, Furlanetto & Sun 2017; Park et al. 2019). Our AllGalaxies (EoS2021) simulation represents the current state-of-knowledge of the evolution of cosmic radiation fields during the first billion years.

Furthermore, we explore how parameters governing star formation and feedback in MCGs impact the 21-cm signal and other observables. We find that MCGs likely dominate the star formation rate density (SFRD) at $z \gtrsim 12$, implying they are important for the timing of the CD but not the EoR (see e.g. Qin et al. 2020a; Wu et al. 2021). We also search for the velocity-induced acoustic oscillations (VAOs) that appear due to the acoustic nature of the streaming velocities (Dalal, Pen & Seljak 2010; McQuinn & O’Leary 2012; Vishal et al. 2012; Fialkov et al. 2013; Muñoz 2019a), and for the first time identify VAOs in a full 21-cm light-cone, as opposed to a co-evol (i.e. fixed-$z$) box. We predict significant VAOs for redshifts as low as $z = 10–15$, which act as a standard ruler to the cosmic dawn (Muñoz 2019b).

This paper is structured as follows. In Section 2, we introduce our model for the first galaxies, both atomic and molecular cooling, which we use in Section 3 to predict how the epochs of cosmic dawn and reionization unfold. Sections 4 and 5 explore how the first galaxies shape the 21-cm signal, where in the former we vary the MCG parameters in our simulations, and in the latter we study the VAOs. Finally, we conclude in Section 6. In this work we fix the cosmological parameters to the best fit from Planck 2018 data (TT, TE, EE + lensing + BAO inaghan et al. 2020), and all distances are comoving unless specified otherwise.

2 MODELLING THE FIRST GALAXIES

We begin by describing our model for the first galaxies. As our seminumerical 21-cmFAST\(^1\) simulations do not keep track of the metallicity and accretion history of each galaxy, we use population-averaged quantities to relate stellar properties to host halo masses. Although these relations can allow for mixed stellar populations, we will make the simplifying assumption (consistent with simulations and merger-tree models; e.g. Xu et al. 2016a; Mebane, Mirocha & Furlanetto 2018) that on average Pop III stars form in (first generation, unpolluted) MCGs, whereas Pop II stars form in (second generation, polluted) ACGs. ACGs are hosted by haloes with virial temperatures of $T_{\text{vir}} \gtrsim 10^4 K$ (Oll & Haiman 2002, corresponding to halo masses $M_h \gtrsim M_{\text{dom}}(z) \sim 10^8 M_\odot$ at the relevant redshifts), whereas MCGs reside in minihaloes with $10^1 K \lesssim T_{\text{vir}} \lesssim 10^4 K$ (Tegmark & Zaldarriaga 2009, and thus $M_{\text{mol}} < M_h < M_{\text{dom}}$, with $M_{\text{mol}}(z) \sim 10^6 M_\odot$ during cosmic dawn). As larger haloes form in the universe, star formation transitions from being dominated by Pop III stars to being dominated by Pop II stars (e.g. Schneider et al. 2002). We now describe how these galaxy populations are modelled and the feedback processes that affect them.

\(^1\)Publicly available at https://github.com/21cmfast/21cmFAST.

\(^2\)EOS light-cones can be downloaded at this link.
2.1 Pop II star formation in atomic-cooling galaxies

For the better-understood Pop II stars forming in ACGs we follow the model from Park et al. (2019), which we now briefly review.

The key ingredient that enters our calculations is the (spatially varying) star formation rate density (SFRD), which for ACGs we calculate as

\[
\text{SFRD}^{(\text{II})} = \int dM_h \frac{dn}{dM_h} M^{(\text{II})}_{\text{SFR}} f_{\text{duty}}, \tag{1}
\]

where \(dn/dM_h\) is the conditional halo mass function (HMF; e.g. Barkana & Loeb 2004), \(f_{\text{duty}}^{(\text{II})}\) is a duty cycle (i.e. halo occupation fraction) described below, and \(M^{(\text{II})}_{\text{SFR}}\) is the star formation rate (SFR). The superscripts \(^{(\text{II})}\) denote that a quantity refers to Pop II-dominated ACGs.

For both ACGs and MCGs, we assume that the SFR is proportional (on average) to the stellar mass divided by a characteristic time-scale:

\[
M^{(\text{II})}_{*} = \frac{M^{(\text{II})}_{*}}{t_\text{H}^{1/(z)}} \tag{2}
\]

where \(t_\text{H}\) is the Hubble expansion rate, and \(t_\text{s}\) is a free parameter, which for simplicity we fix to \(t_\text{s} = 0.5\) (as it will degenerate with the normalization of the stellar fraction).

The SHMR is modelled as a simple power-law:

\[
M^{(\text{II})}_{*} = \frac{\Omega_{\text{baryon}}}{\Omega_{\text{dm}}} \left( \frac{\rho_\text{DM}}{10^{10}\rho_c} \right)^{a_{\text{II}}} M_h, \tag{3}
\]

where \(\Omega_{\text{baryon}}\) are the baryon and matter densities. This SHMR is governed by two parameters, the normalization \(f^{(\text{II})}_{s,10}\) (defined at a characteristic scale \(M_h = 10^{10}\rho_c\)), and a power-law index \(a^{(\text{II})}\), which controls its mass dependence.\(^3\) The value of this index could be set by supernovae (SNe) feedback, which is thought to regulate star formation inside these small galaxies (e.g. Wyithe & Loeb 2013; Dayal et al. 2014; Sung et al. 2019).

This model, without any additional redshift dependence, provides an excellent fit to current EO R observations at \(5 \lesssim z \lesssim 10\), including the Lyman-\(\alpha\) forest opacity fluctuations (Qin et al. 2021b) and the faint-end UV LFs (Park et al. 2019; Rudakovskyi et al. 2021) (or the entire luminosity range if a high-mass turnover is added to characterize AGN-induced feedback in rare, bright galaxies; e.g. Sabit, Muñoz & Blas 2021c).

The final part of equation (1) is the duty fraction,

\[
f_{\text{duty}}^{(\text{II})} = \exp \left(-M^{(\text{II})}_{\text{turn}}/M_h\right) \tag{4}
\]

which accounts for inefficient star formation below a characteristic scale \(M^{(\text{II})}_{\text{turn}}\). Haloes below \(M^{(\text{II})}_{\text{turn}}\) are exponentially less likely to host an ACG, due to inefficient atomic cooling and/or feedback, as discussed below.

2.1.1 Feedback on ACGs

As in previous work (e.g. Qin et al. 2020a), we consider two sources of feedback in ACGs: photoheating and supernovae.

Radiation during cosmic reionization heats the gas around low-mass galaxies (both atomic and molecular cooling), delaying its collapse and subsequent star formation (e.g. Thoul & Weinberg 1996; Noh & McQuinn 2014). We use the 1D collapse-simulation results from Sobacchi & Mesinger (2014) to calculate the (local) critical halo mass below which star formation becomes inefficient due to photoheating (see also Hui & Gnedin 1997; Okamoto, Gao & Theuns 2008; Katz et al. 2020; Ocvirk et al. 2020):

\[
\frac{M^{\text{ion}}_{\text{crit}}}{2.8 \times 10^8 M_\odot} = \left( \frac{\Gamma_{\text{ion}}}{10^{-12} s^{-1}} \right)^{0.17} \left( \frac{10}{1+z} \right)^{2.1} \left[ 1 - \left( \frac{1+z}{1+z_{\text{ion}}} \right)^2 \right]^{2.5}, \tag{5}
\]

where \(\Gamma_{\text{ion}}\) is the local ionizing background, and \(z_{\text{ion}}\) is the reionization redshift of the simulation cell.

Therefore, for ACGs the turn-over mass is

\[
M^{(\text{II})}_{\text{turn}} = \max \left( M^{\text{ion}}_{\text{crit}} / M_{\text{atom}}, \right) \tag{6}
\]

where \(M_{\text{atom}} = 3.3 \times 10^7 M_\odot [(1+z)/21]^{-3/2}\) during the epoch of interest for our fiducial cosmology.

We additionally include feedback from supernovae (SNe) or radiative processes by assuming a smaller fraction of gas is available to form stars inside smaller-mass haloes. Such ‘feedback-limited’ models result in a (positive) power-law scaling of the SHMR, c.f. the power-law index \(a^{(\text{III})}\) introduced in equation (3) (e.g. Wyithe & Loeb 2013). We fix this parameter to \(a^{(\text{III})} = 0.5\) in this work, as it provides an excellent fit to UV LFs at \(z = 6–10\) (Qin et al. 2021b). Different authors infer somewhat different values of \(a^{(\text{III})}\), as they can depend on the assumed star formation histories (e.g. Tacchella et al. 2018; Behroozi et al. 2019). Although \(a^{(\text{III})}\) could also evolve with time, \(z = 2–10\) LFs are consistent with a constant value (e.g. Mason et al. 2015; Tacchella et al. 2018; Mirocha et al. 2021). Our implementation also allows for the possibility that SNe feedback results in an additional characteristic mass scale [effectively adding \(M_{\text{SN}}\) to the RHS of equation (6); c.f. Qin et al. 2020a]. However, we do not consider that possibility in our fiducial model (setting \(M_{\text{SN}} < \max\{M^{\text{ion}}_{\text{crit}} / M_{\text{atom}}\}\); this is supported by \(z = 0\) data that seems to show no cutoff for haloes down to \(M_h \approx 10^8 M_\odot\) (Nadler et al. 2020) roughly at the \(M_{\text{atom}}\) threshold (which however are likely not actively star forming at \(z = 0\)).

2.2 Pop III star formation in molecular-cooling galaxies

The star formation and feedback mechanisms in the first MCG galaxies are still not fully understood. For simplicity, we will assume their SHMR has the same, generic power law form as ACGs, with a few modifications (see Qin et al. 2020a for further details of the implementation in 21-cmFAST). We take the Pop III SFRD\(^4\) to be

\[
\text{SFRD}^{(\text{III})} = \int dM_h \frac{dn}{dM_h} M^{(\text{III})}_{\text{SFR}} f_{\text{duty}}^{(\text{III})} \tag{7}
\]

The changes with respect to equation (1) are in the SFR and the duty cycle. For the former, we assume the same relation between \(M_*\) and \(M_h\) as equation (3), though we allow MCGs to have a unique SHMR:

\[
M^{(\text{III})}_{*} = \frac{f^{(\text{III})}_{s,7} \Omega_{\text{baryon}}}{\Omega_{\text{dm}}} \left( \frac{M_h}{10^{10} M_\odot} \right)^{a^{(\text{III})}} M_h, \tag{8}
\]

which has two new free parameters, \(f^{(\text{III})}_{s,7}\), and \(a^{(\text{III})}\).

\(^4\)We introduce an approximation to analytically calculate the SFRD within 21-cm FAST. This dramatically speeds up the calculation of the SFRD tables, reducing that computational cost by nearly two orders of magnitude. This option can be turned on by setting \texttt{FAST\_FCOLL\_TABLES = True}. We encourage the reader to visit Appendix B for details.
The other difference is the MCG duty cycle
\[ f_{\text{duty}}^{\text{(III)}} = \exp(-M_{\text{turn}}^{\text{(III)}}/M_\odot) \exp(-M_\odot/M_{\text{mol}}), \]
which follows our previous assumption that there is a smooth transition from MCGs (below \( M_{\text{turn}} \)) forming Pop III stars to ACGs (above that mass) forming Pop II stars. Further, there is a lower mass cutoff \( M_{\text{turn}}^{\text{(III)}} \) for MCG stellar formation, set by different feedback processes as detailed below. Our comprehensive treatment of \( M_{\text{turn}}^{\text{(III)}} \) represents the main modelling improvement of this work.

### 2.1.1 Feedback on MCGs

MCGs reside in haloes with shallow gravitational potentials, which makes them highly sensitive to different effects that disrupt their gas distribution, abundance, or \( H_2 \) content. In addition to the previously mentioned feedback, MCGs are also sensitive to: (i) Lyman–Werner photons, and (ii) the DM-baryon relative velocities. As we did for ACGs, we set the turnover mass to be
\[ M_{\text{turn}}^{\text{(III)}} = \max \left( M_{\text{mol}}^{\text{(III)}}, M_{\text{mol}} \right), \]
where the first term accounts for photoionization feedback, and the second accounts for the additional impact of LW and \( v_{\text{ch}} \).

Both the LW and \( v_{\text{ch}} \) feedback effects impede stellar formation, and the simulations from Schauer et al. (2021) and Kulkarni et al. (2021) show that their joint impact is cumulative. We model the molecular-cooling turnover mass as a product of three factors:
\[ M_{\text{mol}} = M_{\odot}(z) f_{\text{ch}}(v_{\text{ch}}) f_{\text{LW}}(J_{21}), \]
where \( J_{21} \) is the LW intensity in units of \( 10^{21} \text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{sr}^{-1} \), and \( M_{\odot}(z) = M_0 (1 + z)^{3/2} \) is the (no-feedback) molecular-cooling threshold, with \( M_0 = 3.3 \times 10^5 M_\odot \) (corresponding to \( T_{\text{vir}} = 10^5 \text{K} \); Tegmark et al. 1997). The two factors \( f_{\text{LW}} \) and \( f_{\text{ch}} \) encode the two new sources of feedback, and our chosen functional form assumes that they add coherently (as assumed previously in Fialkov et al. 2013 and Muñoz 2019a) and are redshift independent. This was shown to be the case by Schauer et al. (2021) for values of \( J_{21} \geq 0.1 \), as well as in Kulkarni et al. (2021) for \( J_{21} = 1 \). For larger values, \( J_{21} \geq 10 \), however, the simulations of Kulkarni et al. (2021) show a deviation from this form (towards less suppression). This large-\( J_{21} \) regime only occurs at \( z \leq 6 \) in our simulations, at which point ACGs completely dominate the global SFRD, making MCGs (and thus their feedback mechanisms) largely irrelevant when computing radiation backgrounds.

### 2.2.2 Lyman–Werner feedback

Photons in the LW band (11.2–13.6 eV) can photodissociate \( H_2 \) molecules, thus impeding the ability of gas to cool on to MCGs (Tegmark et al. 1997; Abel et al. 2002; Bromm & Larson 2004; Haiman & Bryan 2006; Trenti 2010). Machacek et al. (2001) showed that the minimum halo mass for star formation in MCGs has a power-law dependence on \( J_{21} \). This result has been the standard in studies of CD and the EoR (see e.g. Fialkov et al. 2013; Visbal et al. 2014; Mirocha et al. 2018; Muñoz 2019a; Qin et al. 2021a, 2020a). However, recent simulations have shown that self-shielding can be important, which reduces the suppression produced by a given LW background. We include these improved results in our analysis.

