University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Mutch, SJ;Greig, B;Qin, Y;Poole, GB;Wyithe, JSB

Title:
Dark-ages reionization and galaxy formation simulation - XXI. Constraining the evolution of the ionizing escape fraction

Date:
2024-01-01

Citation:
Mutch, S. J., Greig, B., Qin, Y., Poole, G. B. & Wyithe, J. S. B. (2024). Dark-ages reionization and galaxy formation simulation - XXI. Constraining the evolution of the ionizing escape fraction. Monthly Notices of the Royal Astronomical Society, 527 (3), pp.7924-7939. https://doi.org/10.1093/mnras/stad3746.

Persistent Link:
https://hdl.handle.net/11343/345366

License:
cc-by

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **527,** 7924–7939 (2024) https://doi.org/10.1093/mnras/stad3746
Advance Access publication 2023 December 6

Dark-ages reionization and galaxy formation simulation – XXI. Constraining the evolution of the ionizing escape fraction

Simon J. Mutch [1,2,3★] Bradley Greig [1,2] Yuxiang Qin [1,2] Gregory B. Poole[1,4] and J. Stuart B. Wyithe[1,2]

[1]*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*
[2]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*
[3]*Melbourne Data Analytics Platform, The University of Melbourne, Parkville, VIC 3010, Australia*
[4]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*

Accepted 2023 November 28. Received 2023 October 17; in original form 2023 March 13

ABSTRACT

The fraction of ionizing photons that escape their host galaxies to ionize hydrogen in the intergalactic medium (IGM) is a critical parameter in analyses of the reionization era. In this paper, we use the MERAXES semi-analytic galaxy formation model to infer the mean ionizing photon escape fraction and its dependence on galaxy properties through joint modelling of the observed high redshift galaxy population and existing constraints on the reionization history. Using a Bayesian framework, and under the assumption that escape fraction is primarily related to halo mass, we find that the joint constraints of the ultraviolet luminosity function, cosmic microwave background optical depth, and the Ly α forest require an escape fraction of (18 ± 5) per cent for galaxies within haloes of $M \lesssim 10^9\,\mathrm{M}_\odot$ and (5 ± 2) per cent for more massive haloes. In terms of galaxy properties, this transition in escape fraction occurs at stellar masses of $M_\star \sim 10^7\,\mathrm{M}_\odot$, nearly independent of redshift. As a function of redshift, reionization is dominated by the smaller $M_\star \lesssim 10^7\,\mathrm{M}_\odot$ galaxies with high escape fractions at $z \gtrsim 6$ and by the larger $M_\star \gtrsim 10^7\,\mathrm{M}_\odot$ galaxies with lower escape fractions at $z \lesssim 6$. Galaxies with star formation rates of $10^{-2.5}\,\mathrm{M}_\odot\mathrm{yr}^{-1}$ to $10^{-1.5}\,\mathrm{M}_\odot\mathrm{yr}^{-1}$ provide the dominant source of ionizing photons throughout reionization. Our results are consistent with recent direct measurements of a ~ 5 per cent escape fraction from massive galaxies at the end of reionization and support the picture of low mass galaxies being the dominant sources of ionizing photons during reionization.

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

1 INTRODUCTION

The Epoch of Reionization, during which the progenitors of modern day galaxies formed and reionized hydrogen throughout the intergalactic medium (IGM), is the least explored cosmic epoch. However, over the past two decades there has been significant progress in exploring the stellar growth history of these galaxies (e.g. Finkelstein et al. 2015; Bouwens et al. 2017), with a recent step-change in understanding that has arisen through the *JWST* (e.g. Atek et al. 2022; Naidu et al. 2022). At the same time, observations of the evolution of the ionization state of the IGM has been narrowed towards a scenario where reionization completes at, or just below, $z \sim 6$ (e.g. Qin et al. 2021; Bosman et al. 2022). This constraint has been achieved through a combination of the integral measure of the reionization history obtained from cosmic microwave background (CMB) measurements (e.g. Planck Collaboration 2016b), and measurements of absorption signals against sources emitting Ly α photons, including Ly α forest optical depth, dark gaps, and quasar Ly α emission line damping wings (e.g. Becker & Bolton 2013; McGreer, Mesinger & D'Odorico 2015; Bañados et al. 2018; Greig et al. 2022).

Star-forming galaxies are thought to provide the dominant source of ionizing photons to reionize the Universe (e.g. Srbinovsky & Wyithe 2007; Bouwens et al. 2015; Qin et al. 2017). However, we are unable to directly relate the observed star-forming galaxy population to the evolution in the ionization state of the IGM due to the unknown fraction of ionizing photons which escape from the ISM of galaxies. Given the evolution of galaxy density with redshift, and the expected dependence of star formation rate with host-halo mass, we expect that there is likely a dependence of escape fraction on these quantities (e.g. Gnedin, Kravtsov & Chen 2008; Wise & Cen 2009). For example in a scenario where the dominant photon sources are low-mass haloes, reionization starts at higher redshift (e.g. Finkelstein et al. 2019). However, if galaxies in more massive haloes are the more dominant sources, then reionization will be delayed and conclude more rapidly (e.g. Mesinger, Greig & Sobacchi 2016; Naidu et al. 2020). Moreover, since this escape of radiation is likely through channels of low-column density H I and dust, the escape fraction is expected to be direction dependent for individual galaxies and to have a large scatter with galaxy properties (e.g. Cen & Kimm 2015; Yeh et al. 2023).

★ E-mail: smutch@unimelb.edu.au

© 2023 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Given the importance of the escape fraction to our understanding of reionization, there have been a number of studies to (i) measure the escape fraction directly in individual galaxies (e.g. Bassett et al. 2022; Begley et al. 2022; Pahl et al. 2022), (ii) to directly simulate the ISM of star-forming galaxies in a cosmological context (e.g. Wise & Cen 2009; Yeh et al. 2023), and (iii) to infer the average escape fraction properties by comparing observed star formation rates with the reionization history (e.g. Wyithe et al. 2010; Finkelstein et al. 2019; Yung et al. 2020).

In this paper, we follow the third approach, using a semi-analytic model coupled with seminumerical calculations of reionization, which offers a computationally efficient route to calculating the connection between galaxy formation and reionization (e.g. Mutch et al. 2016; Seiler et al. 2019; Visbal, Bryan & Haiman 2020; Hutter et al. 2021). We utilize the MERAXES semi-analytic model (Mutch et al. 2016) to provide a realistic description of the galaxy population at $z \sim 4 - 8$, which can be extended to both redshifts higher than observed and to luminosities fainter than observed based on physical descriptions of galaxy formation physics. MERAXES couples this galaxy evolution to the evolution of the ionization state of the IGM with an assumed escape fraction. Our approach is to provide a flexible relationship between galaxy properties and the escape fraction, and then to constrain the parameters describing this relationship via a Bayesian framework. Efficient implementation of the MERAXES semi-analytic model makes this parameter space exploration possible, even though the model includes on the fly calculation of reionization and is directly run on a dark matter halo population close to the hydrogen cooling limit within a large *N*-body simulation.

The paper outline is as follows. In Section 2, we outline the relevant features of MERAXES. We then describe our implementation of the statistical sampling of the parameter space for MERAXES in Section 3 before constraining our galaxy formation model parameters in Section 4. In Section 5, we describe the reionization observations we employ before presenting our escape fraction model and the resulting constraints in Section 6. We finally provide our discussion and conclusions in Sections 7 and 8, respectively. Throughout this paper, we adopt a standard lambda cold dark matter (ΛCDM) cosmology with $\Omega_{\rm m} = 0.308$, $\Omega_\Lambda = 0.692$, and $H_0 = 67.8$ km s^{-1} Mpc^{-1}.

2 MERAXES

The MERAXES semi-analytic model (Mutch et al. 2016) is designed to explore galaxy formation across cosmic time, with a particular focus on the Epoch of Reionization and the first galaxies. Some of its key features include a delayed supernova feedback scheme which accounts for the finite lifetime of massive stars and the use of 'horizontal' merger trees whereby the entire simulation volume is processed one snapshot at a time (necessary for an efficient treatment of reionization). However, the defining feature of MERAXES is its close coupling to a modified version of the seminumerical reionization code, 21CMFAST (Mesinger, Furlanetto & Cen 2011). This allows coupled modelling of reionization with the growth of the galaxy population, both temporally and spatially. Taking into account both local and external ionizing radiation, we can explore how subsequent star formation is affected and how this feeds back on to the progression of reionization.

As with many other semi-analytic models, MERAXES includes treatments for star formation, supernova feedback, metal enrichment, cooling, active galactic nucleus (AGN) feedback, and galaxy mergers. In this section, we provide a broad overview of the model and any physical prescriptions that are directly relevant for this work. A full description of the model is provided in Mutch et al. (2016) and the subsequent papers referenced below.

2.1 Tiamat

Semi-analytic galaxy formation models work by post-processing dark matter halo merger trees either extracted from cosmological *N*-body simulations or created using Extended Press–Schechter theory. In order to fully model reionization, where the spatial distribution of galaxies directly influences their evolution, the former is required.

In this work, we utilize the Tiamat *N*-body simulation, run as part of the Dark-ages Reionization And Galaxy Observables from Numerical Simulations programme (Poole et al. 2016). Tiamat consists of 2160^3 collisionless particles following a Planck 2015 cosmology (Planck Collaboration 2016a) in a comoving volume of $(100\,{\rm Mpc})^3$, resulting in a particle mass of $3.89 \times 10^6\,{\rm M}_\odot$. This allows us to identify the majority of atomic-cooling mass haloes across all redshifts during reionization within a large enough volume to provide a statistically representative realization of the large-scale process. Tiamat has been run down to a redshift of $z = 1.8$ with 101 snapshots output between $z = 49$ and 5, evenly spaced in time with a cadence of ~ 11 Myr, and a further 64 snapshots at $z < 5$, logarithmically spaced in redshift. This fine snapshot cadence at high-redshift allows us to better track the progression of global reionization, which proceeds rapidly during its mid-phases. It also allows us to more accurately model the stochastic nature of star formation and its impacts on the reionization process at these high-redshifts (Mutch et al. 2016). For more details on the Tiamat simulation, interested readers are referred to Poole et al. (2016) and Angel et al. (2016). Details on the halo merger tree construction can be found in Poole et al. (2017).

2.2 Star formation and supernova feedback

The star formation and supernova feedback prescriptions implemented in MERAXES closely follow those of Croton et al. (2006), with updates presented in Guo et al. (2011).

At each time-step in the simulation, we calculate the amount of gas required to be accreted by each host dark matter halo in order to maintain the required baryon fraction in the presence of the intergalactic ionization field. This gas is assumed to be shock-heated to the virial temperature of the halo before cooling down to the central regions to form a rotationally supported exponential disc. Assuming angular momentum conservation of the gas as it cools, the scale radius of this disc is given by $r_{\rm disc} = R_{\rm vir}\lambda/\sqrt{2}$, where λ is the dimensionless spin parameter of the halo (Bullock et al. 2001).

Following Kennicutt (1989), we assume that star formation in the disc only occurs once the cold gas density reaches a critical threshold, $\Sigma_{\rm crit}$, which can be linked to the instantaneous properties of the host dark matter halo. The resulting star formation rate, $\dot{m}_*$, is calculated by assuming that some fraction, $\alpha_{\rm SF}$, of the cold gas, $m_{\rm cold}$, above this threshold is then turned into stars per disc dynamical time, $t_{\rm dyn}$:

$$\dot{m}_* = \frac{\alpha_{\rm SF}}{t_{\rm dyn}}\left(m_{\rm cold} - 2\pi\,\Sigma_{\rm crit} r_{\rm disc}^2\right), \quad (1)$$

where the second term on the right-hand side corresponds to the cold gas mass resulting in the critical surface density for a disc with scale radius $r_{\rm disc}$.

Galaxy mergers can also drive shocks and turbulence, resulting in enhanced star formation episodes. We model the fraction of cold gas converted into stars in such events following Somerville, Primack &

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Faber (2001):

$$\frac{m_{\rm burst}}{m_{\rm cold}} = \alpha_{\rm burst}\left(\frac{m_{\rm gal}}{m_{\rm parent}}\right)^{\gamma_{\rm burst}}, \quad (2)$$

where $m_{\rm burst}$ is the mass of newly formed stars, $m_{\rm gal}$ and $m_{\rm parent}$ are the total baryonic masses of the infalling and parent galaxies respectively, and $\gamma_{\rm burst}$ and $\alpha_{\rm burst}$ are free parameters. Following Croton et al. (2006), we fix $\gamma_{\rm burst} = 0.7$ and $\alpha_{\rm burst} = 0.57$.

