Inferring the Ionizing Photon Contributions of High-Redshift Galaxies to Reionization with JWST NIRCam Photometry

Nicholas Choustikov[1★], Richard Stiskalek[1], Aayush Saxena[1,2], Harley Katz[3], Julien Devriendt[1], and Adrianne Slyz[1]
[1]Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, United Kingdom
[2]Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
[3]Department of Astronomy & Astrophysics, University of Chicago, 5640 S Ellis Avenue, Chicago, IL 60637, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
JWST observations have the potential to provide unprecedented constraints on the history of reionization and the sources responsible for the ionizing photons due to the detection of large populations of faint galaxies at $z \gg 6$. Modelling reionization requires knowing both the number of ionizing photons that are produced by galaxies and the fraction of those photons that escape into the intergalactic medium. Observational estimates of these values generally rely on spectroscopy for which large samples with well-defined selection functions remain limited. To overcome this challenge, we present an implicit likelihood inference (ILI) pipeline trained on mock photometry to predict the escaped ionizing luminosity of individual galaxies ($\dot{n}_{\rm ion}$) based on photometric magnitudes and redshifts. Compared to traditional SED-fitting methods, the new ILI pipeline is consistently more accurate and significantly faster. We deploy the method on a sample of 4,559 high-redshift galaxies from the JADES Deep survey, finding a gentle redshift evolution of $\log_{10}(\dot{n}_{\rm ion}) = (0.08 \pm 0.01)z + (51.60 \pm 0.06)$, with late-time values for $\dot{N}_{\rm ion}$ consistent with theoretical models and observations. We measure the evolution of the volume-averaged ionized fraction and optical depth to find that observed populations of star-forming galaxies are capable of driving reionization to completion at $z \sim 5.3$ without the need for AGN or other exotic sources. The 20% of UV-brightest galaxies ($M_{\rm UV} < -18.5$) in our sample can reionize only $\sim 30\%$ of the survey volume, demonstrating that faint LyC emitters are crucial for reionization.

Key words: galaxies: evolution – galaxies: high-redshift – dark ages, reionization, first stars – early Universe

1 INTRODUCTION

By the end of the Epoch of Reionization, the Universe had undergone its last major phase-transition, and the intergalactic medium (IGM) became mostly transparent to the Lyman Continuum (LyC: $\lambda \leqslant 912$Å) photons. While current constraints place the mean redshift of reionization at $7.8 \lesssim z \lesssim 8.8$ (Planck Collaboration et al. 2016), various observational studies find that this process was complete by a redshift in the range $z \sim 5-6$ (Fan et al. 2006; Kulkarni et al. 2019a; Becker et al. 2021; Bosman et al. 2022), contributing to the picture that this process was patchy (Iliev et al. 2006; Becker et al. 2015; Puchwein et al. 2023).

Generally, it is believed that the majority of ionizing photons are produced by young, massive stars in galaxies that undergo rapid star formation (e.g. Shapiro & Giroux 1987; Robertson et al. 2015; Hassan et al. 2018; Rosdahl et al. 2018). However, it is still unclear whether this is driven by a small number of massive sources or from more "democratic" contributions from a large number of low-mass galaxies (Paardekooper et al. 2015; Livermore et al. 2017; Mason et al. 2019a; Finkelstein et al. 2019; Naidu et al. 2020; Wu & Kravtsov 2024). Furthermore, certain observational constraints such as the low optical depth to Thompson scattering (Planck Collaboration et al. 2016) and high fraction of broad-line active galactic nucleus (AGN) with large bolometric luminosities among galaxies at redshifts $z \sim 4-6$ (Giallongo et al. 2015, 2019) all suggest that the contribution of AGN to the ionizing photon budget may be important. However, the late reionization of helium (Kriss et al. 2001; Zheng et al. 2004; Shull et al. 2004; Furlanetto & Oh 2008; Shull et al. 2010; Worseck et al. 2016) points to the fact that AGN cannot be a dominant component of hydrogen reionization. Furthermore, difficulties in accurately measuring their masses and accretion rates at high redshifts (e.g. Li et al. 2024) as well as their relative sparsity suggest that they dominate the ionizing photon budget only at lower redshifts $z \lesssim 4$ (e.g. Kulkarni et al. 2019b; Dayal et al. 2020; Trebitsch et al. 2021, 2023).