In particular, we use the results from Kulkarni et al. (2021) and Schauer et al. (2021, see also Skinner & Wise 2020 for similar conclusions), which have independently tackled the issue of jointly simulating the effects of LW (including self-shielding) and relative-velocity feedback (which we will explore later) on the first stars. From equation (11) we can separate the LW feedback factor
\[ f_{\text{LW}} = \frac{M_{\odot}(J_{21})}{M_{\text{mol}}(0)}, \]
and calculate it from simulation results, which factors out differences in the zero-point mass (i.e. on the mass \( M_0 \) in the absence of feedback). We show the simulation results for \( f_{\text{LW}} \) in Fig. 1, where it is clear that the Kulkarni et al. (2021) results are below the Schauer et al. (2021) ones, even for larger values of \( J_{21} \). However, a direct comparison of their results is difficult due to differences in the assumed LW intensities: Kulkarni et al. (2021) considered \( J_{21} \geq 1 \) (or zero), whereas Schauer et al. (2021) considered \( J_{21} \geq 0.1 \).

Rather than use a fit to \( f_{\text{LW}} \) from either group, we use a flexible parametrization inspired by Visbal et al. (2014):
\[ f_{\text{LW}} = 1 + A_{\text{LW}}(J_{21})^{\beta_{\text{LW}}}, \]
with \( A_{\text{LW}} \) and \( \beta_{\text{LW}} \) as free parameters that can be varied within 21 cmFAST. The Kulkarni et al. (2021) simulation results can be well fit setting \( \{A_{\text{LW}}, \beta_{\text{LW}}\} = \{0.8, 0.9\} \), whereas the Schauer et al. (2021) ones require \( \{A_{\text{LW}}, \beta_{\text{LW}}\} = \{3.0, 0.5\} \), which indicates stronger feedback, though a weaker dependence with \( J_{21} \). Both of these fits are shown in Fig. 1. Instead of choosing between these two, we use the flexibility of equation (13) to propose a joint fit, which lands in between both simulation results, with \( A_{\text{LW}} = 2 \) and \( \beta_{\text{LW}} = 0.6 \) (where the round numbers are purposefully chosen to avoid conveying more agreement than the simulations provide). These will be our fiducial parameters for this work. For reference, the old work of Machacek et al. (2001), which did not account for self-shielding, had \( \{A_{\text{LW}}, \beta_{\text{LW}}\} = \{22, 0.47\} \), producing a correction that was nearly a factor of 10 larger (also shown in Fig. 1). We emphasize that we assume that the LW feedback factor only depends on \( J_{21} \), and not on \( z \) or any halo property, as a simplifying assumption, which could however be revisited if required by further simulations.

Figure 1. Feedback factor \( f_{\text{LW}} \) describing how the turnover (minimum) mass for a molecular-cooling halo to form stars grows due to LW feedback, as a function of the LW intensity \( J_{21} \) in units of \( 10^{-21} \text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{sr}^{-1} \). The data points are from simulation data in Kulkarni et al. (2021, green) and Schauer et al. (2021, red). Our fit, which sits in the middle in solid black, follows equation (13) with \( A_{\text{LW}} = 2 \) and \( \beta_{\text{LW}} = 0.5 \). The parameters of equation (13) can be modified to closely follow each simulation, as the red and green lines are obtained with different parameter combinations. The dashed brown line is the fit from Machacek et al. (2001), which did not include self-shielding.
2.2.3 Relative-velocity feedback

The DM-baryon relative velocities impede the formation of stars in the first galaxies in two main ways. The first one, pointed out in Tseliakhovich & Hirata (2010, see also Naoz, Yoshida & Gnedin 2012), is that regions of large velocity show suppressed matter fluctuations at small scales, as the baryons there contribute less to the growth of structure. As a consequence, the abundance of small-mass haloes is modulated by relative velocities: fewer collapsed haloes reside in regions with large streaming motions. This effect is difficult to include in our seminumerical simulations, as it would require altering the power spectrum at each cell. Instead, we include this effect on average by modifying the power spectrum in all cells. We solve for the evolution of the baryon and DM overdensities following Muñoz (2019a, based on Tseliakhovich & Hirata 2010), and find that the impact of $v_{\text{cb}}$ on the power spectrum at a wavenumber $k$ is well captured by

$$\frac{P_{\text{m}}(k, z; v_{\text{cb}})}{P_{\text{m}}(k, z; 0)} = 1 - A_p \exp \left[ - \frac{(\log(k/k_p))^2}{2\sigma_p^2} \right]$$

(14)

where the three free parameters, $A_p$, $k_p$, and $\sigma_p$ depend on $v_{\text{cb}}$, and mildly on $z$. Here, we fix to the values at $z = 20$ and at the root mean square (rms) relative velocity ($v_{\text{cb}} = v_{\text{rms}}$), which are $A_p = 0.24$, $k_p = 300$ Mpc$^{-1}$, and $\sigma_p = 0.9$. This is conservative, in that it will yield no VAOs (as these quantities do not spatially fluctuate) and will suppress MCGs by roughly the average amount. We encourage the interested reader to visit Appendix A for details of how this fit is obtained, as well as to find the fit as a function of $z$ and $v_{\text{cb}}$. This matter-fluctuation suppression also affects ACGs, and therefore should be included even when MCGs are not. However, we find this to be a subdominant effect to the one we will discuss next, and thus our average treatment will suffice. We leave for future work implementing the fluctuations on this suppression.

The second – and largest – effect of the velocities is to suppress star formation in MCGs (Dalal et al. 2010; Tseliakhovich, Barkana & Hirata 2011). Small haloes in regions with large relative velocities have difficulties in accreting gas, as well as cooling gas into stars (e.g. Greif et al. 2011; Naoz et al. 2013; Schauer et al. 2019).

The small-scale (~Mpc) hydro simulations of Schauer et al. (2021) and Kulkarni et al. (2021) were the first to investigate the impact of streaming velocities together with LW feedback. Both groups account for the relative velocities in a similar fashion, and solve for the evolution of gas inside MCGs to find if the conditions for star formation are met. Both works find that the streaming motions impede star formation in the smallest galaxies, raising the minimum mass $M_{\text{mol}}$ required for star formation (in the absence of photoheating feedback). Analogously to $f_{\text{bw}}$, we define a feedback factor also for $v_{\text{cb}}$:

$$f_{\text{cb}} = \frac{M_{\text{mol}}(v_{\text{cb}})}{M_{\text{mol}}(0)}.$$

(15)

We calculate this factor for the different simulation outputs of the two groups, and plot it in Fig. 2, which shows the excellent agreement between the two simulation suites.

---

5We note that the relative velocities are coherent on scales below 3 Mpc (Tseliakhovich & Hirata 2010), and thus can be taken to be constant within each of our cells.

---

Figure 2. Same as Fig. 1 but for the DM-baryon relative velocity $v_{\text{cb}}$ (here divided by its rms value $v_{\text{rms}}$ so it is dimensionless). Our fit, in solid black, follows equation (16), and the red and green lines show the fits from each of the two references. The brown dashed line shows the formula from Fialkov et al. (2012), which underpredicts the star formation suppression from relative velocities.

We fit $f_{\text{cb}}$ as a $z$-independent function, following the functional form in Kulkarni et al. (2021):

$$f_{\text{cb}} = \left(1 + \frac{v_{\text{cb}}}{v_{\text{rms}}} \beta_{\text{cb}} \right)^{-1},$$

(16)

where $v_{\text{rms}} \approx 30$ km s$^{-1}$ is the rms velocity, and we use $A_{v_{\text{cb}}} = 1$, and $\beta_{\text{cb}} = 1.8$, very similar to the values in Kulkarni et al. (2021). We compare this fit with the previous formula from Fialkov et al. (2012), where it was assumed that the halo virial velocity at the suppression scale follows:

$$V_{\text{mol}} = \left[V_0^2 + \alpha_v^2 v_{\text{cb}}^2(z)\right]^{1/2},$$

(17)

where $V_0 = 4$ km s$^{-1}$ is the virial velocity of molecular-cooling haloes in the absence of any feedback, and they find $\alpha_{\text{cb}} = 4$. This translates into a scaling $f_{\text{cb}} = (V_{\text{mol}}(v_{\text{cb}})/V_0)^3$, shown in Fig. 2, which underpredicts the impact of the relative velocities by roughly a factor of 2.

With a prescription for both sources of feedback, we can evaluate $M_{\text{mol}}$ in equation (11). We show this turnover mass at $z = 20$ in Fig. 3 for different values of $v_{\text{cb}}$ (in units of its rms value $v_{\text{rms}}$, which is close to its median) and $J_{21}$ (we mark the values expected at $z = 20$ for two of our fiducial parameter sets: EOS and OPE, see Table 1). We predict $M_{\text{mol}} \approx 10^6 M_\odot$ at $z = 20$, though clearly this quantity depends strongly on $v_{\text{cb}}$. This dependence will imprint $M_{\text{mol}}$ – and thus the SFRD – with the fluctuations of $v_{\text{cb}}$, giving rise to velocity-induced acoustic oscillations, which we study in Section 5.

We find that $M_{\text{mol}}$ depends less strongly on $J_{21}$. In particular, for $J_{21} < 10^{-5}$ the effect of LW feedback is rather weak, and $v_{\text{cb}}$ feedback dominates. The situation is reversed for $J_{21} > 10^{-5}$, though we caution that the two feedback schemes may not add coherently in such a high LW flux regime (Kulkarni et al. 2021). Luckily, this regime does not have a large impact on observables, as it produces $M_{\text{mol}} \approx M_{\text{coll}}$, so Pop III star formation in MCGs would be subdominant compared to Pop II in ACGs. Moreover, as we will see, values of $J_{21} > 10^{-5}$ only appear at very late times ($z \lesssim 10$) in our simulations, where Pop II star formation far dominates.
Figure 3. Value of the minimum-mass $M_{\text{min}}$ for a molecular-cooling halo to host Pop III stellar formation at $z = 20$, as a function of the LW flux ($J_{21}$ in the customary units of $10^{-21} \text{erg s}^{-1} \text{Hz}^{-1} \text{cm}^{-2} \text{sr}^{-1}$) and the streaming velocity ($v_{\text{th}}$ divided by its rms value $v_{\text{rms}}$), following equation (11). We assume negligible photoionization feedback at this redshift. Marked green points represent the expected fluxes $J_{21}$ for our two fiducial simulation sets presented in Table 1. The black solid line shows $M_{\text{atom}}$, above where there is no Pop III stellar formation (see Section 2.2).

Table 1. Summary of our choices for some of the main free parameters related to the SFRD (top part), ionizations (middle), and X-ray emission (bottom). The EOS 2021 (EOS) fiducial model is used throughout the text, except in the last two sections (4 and 5) where we use OPT. Other parameters are fixed to the values motivated in the text.

| Parameter | Fiducial (EOS2021) | Optimistic (OPT) |
|-----------|-------------------|------------------|
| $\log_{10} f_{\ast,10}^{(\text{II})}$ | -1.25 | -1.50 |
| $\log_{10} f_{\ast,7}^{(\text{III})}$ | -2.5 | -1.75 |
| $\alpha_{\ast}^{(\text{II})}$ | 0.5 | 0.5 |
| $\alpha_{\ast}^{(\text{III})}$ | 0 | 0 |
| $\log_{10} f_{\text{esc},10}^{(\text{III})}$ | -1.35 | -1.20 |
| $\log_{10} f_{\text{esc},7}^{(\text{III})}$ | -1.35 | -2.25 |
| $\alpha_{\text{esc}}^{(\text{III})}$ | 0.3 | 0.3 |
| $L_{X}^{(\text{III})}$ | 40.5 | 40.5 |
| $E_{0}$ [keV] | 0.5 | 0.2 |

2.3 Comparison: Pop II versus Pop III

Armed with our two stellar populations (Pop II residing in ACGs, and Pop III in MCGs), and the feedback processes that can affect them (LW and relative velocities for MCGs, and stellar and photoheating feedback for both), we can now compare their relative contribution to different cosmic epochs.

We choose a set of fiducial parameters, which we dub EOS (later in Sections 4 and 5 we will study parameter variations), where the stellar fractions of ACGs and MCGs are set to

$$\log_{10} f_{\ast,10}^{(\text{II})} = -1.25$$

$$\log_{10} f_{\ast,7}^{(\text{III})} = -2.5,$$

with SHMR power-law indices given by

$$\alpha_{\ast}^{(\text{II})} = 0.5$$

$$\alpha_{\ast}^{(\text{III})} = 0.$$

The ACG parameters are very similar to the maximum a posteriori (MAP) from Qin et al. (2021b), which reproduces observations of the UVLFs, the cosmic-microwave background (CMB) optical depth, and the high-$z$ Lyman $\alpha$ forest opacity fluctuations. Our choice of $\log_{10} f_{\ast,7}^{(\text{III})} = -2.5$ is rather conservative (though slightly larger than found in the simulations of Skinner & Wise 2020), but it still allows Pop III stars to dominate the SFRD at early enough times. Given that the Pop III parameters are entirely unknown, we will focus on varying the latter in this work, though we note that the 21 cmMC (Greig & Mesinger 2015) and 21 cmFish (Mason, Munoz & Others) packages allow the user to vary all parameters simultaneously.