Energy injection from supernovæ is thought to be the dominant feedback mechanism for controlling the growth of stellar mass (and hence the production rate of ionizing photons) at the high-redshifts and low-halo masses relevant for reionization (see e.g. Qin et al. 2017). We assume that this feedback mechanism is dominated by massive, short-lived ($\lesssim 40$ Myr) stars, ending in energetic Type-II supernovæ. The total supernova energy injected into the ISM, $\Delta E_{\rm total}$, after a time, t, from past star formation episodes of mass δm_* is given by:

$$\Delta E_{\rm total} = \epsilon \int_0^t \frac{\delta m_*}{{\rm d}t'} {\rm d}\xi(t'){\rm d}t', \quad (3)$$

where $\mathrm{d}\xi$ is the metallicity dependent energy injection from stars of age $t - t'$ which we calculate directly from the stellar population synthesis code STARBURST99 (Leitherer et al. 1999), assuming a Kroupa (2001) initial mass function (IMF). ϵ is a free parameter which controls the efficiency with which the injected energy couples to the surrounding ISM. The amount of gas assumed to be directly heated by this energy injection is modelled as

$$\Delta m_{\rm total} = \eta \delta m_*, \quad (4)$$

where η is a free parameter commonly referred to as the mass-loading factor.

Following Qiu et al. (2019), we assume ϵ and η to take the following functional forms:

$$\epsilon = \epsilon_0 \left(\frac{1+z}{4}\right)^{\alpha_{\rm eject}} \left(\frac{V_{\rm max}}{60\ {\rm km\ s^{-1}}}\right)^{\beta} \text{and} \quad (5)$$

$$\eta = \eta_0 \left(\frac{1+z}{4}\right)^{\alpha_{\rm reheat}} \left(\frac{V_{\rm max}}{60\ {\rm km\ s^{-1}}}\right)^{\beta}, \quad (6)$$

where $V_{\rm max}$ is the maximum circular velocity of the host halo, $\beta = -3.2\,(-1)$ for $V_{\rm max} < (\geq) 60\ {\rm km\ s^{-1}}$ (following Muratov et al. 2015) and $\alpha_{\rm reheat} = 2$ (following Cora et al. 2018). We also assume $\alpha_{\rm eject} = 0$, as suggested by Qiu et al. (2019).

The energy injected into the ISM is used to adiabatically heat the gas in the galaxy to the virial temperature of the host dark matter halo where it joins a hot gas reservoir. The mass of gas added to the hot reservoir, $\Delta m_{\rm hot}$, is given by

$$\Delta m_{\rm hot} = \frac{2\Delta E_{\rm total}}{V_{\rm vir}^2} \Delta m_{\rm total}. \quad (7)$$

If there is more energy available than is required to reheat the full $\Delta m_{\rm total}$ worth of gas to the virial temperature (i.e. $\Delta m_{\rm hot} > \Delta m_{\rm total}$) then the remainder of the energy is used to eject the remaining mass from the system entirely. For more details on the supernova feedback prescription, see Qiu et al. (2019).

2.3 Reionization

One of the key features of MERAXES is its treatment of reionization, which is modelled using a modified version of the seminumerical reionization code 21CMFAST (Mesinger, Furlanetto & Cen 2011). At each time-step in the simulation, Cartesian grids of halo mass, stellar mass, star formation rate, and total matter density are constructed. For this work, we employ a grid resolution of 256^3, corresponding to a side length of 0.39 Mpc. These are then filtered using an excursion set formalism to identify ionized cells where the number of ionizing photons is greater than the number of neutral atoms and recombination events. As discussed in Section 1, the key free parameter controlling this process is the escape fraction of ionizing radiation from the source galaxies, $f_{\rm esc}$.

When a patch of the IGM is reionized by ultraviolet (UV) photons, the temperature of the gas is raised to $\sim 2 \times 10^4$ K. This acts to raise the local Jeans mass, reducing the amount of material accreted by low mass haloes. We model this process using a baryon fraction modifier:

$$\Delta m_{\rm baryon} = f_{\rm b} f_{\rm mod} M_{\rm vir} - m_{\rm baryon}, \quad (8)$$

where $\Delta m_{\rm baryon}$ is the freshly accreted baryon mass, $f_{\rm b} = \Omega_{\rm b}/\Omega_{\rm m}$ is the universal baryon fraction, $m_{\rm baryon}$ is the mass of baryons already present in the halo, and $f_{\rm mod}$ is the baryon fraction modifier. The value of $f_{\rm mod}$ is calculated following Sobacchi & Mesinger (2013) and depends on the mass of the halo, the strength of the local UV ionizing background (which is again strongly dependent on the ionizing escape fraction), and the time over which the halo has been exposed to this background. For further details of the reionization prescription, see Mutch et al. (2016).

2.4 UV luminosities

In order to generate UV luminosities for each galaxy at a time t, we sum the luminosities from each past star formation burst of age τ. These luminosities are calculated by constructing a full spectral energy distribution (SED) and applying a top-hat filter centred on 1600 Å, with a width of 100 Å:

$$L_{\rm UV}(t) = \sum_{\rm i=0}^{n} \int_{1550\text{Å}}^{1650\text{Å}} {\rm d}\lambda \int_0^t {\rm d}t' \times \int_{Z_{\rm min}}^{Z_{\rm max}} {\rm d}Z \psi\left(t' - \tau_{\rm i}, Z\right) S_\lambda(\tau_{\rm i}, Z) T_\lambda(\tau_{\rm i}), \quad (9)$$

where $\psi(t' - \tau_{\rm i}, Z)$ is the metallicity (Z) dependant star formation rate of a burst (i) of age $\tau_{\rm i}$, $S_\lambda(\tau_{\rm i}, Z)$ is the wavelength and metallicity dependent luminosity of a simple stellar population of age $\tau_{\rm i}$, and $T_\lambda(\tau_{\rm i})$ is the dust transmission function. We calculate $S_\lambda(\tau_{\rm i}, Z)$ using the publicly available STARBURST99 stellar population synthesis model (Leitherer et al. 1999), including nebular emission lines and assuming a Kroupa (2001) IMF with $Z_{\rm min} = 0.001$ and $Z_{\rm max} = 0.04$.

Qiu et al. (2019) explored a number of physically motivated parametrizations for the dust transmission, $T_\lambda(\tau)$, by constraining MERAXES to match both the evolution of the observed ultraviolet luminosity function (UVLF) and infrared excess (IRX)-β relation at $z \gtrsim 4$. For this work, we employ the gas column density dust model whereby the optical depth, Γ_λ, is given by

$$\Gamma_\lambda = {\rm e}^{-az} \left(\frac{m_{\rm cold}}{10^{10} h^{-1} {\rm M}_\odot}\right)^{\gamma_{\rm GCD}} \left(\frac{r_{\rm disc}}{h^{-1}{\rm kpc}}\right)^{-2} \left(\frac{\lambda}{1600\text{Å}}\right)^{n}, \quad (10)$$

where a, $\gamma_{\rm GCD}$ and n are free parameters and $r_{\rm disc}$ and $m_{\rm cold}$ are as introduced above. The resulting transmission is calculated using the two-phase model of Charlot & Fall (2000):

$$T_\lambda(t) = \begin{cases} \exp\left(-\left(\tau_{\rm ISM} - \tau_{\rm BC}\right)\Gamma_\lambda\right) & t < t_{\rm BC} \\ \exp\left(-\tau_{\rm ISM}\Gamma_\lambda\right) & t \geq t_{\rm BC} \end{cases}, \quad (11)$$

where $\tau_{\rm ISM}$ and $\tau_{\rm BC}$ are again free parameters. Here, photons emitted by young stars with ages less than $t_{\rm BC}$ must travel through both a dense birth cloud and the more extended interstellar medium (ISM).

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

After this time, the birth cloud is assumed to have been destroyed and the photons are only attenuated by the ISM. For more detail on both the SED construction and dust model see Qiu et al. (2019).

2.5 Unique features for parameter space exploration

One of the key benefits of semi-analytic models over more detailed hydrodynamic simulations is their relatively low computational expense. This opens up the potential for statistical exploration of the underlying free parameter space (e.g. Henriques et al. 2013; Qiu et al. 2019). MERAXES has been designed to maximize this potential by being written to be as efficient as possible without loss of flexibility. Some of the relevant unique features of our model include the pre-loading of input data, the utilization of graphics processing units (GPUs), and the suite of companion tools which have been developed to programmatically interact with the model and its output.

In order to provide a converged sampling of the underlying posterior probability distribution we require many tens-of-thousands of model evaluations for moderate dimensionality problems ($\sim$10 free parameters). In this situation, it is imperative that the model being evaluated can be run as quickly and efficiently as possible. For MERAXES (running on our input N-body simulation, Tiamat) at least $\sim$70 per cent of the time taken for a single model call is devoted to reading in the trees and dark matter density grids from disc, distributing them amongst all available processors, and processing their contents. Therefore, we have modified the model to be able to read in all input data once, store this in memory, and then to run multiple times without the need for any expensive input–output (IO) calls.

A further computational bottleneck for MERAXES is its treatment of reionization using a modified version of the 21CMFAST (Mesinger, Furlanetto & Cen 2011) seminumerical code. This currently requires the model to filter a number of 3D Cartesian grids using an excursion set formalism at every snapshot. Each of these excursion set calculations entails a large number of Fourier transforms and multiple loops over every cell (of which there are 256^3 for this work) in the gridded volume. To reduce the amount of time spent doing this expensive calculation, and thus allow our analysis to be completed in a timely fashion, the reionization portion of MERAXES has been rewritten to take advantage of the GPUs commonly available on the majority of modern high-performance computing systems. This has sped up the reionization calculation by a factor of 8 for typical combinations of grid dimensions and hardware resources.

3 PARAMETER SPACE SAMPLING

In order to carry out a statistical exploration of the MERAXES parameter space, we require a method to calculate the likelihood of the model predictions for a given set of parameters and to then provide the code with a new set of trial parameters for the next run. To facilitate this, as well as future parameter space exploration and calibration studies with MERAXES, we have developed a new companion code called MHYSA which provides a flexible two-way communication layer between the model (written in C) and PYTHON. MHYSA was designed to efficiently facilitate calling MERAXES many times, with different parameter values, and allows us to interface any PYTHON optimization or sampling package with the model. It also allows us to write all of our likelihood calculations in PYTHON and to leverage its many available scientific and numerical libraries, vastly reducing development time. In practice, MHYSA deals with initializing the model, updating its parameters, directly accessing in-memory galaxies and haloes, calling arbitrary PYTHON code at the end of each snapshot, doing calculations on an independent processor in the background whilst the model is running, and orchestrating multiple instances of the model running concurrently.

A further key feature of MHYSA which greatly enhances future science investigations with MERAXES, is that it is not only limited to statistical sampling. Often we wish to simply find the 'best' model which matches our chosen observational constraints, without necessarily investigating parameter degeneracies or uncertainties. MHYSA has been written with a flexible interface that can be used to carry out such a global optimization of the model parameters using any suitable technique, and leveraging the many publicly available packages.

For this work, we make use of the publicly available PYTHON package ULTRANEST (Buchner 2021) to carry out our parameter space sampling using the nested sampling (Skilling 2004) Monte Carlo algorithm MLFRIENDS (Buchner 2014, 2017). The benefit of nested sampling over more traditional Markov Chain Monte Carlo (MCMC) methods for this work is four-fold: Nested sampling does not require us to have an estimate of the global maximum likelihood parameter values in order to avoid a lengthy burn-in phase; it can more efficiently deal with multimodality or complicated degeneracies in probability distributions; it typically requires fewer likelihood evaluations to probe the underlying posterior probability distribution to a satisfactory degree of accuracy; and it naturally provides us with an estimate of the Bayes evidence that can be used to quantitatively compare the support for different escape fraction parametrizations.

For each of the parameter space sampling runs presented in this work, we use a minimum of 200 live points. This is the number of points which are maintained throughout the sampling process and iteratively updated to sample the posterior. More live points result in a higher resolution sampling of the posterior, at the expense of more likelihood evaluations. The default ULTRANEST stopping criteria were employed, with the exception of the maximum fraction of the remaining evidence and minimum effective sample size, which we set to 0.1 and 600, respectively.