Three quantities need to be constrained in order to model the evolution of reionization. First is the UV luminosity function, $\rho_{\rm UV}$, which describes the number density of sources at a given redshift and UV magnitude. This has been measured from deep imaging surveys (e.g. Bowler et al. 2020; Bouwens et al. 2021; Harikane et al. 2022; Robertson et al. 2023; Varadaraj et al. 2023; Donnan et al. 2023, 2024), though the majority of the uncertainty comes from survey completeness (e.g. Robertson et al. 2023). Similarly, while photometric redshift estimates are occasionally known to be a source

★ nicholas.choustikov@physics.ox.ac.uk

of uncertainty[1], these have been found to be generally consistent with spectroscopic confirmations (e.g. Hainline et al. 2024).

Second is the ionizing photon production rate per UV luminosity, $\xi_{\rm ion}$. This can be predicted either by stellar population synthesis models during spectral energy distribution (SED) fitting (e.g. Leitherer et al. 1999; Stanway & Eldridge 2018), or inferred from emission lines such as Hα or Hβ (e.g. Maseda et al. 2020; Saxena et al. 2024) or O III equivalent widths (Chevallard et al. 2018; Tang et al. 2019). Here, uncertainties are primarily driven by differences in stellar population models (e.g. the presence of binaries, initial mass function, gas geometry, etc.) as well as assumptions about physical conditions in the H II regions of sources emitting ionizing photons.

Third, one must account for the fraction of the produced ionizing photons that escape their host galaxy into the IGM ($f_{\rm esc}$). Due to the fact that this depends on complex non-linear physics on small scales in the interstellar medium (ISM) (e.g. Kimm et al. 2019, 2022; Kakiichi & Gronke 2021), $f_{\rm esc}$ is highly line-of-sight dependent (e.g. Fletcher et al. 2019; Choustikov et al. 2024a; Yuan et al. 2024 and references therein), and cannot be directly measured at redshifts $z \gtrsim 4$ due to the increasingly neutral IGM (e.g. Worseck et al. 2014; Inoue et al. 2014),

$f_{\rm esc}$ arguably carries the most uncertainty. The escape fraction of ionizing photons has been studied extensively using both galaxy formation simulations (e.g. Kimm & Cen 2014; Xu et al. 2016; Trebitsch et al. 2017; Rosdahl et al. 2018, 2022; Ma et al. 2020; Saxena et al. 2022a; Giovinazzo et al. 2024) and observations of low-redshift analogues (e.g. Leitherer et al. 2016; Schaerer et al. 2016; Steidel et al. 2018; Izotov et al. 2018b,a; Flury et al. 2022a,b). In the case of simulations, capturing the production and transfer of LyC photons through a multi-phase ISM into a realistic CGM is difficult. To do so requires self-consistently capturing a large dynamical range, along with realistic models for the ISM and feedback processes (Kimm et al. 2019, 2022; Rosdahl et al. 2022). In contrast, it is not clear whether these observed analogues are representative of high-redshift galaxies or plagued by selection effects (e.g. Katz et al. 2022b, 2023b; Brinchmann 2023; Schaerer et al. 2022). In both cases, the general strategy is to derive indirect diagnostics, that trace physically favourable conditions to LyC production and escape from the ISM (Choustikov et al. 2024b). These include a variety of different indirect tracers, including properties of Lyα emission (Jaskot & Oey 2014; Henry et al. 2015; Verhamme et al. 2015, 2017; Steidel et al. 2018; Pahl et al. 2021; Naidu et al. 2022; Choustikov et al. 2024a), high [O III] λ5007/[O II] $\lambda\lambda$3726, 3728 (O_{32}) ratios (Nakajima & Ouchi 2014), particularly negative UV continuum slopes (β) (Chisholm et al. 2022), low amounts of UV attenuation (Saldana-Lopez et al. 2022), Mg II $\lambda\lambda$2796, 2804 doublet ratios (Chisholm et al. 2020), strong C IV $\lambda\lambda$1548, 1550 emission (Schaerer et al. 2022; Saxena et al. 2022b), S II deficits (Chisholm et al. 2018; Wang et al. 2021), relative sizes of resonant line surface brightness profiles (Choustikov et al. 2024a; Leclercq et al. 2024) and multivariate models (Mascia et al. 2023b; Choustikov et al. 2024b).