We begin by showing the SHMR ($f_{\ast} = M_{\ast}/M_{0}$) of Pop II- and Pop III-hosting galaxies in Fig. 4 at $z = 25$ and 7. The impact of the four free parameters is as follows: each index $\alpha_{\ast}^{(i)}$ changes the slope of the corresponding SHMR, whereas the $f_{\ast}^{(i)}$ factors re-scale them up and down. Unlike at lower redshifts (see e.g. Behroozi, Wechsler & Conroy 2013; Mason et al. 2015; Taccella et al. 2018; Wechsler & Tinker 2018; Yung et al. 2019; Rudakovskiy et al. 2021; Sabti, Muñoz & Blas 2021a; Sabti et al. 2021c), for the EoR/CD we are only interested in very low mass haloes ($M_{h} \lesssim 10^{11} M_{\odot}$), as those dominate the photon budget compared to the brighter galaxies, which are rare at the redshifts of interest. Therefore, we can ignore the turnover in the SHMR at $M_{h} \gtrsim 10^{11} M_{\odot}$, commonly attributed to AGN feedback (Qin et al. 2017), and focus on the fainter end of the ACGs that is well characterized by a single power law. Despite their relatively low $f_{\ast}^{(i)}$, MCGs are abundant enough to dominate

\footnotetext[6]{https://github.com/21cmfast/21CMMC\footnotetext[7]{https://github.com/charlottenosam/21cmfish}}
Figure 5. We show different turnover halo mass scales, which enter the SFRDs of Pop II stars in equation (1) and Pop III stars in equation (7). Haloes above the purple line can efficiently cool gas through atomic-line transitions, and we assume they host ACGs forming Pop II stars. The dotted purple line is the prediction from $T_{*)} = 10^4$ K, whereas the solid line is the simulation result that includes photoheating (PH) feedback (which is efficient during reionization, to the left of the red line). The dashed brown line represents the theoretical limit above which haloes hosting MCGs can form Pop III stars (through molecular cooling) in the absence of feedback. To illustrate the strength of different sources of feedback, we include them one at a time. First, in green we add only LW, and in blue only the streaming velocities ($v_{cb}$). The relative velocities have a bigger impact at high $z$, whereas LW feedback dominates at later times. We show their product as the black-dotted line, which represents their total feedback from equation (11). The solid black line is the result from the simulation, which again rises sharply during reionization due to photoheating feedback.

the SFRD at early times ($z \gtrsim 15$), as we quantify below. In our model, MCGs only form stars for a narrow band of halo masses, which varies with $z$ depending on the dominant feedback process. As a consequence of this narrow mass range, the main parameter that controls Pop III star formation in MCGs is $f_{III}^{cb}$, rather than the power-law index $\alpha_{III}^{cb}$.

At later times ($z = 7$), feedback severely suppresses MCG star formation. Furthermore, the evolution of the HMF means that their relative abundance (compared to ACGs) decreases. As a result of these two effects, the contribution of MCGs to cosmic radiation fields becomes subdominant by $z \sim 7$ in our fiducial model.

In order to understand how feedback evolves over time, we show the different characteristic mass scales in Fig. 5. Going from lowest to highest, we first show the mass for molecular cooling of gas in the absence of feedback, $M_{\text{molecular}}$ (in equation (11)), which grows simply as $(1 + z)^{-3/2}$. We then include the impact of relative velocities (through $f_{cb}^{\text{rel}}$), and LW feedback (through $f_{LW}$, with the $\Delta z_{\text{LW}}$ flux self-consistently and inhomogeneously computed in our simulation box) individually, as well as jointly.

The impact of $v_{cb}$ is notable at high redshifts, increasing the turnover mass by nearly an order of magnitude, to $M_{\text{turn}}^{\text{III}} \approx 10^7 M_\odot$ at $z = 25$. At lower redshifts ($z \lesssim 12$ for our fiducial parameters), LW radiation dominates over $v_{cb}$ in setting the MCGs turnover mass. At even later times (below $z \sim 8$), photoheating feedback from reionization steeply increases $M_{\text{turn}}^{\text{III}}$, so by the end of our simulations at $z = 5$ there is essentially no star formation in MCGs. We remind the reader that in our model Pop II stars form above the atomic-cooling threshold, also shown in Fig. 5, and those are also affected by photoheating feedback.

Figure 6. Star formation rate density (SFRD) as a function of $z$ in an average-density region of the universe. The purple and black lines show our prediction for Pop II and Pop III stars hosted in ACGs and MCGs, respectively, and the grey thick line is their sum. The green-dashed line shows the result for Pop III stars if the Machacek et al. (2001) LW feedback prescription was utilized (given the same LW flux), whereas the blue dash–dotted line represents the Pop III result in the absence of relative velocities. The red data-points are the result of extrapolating the Bouwens et al. (2021) UVLFs to $M_{UV, \text{min}} = -13$, corresponding to $M_\odot \approx 10 M_{10}$ (or $\rho_\odot = 10^{10} M_\odot$) in our model. We also show, as black triangles, the result when only integrating down to $M_{UV, \text{min}} = -17$, where we have displaced each triangle by $\Delta z = +0.3$ for better visualization (as well as the lowest-$z$ circle for the same reason).

Given our fiducial choices, we can compute the SFRD for both Pop II and Pop III on a representative (i.e. average-density) patch of the universe (Madau et al. 1996). We do so in Fig. 6, where we show the individual Pop II and Pop III contributions, as well as their sum. As is clear from this figure, Pop III star formation in MCGs dominates over ACGs at higher redshifts, and for our fiducial parameters this transition takes place at $z \approx 15$. We also show two alternative scenarios of Pop III star formation to illustrate the impact of feedback. In one, we turn-off the relative-velocity ($v_{cb}$) effect, and as a result we would overestimate the SFRD of Pop III stars by 50 per cent. In the other, we use the previous feedback formula from Machacek et al. (2001), corresponding to $A_{10} = 22$ in equation (13), which did not include self-shielding. In that case the SFRD is reduced by roughly an order of magnitude, with the discrepancy increasing at lower $z$, where $\Delta z_{\text{LW}}$ is larger (though we note that we use the same $\Delta z_{\text{LW}}$ flux as in the fiducial simulation).

Additionally, we compare our predicted SFRDs with data, obtained by extrapolating the recent measurements from Bouwens et al. (2021). We convert our SFRD to UV luminosities using a constant $\kappa = 1.15 \times 10^{-28} M_\odot \times yr^{-1} erg^{-1}$ (Sun & Furlanetto 2016; Oesch et al. 2018), and extrapolate the Schechter fit from Bouwens et al. (2021, with errors inherited from the uncertainty in the Schechter parameters) to a minimum UV magnitude $M_{UV, \text{min}} = -13$ (Park et al. 2019), which corresponds to haloes with $M_\odot = 10^{10} M_\odot$ for the redshifts of interest. As expected, ACGs fit all the current ($z \lesssim 10$) data well on their own, as observations cannot reach the fainter MCGs (see for instance Sun et al. 2021 for a SPHEREx forecast).

2.4 Cosmic radiation fields

Using the (inhomogeneous) ACG and MCG SFRDs from equations (1) and (7), we compute the corresponding cosmic radiation fields that are relevant for the thermal and ionization evolution of the IGM: soft UV (LW and Lyman series), ionizing and X-ray. 21cmFAST calculates the radiation field incident on each simulation cell through
a combination of excursion-set photon-counting (for ionizing photons) and light-cone integration (for soft UV and X-rays), accounting also for IGM attenuation/absorption. These procedures are described in detail in, e.g. Mesinger et al. (2011), Mesinger, Ferrara & Spiegel (2013), Sobacchi & Mesinger (2014), and we encourage the interested reader to consult these works for further details. Here, we only summarize the free parameters that are the most relevant for our analysis.

We assume the emissivity for all of the above radiation fields scales with the SFR (equations 1 and 7). To calculate UV emission, we take Pop II/Pop III SEDs from Barkana & Loeb (2005), normalized to have $N_{\gamma_{\mathrm{UV}}} = 5000$ and 44 000 ionizing photons per stellar baryon for Pop II and III, respectively. We assume only a fraction $f_{\mathrm{esc}}$ of ionizing photons that are produced manage to escape the host galaxy and ionize the IGM. We take a power-law relation for the typical $f_{\mathrm{esc}}$ as a function of halo mass (Park et al. 2019):

$$f_{\mathrm{esc}}^{(i)} = f_{\mathrm{esc},10}^{(i)} \left( \frac{M_{\odot}}{M_i} \right)^{\alpha_{\mathrm{esc}}^{(i)}}$$

(18)

for both stellar populations ($i = \{ \text{II}, \text{III} \}$), with $M_i = \{ 10^{10}, 10^7 \}$ M$_{\odot}$ and $m_i = \log_{10}(M_i)$ as before. We choose fiducial parameters in broad accordance with the best fits from the latest observations in Qin et al. (2021b), setting

$\log_{10} f_{\mathrm{esc},10}^{(\text{II})} = 0.15$ with respect to Qin et al. (2021b), in order to allow for an increased (though overall small) MCG contribution to reionization. We assume the same scaling with mass

$$\alpha_{\mathrm{esc}}^{(\text{II})} = \alpha_{\mathrm{esc}}^{(\text{III})} = -0.3$$

so the escape fractions from MCGs and ACGs are comparable at their pivot points. We have lowered the ACG escape fraction normalization ($f_{\mathrm{esc},10}^{(\text{II})}$) by 0.15 with respect to Qin et al. (2021b), in order to allow for an increased (though overall small) MCG contribution to reionization. We assume the same scaling with mass for both populations, as these agree with Lyman $\alpha$ forest + CMB data (Qin et al. 2021b). Under these assumptions, we find that ACG-hosted Pop II stars dominate the ionizing photon budget at $z \leq 15$. This is to be expected, as MCGs are subdominant at lower $z$, and we agnostically set a (relatively) low value of $f_{\mathrm{esc}}^{(\text{II})}$ for Pop III stars.

To calculate X-ray emission we assume a power-law SED with a spectral energy index, $\alpha_X$, and a low-energy cutoff $E_0$. X-ray photons with energies below $E_0$ are absorbed within the host galaxies and do not contribute to ionizing and heating the IGM. This was shown to be an excellent characterization of the X-ray SED from either the hot interstellar medium (ISM) or high-mass X-ray binaries (HMXBs), when emerging from simulated, metal-poor, high-redshift galaxies (Fragos et al. 2013; Pacucci et al. 2014; Das et al. 2017). Both $\alpha_X$ and $E_0$ control the hardness of the emerging X-ray spectrum, and thus the patchiness of IGM heating. Here, we fix $\alpha_X = 1$, and vary $E_0$. For our fiducial value, we choose $E_0 = 0.5$ keV, based on the ISM simulations of Das et al. (2017), though we also explore a more optimistic model with a softer SED. The normalization of the X-ray SED is determined by the soft-band (with energies less than 2 keV)$^9$ X-ray luminosity to SFR parameter, $L_{X,\lesssim2keV}/\mathrm{SFR}$. Consistent with simulations of HMXBs in metal-poor environments (Fragos et al. 2013), here we take $\log_{10}(L_{X,\lesssim2keV}/\mathrm{SFR}) \approx 40.5$ erg s$^{-1}$ per unit SFR (M$_{\odot}$ yr$^{-1}$) for both ACGs and MCGs. Such high values are further supported by recent 21-cm power spectrum upper limits at $z = 8$ from HERA The HERA Collaboration (2022).

3 OBSERVABLES DURING THE CD AND EOR

We now use the models for ACGs and MCGs outlined above to predict the evolution of the thermal and ionization state of the IGM at high redshifts. We focus on the contribution of Pop III star-forming MCGs, as their corresponding feedback is the main improvement of this work.

We use two sets of galaxy parameters, Fiducial (EOS) and Optimistic (OPT), summarized in Table 1. The fiducial (EOS) uses the same ACG parameters as in Qin et al. (2021b), where such a model was shown to reproduce EoR observables (at redshifts $z \leq 10$ when the MCG contribution is negligible). However, we do lower the escape fraction of ACGs slightly ($\Delta \log f_{\mathrm{esc}}^{(\text{II})} = -0.15$), so as to allow for a modest contribution of MCGs to the high-redshift tail of the EoR (as could be slightly preferred by the CMB EE PS at $l \sim 20$–30; e.g. Qin et al. 2020b; Ahn & Shapiro 2021; Wu et al. 2021).

Our Optimistic (OPT) parameter set was chosen in order to enhance the relative contribution of MCGs to the SFRD. Although the actual parameter values are fairly arbitrary (given the lack of MCG constraints), we tuned the OPT model to reproduce the timing of the putative EDGES global 21-cm detection at $z \sim 17$ (Bowman et al. 2018). This is mainly done by increasing the SFRD of minihaloes (through $\phi_{\mathrm{M},\nu}(z)$) as well as the softness of the X-ray SED emerging from galaxies (through $E_0$).

In this section, we introduce the main high-redshift observables using our Fiducial (EOS) model. We will show the impact of Pop III stars and present the 2021 installment of the Evolution Of 21-cm Structure (EOS) project, whose goal is to show the state of knowledge of the astrophysics of cosmic dawn and reionization. Later in Sections 4 and 5, when we study parameter variations and VAs, we will mainly focus on the optimistic (OPT) model.

3.1 UV luminosity functions

The first observable we show are UVLFs. Although limited to comparably brighter galaxies during the EoR/CD, UVLFs detected with the HST provided invaluable insights into galaxy formation and evolution at $z \leq 10$.

We follow Park et al. (2019), where the 1500 Å luminosity is obtained from the SFR with a conversion factor of $\kappa = 1.15 \times 10^{-28}$ M$_{\odot}$ s yr$^{-1}$ erg$^{-1}$ (Sun & Furlanetto 2016; Oesch et al. 2018), and for simplicity we take $\kappa$ to be the same for both Pop II and Pop III populations. Detailed population-synthesis models suggest that there could be a factor of $\sim 2$ variation in this conversion, based on the IMF, metallicity, and star formation history (e.g. Wilkins, Lovell & Stanway 2019).

We show the predicted UVLFs for our fiducial (EOS2021) parameters in Fig. 7, which matches very well the observational data from Bouwens et al. (2015, 2016) and Oesch et al. (2018). This is by construction, as our parameters are motivated by the maximum a posteriori (MAP) model from Qin et al. (2021b), which included UVLFs (in addition to other EoR observables) in the likelihood.

We note that the atomic-cooling threshold $M_{\mathrm{cool}} = 2 \times 10^7$ M$_{\odot}$ at $z = 6$ corresponds to $M_{\mathrm{UV}} \approx -9$ in our model, making it difficult to directly observe the even fainter MCGs. This is clear in our Fig. 7, where MCGs only dominate at fainter magnitudes, beyond the reach of even JWST. We note that previous work (e.g. O’Shea et al. 2015;
The first galaxies during CD and the EoR

Figure 7. UVLFs at five redshifts for our EOS fiducial parameters, where the solid and dashed lines represent the ACG and MCG contribution, respectively. The data points are from Bouwens et al. (2015, 2016) and Oesch et al. (2018), where only galaxies with $M_{\text{UV}} \geq -20$ are shown to dominate the cosmic radiation backgrounds at high redshifts of interest. The top panel shows $z = 6–8$, where the data covers a broader magnitude range, whereas the bottom panel focuses on earlier epochs, where only some data are available at $z = 10$. The vertical dotted line shows the magnitude expected of a halo with $M_h = M_{\text{atom}}$ at $z = 6$, illustrating that in our model MCGs are far too faint to be directly detectable through the UVLF. Nevertheless, these data allow us to constrain the SHMR of ACGs, and thus isolate the contribution of MCGs to cosmic radiation fields using 21-cm observations.