4 THE GALAXY UV LUMINOSITY FUNCTION AND $\beta - M_{\rm UV}$ RELATION

The galaxy UV luminosity function is the most direct observational probe of the reionizing galaxy population at $z \geq 5$. It traces the recent star formation at a fixed number density and hence, the total ionizing photon production rate and its spatial distribution (under the assumption that stars provide the bulk of ionizing photons during the Epoch of Reionization; e.g. Qin et al. 2017). Therefore, if we are to carry out a quantitative investigation of allowed escape fraction models, we must first calibrate our model against the observed UVLF.

The free parameters of the model which we allow to vary when fitting the UVLF are listed in Table 1. These are identical to the free parameters used by Qiu et al. (2019) and include four galaxy formation parameters controlling star formation and supernova feedback, as well as four parameters which control dust extinction and how it varies with cold gas mass. These parameters have been shown in previous works to be the most important for controlling the growth of early galaxies in our model (e.g. Mutch et al. 2016) and reproducing the evolution of the UVLF (Qiu et al. 2019).

The assumed prior distributions for each parameter are presented in Appendix A. They are selected to weakly encode our past experience and physical intuition, however, we have ensured that our conclusions are insensitive to our precise choices except where explicitly stated in the text.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Table 1. The free galaxy formation and dust model parameters used in this work when constraining against the observed UVLF and $\beta - M_{\rm UV}$ relations. See Section 4 for details.

Parameter	Description	Equation
Galaxy formation		
$\log_{10}(\alpha_{\rm SF})$	Star formation efficiency	(1)
$\log_{10}(\Sigma_{\rm n})$	Surface density threshold for star formation	(1)
$\log_{10}(\epsilon_0)$	Supernova energy coupling efficiency	(5)
$\log_{10}(\eta_0)$	Mass loading factor	(6)
Dust		
$\tau_{\rm ISM}$	ISM optical depth normalization	(11)
$\tau_{\rm BC}$	Birth cloud optical depth normalization	(11)
$\gamma_{\rm GCD}$	Cold gas mass dependence	(10)
n	Reddening slope	(10)
a	Redshift dependence	(10)

For this work, we utilize the observational luminosity functions from Bouwens et al. (2015). Following Gillet, Mesinger & Park (2020; and as originally suggested by Bouwens et al. 2017) we enforce a minimum uncertainty of 0.08 dex on all data points. Since reionization is known to be largely complete to within $x_{\rm HI} \sim 0.05$ by $z \sim 5$ (e.g. Qin et al. 2021; Bosman et al. 2022), we only consider redshifts greater than this. The volume of our input N-body simulation $(100\,{\rm Mpc})^3$ also limits the brightest galaxies for which MERAXES can provide statistically significant predictions, due to the low number density of these biased objects. Therefore, at each redshift, we only consider luminosities below which observations indicate there should be at least 10 galaxies.

We calculate the likelihood of a model parameter set, θ, assuming uncorrelated, (log-)Normally distributed, errors:

$$\ln\left(\mathcal{L}_{\rm UVLF}(\theta)\right) = -\frac{1}{2}\sum_{\rm i=0}^{N_{\rm d}}\left(\frac{(\log_{10}(\phi_{\rm i,obs}) - \log_{10}(\phi_{\rm i,mod}))^2}{\sigma_{\rm i}^2} - \ln(2\pi\sigma_{\rm i})\right), \quad (12)$$

where $\phi_{\rm i,\,obs}$ and $\phi_{\rm i,\,mod}$ are the observations and model predictions for data point i, respectively, $\sigma_{\rm i}^2 = \sigma_{\rm i,obs}^2 + \sigma_{\rm i,mod}^2$ is the corresponding variance, and $N_{\rm d}$ is the total number of data points. Note, that we use the appropriate value of $\sigma_{\rm i,\,obs}$ when the model prediction is greater than or less than the observational data point if the quoted observational uncertainty is non-symmetric. Our choice of uncorrelated log-normal uncertainties is simply our best estimate of the true uncertainty distribution in the absence of further information (such as a covariance matrix) and is a commonly made assumption. For the model variance, we assume a Poisson distribution with a mean equal to the galaxy number counts in each bin.

In order to fully constrain all eight free model parameters and break degeneracies in the dust extinction model, Qiu et al. (2019) found it necessary to additionally constrain against the evolution in the UV slope (β) as a function of magnitude ($M_{\rm UV}$). We do similarly here, utilizing the biweight mean observations of Bouwens et al. (2014) and again assuming uncorrelated, (log-)Normally distributed errors as per equation (12).

4.1 Constraints on galaxy formation parameters

We now explore the constraining power of the UVLF and $\beta - M_{\rm UV}$ relation on the galaxy formation parameters of our model. In Mutch et al. (2016), we demonstrated that reionization feedback only plays a minor role in shaping the evolution of the observable $z \gtrsim 5$ galaxy population. We therefore ignore our reionization constraints and turn- off reionization feedback for this section, simplifying our interpretation and speeding up our calculations.

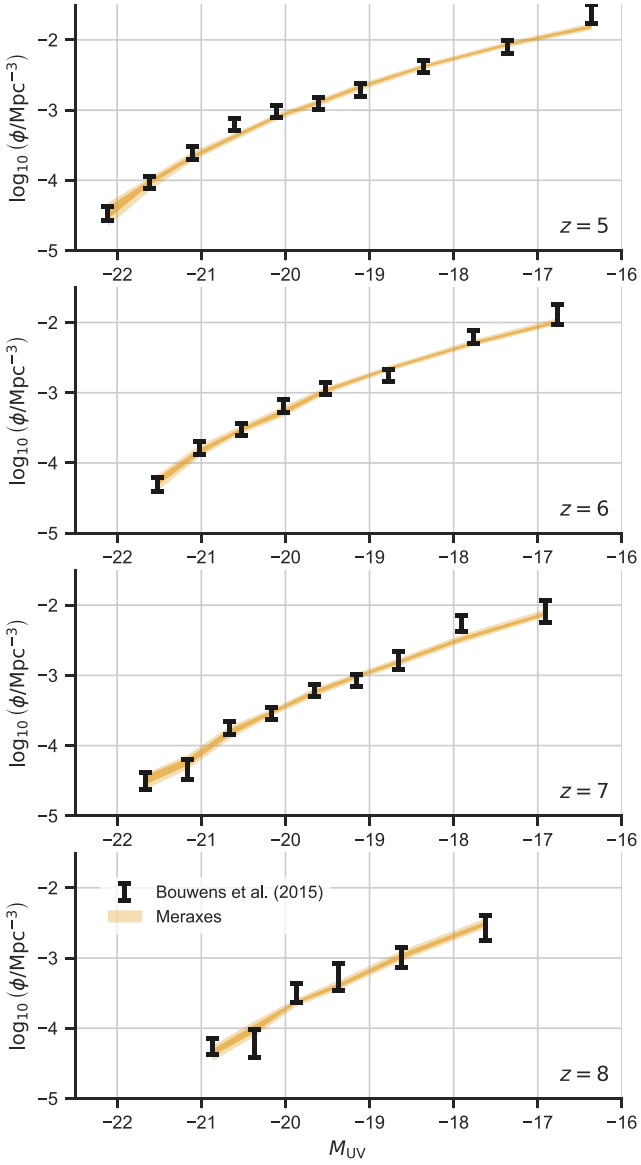


Figure 1. The best-fitting galaxy UVLF (shaded regions) obtained by fitting against the observed evolution of the UVLF alone (error bars). Dark and light shading indicates the 68 per cent and 95 per cent confidence intervals, respectively. Our model does a reasonable job of matching the observations at all fitted redshifts, providing confidence that we can correctly model the evolution of the observed faction of the total ionizing photon budget. See Section 4.1 for more details.

The model UVLFs are presented in Fig. 1. Shaded bands indicate the 68 per cent and 95 per cent confidence intervals of the model results, sampled from the underlying posterior distribution. The model is able to replicate the observational data reasonably well, across all redshifts, however, the 68 per cent range of the model ϕ at fixed $M_{\rm UV}$ is noticeably smaller than that of the constraining observations. This stems from a tension in matching all four redshifts simultaneously, especially around $-21 \lesssim M_{\rm UV} \lesssim -20$, where the observed d$\phi$/d$z$ varies non-monotonically. When the model is constrained against a single redshift, the freedom afforded by the

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

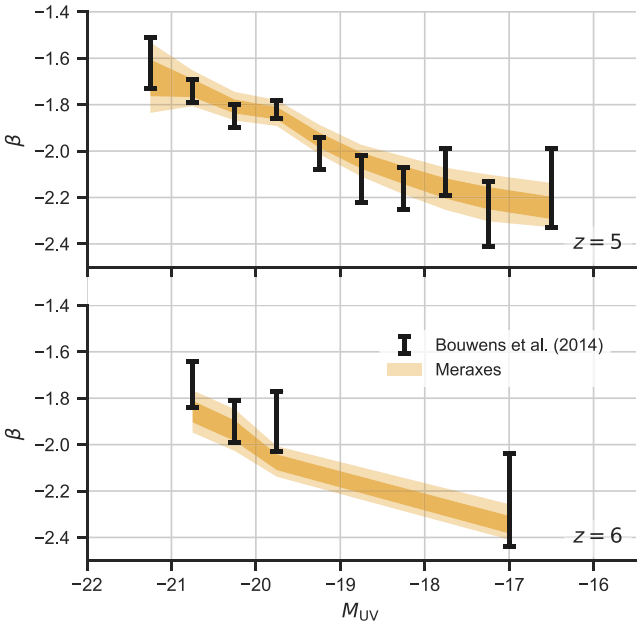


Figure 2. The best-fitting galaxy β-M_{UV} relation (shaded regions) obtained by fitting against the observed evolution of the UVLF and β-M_{UV} relation alone (error bars). Dark and light shaded regions indicate the 68 per cent and 95 per cent confidence intervals, respectively. See Section 4.1 for more details.

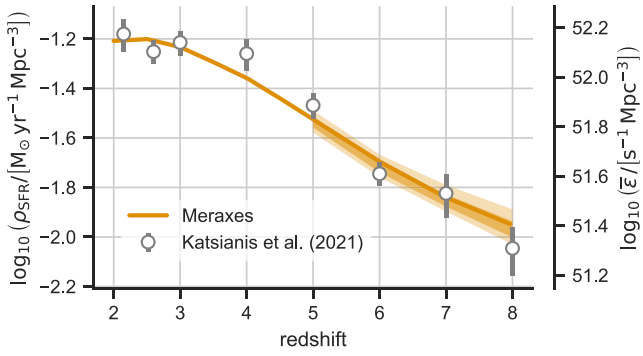


Figure 3. The predicted cosmic star formation rate density evolution obtained by fitting against the observed evolution of the UVLF and β–M_{UV} relation at $z \geq 5$ (solid line). The right hand axes shows the corresponding volume averaged intrinsic ionizing emissivity. Dark and light shaded regions indicate the 68 per cent and 95 per cent confidence intervals at $z \geq 5$. Black data points show the observed star formation rate density from Katsianis, Yang & Zheng (2021) for comparison. Note, that the model was not calibrated to reproduce the star formation rate density evolution specifically (see Section 4.2).

model and our choice of free parameters means that the magnitude of our uncertainties are similar to that of the constraining observations. When fitting against multiple redshifts simultaneously, the volume of acceptable model parameters reduces such that we are no longer able to reproduce the full variance of a single redshift. In Fig. 2, we show the corresponding $\beta - M_{UV}$ relation. Here, the uncertainty in the model results is more similar to that of the constraining observations at $z = 5$. We also show in Fig. 3 the corresponding prediction for the star formation rate density (and volume averaged intrinsic ionizing emissivity) evolution together with their 68 per cent and 95 per cent confidence intervals. At redshift 8, the predicted emissivity can vary by ∼25 per cent, decreasing to ∼10 per cent by $z \lesssim 5$.

The maximum a posteriori (MAP) parameter values, along with their marginalized posterior modes and highest density intervals (HDIs), are presented in Table 2. The corresponding 1D and 2D posteriors are also shown in Appendix A. With the exception of the critical surface density for star formation (Σ_n) and supernova energy coupling efficiency (ϵ_0), all of the parameters possess a well-defined peak in their 1D marginalized posteriors which is independent of the imposed priors. There are a number of degeneracies between the galaxy formation and dust model parameters. For example a higher critical density for star formation (Σ_n) leaves more cold gas available for efficient star formation in massive systems and hence requires more dust attenuation there (controlled by γ_{GCD}) to match the bright end of the UVLF. The similarity of the ϵ_0 posterior distribution to its prior at all but the smallest values indicates that the model's supernova feedback energy coupling has a minimum viable efficiency, beyond which there is little change to the predicted UVLF. This is due to saturation effects. Once ϵ_0 is large enough, all available cold gas (as determined by the mass-loading factor, η_0, and available cold gas reservoirs in the galaxy) is ejected and increasing the value of ϵ_0 has no effect.