The primary limitation is that the vast majority of the methods used to infer these properties on a case-by-case basis require spectroscopic information about a given galaxy, which is expensive (particularly in comparison to photometric surveys). Furthermore, studies that are able to make use of photometric observations to constrain certain parameters (primarily by using SED fitting) often require making assumptions about the others (particularly $f_{\rm esc}$) being constant or evolving on a population level only (e.g. Boyett et al. 2022; Simmonds et al. 2023, 2024b). Finally, performing advanced SED fitting over a large galaxy sample is a very computationally expensive and time-consuming exercise.

Given the availability of unprecedented photometric data from *JWST*, the objective of the present work is to develop a model to infer the total escaping output of ionizing photons of a given source based on *JWST* NIRCAM photometric measurements. To do this, we build an implicit likelihood inference (ILI) pipeline developed using LTU-ILI (Ho et al. 2024) trained on dust-attenuated mock photometry of a statistical sample of representative high-redshift galaxies from the SPHINX[20] simulation (Rosdahl et al. 2022; Katz et al. 2023a). This pipeline is able to make accurate and fast predictions for the ionizing photon contribution ($\dot{n}_{\rm ion}$) of individual sources with reliable uncertainties, based on filters used by the *JWST* Advanced Deep Extragalactic Survey[2] (JADES: Eisenstein et al. 2023a). Using public data from JADES, we aim to explore the redshift evolution of $\dot{n}_{\rm ion}$ for a sample of 4,559 photometrically selected galaxies at high redshift. Finally, we will combine these measurements to constrain the evolution of the global ionizing photon production rate ($\dot{N}_{\rm ion}$), allowing us to investigate the redshift evolution of reionization in the GOODS-S field.

This paper is arranged as follows. First, in Section 2 we outline the pipeline that we have built to predict $\dot{n}_{\rm ion}$ based on *JWST* photometry. In Section 3 we benchmark our inference pipeline and compare it with another SED-fitting method. Next, in Section 4 we apply this pipeline to a sample of JADES galaxies imaged using *JWST* NIRCam and characterise the ionizing photon contributions of this population of galaxies. Using this, we then compute the evolution of the ionized fraction of the IGM. Finally, we conclude in Section 5.

Throughout this paper, we assume a flat ΛCDM cosmology with cosmological parameters compatible with Planck Collaboration et al. (2014)[3] as well as a primordial baryonic gas of hydrogen and helium, with mass contents of $X = 0.75$ and $Y = 0.25$, respectively.

2 A MODEL FOR ESCAPING IONIZING LUMINOSITY WITH IMPLICIT LIKELIHOOD INFERENCE

ILI, also known as simulation-based inference (SBI) or likelihood-free inference (LFI), is a class of methods to infer the statistical relationship between the observed data ($\boldsymbol{X}$) and the underlying parameters of a model that generated the data ($\boldsymbol{\theta}$). For a thorough review see for example, Marin et al. (2011) or Cranmer et al. (2020). To infer $\boldsymbol{\theta}$ from $\boldsymbol{X}$, the Bayes theorem states that the posterior distribution of $\boldsymbol{\theta}$ is given by

$$\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{X}, I) \propto \mathcal{L}(\boldsymbol{X}|\boldsymbol{\theta}, I)\pi(\boldsymbol{\theta}|I), \tag{1}$$

where $\mathcal{L}(\boldsymbol{X}|\boldsymbol{\theta}, I)$ is the likelihood of the data given the model, $\pi(\boldsymbol{\theta}|I)$ is the prior distribution of the model parameters, and I denotes the remaining information required to specify the model. In many applications, the likelihood function may be unknown or computationally intractable while the mapping $\boldsymbol{\theta} \rightarrow \boldsymbol{X}$ is available. Thus, ILI relies on a "simulator" which can or has generated such synthetic data to populate a high-dimensional space of model parameters and observed

[1] Particularly in the case of sources at apparently extreme redshifts (e.g. Donnan et al. 2023).