Xu et al. 2016b; Qin et al. 2021a), found a turnover for MCGs at even fainter magnitudes ($M_{\text{UV}} \gtrsim -7–8$). This difference is due to our feedback prescriptions, including relative velocities that dominate at early times (cf. Fig. 5). As a consequence, the UVLF for MCGs is always below (or comparable to) that of ACGs in Fig. 7. We stress that even though we are unlikely to observe such ultrafaint magnitudes, UVLFs provide an invaluable data set by allowing us to anchor the SHMR scaling relations at the brighter end that is well probed by observations ($–20 \leq M_{\text{UV}} \leq -15$).

Through the rest of this section we will study how the inclusion of Pop III-hosting MCGs – and the $v_{\text{fb}}$ feedback on them – affects reionization and the 21-cm signal. We will do so by comparing our best-guess EOS fiducial simulation to one without MCGs, as well as one with both ACGs and MCGs but no relative velocities (similar to Qin et al. (2020a) though with an updated LW feedback prescription).

3.2 EoR history

We show the evolution of the EoR from our Fiducial model (EOS2021) in Fig. 8. We plot both $x_{\text{HI}}$ and the optical depth $\tau_{\text{CMB}}$ of the CMB due to reionization, as a function of $z$. Given our fiducial parameters, MCGs only make a small contribution to cosmic reionization, and chiefly at high $z$. The overall evolution of $x_{\text{HI}}$ agrees broadly with current measurements from McGreer, Mesinger & D’Odorico (2015), Greig et al. (2017, 2019), Mason et al. (2019a), Whittet et al. (2020), Wang et al. (2020), and Planck 2018 temperature and polarization (reanalysed in de Belsunce et al. 2021). This is mostly by design, as our ACG (Pop II) parameters were chosen to be consistent with Qin et al. (2021b), who used various EoR observables to constrain the EoR history. In particular, the final overlap stages of reionization are constrained by observations of the Lyman $\alpha$ opacity fluctuations. Qin et al. (2021b) found that the latest forest spectra require a late end to reionization at $z \sim 5$ (see also e.g. Kulkarni et al. 2019; Keating et al. 2020; Nasir & D’Aloisio 2020; Choudhury et al. 2021).

The contribution of MCGs would however be largest during the earliest stages of the EoR. In our fiducial model Pop III-hosting MCGs drive a modest increase of the CMB optical depth: $\Delta \tau_{\text{CMB}} \approx 0.015$ (though the precise value is sensitive to our fairly arbitrarily chosen MCG parameters). As seen from the figure, this contribution is mostly sourced at $z \sim 8–12$, with the very high-$z$ tail only contributing $\tau(z = 15–30) \approx 2 \times 10^{-3}$, well below the limit of 0.02 from Planck 2018 (Millea & Bouchet 2018; Heinrich & Hu 2021). While further data from CMB experiments will more precisely pinpoint $\tau_{\text{CMB}}$ (Abazajian et al. 2016; Ade et al. 2019), it will be difficult to isolate the high-$z$ contribution that could be caused by MCGs (Wu et al. 2021). We note that ignoring $v_{\text{fb}}$ feedback roughly doubles the contribution of MCGs to reionization, for our fiducial parameter choices.

In Fig. 9, we compare our fiducial, EOS2021 EoR history to that of the EOS2016 release (Mesinger et al. 2016). The 2016 release was comprised of two models, FaintGalaxies and BrightGalaxies, both of which only included Pop II-hosting ACGs but assumed different turnover mass scales for SNe feedback. In this work, we assume SNe feedback does not induce a turnover, and include also Pop III-hosting MCGs with the associated LW and $v_{\text{fb}}$ feedback followed self-consistently. To make this distinction explicit, we denote our EOS2021 model as AllGalaxies in the figure.
All EOS releases are ‘tuned’ to reproduce the current state of knowledge. In 2016, our estimate of $\tau_{\text{CMB}}$ and forest data suggested an earlier middle/end of reionization, as is reflected in this comparison plot. The shapes of the EoR histories are also notably different. By including MCGs, AllGalaxies (EOS2021) produce a more extended tail to higher redshifts, with percent level ionization up to $z \sim 20$. Despite this earlier start, the ACG-driven mid and late stages of reionization occur more rapidly in AllGalaxies than in FaintGalaxies. This is because both EOS2016 models assumed a constant mass-to-light ratio ($q_{\text{UV}}^{10} = 0$). This assumption, albeit common, is inconsistent with the latest UVLF observations and overestimates star formation in small galaxies, thus resulting in a slower evolution of the SFRD and associated cosmic epochs (see also Mirocha et al. 2017; Park et al. 2019). This figure highlights the importance of using the latest observations to guide our models of the early universe.

### 3.3 The 21-cm line during the cosmic dawn

The biggest impact from MCGs will be to the cosmic-dawn epoch, which we mainly observe through the 21-cm line of neutral hydrogen. We now explore this observable.

We first show our definitions, and refer the reader to Furlanetto, Oh & Briggs (2006) and Pritchard & Loeb (2012) for detailed reviews of the physics of the 21-cm line. We use the full expression for the 21-cm brightness temperature (Barkana & Loeb 2001)

$$T_{21} = T_s - T_{\text{CMB}} \frac{1}{1+z} (1 - e^{-\tau_{21}}),$$

where $T_{\text{CMB}}$ is the temperature of the CMB (which acts as the radio back light), $T_s$ is the spin temperature of the IGM, and

$$\tau_{21} = (1 + \delta) \tau_{10} \frac{H(z)}{T_s} \frac{\Delta v_s}{\Delta v_{00}} (1+z),$$

where $\Delta v$ is the line-of-sight gradient of the velocity. We have defined a normalization factor

$$T_0 = 34 \text{ mK} \left( \frac{1+z}{16} \right)^{1/2} \left( \frac{\Omega_b h^2}{0.022} \right) \left( \frac{\Omega_m h^2}{0.14} \right)^{-1/2}$$

anchored at our Planck 2018 cosmology.

The spin temperature of hydrogen, which determines whether 21-cm photons are absorbed from the CMB (if $T_s < T_{\text{CMB}}$) or emitted (for $T_s > T_{\text{CMB}}$), is set by competing couplings to the CMB and to the gas kinetic temperature $T_K$, and can be found through

$$T_s^{-1} = \frac{T_{\text{CMB}}^{-1} + x_h T_s^{-1} + x_e T_K^{-1}}{1 + x_h + x_e},$$

where $T_s$ is the colour temperature (which is closely related to $T_K$, Hirata 2006), and $x_h$ are the couplings to $T_K$ due to Lyman-$\alpha$ photons through the Wouthuysen–Field effect (Wouthuysen 1952; Field 1959) and through collisions ($x_e$, which are only relevant in the IGM at $z \gtrsim 30$, Loeb & Zaldarriaga 2004).

There are two main avenues for measuring the 21-cm signal. The first is by through its monopole against the CMB, usually termed the global signal (GS hereafter). The second is through the fluctuations of the signal, commonly simplified into the Fourier-space two-point function or power spectrum (PS hereafter). We now explore each in turn.

#### 3.3.1 Global Signal

The 21-cm GS (denoted by $T_{21}$) is currently targeted by experiments such as EDGES (Bowman et al. 2018), LEDA (Price et al. 2018), SARAS (Singh et al. 2018), Sci-Hi (Voytek et al. 2014), and Prizm (Philip et al. 2019). We note however that the lack of angular resolution makes the GS especially difficult to disentangle from foregrounds, which can be several orders of magnitude stronger than the cosmic signal.

We show our predictions for the 21-cm GS in Fig. 10. The 21-cm signal is characterized three well-known eras at high $z$. First, there is the epoch of coupling (EoC, $z \sim 15$–25 for our EOS parameters), where the GS $T_{21}$ becomes more negative due to the Wouthuysen–Field (WF) coupling sourced by the Lyman-series photons from the first galaxies (Wouthuysen 1952; Field 1959). Then, there is the epoch of X-ray heating (ExH), where X-rays emitted by galaxies heat up the IGM, slowly increasing $T_{21}$ until it is above zero. Finally, during the EoR $T_{21}$ is driven towards zero following $T_{\text{HI}}$.

The difference between models with Pop II only and with Pop III stars is dramatic, as shown in the top panel of Fig. 10. This is to be expected, as the high-$z$ SFRD is dominated by the smaller (and thus more abundant) MCGs. Ignoring Pop III-dominated MCGs delays the minimum in the GS from $z \sim 15$ to $z \approx 12.5$, and results in a more rapid evolution of all epochs (ExC, ExH, and ExE). The top panel of Fig. 10 also shows that neglecting $v_{\text{gs}}$ feedback shifts the ExC and ExH earlier by $\Delta z \approx +2$, which highlights the importance of this effect during the cosmic dawn.

We also compare the GS of the fiducial EOS2021 simulation against the previous two models of EOS2016. The most striking distinction between these three simulations is that the depth for the new EOS2021 (AllGalaxies) model is a factor of 2 shallower than for both EOS 2016 simulations, only reaching values of $T_{21} \approx -70$ mK. As already discussed, the emissivity in the 2016 simulations did not follow the SHMR implied by UUFLs but was instead proportional to the collapsed-fraction (equivalent to a constant $f_0$ in Fig. 4). As a consequence, the SFRD in those models evolved more rapidly, producing more distinct ExC and ExH epochs, whereas in the realistic AllGalaxies model these overlap (see also Mirocha et al. 2017; Park et al. 2019; Qin et al. 2020a).

Such a shallower absorption trough also has implications for exotic cosmic explanations of the recent EDGES detection (Bowman et al. 2018); though we caution that a cosmological interpretation of the detected signal remains very controversial (Hills et al. 2018;
3.3.2 Power spectrum

We now study the 21-cm fluctuations. For simplicity, we will focus on the spherically averaged PS summary statistic, defined through

$$\langle \delta T_21(k)\delta T_21(k') \rangle = (2\pi)^3 \delta D(k + k') P_21(k),$$

(23)

although in practice we will employ the reduced power spectrum

$$\delta^2_21 = k^3 P_21(k)/(2\pi^2),$$

with units of mK$^2$, for convenience. Interferometers can measure many 21-cm modes at each $z$, and thus the PS (and other spatially dependent statistics) can provide more detailed insights into the early universe, compared to the GS (Pritchard & Furlanetto 2007; Parsons et al. 2012; Pober et al. 2013b; Fialkov & Barkana 2014; Cohen, Fialkov & Barkana 2018; Muñoz, Dvorkin & Cyr-Racine 2020; Jones et al. 2021). Experiments such as HERA (DeBoer et al. 2017), LOFAR (van Haarlem et al. 2013), MWA (Tingay et al. 2015), LWA (Eastwood et al. 2019), and the SKA (Koopmans et al. 2015) are aiming to measure the 21-cm PS.

We show the evolution of the 21-cm PS at two different Fourier-space wavenumbers $k$ in Fig. 10 (along with the GS for visual aid). The large-scale ($k = 0.15 \text{Mpc}^{-1}$) power has three bumps, corresponding to the three eras outlined above, when fluctuations in the Lyman $\alpha$ background, IGM temperature, and ionization fractions dominate the 21-cm PS, respectively. Between these, the negative contribution of the cross power between these fields gives rise to relative troughs in the PS (Pritchard & Furlanetto 2007; Lidz et al. 2008; Mesinger et al. 2013) and as a result $\delta^2_21 \propto dT_21/dz$ on large scales (Muñoz & Cyr-Racine 2021). At smaller scales ($k = 0.15 \text{Mpc}^{-1}$), however, this cancellation does not take place, and the power is larger overall. The power is larger for the Pop II-only model at both large and small scales, as the smaller-mass MCGs that host Pop III stars are less biased, producing smaller 21-cm fluctuations. The absence of $v_{\text{th}}$ feedback shifts all curves towards earlier times. Moreover, as we will see in Section 5, the $v_{\text{th}}$ fluctuations become imprinted on to the 21-cm PS (Dalal et al. 2010; Visbal et al. 2012; Muñoz 2019a), giving rise to sizable wiggles on the 21-cm power spectrum.

In Fig. 12, we compare the PS from our fiducial EOS2021 model ($\text{AllGalaxies}$), to the previous EOS2016 models ($\text{BrightGalaxies}$ and $\text{FaintGalaxies}$). The newer $\text{AllGalaxies}$ model shows significantly smaller power during the cosmic dawn than both EOS 2016

---

Sims & Pober 2020). The lowest point of our trough ($T_{21} \approx -70$ mK) is roughly a factor of 7 shallower than claimed by EDGES ($T_{21} \approx -500$ mK), requiring either stronger dark matter electric charges (e.g. Barkana 2018; Muñoz & Loeb 2018), or a brighter extra radio background (e.g. Ewall-Wice et al. 2018; Pospelov et al. 2018) than previously assumed. However in terms of timing, our fiducial EOS model (with Pop III stars) peaks at $z \approx 15$, only slightly later than the timing of the first claimed EDGES detection (at $z \approx 17$). We will study a different set (OPT) of MCG parameters in Sections 4 and 5, which give rise to an absorption trough at an earlier $z \approx 17$.
predictions. That is because of both the shallower absorption – and slower evolution – of the 21-cm global signal. This reduction reaches an order of magnitude at \( z \geq 10 \), and will hinder the observation of the 21-cm power spectrum with interferometers.

Nevertheless, our AllGalaxies 21-cm power spectrum is still significantly above the thermal noise level forecasted for upcoming interferometers like HERA and SKA. We show in Fig. 12 the expected noise after 1 yr (1080 h) of integration with HERA, calculated with 21cmSense\(^{10}\) (Pober et al. 2013a; Pober et al. 2014) under the moderate-foreground assumption with a buffer \( a = 0.1 \) hMpc\(^{-1} \) above the horizon. We assume a fixed bandwidth of 8 MHz (corresponding to \( \Delta z = 0.7 \) at \( z = 10 \)), spherical bins of \( \Delta k = 0.1 \) Mpc\(^{-1} \), and a system temperature (DeBoer et al. 2017)

\[
T_{\text{sys}}(\nu) = 100 \, \text{K} + 120 \, \text{K} \times \left( \frac{\nu}{150 \, \text{MHz}} \right)^{-2.55}.
\]

This results in a noise that is below the signal up to \( z \approx 20 \) at large scales, and comparable to the signal at small scales. At low \( k \), the noise is dominated by cosmic variance, which tracks the amplitude of the PS, whereas for at high \( k \) it is largely thermal.