4.2 Low-redshift predictions

For this work, we choose to constrain our galaxy formation model parameters against the $z \geq 5$ observations alone. Whilst MERAXES is capable of running all the way to $z = 0$, there are numerous biases and complications associated with simultaneously constraining against high- and low-redshift observations. The abundance of available observations at low-redshift, coupled with the typically smaller measurement uncertainties, means that our fit would prefer matching them, at the expense of matching potentially inconsistent high-redshift data.

Having said that, any model which completely fails to reproduce relevant low-redshift observations should be deemed invalid. For this reason, we ran our model with the MAP galaxy formation parameters down to $z = 1.8$ (the lowest redshift available from the Tiamat simulation) to ensure our predictions look sensible. Fig. 3 shows the predicted evolution of the cosmic star formation rate density compared against the UV-derived observational results reported by Katsianis, Yang & Zheng (2021). Fig. 4 further shows the predicted $z = 2$ UVLF compared against the observations of Bouwens et al. (2021). In both cases there is a reasonable level of agreement between our best-fitting model and the observations, despite the fact that the model was not directly calibrated to achieve this. We are therefore satisfied that our best-fitting galaxy formation parameters are valid, at least for the purposes of reproducing the UV evolution of the galaxy population.

5 OBSERVATIONAL CONSTRAINTS ON REIONIZATION

Whilst the evolution of the observed UV luminosity function allows us to constrain the ionizing photon budget as a function of time, it does not provide any direct constraint on the fraction of these photons which escape their source galaxy and participate in reionization (f_{esc}). For this, we turn to the following three observations.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Table 2. The maximum a posteriori (MAP) for each galaxy formation parameter when constraining against the observed evolution of the UVLF and β-$M_{\rm UV}$ relation. The 1D marginalized MAP, mean, and 68 per cent and 95 per cent highest density intervals (HDI) are also presented for each parameter. Plots of the 1D and 2D marginalized posterior distributions are presented in Fig. A3.

				Marginalized	
	MAP	Mode	Mean	68 per cent HDI	95 per cent HDI
$\log_{10}(\alpha_{\rm SF})$	−1.11	−1.18	−1.20	[−1.26, −1.12]	[−1.35, −1.05]
$\log_{10}(\Sigma_{\rm n})$	−0.85	−1.10	−1.50	[−1.76, −0.61]	[−2.75, −0.57]
$\log_{10}(\epsilon_0)$	0.65	1.47	2.43	[0.62, 2.96]	[0.51, 4.88]
$\log_{10}(\eta_0)$	0.68	0.64	0.68	[0.60, 0.75]	[0.51, 0.84]
$\tau_{\rm ISM}$	14.00	9.41	18.66	[0.17, 22.08]	[0.01, 46.66]
$\tau_{\rm BC}$	289.77	282.73	671.60	[105.27, 801.06]	[65.63, 1650.87]
$\gamma_{\rm GCD}$	1.52	1.35	1.39	[1.15, 1.53]	[1.06, 1.83]
n	−2.43	−2.55	−2.46	[−2.88, −2.09]	[−3.17, −1.64]
a	0.69	0.88	0.87	[0.72, 0.97]	[0.64, 1.11]

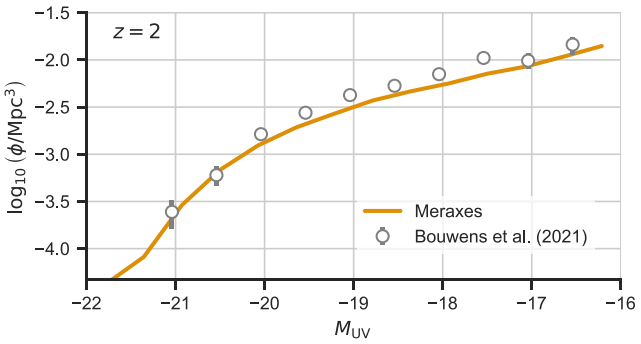


Figure 4. The predicted $z = 2$ UVLF found by constraining against the observed evolution of the UVLF and β–$M_{\rm UV}$ relation at $z \geq 5$ (solid line). The observational results of Bouwens et al. (2021) are shown for comparison (error bars). Note that the model was not calibrated to reproduce the UVLF at $z < 5$ specifically (see Section 4.2).

5.1 Thompson scattering optical depth

The integrated optical depth due to Thompson scattering of CMB photons by free electrons can be expressed as:

$$\tau_e = \int_{z=0}^{\infty} \frac{c\,{\rm d}t}{{\rm d}z}(1+z)^3\sigma_{\rm T} \times \left[Q^{\rm m}_{\rm HII}\,\langle n_{\rm H}\rangle + (Q^{\rm m}_{\rm HeII} + 2Q^{\rm m}_{\rm HeIII})\,\langle n_{\rm He}\rangle\right]{\rm d}z\,, \quad (13)$$

where $\sigma_{\rm T} = 6.652 \times 10^{-25}\,{\rm cm}^2$ is the Thomson scattering cross-section, $Q^{\rm m}_X$ is the mass-weighted global ionized fraction of species X, and $\langle n_{\rm H}\rangle = 1.88 \times 10^{-7}(\Omega_{\rm b}h^2/0.022)\,{\rm cm}^{-3}$ and $\langle n_{\rm He}\rangle = 0.148 \times 10^{-7}(\Omega_{\rm b}h^2/0.022)\,{\rm cm}^{-3}$ are the average comoving density of hydrogen and helium, respectively (Wyithe & Loeb 2003). In order to calculate $\tau_{\rm e}$ from our model, we follow the assumptions outlined in Mutch et al. (2016). In particular, we assume that helium and hydrogen are singly ionized at the same rate [i.e. $Q^{\rm m}_{\rm HeII}(z) = Q^{\rm m}_{\rm HII}(z)$ and $Q^{\rm m}_{\rm HeIII}(z > 4) = 0.0$] until $z = 4$, after which helium is assumed to be doubly ionized [Kuhlen & Faucher-Giguère 2012, i.e. $Q^{\rm m}_{\rm HeIII}(z \leq 4) = 1.0$].

We utilize the Planck Collaboration (2016b) measurement of $\tau_{\rm e} = 0.058 \pm 0.012$ as our single data point. As with the ionizing emissivity constraint above, we use use a standard χ^2 likelihood and assume Normally distributed errors. We further assume that the uncertainty on the model prediction is negligible and that there is no missing or underestimated systematic uncertainties associated with the observational data.

5.2 Ionizing emissivity

The instantaneous ionizing emissivity (the number of metagalactic ionizing photons per unit time, per hydrogen atom[1]) at the latter stages of reionization, can be inferred from Ly α forest observations (e.g. Bolton & Haehnelt 2007; McQuinn, Oh & Faucher-Giguère 2011; Becker & Bolton 2013). Following Mutch et al. (2016), we calculate this quantity from MERAXES using:

$$\dot{N}_\gamma = \frac{\dot{m}_* N_{\gamma/m_\star} f_{\rm esc}}{f_{\rm b} M_{\rm tot}(1 - 0.75 Y_{\rm He})}\,, \quad (14)$$

where $\dot{m}_*$ is the total star formation rate in the full simulation volume, $N_\gamma/m_\star$ is the number of ionizing photons per stellar baryon (fixed at 6000 to be consistent with our assumed Kroupa 2001) IMF, $f_{\rm b} = \Omega_{\rm b}/\Omega_{\rm m}$ is the universal baryon fraction, and $Y_{\rm He}$ is the helium mass fraction.

In this work, we constrain against the emissivities calculated by D'Aloisio et al. (2018). These were derived by combining the most recent measurements of the IGM opacity and gas temperature out to $z \sim 6$ with a suite of high-resolution hydrodynamic simulations covering a broad range of IGM thermal histories. For more details on both the methodology used to obtain the observed ionizing emissivity values and the treatment of the associated errors, we refer the reader to D'Aloisio et al. (2018). We utilize a standard χ^2 likelihood assuming Normally distributed and uncorrelated errors (c.f. equation 12). We choose to restrict ourselves to $z > 5$ in this work. While observations are available at $z < 5$, and these ionizing emissivity data would place tighter constraints, the contribution from AGN can become non-negligible at $z \lesssim 4$ (e.g. Faucher-Giguère 2020; Trebitsch et al. 2021). Furthermore, at $z = 5$ the snapshot cadence of the Tiamat N-body simulation changes from constant in time to logarithmic in redshift, introducing complications in any analysis between $z \sim 4.5$ and 5.

5.3 Global neutral fraction

The volume-averaged global neutral fraction can be probed observationally by a number of indirect methods. These include Ly α source counts (e.g. McGreer, Mesinger & D'Odorico 2015), Ly α and Lyman-break galaxy clustering (e.g. Mason et al. 2018; Hoag et al. 2019), and QSO damping-wing absorption features (e.g. Greig et al. 2017; Davies et al. 2018; Greig, Mesinger & Bañados 2019;

[1]Often commonly expressed as the number of ionizing photons per unit time, per unit volume.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Wang et al. 2020). Due to the small available sample sizes and model-dependent nature of these probes, the associated uncertainties on individual measurements are typically large. However, a direct measure of the volume-averaged global neutral fraction, $\bar{x}_{\rm HI}$, can also be obtained from the fraction of pixels with zero flux in high-z quasar spectra (Mesinger 2010). The appearance of these 'dark' pixels in the Ly α and/or β forests indicates the presence of intervening neutral hydrogen, providing an upper limit to the cosmological neutral hydrogen fraction.

In this work, we utilize a number of different neutral fraction constraints assembled from the recent literature, including both model dependent and independent sources (McGreer, Mesinger & D'Odorico 2015; Greig et al. 2017; Davies et al. 2018; Mason et al. 2018; Greig, Mesinger & Bañados 2019; Hoag et al. 2019; Wang et al. 2020). With the exception of the model independent dark pixel constraints, we restrict ourselves to only utilizing observational values with error bars, rather than upper or lower limits. This is largely motivated by the difficulty in sensibly assigning uncertainties to such limits. However, limited investigation has further suggested that the inclusion of these limits, with the most conservative assumption of a step function likelihood, leads to very little additional constraining power on the evolution of our model escape fraction in any case.

We use the 1σ uncertainties on each data point as presented and, following Greig & Mesinger (2017), we model upper and lower limits as a one-sided Gaussian with a likelihood of 1 below (/above) the observed value. The model predictions are obtained by calculating the fraction of cells in our reionization grid with a neutral fraction $<10^{-4}$ at the central redshift of each data point.

6 RESULTS

In this section, we use our best-fitting model (MAP) combined with a flexible halo mass dependent model for escape fraction to constrain the relation between escape fraction and galaxy properties, and to compute their contribution to the ionizing photon budget.

6.1 Constraints on escape fraction evolution

Recent numerical studies have indicated that there is a likely correlation between halo mass and the escape fraction of ionizing photons (e.g. Paardekooper, Khochfar & Dalla Vecchia 2015; Kimm et al. 2017; Yeh et al. 2023) This is thought to be a result of low-mass haloes having shallower potential wells, allowing supernova to more efficiently clear high-density gas in the IGM and provide channels for ionizing photons to escape the galaxy. In this section, we thus investigate the constraining power of the Thompson scattering optical depth, ionizing emissivity, and neutral fraction evolution on $f_{\rm esc}$ as a function of halo mass.

We fix all galaxy evolution parameters to their MAP values found in Section 4, where we constrained the model against the evolution of the UVLF. We note that this removes any extra scatter in the recovered escape fraction introduced by uncertainties in the evolution of ionizing source population, simplifying our interpretation. However, the observed star formation rate density is constrained by the luminosity function to a ~ 10 per cent level at $z \sim 6$ (see Fig. 3), which is more precise than the observed uncertainties in the ionization rate in the IGM. The 10 per cent uncertainty on the star formation rate density (which is linearly related to the escape fraction with respect to the timing of reionization) is also smaller than the fractional constraints on the escape fraction derived below. We therefore find that the effect of simultaneously allowing our galaxy formation parameters to be free, in addition to the reionization parameters, would be minimal.

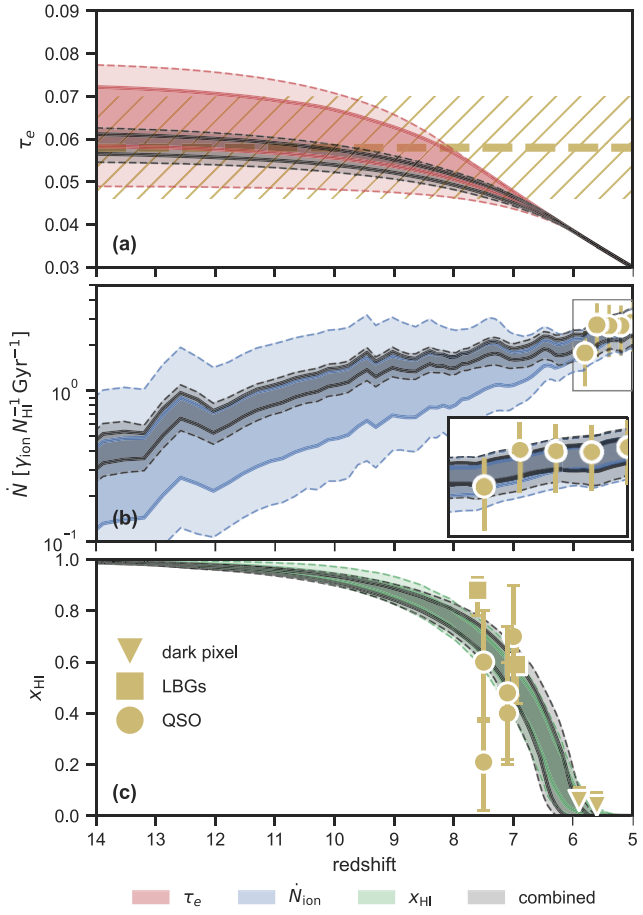


Figure 5. The reionization constraints used in this work (tan coloured points and hatched area), along with the mass dependent escape fraction model constrained against each observation independently (red, blue, and green shading) and simultaneously (black shading). Dark and light shading indicate the 68 per cent and 95 per cent confidence intervals, respectively. The MERAXES galaxy formation parameters are fixed to their MAP values as presented in Table 2. See Section 6.1 for discussion of these results. (a): The Thompson scattering optical depth (Section 5.1). (b): The evolution of the total ionizing emissivity (Section 5.2). The inset panel shows a zoom-in of z = 5-6. (c): The evolution of the volume weighted neutral fraction (Section 5.3).

We choose a logistic functional form for the escape fraction:

$$f_{\rm esc}(M_{\rm vir}) = f_{\rm min} + \frac{f_{\rm max} - f_{\rm min}}{1 + \exp\left(-\gamma_{\rm esc} \log_{10}\left(\frac{M_{\rm vir}}{M_{\rm offset}}\right)\right)}, \quad (15)$$

where

$$\gamma_{\rm esc} = \frac{4 \frac{{\rm d}f}{{\rm d}\log_{10} M_{\rm vir}}}{(f_{\rm max} - f_{\rm min})}. \quad (16)$$

This flexible, four parameter function allows the escape fraction to vary between $f_{\rm min}$ and $f_{\rm max}$ with a maximum rate of df/(d$\log_{10}M_{\rm vir}$) mid-way between the two, at a mass of $M_{\rm offset}$. We use non-flat priors for these parameters, both to encode our physical intuition, and to improve the efficiency of our sampling. More details can be found in Appendix A, however, we have ensured that the precise choice of priors does not influence results except where explicitly discussed below.

The results of the model after constraining against the three observations are presented in Fig 5 and the 1D marginalized posterior

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

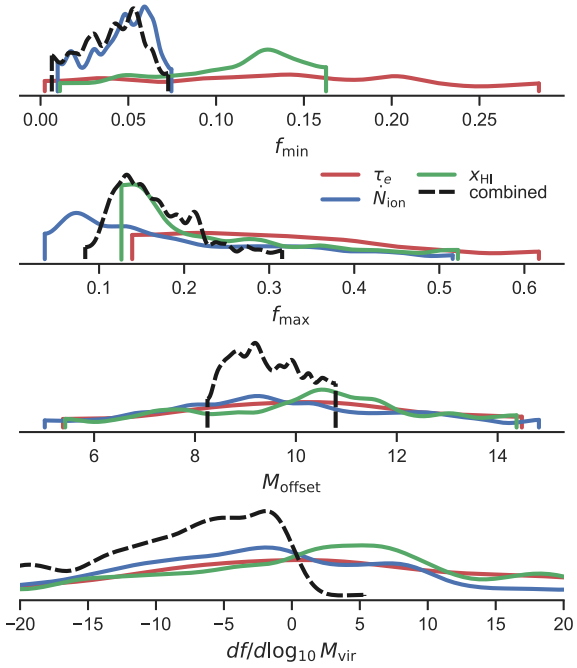


Figure 6. Kernel density estimate plots of the 1D marginalized constraints on the escape fraction parameters (equation 15). The red, green, and blue curves show the constraints from the CMB optical depth, neutral fraction, and ionizing photon rate observations, respectively. The black curves show the combined constraints.

distributions of the escape fraction parameters are presented in Fig. 6. Coloured lines and shaded regions show the results of fitting the escape fraction to match each observation independently, whilst black lines and shaded regions show the results of constraining against all three observational data sets simultaneously.

6.2 Thompson scattering optical depth

The Thompson scattering optical depth is an integrated quantity spanning the full history of the IGM back to the surface of last scattering. As such, it is sensitive to both the timing and effective mid-point of reionization. From panel (a) of Fig. 5 we see that the model is able to reproduce the Planck Collaboration (2016a) observations (dashed line and hatched region) when using a galaxy growth history consistent with observations.

The kernel density estimate (KDE) of the 1D marginalized constraints on the escape fraction parameters are shown by the red lines in Fig. 6. From here, we can see that the optical depth provides very little constraining power on the escape fraction. It merely demands that the escape fraction value be between $\sim 5-35$ per cent across the full range of contributing halo masses present in the model ($10^8 \leq \mathrm{M}_\odot \lesssim 10^{13}$).

Interestingly, our model disfavours $\tau_e \lesssim 0.52$, even when not requiring a simultaneous match to the observed ionizing emissivity and neutral fraction evolution (red shaded regions). This suggests that achieving lower optical depths requires an escape fraction model which has a stronger scaling with redshift than is possible with halo mass alone.

More recent determinations of the optical depth from Planck Collaboration (2020) suggest $\tau_e = 0.0544^{+0.007}_{-0.008}$. This is lower than the constraint used in this work (see Section 5.1) and would require a more rapid reionization with a larger minimum f_{max} value than we see in Fig. 6. The result would be an increasing tension with our constraints on $\dot{N}_{ion}$ and $x_{H\,I}$, both of which prefer lower f_{max} values.

6.3 Ionizing emissivity

The total ionizing emissivity ($\dot{N}$) directly constrains the number of ionizing photons entering the IGM and thus, when coupled with a fixed UVLF evolution, places the strongest constraints on the escape fraction. From the top two panels of Fig. 6 we see that ionizing emissivity prefers a mass independent escape fraction ($f_{min} \approx f_{max} \approx 7$ per cent). For escape fractions parametrizations which vary with mass, there is a preference for a large positive slope ($df/d\log_{10}M_{vir}$) centred around $M_{vir} \sim 10^9 - 10^{10}\,\mathrm{M}_\odot$. This corresponds to a rapidly increasing escape fraction with decreasing halo mass in our parametrization, resulting in a constraint at $M_{vir} = 10^8\,\mathrm{M}_\odot$ of $f_{esc} \sim 5-25$ per cent. This is driven by the observed flattening in the ionizing emissivity beyond $z \lesssim 6$, requiring both a rapid increase in the emissivity using early-forming, low-mass haloes, and a reduced contribution from high-mass haloes to keep the emissivity as flat as possible at low redshift [see blue shaded regions in panel (b) of Fig. 5].

6.4 Neutral fraction evolution

The range of allowed neutral fraction evolution models is shown by the green shaded regions in panel (c) of Fig. 5. The dark pixel fraction upper limits of McGreer, Mesinger & D'Odorico (2015) ensure that reionization is all but complete by $z \sim 5.5$. Combined with the higher redshift constraints, particularly that of Hoag et al. (2019), the escape fraction parameters are driven towards producing a late and rapid reionization scenario, requiring a minimum escape fraction of ~ 12 per cent across all halo masses (68 per cent confidence). The KDEs of the 1D marginalized parameter distributions in Fig. 6 again indicate that these observations are consistent with a mass independent escape fraction. However, there is a weak preference for a positive slope ($df/d\log_{10}M_{vir}$) and an increase in escape fraction with increasing halo mass. This is contrary to our physical intuition and results from the desire to have as rapid evolution in the neutral fraction as possible, whilst maintaining a late start by keeping the contribution from early-forming, low-mass haloes as small as allowed.

6.5 Combined reionization constraints

The black shaded regions in Fig. 5 and black dashed lines in Fig. 6 show the result of calibrating the escape fraction model using all three reionization constraints together. For these combined constraints we also present the 2D marginalized distributions for escape fraction model parameters in Fig. 7 and the corresponding summary statistics in Table 3. From Fig. 5, we can see that the model is able to successfully match all of the observational constraints simultaneously. By including all constraints, we also greatly reduce the allowed range of model results, most notably those of the Thompson scattering optical depth (panel a) and ionizing emissivity (panel b). Figs 6 and 7 show that a negative $df/d\log_{10}M_{vir}$ (increasing escape fraction with decreasing halo mass) is strongly favoured. The midpoint of the transition between f_{min} and f_{max} is also restricted to a much narrower range than was the case when constraining against any one observation alone.

In Fig. 8 we show the resulting constraint on the escape fraction as a function of halo mass. At $M_{vir} \gtrsim 10^{10.5}\,\mathrm{M}_\odot$ the escape fraction is ~ 5 per cent This is driven largely by a low escape fraction in high

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

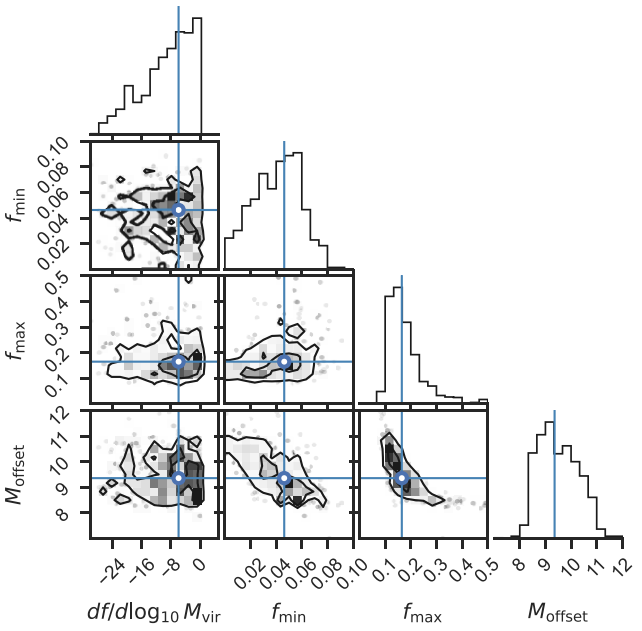


Figure 7. The 1D and 2D marginalized posterior distributions for our escape fraction model parameters when fitting to all three observational constraints simultaneously. Blue cross-hairs indicate the MAP parameter set. The 2D contours represent 1 and 2 σ probability density regions.

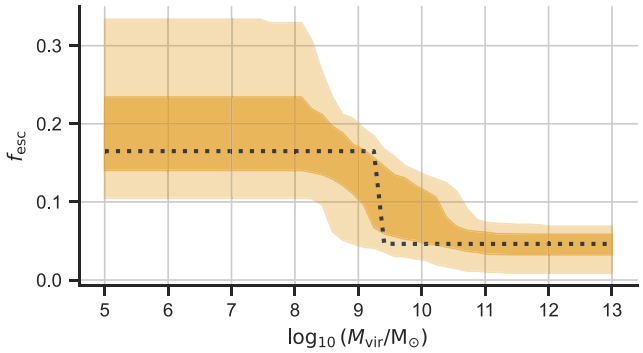


Figure 8. The predicted escape fraction evolution when constrained against the Thompson scattering optical depth, ionizing emissivity, and neutral fraction evolution. The galaxy formation parameters are fixed to the MAP values found in Section 4. Dark and light shaded regions indicate the 68 per cent and 95 per cent confidence intervals, respectively. The dotted line shows the MAP escape fraction evolution. A general trend of increasing escape fraction with decreasing halo mass can be seen, in good agreement with both theory and other works.