We further quantify the detectability of our EOS 2021 model (AllGalaxies), by computing the signal-to-noise ratio (SNR). We calculate this quantity for each wavenumber \( k \) and redshift bin \( z \) considered, and add them in quadrature. We consider two values for \( T_{\text{sys}} \), that of equation (24) and a more pessimistic one of

\[
T_{\text{sys}}^{\text{pess.}}(\nu) = 100 \, \text{K} + 400 \, \text{K} \times \left( \frac{\nu}{150 \, \text{MHz}} \right)^{-2.55}.
\]

\(^{10}\)https://github.com/jpober/21cmSense

**Table 2.** SNR for the two interferometers we consider, under a regular system-noise assumption, given by equation (24), and a pessimistic one, from equation (25). The three epochs cover the ranges of \( z \leq 10 \) for the EoR, \( 10 < z < 15 \) for the EoH, and \( z \geq 15 \) for the EoC.

|               | EOS2021 | EoR | EoH | EoC |
|---------------|---------|-----|-----|-----|
| HERA (pess.)  | 87      | 86  | 8   | 1   |
| HERA          | 186     | 183 | 34  | 9   |
| SKA           | 164     | 157 | 41  | 22  |
| SKA (pess.)   | 85      | 84  | 12  | 4   |

**Figure 12.** The 21-cm power spectrum at two wavenumbers \( k = 0.2 \) and \( 1.0 \) Mpc\(^{-1} \) for all EOS simulations, with the same colour coding as Fig. 11. We additionally show the expected noise from HERA in different \( z \) bands (all corresponding to a bandwidth of \( B = 8 \) MHz), where we assume 180 d of observation, moderate foregrounds, and \( \Delta k = 0.1 \) Mpc\(^{-1} \). We have taken AllGalaxies as the fiducial for the cosmic-variance noise, and find a total SNR = 183 added in quadrature over \( k \) and \( z \).

**Figure 13.** Power spectrum of 21-cm fluctuations as a function of wavenumber \( k \) for our AllGalaxies (EOS2021) simulation. The four redshifts are chosen to be during the EoC (\( z = 16 \)), half way through the EoH (\( z = 13 \)), in the transition between the EoH and the EoR (\( z = 9 \)), and finally during the EoR (\( z = 6 \)). The small error bars come from the Poisson noise in our box, which is 1.5 Gpc comoving in size. The lines represent a fit (using \( k = 0.02-0.5 \) Mpc\(^{-1} \)) to a smooth polynomial added to the wiggles from the VAOs (sourced by the streaming velocities \( v_{SB} \)) as we will explain in Section 5.

following Dewdney et al. (2016), which results in a noise larger by a factor of \( \sim 3 \) at the redshifts of interest. We also calculate the SNR for the fiducial SKA-LOW 1 design Dewdney et al. (2016) using a tracked observing strategy. Specifically, we assume a 6 h per-night tracked scan for a total of 1000 h. Table 2 shows the SNRs for the different setups, where using equation (24) we find SNR = 186 for HERA and SNR = 164 for the SKA, both of which would provide detections at high significance. These SNRs would be reduced by a factor of \( \sim 2-3 \) for the pessimistic \( T_{\text{sys}} \) from equation (25). Divided into epochs, the SNR is significantly dominated by the EoR, with the EoH contributing a factor of \( \sim 5 \) less, and the EoC only showing SNR \( \sim 10 \). Interestingly, for our fiducial ‘narrow and deep’ SKA survey, the SKA can reach larger SNR at high \( z \) where thermal noise dominates, while HERA performs better at lower \( z \) where cosmic variance can dominate the noise. Assuming different SKA observing strategies can shift the balance between cosmic-variance and thermal-noise errors by considering either larger observing volumes or deeper integration times (see e.g Greig, Mesinger & Koopmans 2020a). We will explore in Mason et al. (in preparation) the range of constraints that such a detection would provide for astrophysical and cosmological parameters.

Finally, we show the scale dependence of the 21-cm PS at four redshifts in Fig. 13. These are chosen to illustrate the power spectrum during each of the three epochs of interest (EoR, EoH, and EoC), as well as in the transition between the EoR and EoH. The power is
relatively flat with $k$ except in the transition case ($z=9$), where the large-scale power drops dramatically due to the negative contribution of the cross-terms (Pritchard & Furlanetto 2007; Lidz et al. 2008; Mesinger et al. 2013; Muñoz & Cyr-Racine 2021). Interestingly, at $z = 13$ the power peaks at $k = 0.1\ \text{Mpc}^{-1}$, where there are wiggles in the 21-cm power spectrum. These are due to the streaming velocities $v_{cb}$, which have acoustic oscillations that become imprinted on to the SFRD (through the feedback described in Section 2), and thus on the 21-cm signal. We will describe these velocity-induced acoustic oscillations (VAOs) in detail in Section 5.

3.4 Visualizations

We begin with Fig. 14, which shows 2D slices through cosmic light-cones from our fiducial EOS 2021 simulations. The horizontal axis shows evolution with cosmic time, while the vertical axis corresponds to a fixed comoving length (here taken to be half of the full EOS size in order to more easily identify small-scale features). While we only track a few variables in that figure, we note that 21cmFAST can output relevant quantities such as the local recombination rate, the intensity of the UV, X-ray, Lyman $\alpha$ backgrounds, the velocity fields, and kinetic and spin temperatures. We describe each of the panels in turn.

(i) The top panel shows the matter over/underdensities, which grow due to gravity as the universe evolves, forming the cosmic web that we see today.

(ii) The second panel of Fig. 14 shows the LW flux, which dissociates H$_2$ molecules and thus impedes star formation in MCGs.
The overall LW flux $J_{21}$ grows rapidly over time (roughly following the SFRD), with notable spatial fluctuations. The LW flux is largest in regions corresponding to the largest matter overdensities, which host the first generations of (highly biased) galaxies. Therefore these anisotropies result in a stronger suppression of Pop III star formation, than would be expected assuming a homogeneous background, and imprint structure in the 21-cm signal.

(iii) The third and fourth panels of Fig. 14 show the evolution of the two quantities most directly observable: the neutral hydrogen fraction $x_{\text{HI}}$ and the 21-cm brightness temperature $T_{21}$. The neutral fraction $x_{\text{HI}}$ is homogeneously close to unity until $z \approx 10$, as chiefly MCGs form at those early times, which are not very efficient at reionizing hydrogen. The EoR takes place during $z = 5$–10, accelerating at later times. This panel reveals large-scale neutral patches at $z \approx 5.5$–6, as required by recent Lyman $\alpha$ data (Becker et al. 2015; Bosman et al. 2018). These last neutral regions trace underdense environments.

(iv) The 21-cm signal $T_{21}$, in the last panel, shows a dramatic evolution during cosmic dawn, beginning with absorption at $z \approx 15$–20 due to the Wouthuysen–Field effect, followed by a transition to emission at $z \approx 12$ as the IGM is heated by X-rays from the first galaxies, and finishing with a slow decay towards zero as reionization takes place. The late-time ($z \leq 10$) behaviour of the 21-cm signal is dominated by the reionization bubbles, as the $x_{\text{HI}}$ and $T_{21}$ fields are clearly correlated. The early-time ($z \gtrsim 10$) behaviour, however, is related to the UV and X-ray flux from the first galaxies, which depends on the densities $\delta$ in a non-linear and non-local way.

In Fig. 15, we show a zoom-in slice through the 21-cm map at $z = 11$. This redshift corresponds roughly to the transition between the EoH and the EoR. It is clear that there is a large overlap between these two eras, as the overdense heated regions (with $T_{21} > 0$, shown in red), are beginning the process of reionization from the inside (producing $T_{21} \rightarrow 0$, in white). We note that the underdense regions, exposed to a smaller X-ray flux, are remain colder than the CMB at this redshift ($T_{21} < 0$, in blue).

Fig. 15 also shows the dynamic range of our simulations, which can resolve structure as large as the simulation box (1.5 Gpc) and down to the cell size (1.5 Mpc). A zoom-in animation of the entire cosmic dawn and EoR evolution is provided at this url.

4 LEARNING ABOUT THE FIRST GALAXIES
In the previous sections we have demonstrated that MCGs drive the 21-cm signal from the early cosmic dawn, given our fiducial set of parameters. As a consequence, 21-cm studies are a promising avenue to learn about the properties of the first Pop III stars and their host galaxies. Here, we perform a brief exploratory study of how the 21-cm signal varies as a function of the different MCG stellar and feedback parameters in our model.

Throughout this section, we will vary parameters around a more optimistic set of galaxy properties, labelled OPT in Table 1. This set allows for a more significant contribution from MCGs compared to EOS, making it easier to learn about these first galaxies. In particular, we increase the stellar efficiency of MCGs by $\Delta \log_{10} f_{*10}^{(\text{OPT})} = -0.75$.
(roughly a factor of ≈ 5), and decrease it for ACGs by $\Delta \log_{10} f_{\text{esc}}^{(\text{III})} = -0.25$, in both cases compensating their $f_{\text{esc}}$ to keep a similar EoR evolution. This higher MCG contribution pushes the trough of the 21-cm GS to an earlier $z \sim 17$ (in line with the central redshift claimed by EDGES in Bowman et al. 2018), rather than the $z \sim 15$ of our EOS parameters. We also find that in OPT, Pop III stars dominate the SFRD for $z \gtrsim 10$, thus dictating the evolution of the entirety of cosmic dawn.

We begin by studying the slope of the SHMR, here parametrized with the $\alpha_{*}^{(\text{III})}$ parameter. The SHMR slope likely holds clues about SNe and other galactic feedback mechanisms. For example, assuming star formation is regulated by SNe-driven outflows with a constant energy coupling efficiency, can result in $\alpha_{*} \sim 2/3$ (e.g. Wyithe & Loeb 2013). This is remarkably close to the empirically determined value from $z \gtrsim 6$ UVLFs of $\alpha_{*}^{(\text{III})} \sim 0.5$ (for ACGs hosting Pop II stars).

As discussed previously, ACGs are too faint to allow us to directly measure $\alpha_{*}^{(\text{III})}$ from UVLFs. Although it is tempting to assume the same SHMR slope for ACGs and MCGs, this could be incorrect due to, e.g. their different IMFs and associated SNe energies. For example, a fixed mass of stars forming in each MCG (e.g. Kulkarni et al. 2019) would result in $\alpha_{*}^{(\text{III})} < 0$, which would be very different from the $\alpha_{*}^{(\text{III})} \sim 0.5$ that is empirically determined from $z \gtrsim 6$ UVLFs.

In Fig. 16, we show the impact of varying the power-law index $\alpha_{*}^{(\text{III})}$ on the 21-cm CD signal (global and power spectrum evolution). We consider $\alpha_{*}^{(\text{III})}$ in the $-0.2–0.5$ range. It is clear that the 21-cm signal does not vary dramatically over this range of SHMR slopes. Steeper indices (larger $\alpha_{*}^{(\text{III})}$) effectively result in steeper SFRD evolutions at very high redshifts. Although initially there are fewer Pop III stars in those models, delaying the 21-cm signal, subsequently the cosmic evolution accelerates.

Interestingly, all models in Fig. 16 agree at $z \approx 15$, when the contribution from ACGs starts being relevant. However, we do see that the $J_{21}$ evolution of some models crosses-over around this redshift (c.f. the green curve). This is due to LW feedback: models that initially have a larger $J_{21}$ are subsequently able to more strongly quench MCG star formation, thus resulting in a weaker $J_{21}$ at $z < 15$.

We now study the impact of the other parameters regulating the UV and X-ray emissivities of MCGs. Specifically, in Fig. 17 we show how the 21-cm GS and PS (at $k = 0.15$ Mpc$^{-1}$) vary with: (i) $A_{\text{UV}}$, the amplitude of LW feedback; (ii) $f_{\text{esc}}^{(\text{III})}$, the normalization of the SHMR; (iii) $f_{\text{esc}}^{(\text{III})}$, the ionizing escape fraction; and (iv) $L_{X,2\text{keV}}^{(\text{III})}$/SFR, the X-ray luminosity to SFR relation. Unlike for $\alpha_{*}^{(\text{III})}$ discussed above, the uncertainty on these parameters is better sampled in log space. Therefore in Fig. 17 we show results when increasing or decreasing each parameter by a factor of three around the OPT values from Table 1.

The first panel of Fig. 17 shows that stronger LW feedback (larger $A_{\text{UV}}$) translates into a delayed 21-cm GS and PS, especially at high $z$. The larger impact of other feedback sources, chiefly the relative velocities, makes the signal depend only weakly on $A_{\text{UV}}$, though this parameter can still delay the cosmic-dawn milestones by $\Delta z \approx 1$ within the range of values we study.

Changing $f_{\text{esc}}^{(\text{III})}$ also has a modest impact, as MCGs are generally negligible contributors to the EoR in our models. However, the largest values of the escape fraction shown here do result in an earlier start to the EoR (driven by MCGs), but with a similar end (driven by ACGs).

On the other hand, varying the stellar fraction $f_{*}^{(\text{III})}$ (second panel) or $L_{X,2\text{keV}}^{(\text{III})}$/SFR (fourth panel) notably changes the signal during the two CD epochs driven by MCGs: the EoH and EoC. Changing the stellar fraction impacts both epochs, as star formation drives all of the cosmic radiation fields in our models. Higher stellar fractions shifts the CD to earlier times, resulting in a higher effective bias of the sources driving each epoch, and thus a higher 21-cm PS on large scales.

Changing $L_{X,2\text{keV}}^{(\text{III})}$/SFR only impacts the relative timing of the EoH. Increasing the X-ray luminosity of the first galaxies results in a larger overlap of the EoH and EoC, as the coupling is not completed before the IGM is heated. Consequently, the GS absorption trough is shallower, and the large-scale power decreases from the increased negative contribution of the cross-power in these two fields (e.g. Pritchard & Furlanetto 2007; Mesinger et al. 2013; Schneider et al. 2021).

Each of the parameters impacts the signal differently as a function of redshift and scale, which may allow us to distinguish between them. However, in order to forecast parameter uncertainties, one has to capture the correlations between them, for instance through an MCMC (Greig & Mesinger 2015) or Fisher matrix (Mason et al., in preparation). We leave this question for future work. We note that the expected SNRs for the OPT model are similar to the EOS ones reported in Section 3. That is because the OPT model shows slightly
larger fluctuations, though at lower frequencies where noise is larger. Table 3 contains our forecasted SNRs for the OPT parameters under each of the assumptions considered.

5 VELOCITY-INDUCED ACOUSTIC OSCILLATIONS

The last study we perform is on the unique signature of the DM-baryon relative velocities on the 21-cm fluctuations. We quantify to what extent the streaming velocities produce velocity-induced acoustic oscillations (VAOs) on the 21-cm signal in our simulations, for the first time jointly including inhomogeneous LW feedback with self-shielding. Throughout this section we will assume astrophysical parameters from the Optimistic (OPT) set, unless otherwise specified.