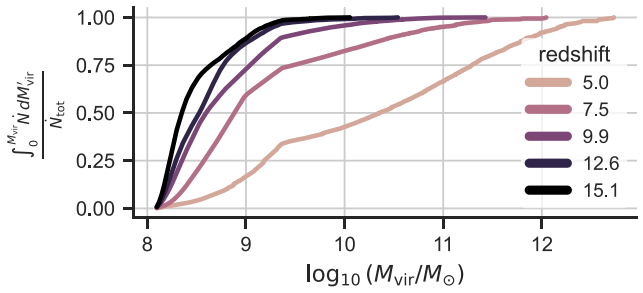


Figure 9. The cumulative contribution to the instantaneous ionizing photon rate from haloes below a given mass at different redshifts resulting from the best-fitting $f_{\rm esc}(M_{\rm vir})$ model constrained to match all reionization data sets simultaneously.

mass halos being needed to keep the ionizing emissivity as flat as possible at $z = 5 - 6$. The lack of ionizing photons entering the IGM from high mass haloes reduces the ability of the model to achieve a rapid reionization, producing optical depths $\gtrsim 0.062$ and a slower neutral fraction evolution [see panels (a) and (c) of Fig. 5]. At lower masses, the escape fraction rises to ~ 15 per cent $- 30$ per cent, required to start reionization early enough to match all three constraints.

The dotted line in Fig. 8 shows the MAP escape fraction parametrization (see Table 3 for the corresponding parameter values). It is tempting to draw conclusions about physical processes colluding to put some particular significance on halo masses of $M_{\rm vir} \approx 10^{9.35}\,{\rm M}_\odot$, where the sharp transition in $f_{\rm esc}$ occurs. However, the marginalized 68 per cent highest density interval for the $M_{\rm offset}$ parameter is broad ($10^{8.38} - 10^{9.97}\,{\rm M}_\odot$; Table 3). From Fig. 7, we can also see that the position of the step is correlated with the maximum and minimum escape fraction parameters. Indeed, this correlation is what drives the appearance of a gradual evolution in 68 per cent and 95 per cent confidence intervals of $f_{\rm esc}$ versus $M_{\rm vir}$ in Fig. 8.

Fig. 9 shows the corresponding cumulative fractional contribution of haloes to the instantaneous ionizing photon rate from haloes below a given mass at different redshifts. As expected the typical halo mass producing ionizing photons increases during reionization. However, as noted above, the best-fitting $f_{\rm esc}(M_{\rm vir})$ model favours an escape fraction that is dominated by low-mass galaxies, while reionization also supresses star formation in low-mass haloes. Together these effects result in a very rapid transition in the halo mass above which 50 per cent of photons are produced from $M_{\rm vir} \gtrsim 10^{8.5}\,{\rm M}_\odot$ at $z \sim 9$ to $M_{\rm vir} \gtrsim 10^{10.5}\,{\rm M}_\odot$ at $z \sim 5$.

Table 3. The MAP for each escape fraction model parameter (equation 15) when constraining against the observed Thompson scattering optical depth, ionizing emissivity evolution, and global neutral fraction evolution. The 1D marginalized mode, mean, and 68 per cent and 95 per cent HDIs are also presented for each parameter. Plots of the 1D and 2D marginalized posterior distributions are presented in Figs 6 and 7, respectively.

	MAP	Marginalized			
		Mode	Mean	68 per cent HDI	95 per cent HDI
$df/d\log_{10}M_{\rm vir}$	-5.89	-2.63	-21.73	[−0.11,−15.74]	[−0.11,−100.26]
$f_{\rm min}$	0.05	0.05	0.04	[0.03,0.06]	[0.00,0.07]
$f_{\rm max}$	0.16	0.13	0.18	[0.10,0.20]	[0.08,0.34]
$M_{\rm offset}$	9.36	9.18	9.52	[8.38,9.97]	[8.33,10.94]

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

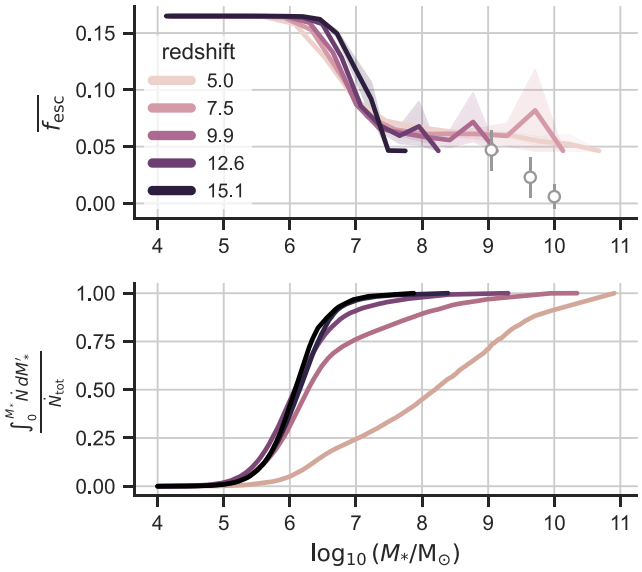


Figure 10. *Upper panel*: The evolution of mean escape fraction versus galaxy stellar mass resulting from the best-fitting $f_{\rm esc}(M_{\rm vir})$ model constrained to match all reionization data sets simultaneously. Results are shown at five values of redshift. Shaded regions represent the 95 per cent confidence intervals of the mean, calculated from 1000 bootstrapped samples. Data points are the observed $z \sim 3$ results of Pahl et al. (2022). We highlight that these measurements were not included as constraints on our model. *Lower panel*: The corresponding cumulative contribution to the instantaneous ionizing photon rate from haloes below a given stellar mass.

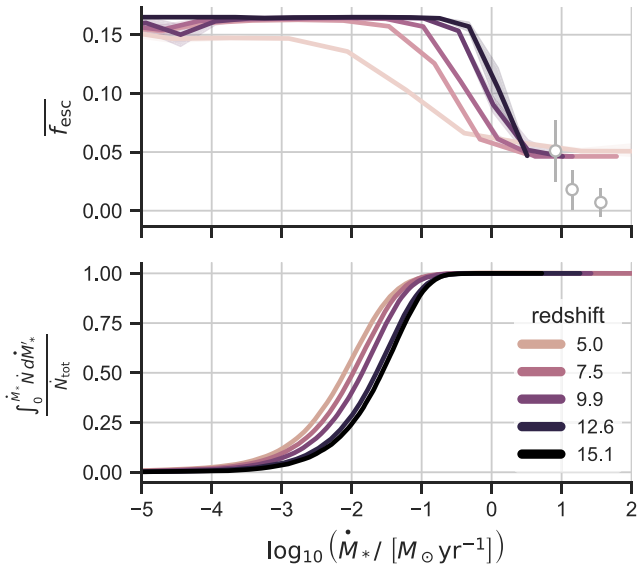


Figure 11. *Upper panel*: The evolution of mean escape fraction versus instantaneous star formation rate resulting from the best-fitting $f_{\rm esc}(M_{\rm vir})$ model constrained to match all observational datasets simultaneously. Results are shown at five values of redshift. Shaded regions (only visible for $z = 15$) represent the 95 per cent confidence intervals of the mean, calculated from 1000 bootstrapped samples. Data points are the observed $z \sim 3$ results of Pahl et al. (2022). We highlight that these measurements were not included as constraints on our model. *Lower panel*: The corresponding cumulative contribution to the instantaneous ionizing photon rate from haloes below a given stellar mass.

6.6 Escape fraction evolution with galaxy properties

Previous studies of escape fraction evolution, both using simulations and observations, have found mean trends with galaxy properties such as stellar mass and colour (e.g. Begley et al. 2022; Pahl et al. 2022; Yeh et al. 2023). In this section, we use the MAP $f_{\rm esc}(M_{\rm vir})$ parameter values to investigate the predicted evolution of the escape fraction with stellar mass and star formation rate.

The upper panel of Fig. 10 shows the mean escape fraction as a function of galaxy stellar mass resulting from the $f_{\rm esc}(M_{\rm vir})$ model (equation 15) with the MAP parameter values listed in Table 3. The shaded region indicates the 95 per cent confidence intervals on the mean, generated from 1000 bootstrap samples. Since stellar mass correlates positively with halo mass in our model (Mutch et al. 2016), the results are qualitatively similar to the dependance on halo mass. The typical escape fraction decreases from its maximum of ~ 20 per cent at $M_\star \lesssim 10^6\,{\rm M}_\odot$ to ~ 7 per cent at $M_\star \gtrsim 10^8\,{\rm M}_\odot$. This decrease begins at $M_\star \sim 10^{6.5}\,{\rm M}_\odot$ at $z \sim 15$ and shifts to smaller masses as reionization progresses due to a mild evolution of the typical stellar mass hosted by $M_{\rm offset} = 10^{9.25}\,{\rm M}_\odot$ haloes, the pivot point of our MAP $f_{\rm esc}(M_{\rm vir})$ model.

The lower panel of Fig. 10 shows the corresponding cumulative fractional contribution of galaxies below a given stellar mass to the instantaneous ionizing photon rate. Since stellar mass correlates positively with halo mass we see a similar trend as Fig. 9 where the stellar mass above which galaxies dominate the ionizing photon budget increases during reionization. The stellar mass above which 50 per cent of photons are produced is predicted to be close to a constant during the reionization era ($M_\star \sim 10^6\,{\rm M}_\odot$ at $z \gtrsim 7.5$), but then increases rapidly to $M_\star \sim 10^{8.5}\,{\rm M}_\odot$ at $z \lesssim 6$.

Fig.11 shows how the mean escape fraction instead varies by star formation rate for our MAP model. We see a similar evolutionary trend as was found with stellar mass. The escape fraction has a more extended evolution with star formation rate as redshift decreases. The lower panel of Fig. 11 illustrates that galaxies with star-formation rates between $\dot{M}_\star \sim 10^{-1.5}\,{\rm M}_\odot\,{\rm yr}^{-1}$ and $\dot{M}_\star \sim 10^{-2.5}\,{\rm M}_\odot\,{\rm yr}^{-1}$ contribute the dominant sources of ionizing photons throughout reionization.

7 DISCUSSION

Before concluding we compare our population based findings with other recent theoretical results and observational studies.

Simulations typically show values of escape fraction that vary between a few and a few 10s of per cent (e.g. Paardekooper, Khochfar & Dalla Vecchia 2015; Rosdahl et al. 2022), which is in agreement with our results and with prior empirical comparisons of galaxy luminosity functions and Ly α forest constraints (e.g. Bolton & Haehnelt 2007; Wyithe et al. 2010). Hydrodynamic radiative transfer simulations further indicate that escape fractions decrease with halo mass for haloes with $M_{\rm vir} \lesssim 10^9\,{\rm M}_\odot$ (e.g. Kimm et al. 2017; Lewis et al. 2020). Most recently Yeh et al. (2023) have computed the escape fractions of reionizing galaxies within the THESAN project. They find that low-mass galaxies having $M_\star \lesssim 10^7\,{\rm M}_\odot$ are the main drivers of reionization above $z \gtrsim 7$, while higher-mass galaxies with $M_\star \gtrsim 10^8\,{\rm M}_\odot$ dominate the escaped ionizing photon budget at lower redshifts. When calculating the dependence on mean escape fraction as a function of halo mass, Yeh et al. (2023) report a value of a few per cent in larger haloes with a sharp rise towards halo masses below $M_{\rm vir} \sim 10^9\,{\rm M}_\odot$ (see also Kostyuk, Ciardi & Ferrara 2023). Inference studies by Park et al. (2019) also consider similar observational constraints to those used in this work (the UV luminosity function, CMB optical depth, and dark pixels). Although

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

their UV ionizing escape fraction has a simpler, power-law relation with the host haloes mass, they also find $f_{\rm esc}$ around 5–30 per cent for low-mass haloes, with no significant preference on the power-law index. All of these results are in qualitative agreement with our findings, especially when the different simulation techniques and mode of analysis are considered.

In addition to theoretical results, we can compare our results in light of observational estimates of the escape fraction. Using stacked spectra of Lyman break galaxies at $z \sim 3$ Pahl et al. (2021; following Steidel et al. 2018) find an average value for the escape fraction of 0.06 ± 0.01. Subsequently Pahl et al. (2022) find an anticorrelation between the escape fraction and stellar mass. In Figs 10 and 11, we reproduce the measurements of escape fraction from Pahl et al. (2022) for comparison with our model constraints (we stress that these measurements were not included as constraints on our model). The masses of these galaxies are at the upper end of those thought to be responsible for reionization ($M_\star \gtrsim 10^8\,{\rm M}_\odot$). However, our results overlap at the lowest stellar masses and star formation rates present in Pahl et al. (2021), as well as with many of the simulation studies referenced above. Whilst the correlations at high masses and star formation rates appear different, we note that we have constrained our escape fraction to follow a redshift-independent, halo mass dependent form. We have very few galaxies in these high-stellar mass/star formation rate regimes due to our limited simulation volume and therefore they provide almost no contribution to our fits.

Finally, we note that, whilst our *N*-body simulation, Tiamat, provides a good balance between mass resolution and simulated volume, it does not fully resolve all haloes down to the redshift-dependent atomic cooling mass threshold. However, we miss <3 per cent of stellar mass at $z \approx 6$ and <35 per cent at $z \lesssim 12$, therefore we do not expect this to significantly alter our conclusions (see section A1 of Mutch et al. 2016).

8 CONCLUSIONS

We have used the MERAXES semi-analytic model to infer the mean ionizing photon escape fraction and its dependence on galaxy properties through joint modelling of the observed high redshift UV galaxy luminosity function and the available constraints on the reionization history with the MERAXES semi-analytic model. MERAXES couples galaxy formation to a seminumerical description of reionization, allowing the exploration of reionization and the growth of the galaxy population both temporally and spatially.

In this paper, we present the application of nested sampling to MERAXES, allowing statistical exploration of the underlying free parameter space, including identification of the best-fitting model as well as statistical confidence intervals. Using a flexible model that relates the escape fraction to halo mass, we use the joint constraints of the UV luminosity function, CMB optical depth, and the Lyα forest to evaluate the allowed range of escape fraction values and halo mass dependence. Our results show that available constraints require an escape fraction of (18 ± 5) per cent for galaxies within haloes of $M \lesssim 10^9\,{\rm M}_\odot$ and (5 ± 2) per cent for larger mass haloes.

Because MERAXES provides the star-formation rates and stellar masses for each dark matter halo at each simulation snapshot, we can use these results to estimate the dependence of escape fraction on stellar mass and redshift. We find that there is a transition from an escape fraction of ~ 18 per cent to a smaller value of ~ 5 per cent which occurs at stellar masses of $M_\star \sim 10^7\,{\rm M}_\odot$. We find that this transition value is quite insensitive to redshift. Early in reionization it corresponds to star formation rates of $\dot{M}_\star \sim 10^{-1}\,{\rm M}_\odot\,{\rm yr}^{-1}$, and decreases towards $\dot{M}_\star \sim 10^{-2}\,{\rm M}_\odot\,{\rm yr}^{-1}$ by the end of reionization at $z \sim 5$.

When considering the contribution of galaxies to the ionizing photon budget as a function of redshift, we find that reionization is dominated by the smaller $M_\star \lesssim 10^7\,{\rm M}_\odot$ galaxies with high escape fractions at $z \gtrsim 6$, and by the larger galaxies $M_\star \gtrsim 10^7\,{\rm M}_\odot$ with lower escape fractions at $z \lesssim 6$. These stellar masses correspond to galaxies with star formation rates of between $\dot{M}_\star \sim 10^{-1.5}\,{\rm M}_\odot\,{\rm yr}^{-1}$ and $\dot{M}_\star \sim 10^{-2.5}\,{\rm M}_\odot\,{\rm yr}^{-1}$, which represent the dominant sources throughout reionization. In sum, these results agree with recent direct measurements of a ~ 5 per cent escape fraction from massive galaxies at the end of reionization, and support results from hydrodynamic modelling and direct observation indicating that the escape of ionizing photons is dominated by low-mass galaxies during reionization.

ACKNOWLEDGEMENTS

This research was supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3D (ASTRO 3D), through project number CE170100013. Parts of this work were performed on the OzSTAR national facility at Swinburne University of Technology. This work was supported by software support resources awarded under the Astronomy Data and Computing Services (ADACS) Merit Allocation Program. ADACS is funded from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government and managed by Astronomy Australia Limited (AAL). This research relies heavily, and with great thanks, on the PYTHON open source community, in particular NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), CYTHON (Behnel et al. 2011), MATPLOTLIB (Hunter 2007), H5PY, ULTRANEST (Buchner 2021), SEABORN (Waskom 2021), MPI4PY (Dalcín, Paz & Storti 2005), JUPYTER (Granger & Pérez 2021), JINJA2, CLICK, PANDAS (pandas development team 2020), and XARRAY (Hoyer & Hamman 2017).

DATA AVAILABILITY

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

REFERENCES

Angel P. W., Poole G. B., Ludlow A. D., Duffy A. R., Geil P. M., Mutch S. J., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 459, 2106
Atek H. et al., 2022, MNRAS, 000
Bañados E. et al., 2018, Nature, 553, 473
Bassett R. et al., 2022, MNRAS, 511, 5730
Becker G. D., Bolton J. S., 2013, MNRAS, 436, 1023
Begley R. et al., 2022, MNRAS, 513, 3510
Behnel S., Bradshaw R., Citro C., Dalcin L., Seljebotn D., Smith K., 2011, Computing in Science Engineering, 13, 31
Bolton J. S., Haehnelt M. G., 2007, MNRAS, 382, 325
Bosman S. E. I. et al., 2022, MNRAS, 514, 55
Bouwens R. J. et al., 2014, ApJ, 793, 115
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2017, ApJ, 843, 129
Bouwens R. J. et al., 2021, AJ, 162, 47
Buchner J. 2014, Stat. Comput., 26, 383
Buchner J., 2017, Astron. Soc. Pac., 131, 108005
Buchner J., 2021, JOSS, 6, 3001

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Cen R., Kimm T., 2015, ApJ, 801, L25
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Cora S. A. et al., 2018, MNRAS, 479, 2
Croton D. J. et al., 2006, MNRAS, 365, 11
D'Aloisio A., McQuinn M., Davies F. B., Furlanetto S. R., 2018, MNRAS, 473, 560
Dalcín L., Paz R., Storti M., 2005, J Parallel Distrib. Comput., 65, 1108
Davies F. B. et al., 2018, ApJ, 864, 142
Faucher-Giguère C.-A., 2020, MNRAS, 493, 1614
Finkelstein S. L. et al., 2015, ApJ, 810, 71
Finkelstein S. L. et al., 2019, ApJ, 879, 36
Gillet N. J. F., Mesinger A., Park J., 2020, MNRAS, 491, 1980
Gnedin N. Y., Kravtsov A. V., Chen H.-W., 2008, ApJ, 672, 765
Granger B. E., Pérez F., 2021, Comput. Sci. Eng., 23, 7
Greig B., Mesinger A., 2017, MNRAS, 465, 4838
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
Greig B., Mesinger A., Bañados E., 2019, MNRAS, 484, 5094
Greig B., Mesinger A., Davies F. B., Wang F., Yang J., Hennawi J. F., 2022, MNRAS, 512, 5390
Guo Q. et al., 2011, MNRAS, 413, 101
Harris C. R. et al., 2020, Nature, 585, 357
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Hoag A. et al., 2019, ApJ, 878, 12
Hoyer S., Hamman J., 2017, JOSS, 5, 10
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Hutter A., Dayal P., Yepes G., Gottlöber S., Legrand L., Ucci G., 2021, MNRAS, 503, 3698
Katsianis A., Yang X., Zheng X., 2021, ApJ, 919, 88
Kennicutt Robert C. J., 1989, ApJ, 344, 685
Kimm T., Katz H., Haehnelt M., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 466, 4826
Kostyuk I., Ciardi B., Ferrara A., 2023, preprint (arXiv:2308.01476)
Kroupa P., 2001, MNRAS, 322, 231
Kuhlen M., Faucher-Giguère C.-A., 2012, MNRAS, 423, 862
Leitherer C. et al., 1999, ApJS, 123, 3
Lemoine N. P., 2019, Oikos, 128, 912
Lewis J. S. W. et al., 2020, MNRAS, 496, 4342
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499
McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, ApJ, 743, 82
Mesinger A., 2010, MNRAS, 407, 1328
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Mesinger A., Greig B., Sobacchi E., 2016, MNRAS, 459, 2342
Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, MNRAS, 454, 2691
Mutch S. J., Geil P. M., Poole G. B., Angel P. W., Duffy A. R., Mesinger A., Wyithe J. S. B., 2016, MNRAS, 462, 250
Naidu R. P., Tacchella S., Mason C. A., Bose S., Oesch P. A., Conroy C., 2020, ApJ, 892, 109
Naidu R. P. et al., 2022, ApJ, 940, L14
Paardekooper J.-P., Khochfar S., Dalla Vecchia C., 2015, MNRAS, 451, 2544
Pahl A. J., Shapley A., Steidel C. C., Chen Y., Reddy N. A., 2021, MNRAS, 505, 2447
Pahl A. J., Shapley A., Steidel C. C., Reddy N. A., Chen Y., Rudie G. C., Strom A. L., 2022, MNRAS, 521, 3247
Pandas Development Team T., 2020, pandas-dev/pandas: Pandas, https://doi.org/10.5281/zenodo.3509134(accessed March 2)
Park J., Mesinger A., Greig B., Gillet N., 2019, MNRAS, 484, 933
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck CollaborationXLVII, 2016b, A&A, 596, A108
Planck CollaborationVI, 2020, A&A, 641, A6
Poole G. B., Angel P. W., Mutch S. J., Power C., Duffy A. R., Geil P. M., Mesinger A., Wyithe S. B., 2016, MNRAS, 459, 3025
Poole G. B., Mutch S. J., Croton D. J., Wyithe S., 2017, MNRAS, 472, 3659
Qin Y. et al., 2017, MNRAS, 472, 2009
Qiu Y., Mutch S. J., da Cunha E., Poole G. B., Wyithe J. S. B., 2019, MNRAS, 489, 1357
Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, MNRAS, 506, 2390
Rosdahl J. et al., 2022, MNRAS, 515, 2386
Seiler J., Hutter A., Sinha M., Croton D., 2019, MNRAS, 487, 5739
Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V.eds, AIP Conf. Ser. Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p.395
Sobacchi E., Mesinger A., 2013, MNRAS, 432, 3340
Somerville R. S., Primack J. R., Faber S. M., 2001, MNRAS, 320, 504
Srbinovsky J. A., Wyithe J. S. B., 2007, MNRAS, 374, 627
Steidel C. C., Bogosavljević M., Shapley A. E., Reddy N. A., Rudie G. C., Pettini M., Trainor R. F., Strom A. L., 2018, ApJ, 869, 123
Trebitsch M. et al., 2021, A&A, 653, A154
Virtanen P. et al., 2020, Nature Methods, 17, 261
Visbal E., Bryan G. L., Haiman Z., 2020, ApJ, 897, 95
Wang F. et al., 2020, ApJ, 896, 23
Waskom M. L., 2021, JOSS, 6, 3021
Wise J. H., Cen R., 2009, ApJ, 693, 984
Wyithe J. S. B., Loeb A., 2003, ApJ, 586, 693
Wyithe J. S. B., Hopkins A. M., Kistler M. D., Yüksel H., Beacom J. F., 2010, MNRAS, 401, 2561
Yeh J. Y. C. et al., 2023, MNRAS
Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., Venkatesan A., Behroozi P., Ferguson H. C., 2020, MNRAS, 496, 4574

APPENDIX A: PRIOR AND POSTERIOR PARAMETER DISTRIBUTIONS

In this Appendix, we present the 2D marginalized probability distributions (Fig. A3) for the fits of the MERAXES model to the galaxy UV luminosity function and $\beta - M_{\rm UV}$ relation discussed in Section 4.1. It is worth noting that Qiu et al. (2019) also used MERAXES with the same free galaxy formation and dust parameters as we do, but recovered different posterior distributions for some parameters (see their fig. A3). This can largely be attributed to the different redshift range utilized for their constraints ($z = 7 - 4$). The inclusion of lower redshift data, in particular, provides significant extra-constraining power. The remaining differences are due to our different choice of priors which can have an impact on poorly constrained parameters.