5.1 VAOs across cosmic dawn

The interactions between baryons and photons give rise to the well-known baryon acoustic oscillations (BAOs), which at low redshift z are observed in the matter distribution as an excess correlation at the baryon acoustic scale (here used interchangeably with the baryon drag scale \( r_{\text{drag}} \); Eisenstein & Hu 1998). The dark matter, however, does not partake in these BAOs, which gives it different initial conditions than baryons at recombination. This produces a relative (or streaming) velocity between dark matter and baryons, which fluctuates spatially with the same \( r_{\text{drag}} \) scale, due to their acoustic origin (Tseliakhovich & Hirata 2010). In Fourier space, the power spectrum of \( v_{\text{ch}} \) presents large wiggles, which are inherited by the radiation fields, as regions of large relative velocity suppress the formation of the first stars (chiefly PopIII, see Section 2.2). Consequently, the 21-cm signal becomes modulated by these streaming velocities during cosmic dawn, giving rise to velocity-induced acoustic oscillations (VAOs; Dalal et al. 2010; Visbal et al. 2012; Fialkov et al. 2013; Muñoz 2019a), with the same acoustic origin as the BAOs, though sourced by velocity – rather than density – fluctuations. Muñoz (2019b) showed that these VAOs provide us with a standard ruler to measure physical cosmology during cosmic dawn, independently of galaxy astrophysics.

Until recently it was not known how the feedback from \( v_{\text{ch}} \) interacted with the LW dissociation of molecular hydrogen. As we showed in Section 2, however, recent hydrodynamical simulations from Kulkarni et al. (2021) and Schauer et al. (2021) indicate that for the regime of interest (\( J_{21} \leq 1 \)) these two processes act coherently. Furthermore, self-shielding in the first galaxies produces weaker LW feedback, and thus a larger impact of the relative velocities. Together, these two effects give rise to sizable VAOs, as we now show.

5.1.1 Slices

We begin by showing the impact of \( v_{\text{ch}} \) directly on the 21-cm maps. For that we compare a standard simulation (with OPT parameters) against one with no fluctuating \( v_{\text{ch}} \) (achieved by setting \( \text{FTX}_V \text{AVG} = \text{True} \) in 21-cm FAST). The latter simulation just uses a homogeneous value of \( v_{\text{ch}} = v_{\text{avg}} \approx 26 \text{ km s}^{-1} \), corresponding to the mean of its distribution. The reason for this choice, rather than setting \( v_{\text{ch}} = 0 \), is that the background evolution in the latter case would be significantly different (see e.g. Fig. 10), making it difficult to compare results at a fixed redshift.

We plot slices (1.5 Mpc thick) through our simulations at \( z = 11 \) (during the EoH) in Fig. 18. The slice through the relative-velocity field clearly shows large-scale acoustic structure, with islands of large \( v_{\text{ch}} \) separated by roughly \( r_{\text{drag}} \approx 150 \text{ Mpc} \). In contrast, the matter field (\( \delta \)) has power on all visible scales, down to our cell size. We also show the 21-cm map resulting from our two simulations with and without fluctuating \( v_{\text{ch}} \) (but with otherwise identical OPT parameters). Regions of large \( v_{\text{ch}} \) have a colder IGM, as they form fewer stars, and thus emit fewer X-rays. In order to illustrate this effect, we zoom into a patch 100 Mpc in size near a region of large \( v_{\text{ch}} \), where the full-physics simulation clearly presents deeper absorption correlated with the velocity map.
Figure 18. Effect of relative velocities on the 21-cm signal with our OPT parameters and the most up-to-date feedback prescriptions. We show a slice through our simulations, 1.5-Mpc deep and 600 Mpc on a side, at \( z = 11 \), where different panels show different quantities. Top left shows the matter density, and bottom left the DM-baryon relative velocity. The two central panels plot the 21-cm brightness temperature, on the top without fluctuating relative velocities (i.e. \( v_{cb} = \nu_{avg} \), and in the bottom with the full \( v_{cb} \) effect. We also show a zoom-in (100 by 100 Mpc) region of large relative velocity on the right, where the suppressive effect of \( v_{cb} \) on the first galaxies – and thus on \( T_{21} \) – is readily apparent.

We now move on to study the effect of \( v_{cb} \) during other cosmic eras. In Fig. 19, we show a 100 Mpc zoom-in of our simulations, at the same location as in Fig. 18. Rather than showing maps of \( \delta T_{21} \) with and without VAOs, we plot the difference (Diff \( \equiv \delta T_{21}^{\text{full}} - \delta T_{21}^{\text{no VAO}} \)) between the two cases, which allows for a closer comparison, at three redshift snapshots. The first is at \( z = 7 \), during the EoC, where the effect of the relative velocities is rather small, affecting the signal at the \( \sim 1\text{-}2 \) mK level. The second is at \( z = 15 \), at the peak of heating, and shows a large impact of \( v_{cb} \), causing Diff \( \approx \pm 5 \text{ mK} \), with the largest differences taking place in the lowest-\( v_{cb} \) regions, which heated more slowly. The last snapshot is at \( z = 20 \), which is during the EoC and where \( v_{cb} \) impacts the signal moderately, but in the opposite direction (as fewer photons produce less coupling, and thus more positive \( \delta T_{21} \)). As is clear from Fig. 18, the profile of the relative velocity (left-hand panel) is smeared when observed in the 21-cm signal due to photon propagation. This is especially true when \( z = 20 \) panel, where the difference has a homogeneous value of \( \sim 2 \) mK in the entire (100 Mpc) zoom-in region. This is because the photons just redward of Lyman \( \beta \) that drive WF coupling have mean free paths comparable to the 100-Mpc scale of these zoom-ins. Below we provide a simple analytical expression for the 21-cm PS including VAOs and accounting for such photon diffusion via window functions.

5.1.2. Power spectrum

The simulation slices studied above give us an idea of the impact of the relative velocities on the 21-cm signal at different epochs. We now calculate the 21-cm PS as a function of \( k \), quantifying the observable signature of the VAOs. We plot this observable from our simulations in Fig. 20 at the same redshifts as were shown in Figs. 18 and 19. The amplitude of those power spectra trace the overall redshift evolution that we studied in Section 3. However, more interestingly from the point of view of VAOs is the shape of the PS with \( k \). The simulation data points show marked acoustic oscillations (i.e. wiggles) at \( k = 0.05\text{-}0.5 \text{ Mpc}^{-1} \), inherited from the \( v_{cb} \) fluctuations. These VAOs are most pronounced during the X-ray heating era, increasing the power spectrum by an \( O(1) \) factor both at \( z = 11 \) and \( z = 15 \). They also appear during the EoC, at \( z = 20 \), and to a much lesser degree in the EoR at \( z = 7 \) (though we do not consider the effect of \( v_{cb} \) on ionizing sinks, as described in Cain et al. (2020) and Park et al. (2021), which may enhance the late-time VAOs).

The relative velocity \( v_{cb} \) is a vector field, so due to isotropy it can only affect observables through \( v_{cb}^2 \) to first order. We define the VAO shape \( \Delta_{v_{cb}} \) to be the power spectrum of

\[
\delta_{v_{cb}} = \sqrt{ \frac{3}{2} \left( \frac{v_{cb}^2}{v_{cm}^2} - 1 \right) } .
\]

This quantity has unit variance, zero mean, and is redshift independent. In Muñoz (2019a), we showed that the shape of the VAOs is unaltered by the complex astrophysics of cosmic dawn,\(^{11}\) although

\(^{11}\)This is not true for all scenarios, as for instance the sharp cutoffs in the primordial-black hole accretion model of Jensen & Ali-Haimoud (2021) do not always follow the VAO shape.
Figure 19. Zoom-in slice, 100 Mpc on a side and 1.5 Mpc in depth, from the same region as Fig. 18. We show the zoomed-in density $\delta$ and relative velocity $v_{cb}$ at $z = 11$ in the left two panels. The rest of top panels show the 21-cm fluctuations $\delta T_{21}$ for no VAOs ($v_{cb} = v_{\text{avg}}$), whereas the bottom panels show its difference with the full (i.e. with VAO) case, defined as Diff = $\delta T_{21}^{\text{full}} - \delta T_{21}^{\text{no VAO}}$. The effect of VAOs is subtle during the EoR ($z = 7$, second column), as differences are at the $\pm 2$ mK level, but more noticeable during the EoH ($z = 15$, third column) and the EoC ($z = 20$, fourth column).

Figure 20. The 21-cm power spectrum as a function of wavenumber $k$ at four redshifts $z$. The data-points show our simulation results (and Poisson noise) for the OPT parameters, and the lines present fits obtained using equation (27). The solid line contains VAOs, whereas the dashed line does not (i.e. it has $b_{vcb} = 0$), which has lower power at large scales and no wiggles. Simulation data-points in the grey shaded regions have not been included in the fit. The VAOs are most obvious during the EoH ($z \approx 10-15$), are somewhat visible during the EoC ($z \gtrsim 15$, though suppressed at high $k$), and very small during the EoR $z \lesssim 10$.

Its amplitude is damped if the X-ray or UV photons that affect the 21-cm signal travel significant distances (comparable to $1/k$). Thus, not only does the amplitude of the VAOs change between eras, but also the $k$ range where they are visible. This is clear from Fig. 20, as the $z = 15$ power spectrum has $\sim 3-4$ visible wiggles, whereas at $z = 20$ only 2 can be distinguished. That is because during the EoC the mean-free path of the relevant Lyman band photons is rather large ($\sim 100$ Mpc; Dalal et al. 2010), whereas during the X-ray heating era it is much shorter for realistic SEDs (Pacucci et al. 2014; Das et al. 2017). This is especially true for our OPT set of parameters, which has an optimistic value of $E_0 = 0.2$ keV and thus X-rays travel shorter distances. For the EOS parameter set on the other hand, we set $E_0 = 0.5$ keV (see Das et al. 2017), resulting in longer X-ray mean free path and thus more damped VAOs (c.f. Fig. 13, where only 2–3 wiggles are apparent).

In order to analytically model the VAOs we follow the approach of Muñoz (2019a,b, see also Hotinli et al. 2021), and use the fact that density- and $v_{cb}$-sourced fluctuations are uncorrelated to first order to write

$$\Delta_{21}^2 = \mathcal{P}_{\text{av}}(k) + b_{vcb}^2 W_i(k) \Delta_{i, vcb}^2(k), \quad (27)$$

where $\Delta_{i, vcb}^2$ is the power spectrum of $\delta_i$ in equation (26) and contains the VAO shape, $b_{vcb}$ is a bias factor that determines its amplitude, and $W_i$ is a window function that accounts for photon propagation. The $\mathcal{P}_{\text{av}}(k)$ 'no-wiggle' term, instead, accounts for the usual density-sourced 21-cm fluctuations, and we model it as a simple 4th order polynomial,

$$\log \mathcal{P}_{\text{av}}(k) = \sum_{j=0}^{4} c_j \left[ \log(k) \right]^j, \quad (28)$$

which suffices to capture its behaviour in the region of interest ($k = 0.02-0.5$ Mpc$^{-1}$).

For the ‘wiggle’ VAO part we know $\Delta_{i, vcb}^2$, but need to find both the window function $W_i$ that accounts for damping of VAOs, and the
bias $b_{\text{v}_b}$ that parametrizes their amplitudes. Let us begin with the window function.

We follow the approach of Muñoz (2019a) and define separate window functions $W_i(k)$ for the EoC ($i = \alpha$) and the EoH ($i = \lambda$). Rather than computing them numerically, as is done in Dalal et al. (2010) and Muñoz (2019a), here we use a parametrized form that provides a good fit across the entire range of interest. We write

$$W_i(k) = \frac{1}{1 + (k/k_{\text{cut},i})^{\beta_i}},$$

(29)

which has two free parameters ($\beta_i$ and $k_{\text{cut},i}$) for each era, fitted to the results of Muñoz (2019a) but kept independent of $z$ otherwise. We find that for Lyman $\alpha$ photons $\beta_\alpha = 2$ and $k_{\text{cut},\alpha} = 0.015 \text{ Mpc}^{-1}$ provide an excellent fit. For X-rays, the damping depends on the assumed energy cutoff scale (i.e. the minimum X-ray energy $E_0$ escaping the ISM of host galaxies). First, for $E_0 = 0.2 \text{ keV}$ (as set in our OPT parameters), a good fit is provided by $\beta_\lambda = 1$ and $k_{\text{cut},\lambda} = 0.3 \text{ Mpc}^{-1}$. Second, when setting $E_0 = 0.5 \text{ keV}$ (for the EOS fiducial) the X-rays have longer mean free paths, and we find $\beta_\lambda = 1.5$ and $k_{\text{cut},\lambda} = 0.04 \text{ Mpc}^{-1}$. We show the VAO power spectrum in Fig. 21 multiplied by each of these window functions, together with the numerical result from Muñoz (2019a), finding good agreement. This figure also shows that, as expected, longer distance transitions yield more suppression of VAO amplitudes. The Lyman $\alpha$ VAOs are more damped than X-rays with $E_0 = 0.5 \text{ keV}$, which in turn are more damped than X-rays with a lower energy cutoff at $E_0 = 0.2 \text{ keV}$. Numerically, at $z = 15$ the distance a Lyman $\alpha$ photon travels until entering the Lyman $\alpha$ resonance is roughly 300 Mpc comoving, whereas the mean-free path of X-rays is a significantly shorter 30 Mpc for $E_0 = 0.5 \text{ keV}$, or 3 Mpc for 0.2 keV (McQuinn 2012). As a consequence, these latter cases have higher $k_{\text{cut},i}$, and a shallower suppression index $\beta_i$. We emphasize that the parameters of this fit would technically vary with $z$, and have not been fit to high precision, but instead to round numbers, as that suffices for our purposes of studying the detectability of VAOs and their extraction from simulated power spectrum data.

The amplitude of the VAOs, parametrized in our analytical PS expression through the bias $b_{\text{v}_b}$, depends on the astrophysics driving the 21-cm signal at any given epoch. Star formation feedback from streaming velocities preferentially impacts smaller scales, and thus MCGs are more affected than ACGs (though see Section 2 for their impact on ACGs). As a consequence, $b_{\text{v}_b}$ is larger (i.e. with more prominent VAOs) for models in which the SFRD is driven by the smallest haloes. In Fig. 20, we assume the optimistic (OPT) parameters for the MCG SFR from Table 1, and thus VAOs are evident during both the EoH and EoC.

We fit for the bias parameter $b_{\text{v}_b}$ for each redshift independently over the $k$-ranges shown in white in Fig. 20. The results are shown in Fig. 22. The amplitude of VAOs has two peaks, corresponding to the EoC ($z \sim 20$ for our OPT parameters) and the EoH ($z \sim 15$), and all but vanishes in the transition between the two, as the effect of $v_b$ goes in opposite directions between these two eras, producing less coupling at earlier times (and thus higher $T_{21}$), and less heating at late times (lower $T_{21}$). We also show $b_{\text{v}_b}$ for our EOS fiducial (AllGalaxies) simulation in that Figure, which shows somewhat smaller VAOs, delayed to later times; as well as a null test $b_{\text{v}_b}$ for a simulation with no VAOs as an error-bar estimation of our fitting procedure. For comparison, the HERA Collaboration (2022) found that $b_{\text{v}_b} < [50, 180] \text{ mK}^2$ at 95 percent CL, using their phase-1 limits at $z = [8, 10]$. These data only cover lower redshifts, where VAOs are not expected to be important (c.f. Fig. 22), but they highlight the need for further sensitivity to reach the level of VAOs ($b_{\text{v}_b} \approx 10 \text{ mK}^2$) predicted in our models.

In the bottom panel of Fig. 22, we plot the $v_b$-only component of the 21-cm power spectrum, defined as $\Delta_{21,v_b}^2 = b_{\text{v}_b}^2 W^2(k) \Delta_{\text{v}_b}^2(k)$, at a scale $k = 0.15 \text{ Mpc}^{-1}$ (roughly corresponding to a ‘sweet spot’ in terms of foreground contamination and thermal noise for interferometers; e.g. Greig et al. 2020a; The HERA Collaboration 2022; Tingay et al. 2013; van Haarlem et al. 2013). This VAO power is relatively high during the EoH, reaching $\Delta_{21,v_b}^2 \approx 10 \text{ mK}^2$. During the EoC, however, it only has values $\Delta_{21,v_b}^2 \approx 0.5 \text{ mK}^2$, as the photon propagation in the latter strongly suppresses large-$k$ fluctuations. To illustrate this point, we also plot the
VAO-only power for \( k = 0.05 \, \text{Mpc}^{-1} \) (where the deepest LOFAR limits lie Mertens et al. 2020) in that Figure, which grows by nearly two orders of magnitude during the EoC (and one during the EoH), showing that reaching lower \( k \) by careful foreground cleaning is ideal for detecting acoustic wiggles in the high-\( z \) 21-cm signal. We predict a smaller VAO power across all of cosmic dawn than our previous work. In Muñoz (2019a) we had found \( \Delta_{21,\text{vb}}^2 \approx 50 \, \text{mK}^2 \) during the EoH, a factor of a few larger. Part of the reason is the inclusion of inhomogeneous LW feedback, which tends to suppress VAOs (Fialkov et al. 2013). The largest factor, however, is the new parametrization of the SFRD (see equations 1, 7). In Muñoz (2019a), we considered a mass-independent SHMR shared for PopII and PopIII stars (i.e. \( \alpha_i^\text{SHMR} \) = 0 for \( i = \text{II} \) and \( \text{III} \)), which produces a much faster evolution of CD and larger 21-cm fluctuations (see discussion in Section 3). With the more-realistic SHMR considered here both the overall 21-cm PS and the VAOs are smaller, so VAOs are still an \( \mathcal{O}(1) \) component of the large-scale 21-cm power spectrum in Fig. 20.

5.2 VAOs and the SHMR of PopIII hosts

So far we have shown VAOs in simulations with either our OPT and EOS parameter sets. Given our lack of knowledge about cosmic dawn, however, the first galaxies could have much different parameters than we expected. We now perform a brief exploratory study of how the amplitude of the VAOs can be used to learn about the astrophysics of cosmic dawn.

We focus on the slope of the SHMR for MCGs, parametrized with \( \alpha_i^{\text{SHMR}} \), which we showed in the previous section has a very modest impact on the redshift dependence of the PS (at fixed \( k \)) and the GS. Here, we study its impact on the VAO component of the PS, shown in Fig. 22 for the same values of \( \alpha_i^{\text{SHMR}} \) as in Fig. 16. We see that steeper SHMRs (larger \( \alpha_i^{\text{SHMR}} \)) suppress the VAO amplitude, especially at high \( z \). This is understandable since steeper SHMRs decrease the relative contribution of the smallest haloes to the SFRD, and these smallest haloes are the most sensitive to the streaming velocities. Interestingly, the impact of \( \alpha_i^{\text{SHMR}} \) on the VAO component of the power appears more noticeable than in the overall 21-cm power spectrum or global signal (c.f. Fig. 16). Therefore, the amplitude of VAOs can provide a cleaner view of the halo–galaxy connection of MCGs.

We also study variations of the amplitude \( A_{1,\text{LW}} \) of the LW feedback. We find that increasing or decreasing \( A_{1,\text{LW}} \) by a factor of up to 3 does not change the amplitude of the VAOs, only its \( z \) dependence. This is because the LW and \( v_{\text{vb}} \) feedback effects multiply coherently, so they are rather independent. Given that the 21-cm GS and the full PS amplitude are also insensitive to \( A_{1,\text{LW}} \) (see Fig. 17), the best avenue for studying this parameter may be further hydrodynamical simulations, instead of inferring it from 21-cm data.

5.3 Detectability

The VAOs that we study here have been shown to be a robust standard ruler during cosmic dawn, allowing 21-cm interferometers to measure the cosmic expansion rate at \( z \sim 10–20 \) (Muñoz 2019b). However, we would first need to detect them.

As opposed to Section 3, where we considered the entire 21-cm power spectrum, here we forecast SNRs for the \( \Delta_{21,\text{vb}}^2 \) component. For that we use the same noise as before, which includes the cosmic variance from the full \( \Delta_{21}^2 \) signal. This noise is shown in Fig. 22 along with the VAO-only power spectrum. We find a SNR = 5 for the EOS parameters assuming 1080 h of HERA data at moderate foregrounds. For the OPT parameters, instead, we find more optimistic estimates, with SNR = 9. In both cases, the SNR is only above unity over the range \( z \sim 10–15 \), showing that the EoH is the most promising epoch to detect VAOs, and thus to measure \( H(z) \) (Muñoz 2019b).

5.4 VAOs in the lightcone

All work on VAOs thus far has been on co-eval boxes (i.e. at fixed \( z \)). In reality, however, 21-cm fluctuations are measured in the light-cone, as the fluctuations along the line-of-sight (LoS) direction evolve with \( z \). This is particularly important for using VAOs as a standard ruler, as mainly LoS modes are observed by interferometers, which are then used to measure \( H(z) \).

It is expected that LoS effects slightly change the large-scale 21-cm power (Datta et al. 2012; La Plante et al. 2014; Ghara, Datta & Choudhury 2015; Greig & Mesinger 2018). In order to include these effects we divide our light-cone, which is 600 Mpc on a side and 3830 Mpc in length, into 10 blocks (each 383 Mpc along the LoS), and compute the power spectrum in each of them. We show the resulting full-light-cone 21-cm power spectra in Fig. 23, where the VAOs are still clearly apparent (with larger Poisson noise from the simulations as each \( z \) box corresponds to a smaller comoving volume). While a full comparison of light-cone and co-eval boxes is beyond the scope of this work, we find that during the EoH the bias \( b_{\text{vb}} \) is reduced by 20 per cent, showing a small but not negligible LoS effect on the amplitude of VAOs, though not on their shape. This is an important cross-check for using VAOs as a standard ruler. For a brief study of how to measure VAOs without Poisson variance from the simulations we encourage the reader to visit Appendix C.

6 DISCUSSION AND CONCLUSIONS

The first generation of PopIII stars heralded the transition from the dark ages to the cosmic dawn. In this work, we have improved the treatment of these stars in the public code 21cmFAST, including the combined impact on star formation from baryon-DM streaming velocities \( (v_{\text{vb}}) \) and the inhomogeneous LW background. In our model, PopIII stars are hosted in molecular-cooling galaxies (MCGs), with halo masses \( M_h < M_{\text{atom}} \sim 10^{7–8} \, \text{M}_\odot \) whereas PopII stars form in atomic-cooling galaxies (ACG, with masses above \( M_{\text{atom}} \)). Thus, PopIII stars dominate the photon budget in the early universe, and are expected to set the timing of the cosmic-dawn 21-cm signal at \( z \sim 10–20 \). Later on, however, feedback and the natural appearance of PopII stars may increase the signal strength and set the timing of the cosmic dawn.
of heavier haloes will make MCGs subdominant, and PopII stars are expected to drive cosmic reionization, at $z \sim 5$–10.

Many questions remain about the formation of the first galaxies. As such, we implement flexible models that account for two distinct (PopII and III) stellar populations with different SHMRs, building upon Park et al. (2019), Muñoz (2019a), and Qin et al. (2020a). We provide generic fitting formulae for the impact of LW and $v_{\text{th}}$ feedback (in Section 2), and calibrate them with results from state-of-the-art hydrodynamical simulations. This allows us to generate a new EOS 2021 model covering the evolution of the 21-cm signal across cosmic dawn and reionization, which encapsulates our current knowledge of these epochs. We have dubbed this model AllGalaxies, and it enhances the previously available Bright and FaintGalaxies (from EOS 2016) by including PopIII-hosting MCGs.

The parameters of this AllGalaxies model are chosen to give rise to late reionization (finishing at $z \approx 5.5$. as expected from recent Lyman $\alpha$ forest data), and its CMB optical depth $\tau_{\text{CMB}} \approx 0.063$, in line with recent determinations from Planck. The 21-cm signal in our AllGalaxies simulation is shallower (only reaching $T_2 \approx 75$ mK during cosmic dawn) than in the previous EOS models. This is due to both the inclusion of MCGs and a steeper SHMR for ACGs, the latter being required to match UVLF observations. As a consequence, both the expected 21-cm global signal and its fluctuations will be more difficult to detect (see also Mirocha et al. 2017; Park et al. 2019).

Nevertheless, we expect both the HERA and SKA interferometers to reach a 21-cm power spectrum detection of this model at an SNR $\approx 200$ with 1000 h of integration. We make public the detailed light-cones of this model, as well as associated visualizations.

We also performed an exploratory study of how PopIII stars affect the cosmological 21-cm signal. We have found that for the redshifts of interest for MCGs ($z \geq 12$), $v_{\text{th}}$ feedback likely dominates over LW feedback (cf. Fig. 5). As a consequence, the amplitude $A_{\text{LW}}$ of LW feedback only has a modest impact on the 21-cm signal. Similarly, the signal is not sensitive to the escape fraction $f_{\text{esc}}^{(\text{III})}$ of ionizing photons from MCGs, as MCGs do not significantly contribute to reionization in our model. On the other hand, the star formation efficiency and the X-ray luminosities of MCGs do impact the 21-cm signal significantly.

We found that for the redshifts of interest for MCGs ($z \geq 12$), $v_{\text{th}}$ feedback likely dominates over LW feedback (cf. Fig. 5). As a consequence, the amplitude $A_{\text{LW}}$ of LW feedback only has a modest impact on the 21-cm signal. Similarly, the signal is not sensitive to the escape fraction $f_{\text{esc}}^{(\text{III})}$ of ionizing photons from MCGs, as MCGs do not significantly contribute to reionization in our model. On the other hand, the star formation efficiency and the X-ray luminosities of MCGs do impact the 21-cm signal significantly.

We also performed an exploratory study of how PopIII stars affect the cosmological 21-cm signal. We have found that for the redshifts of interest for MCGs ($z \geq 12$), $v_{\text{th}}$ feedback likely dominates over LW feedback (cf. Fig. 5). As a consequence, the amplitude $A_{\text{LW}}$ of LW feedback only has a modest impact on the 21-cm signal. Similarly, the signal is not sensitive to the escape fraction $f_{\text{esc}}^{(\text{III})}$ of ionizing photons from MCGs, as MCGs do not significantly contribute to reionization in our model. On the other hand, the star formation efficiency and the X-ray luminosities of MCGs do impact the 21-cm signal significantly.

The streaming velocities $v_{\text{th}}$ fluctuate spatially with an acoustic signature inherited from recombination. As a consequence, the distribution of the first galaxies (and thus the 21-cm signal) shows velocity-induced acoustic oscillations (VAOs); large wiggles in their power spectrum at $k \approx 0.1$ Mpc$^{-1}$. We showed that VAOs are present and detectable even when including inhomogeneous LW feedback, and considering light-cone effects. This will allow us to use the 21-cm power spectrum as a standard rater during cosmic dawn. Moreover, the amplitude of the 21-cm VAO oscillations can be used to study the SHMR of MCGs. The slope of the SHMR can provide insight into the stellar content of MCGs and the associated feedback mechanisms; without VAOs, this important quantity would be very difficult to detect.

Our results and public simulations can be used to guide 21-cm observing strategies and data pipelines. With our current state of knowledge we expect the 21-cm power spectrum to be detected at high significance by upcoming interferometric observations. Such a detection will provide us with a new window on the stellar and energy content of our cosmos at unprecedented early times.

ACKNOWLEDGEMENTS

We are grateful to Greg Bryan, Daniel Eisenstein, and Jordan Mirocha for enlightening discussions. JBM is supported by a Clay fellowship at the Smithsonian Astrophysical Observatory. AM acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 638809 – AIDA). The results presented here reflect the authors’ views; the ERC is not responsible for their use. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. We gratefully acknowledge computational resources of the Center for High Performance Computing (CHPC) at Scuola Normale Superiore (SNS). This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. This work was partially supported by a grant from the Simons Foundation. CAM acknowledges support by the VILLUM FONDEN under grant 37459 and from NASA Headquarters through the NASA Hubble Fellowship grant HST-HF2-51413.001. A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555.