We also provide detail of the prior probability distributions employed for all free parameters in this work. These encode the range in allowed values (e.g. efficiency parameters must be in the range 0–1). In some cases, they also weakly encode physical intuition and the authors experience from past studies. We have ensured that the precise choice of priors makes no significant qualitative or quantitative difference to our findings except where explicitly noted in the text. In Table A1, we present a tabulated description of the priors, and show these graphically in Figs A1 and A2.

Note, that we do not use uniform priors as commonly employed in the galaxy formation literature. Our reasoning is that flat priors give extra, undue, weight to parameter values which we know from experience and physical intuition to be unlikely. When using such priors, one also must choose bounds for the parameter values, which in itself often imposes prior knowledge or intuition anyway. The use of weakly informative priors, such as the ones we employ, additionally provides numerous other statistical benefits (e.g. Lemoine 2019).

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

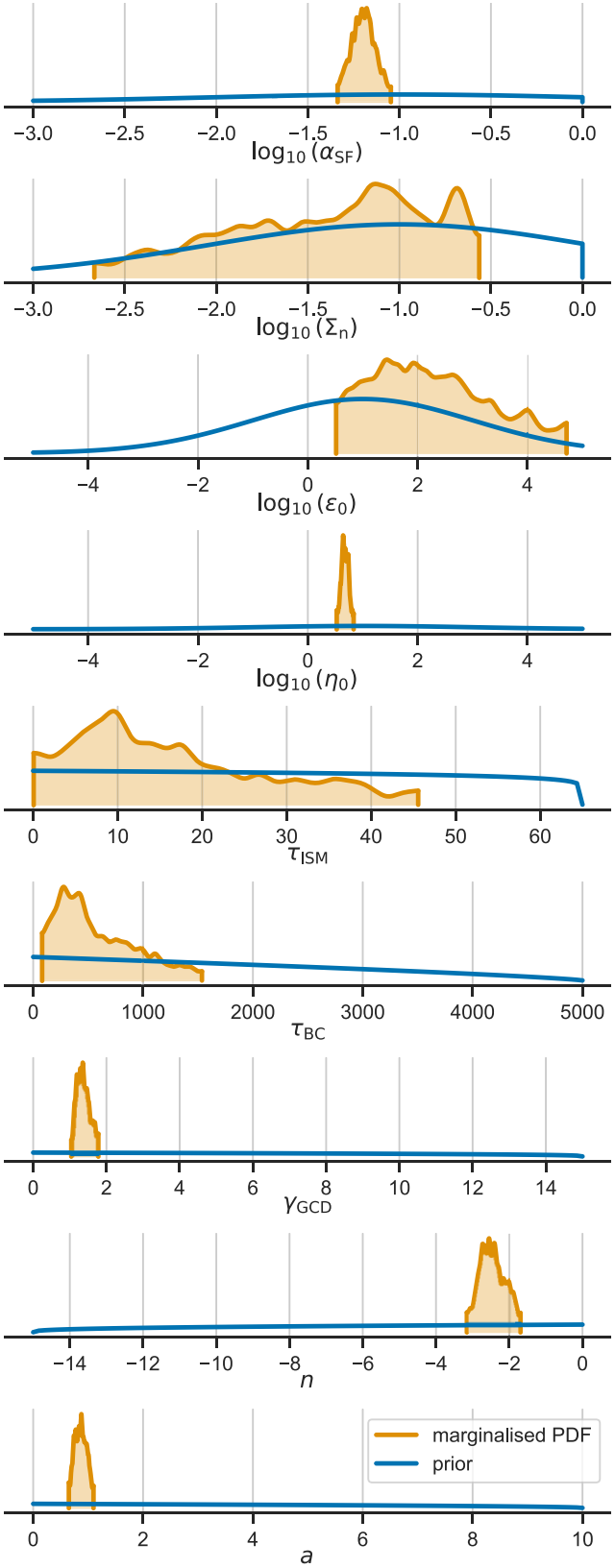


Figure A1. Kernel density estimates of the marginalized PDFs for each galaxy formation parameter (filled curves; see Section 4.1) compared with the corresponding prior distributions (lines; see Tables 2 and A1). A similarity between posterior distributions and priors indicates a lack of constraining power [e.g. $\log_{10}(\epsilon_0)$].

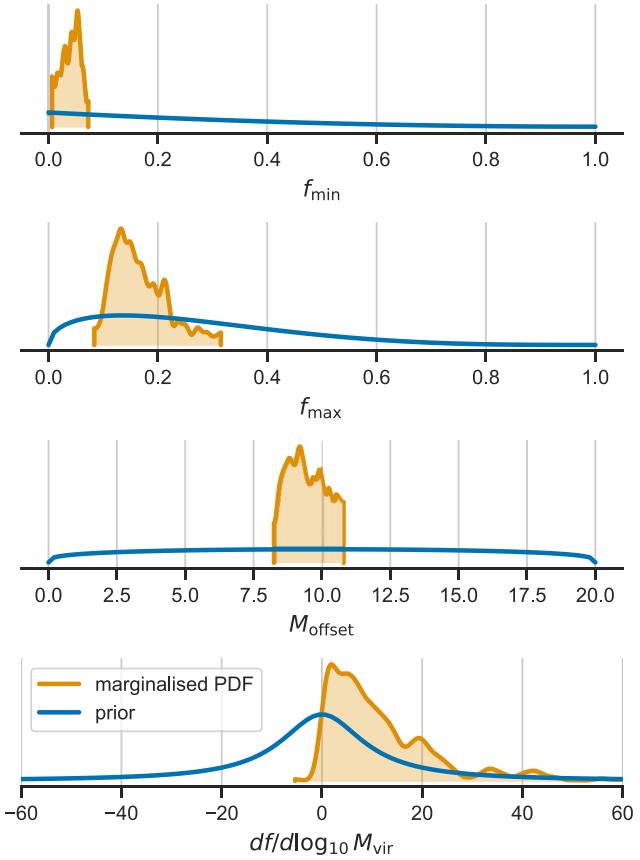


Figure A2. Kernel density estimates of the marginalized PDFs for each f_{esc} model parameter when fitting against the combined reionization constraints (filled curves; see Section 6.1). These are compared with the corresponding prior distributions (lines; see Tables 3 and A1). A similarity between posterior distributions and priors would indicate a lack of constraining power.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

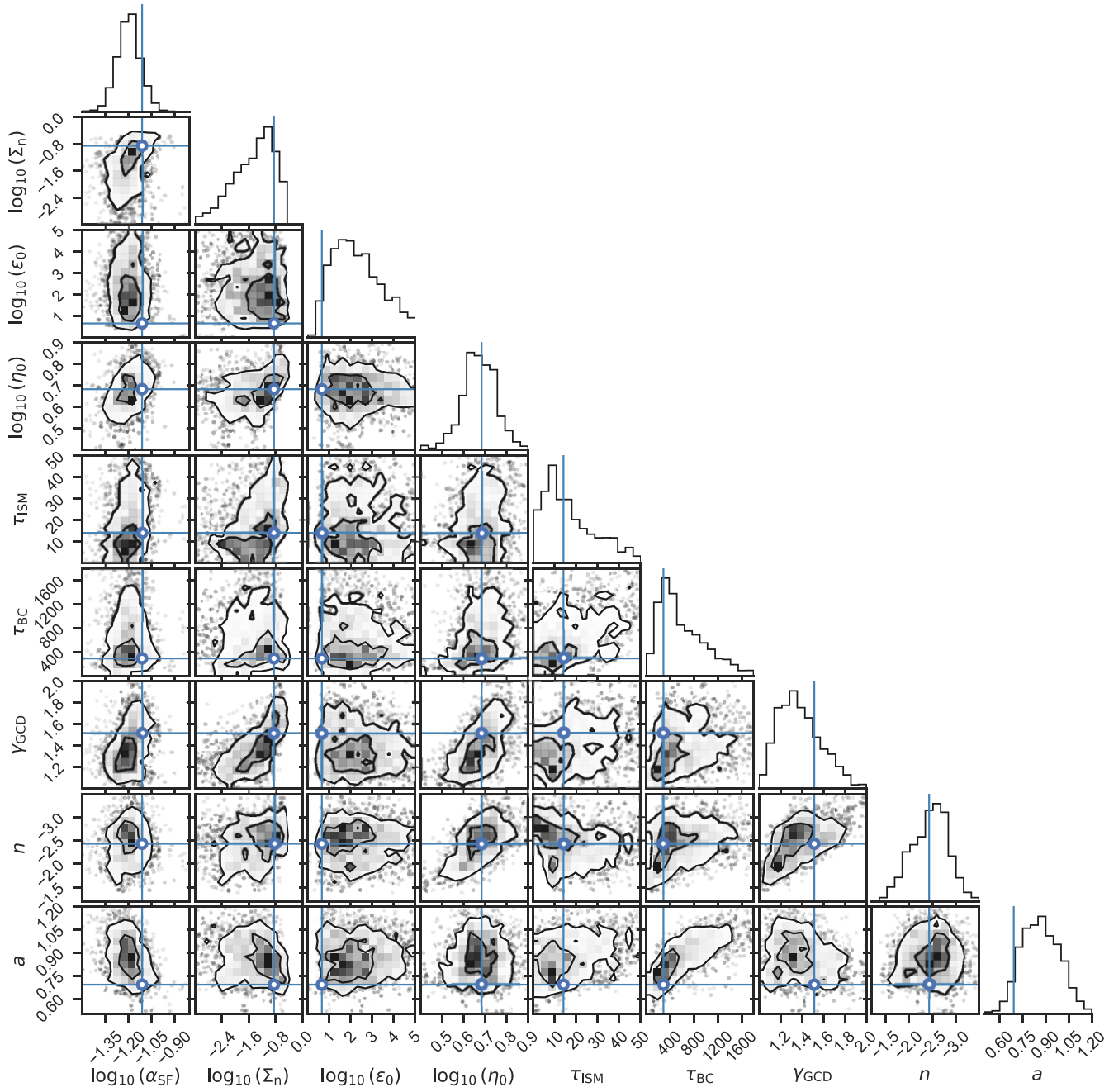


Figure A3. The 1D and 2D marginalized posterior distributions for our galaxy formation model parameters when constraining against the observed evolution of the UVLF and β-$M_{\rm UV}$ relation. Cross-hairs indicate the MAP parameter set. The 2D contours represent 1 and 2 σ probability density regions.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

Table A1. Description of the priors employed for all free parameters.

Parameter	Prior
	Galaxy formation
$\log_{10}(\alpha_{\rm SF})$	$f(x) = \text{Truncated Normal}(x, x_{\rm min} = -3.0, x_{\rm max} = 0.0, \mu = -1.0, \sigma = 1.05)$
$\log_{10}(\Sigma_{\rm n})$	$f(x) = \text{Truncated Normal}(x, x_{\rm min} = -3.0, x_{\rm max} = 0.0, \mu = -1.0, \sigma = 1.05)$
$\log_{10}(\epsilon_0)$	$f(x) = \text{Normal}(x, \mu = 1.0, \sigma = 2.0)$
$\log_{10}(\eta_0)$	$f(x) = \text{Normal}(x, \mu = 1.0, \sigma = 2.0)$
	Dust model
$\tau_{\rm ISM}$	$f(x) = \text{Beta}(x/65, \alpha = 1.0, \beta = 1.1)/65$
$\tau_{\rm BC}$	$f(x) = \text{Beta}(x/5000, \alpha = 1.0, \beta = 1.8)/5000$
γ	$f(x) = \text{Beta}(x/15, \alpha = 1.0, \beta = 1.35)/15$
n	$f(x) = \text{Beta}(x/15, \alpha = 1.0, \beta = 1.35)/15$
a	$f(x) = \text{Beta}(x/10, \alpha = 1.0, \beta = 1.5)/10$
	Escape fraction model
$f_{\rm min}$	$f(x) = \text{Beta}(x, \alpha = 1.0, \beta = 3.0)$
$f_{\rm max}$	$f(x) = \text{Beta}(x, \alpha = 1.5, \beta = 4.2)$
$M_{\rm offset}$	$f(x) = \text{Beta}(x/20, \alpha = 1.3, \beta = 1.3)/20$
$\mathrm{d}f/\mathrm{d}\log_{10}M_{\rm vir}$	$f(x) = \text{Student-t}(x/10, \nu = 1)/10$

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/527/3/7924/7459932 by guest on 02 June 2024

© 2023 The Author(s).
Published by Oxford University Press on behalf of Royal Astronomical Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.