DATA AVAILABILITY

All the simulations presented in this work are freely available for download (see footnote 2). We make light-cones from the AllGalaxies simulation (initial conditions, perturbed densities, relative velocities, spin and kinetic temperatures, ionization fractions, and 21-cm brightness temperatures) publicly available. In addition to the AllGalaxies simulation (1.5 Gpc in size), we share the output of the simulations used in Section 3 (which are either 600 or 400 Mpc in size, at the same resolution). We compile the names of the simulations, and their parameters, in Table 4.
The first galaxies during CD and the EoR 3679

Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1
The HERA Collaboration, 2022, ApJ, 924, 51
Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608
Tingay S. J. et al., 2013, Publ. Astron. Soc. Aust., 30, e007
Trenti M., 2010, AIP Conf. Proc., 1294, 134
Tseliakhovich D., Hirata C., 2010, Phys. Rev., D82, 083520
Tseliakhovich D., Barkana R., Hirata C. M., 2011, MNRAS, 418, 906
van Haarlem M. P. et al., 2013, A&A, 556, A2
Visbal E., Barkana R., Fialkov A., Tseliakhovich D., Hirata C. M., 2012, Nature, 487, 70
Visbal E., Haiman Z., Terrazas B., Bryan G. L., Barkana R., 2014, MNRAS, 445, 107
Voytek T. C., Natarajan A., Jáuregui García J. M., Peterson J. B., López-Cruz O., 2014, ApJ, 782, L9
Wang F. et al., 2020, ApJ, 896, 23
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Whiter L. R., Mason C. A., Ren K., Dijkstra M., Mesinger A., Pentericci L., Trenti M., Treu T., 2020, MNRAS, 495, 3602
Wilkins S. M., Lovell C. C., Stanway E. R., 2019, MNRAS, 490, 5359
Wise J. H., Abel T., 2008, ApJ, 685, 40
Wouthuysen S. A., 1952, AJ, 57, 31
Wu X., McQuinn M., Eisenstein D., Iiričić V., 2021, MNRAS, 508, 2784
Wyithe J. S. B., Loeb A., 2013, MNRAS, 428, 2741
Xu H., Norman M. L., O’Shea B. W., Wise J. H., 2016a, ApJ, 823, 140
Xu H., Wise J. H., Norman M. L., Ahn K., O’Shea B. W., 2016b, ApJ, 833, 84
Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2019, MNRAS, 483, 2983

APPENDIX A: STREAMING-VELOCITY SUPPRESSION ON THE MATTER POWER SPECTRUM

In this appendix we show a simple but accurate fit to the suppression on the small-scale matter power spectrum induced by the relative velocities.

Following Tseliakhovich & Hirata (2010) and Muñoz (2019a), we solve for the evolution of the DM and baryon densities for different initial values of $v_{rb}$ and its inner-product cosine $μ$ with each wavenumber $k$, in order to compute the density fluctuations of baryons and DM at each $z$ and $k$. From there, we can calculate the matter power spectrum as a function of $v_{rb}$ by averaging over $μ$, to find $P_m(k, z; v_{rb})$ (Tseliakhovich & Hirata 2010). We show in Fig. A1 the ratio of this quantity to its no-velocity counterpart

$$T_m = \frac{P_m(k, z; v_{rb})}{P_m(k, z; 0)} \quad (A1)$$

as well as the fit defined in equation (14). The main feature of these curves is that larger values of $v_{rb}$ produce a bigger drop in the power spectrum, as the larger relative velocity puts DM and baryons out of phase (or equivalently, it does not allow baryons to collapse to DM overdensities, slowing growth at small scales). However, at much smaller scales the baryonic fluctuations are damped due to the Jeans pressure, driving $δ_b$ down for $k > k_J \approx 500$ Mpc$^{-1}$, and recovering the DM-only fluctuations regardless of $v_{rb}$. Put together, these two effect produce a dip at $k \approx 300$ Mpc$^{-1}$ in Fig. A1, which turns back to unity at larger $k$. We have used equation (14) to fit this dip as a Gaussian with three parameters (as described in Section 2 of the main text), where

$$T_m(v_{rb}) = 1 - A_v \exp \left[ -\frac{(\log[k/k_J])^2}{2\sigma_v^2} \right], \quad (A2)$$

$\sigma_v$ and $A_v$ are presented in Table A1.
and we find the following fits

\begin{align}
A_p(v_{cb}, z) &= 0.24 \times \left(1 + z\right)^{1/8} \left(v_{cb}/v_{rms}\right)^{3/2}, \\
B_p(v_{cb}) &= 300 \, \text{Mpc}^{-1} \exp\left[-0.2 \left(v_{cb}/v_{rms} - 1\right)\right], \\
\sigma_p(v_{cb}) &= 0.8 \left(v_{cb}/v_{rms}\right)^{1/3},
\end{align}

(A3)

for the Planck 2018 cosmology, where \(A_p = 1\) for \(v_{cb} \leq v_{rms}\) and \(A_p = 0.5\) otherwise. This heuristic fit was calibrated at \(z = 20\), and while neither \(B_p\) nor \(\sigma_p\) depend on \(z\), see it provides a good fit during the relevant \(z\) at cosmic dawn. To illustrate this point, we show the real transfer function \(T_m(v_{cb})\) in Fig. A1 at \(z = 15\) and 30, as well as our fit, where the two are in good agreement.

**APPENDIX B: FAST SFRD TABLES**

The excursion-set algorithm in 21-cmFAST takes advantage of the extended Press–Schechter approach (Bond et al. 1991) to find the SFRD (and its derived quantities such as the ionizing flux) at a given scale and overdensity. This requires tabulating the SFRD for many values of the overdensity \(\delta_R\) and variance \(\sigma_R\) for different radii \(R\), so as to not compute it in every cell at every \(z\). This is not extremely computationally expensive for runs with only atomic-cooling galaxies (AGCs, taking approximately \(\sim 2\) min to generate the tables down to \(z = 6\)), though it is for runs with minihaloes hosting molecular-cooling galaxies (MCGs, taking \(\sim 4\) h for the same settings), as the tables ought to be generated for many different values of \(M_{\text{min}}\).

Here, we show an analytical approximation that allows us to compute those tables much faster, cutting the generation time of tables by a factor of \(\approx 30\).

In 21-cmFAST the SFRD (and derived quantities such as UV and X-ray emissivities) is modulated using extended PS theory, where we have

\[
\text{SFRD}^{\text{PS}}(R, x) \propto \int dM_h \frac{dn}{dM_h}(x) M_\star(M_h) f_{\text{dark}}(M_h),
\]

(B1)

for each cell at \(x\) and radius \(R\) around it (which will enter an integral over previous times). The HMF depends on position through the overdensity \(\delta_R(x)\) and the variance \(\sigma_R\), and the pre-factor is

![Figure A1. Transfer ratio of matter fluctuations at small scales due to the effect of the DM-baryon relative velocities \(v_{cb}\), defined in equation (A1). The solid lines show the exact result from solving the ODEs (as in Muñoz 2019a), whereas the dashed lines show our Gaussian approximation from equation (A3). The red and green lines are evaluated at \(z = 30\), whereas the orange and purple are at \(z = 15\). Upper (red and orange) lines have \(v_{cb} = v_{rms}\), whereas lower (purple and green) ones have \(v_{cb} = 2 \times v_{rms}\). In the main text we use the \(z = 20\) result, described in equation (14).](image1)

![Figure A2. Top: Relation between the halo mass \(M_h\) and the rms fluctuation \(\sigma(M_h)\) at their scales. The red line is the result using the full power spectrum, whereas the black dashed line is the broken power law from equation (B7). Note that we need to go to low \(M\) values (large \(\sigma\)) since in some cells \(\delta_R \approx \delta_{\text{crit}}\), which corresponds to \(v \to 0\) (or \(\sigma \to 1\)). All \(\sigma\) are linearly extrapolated to \(z = 0\). Bottom: Light-cone power spectra with the FAST_PCOLL_TABLES turned on (points with errors) and off (solid line) at three redshifts, for our OPT parameters. The difference between the two is negligible at high redshifts, and below 10 per cent at low redshifts. The global signal is identical to the percent level between the two cases.](image2)

![Figure B1. Difference between the full relative-velocity case and one with fixed \(v_{cb} = v_{\text{avg}}\) (FTX_VAVG in the code), in solid blue at \(z = 15\) and red at \(z = 20\). These two cases have a similar background, but the VAOs are absent on the average case, as the velocities do not fluctuate. The dashed lines correspond to the fit from equation (27), with the biases and window functions from that section. We do not show error-bars on the simulation results here as it is the difference of two runs with the same initial conditions.](image3)
position-independent so it cancels out when dividing by the average over all cells. Our assumption is that feedback produces a power-law SHMR, equations (1), (7), so
\[ M_*(M_h) \propto M_h \times (M_h / 10^{10} M_\odot)^\alpha \]  
for a power-law index \( \alpha \) (which we can identify with \( \alpha_\sigma \) for X-ray and non-ionizing UV fluxes), but for ionizations is \( \alpha = \alpha_\sigma + \alpha_{\text{esc}} \).
In these calculations the HMF \( dn/dM \) is given by the standard old PS formula, as that is the only one for which the EPS formalism has been properly tested.

We will speed up the computation by using an approximate analytical solution. We know how to compute the integral in equation (B1) analytically for the case of \( \alpha = 0 \) (so \( M_*(M_h) \propto M_h \)) with a sharp cutoff (i.e. a Heaviside rather than exponential functional form for \( f_{\text{iay}}(M_h) \)). In that case we can use the variable \( v = \delta_{\text{crit}}/\sigma^2 \) to re-write the integral as
\[ \text{SFRD}^{\text{EPS}}(x, \sigma) \propto \int_{v_{\text{min}}}^{\infty} \frac{dv}{\sqrt{v}} e^{-v/2} \sqrt{\frac{v_{\text{min}}}{2}} \text{erfc} \left( \sqrt{\frac{v_{\text{min}}}{2}} \right) , \]  
where erfc is the complementary error function, and we have defined
\[ v_{\text{min}} = \delta_{\text{cut}}^2/\sigma(M_{\text{mum}})^2 . \]  
The EPS formalism (Bond et al. 1991) shows that for regions of radius \( R \) that are overdense by \( \delta_R \), given a variance of matter fluctuations of \( \sigma_R^2 \), then all the \( v \) are corrected to be \( \tilde{v} \), defined as
\[ \tilde{v} = \delta_{\text{cut}}/\sigma^2 . \]  
with \( \delta_{\text{cut}} = \delta_{\text{cut}} - \delta_R \), and \( \sigma^2 = \sigma_R^2 - \sigma_R^2 \). This allows us to compute the SFRD at an arbitrary point through a simple erfc, as in equation (B3). This expression is exact, and extremely fast to evaluate (compared to performing many numerical integrals). So, while it cannot be used directly in our tables (as we do not have a Heaviside \( f_{\text{iay}} \) or \( \alpha = 0 \) always), it gives us a good scaffolding upon which to build our analytical solution.

First, we will use a Heaviside \( f_{\text{iay}} \) in our EPS, though we keep the usual (exponential) duty factor in the average result that normalizes at each \( z \).

Secondly, we can use the following analytic integral,
\[ \int_{v_{\text{min}}}^{\infty} \frac{dv}{\sqrt{v}} e^{-v/2} v^\beta = 2^{\beta+1/2} \Gamma(\beta + 1/2, v_{\text{min}}/2) , \]  
which is a generalization of equation (B3) adding a power-law \( v^\beta \). Then, the trick is to approximate the function
\[ \sigma(M_h)/\sigma_p = (M_h/M_p)^{1/\gamma} , \]  
for some power-law index \( \gamma \) and pivot scale \( M_p \) (for which \( \sigma(M_p) \equiv \sigma_p \)). In practice, this can only be done over a small range of variances \( \sigma \) (or halo masses \( M_h \)), as otherwise the functional form does not work. For LCDM models, and the mass range of interest, a good fit is a broken power law, with \( \gamma = 9 \) for masses below a first pivot (\( M_h < M_{p1} = 5.3 \times 10^9 M_\odot \)), \( \gamma = 21 \) for masses above the second pivot (\( M_h > M_{p2} = 1.5 \times 10^9 M_\odot \)), and where the index between the two pivots is set to \( \gamma = 13.6 \) by continuity. This approximates \( \sigma(M_h) \) as a broken power law of \( M_h \) to \( \sim 10 \) per cent precision, as shown in Fig. A2 which suffices for our purposes. For each of the power-law indices, we take \( \beta = \alpha / \gamma / 2 \) in equation (B6).

A subtle point is that the regions of radius \( R \) at which \( \sigma \) is used for \( \sigma(R) \) with \( \alpha \) an arbitrary value. We approximates \( \sigma(R) \) for the second pivot (\( R > R_{p2} \)) and \( \sigma(R) \) for the first pivot (\( R > R_{p1} \)). The result of this approximation is to break \( (1 + v/v_p) - \beta \) into the two regimes of 1 and \( (v/v_p)^\beta \) for \( v < v_p \) and \( v > v_p \), respectively. Therefore, for \( v \geq v_p \) the power laws cancel (recovering an erfc), whereas below \( v_p \) follows equation (B6). This allows us to write the collapsed fraction through this broken power law as a sum of \( \Gamma \) functions.

The code allows the user to turn on this functionality with the flag FAST_FCOLL_TABLES. By default they are not used for MCGs, as the speed boost is only \( \sim 2 \), but for MCGs it can reach \( \times 30 \), so it is recommended. The calculation for the ionization tables proceeds identically, with \( \alpha = \alpha_\sigma + \alpha_{\text{esc}} \). We show an example of how the SFRD compares to the exact calculation in Fig. A2, which for the MCG regime of interest agrees to \( \sim 10 \) per cent, and \( \approx 5 \) per cent for \( \alpha = 0 \), which is the MCG-assumed value. The agreement is even closer in practice, since we only use these tables for the EPS part of the calculation, so an overall offset is cancelled out, and only the \( \beta \) behaviour remains. We note that we are not explicitly enforcing \( f_\sigma \) or \( f_{\text{esc}} \) to be less than unity in these approximations. Although models that violate that condition are disfavoured by observations, we caution the reader to turn off the FAST_FCOLL_TABLES when exploring a wide parameter space of extreme models.

**APPENDIX C: VAOS WITHOUT POISSON NOISE**

Here, we show a simple check that the VAOs can be recovered from our simulations with small Poisson variance. Our goal is to show that one need not run very large 21-cm boxes to obtain the large-scale VAOs, if two simulations are compared: one with full relative velocities and one with a fixed \( v_{\text{rb}} = v_{\text{av}} \). These two cases share a similar background evolution but have different large-scale power spectrum, as the \( v_{\text{rb}} \) anisotropies become imprinted on to the 21-cm fluctuations. We show the difference in the power spectra during the EoH and EoC in Fig. B1, as well as the fit we found in Section 5, which provides a reasonable approximation to this difference. Notice that the high-\( k \) part of the power spectrum is still deviates in the two cases, as the background matching is not perfect for the average-velocity case. Subtracting these two cases severely reduces the simulation Poisson error (as the two simulations share initial conditions), though it does not fully cancel it, as clear from the data variation in Fig. B1. This helps extracting VAOs from more modestly sized simulations.

This paper has been typeset from \( \TeX \)/\( \LaTeX \) file prepared by the author.