[2] In principle, the method outlined in this paper is extendable to almost any other *JWST* survey. However, we have focused on JADES because it is a particularly deep survey with a large number of filters, making it an ideal proving ground.

[3] This is chosen to be consistent with the training data from the SPHINX[20] simulation (Rosdahl et al. 2022).

data (see Figure 1 in Ho et al. 2024). In turn, this can be used to infer the *distribution* of plausible model parameters that may have generated the observed data by slicing the space at the observed data.

In this work, we opt for the neural posterior estimation method (Papamakarios & Murray 2016; Greenberg et al. 2019), which directly emulates the posterior distribution. This is particularly suitable because in our case we have a single model parameter ($\dot{n}_{\rm ion}$) and a 13-dimensional space of observed data. Specifically, to be consistent across all sources, we use photometric magnitudes in the F115W, F150W, F200W, F277W, F335M, F356W, F410M and F444W filters normalised by the apparent UV magnitude ($m_{\rm AB}^{1500}$), three colours (F115W-F150W, F150W-F277W, and F277W-F444W) as well as $m_{\rm AB}^{1500}$ and redshift. However, other flavours of ILI exist such as the neural likelihood estimation (Alsing et al. 2018; Papamakarios et al. 2018) or the neural ratio estimation (Hermans et al. 2019).

In case of neural posterior estimation, we wish to approximate the "true" posterior $\mathcal{P}(\boldsymbol{\theta}|\boldsymbol{X}, I)$ with the neural posterior $\hat{\mathcal{P}}(\boldsymbol{\theta}|\boldsymbol{X}, I)$ while only having access to samples $\mathcal{D}_{\rm train} = \{\boldsymbol{X}_i, \boldsymbol{\theta}_i\}$ from the simulator. The neural posterior may be decomposed as

$$\hat{\mathcal{P}}(\boldsymbol{\theta}|\boldsymbol{X}, I) = \frac{\pi(\boldsymbol{\theta}|I)}{p(\boldsymbol{\theta}|I)} q_{\boldsymbol{w}}(\boldsymbol{\theta}|\boldsymbol{X}, I), \tag{2}$$

where $p(\boldsymbol{\theta}|I)$ is the proposal prior representative of the distribution of $\boldsymbol{\theta}$ in the simulated (training) data and $q_{\boldsymbol{w}}(\boldsymbol{\theta}|\boldsymbol{X}, I)$ is the neural network output. Although typically $q_{\boldsymbol{w}}$ is modelled with a normalizing flow (Papamakarios et al. 2019), in our case $\boldsymbol{\theta}$ is only 1-dimensional and thus we opt for a mixture density network (Bishop 1994). Specifically, we use a Gaussian mixture density network to model $q_{\boldsymbol{w}}$, where the neural network with weights and biases $\boldsymbol{w}$ outputs the parameters of the mixture (mean and standard deviation of each component of the mixture). Furthermore, we also assume the prior and proposal distributions to be identical. During training, the network parameters $\boldsymbol{w}$ are optimized using a loss function

$$L = -\sum_{i \in \mathcal{D}_{\rm train}} \log \hat{\mathcal{P}}(\boldsymbol{\theta}_i|\boldsymbol{X}_i, I), \tag{3}$$

introduced by Papamakarios & Murray (2016). We implement the neural posterior estimator using LTU-ILI[4] pipeline introduced by Ho et al. (2024).

In order to train the model, we use 13,800 mock line-of-sight dust-attenuated photometric observations of star-forming galaxies from SPHINX[20] (Rosdahl et al. 2018, 2022), a cosmological radiation hydrodynamical simulation of reionization in a 20 cMpc box with sufficient resolution to resolve the multi-phase ISM in a large population of constituent galaxies. Specifically, this data-set consists of a sample of 1,380 star-forming galaxies at $z = 10, 9, 8, 7, 6, 5,$ and 4.64. These galaxies were selected to have 10 Myr averaged SFR $\geqslant 0.3\,{\rm M_\odot yr^{-1}}$, so that they form a representative sample of galaxies that could be observed by a flux-limited *JWST* survey (Choustikov et al. 2024b). Available as part of the SPHINX[20] Public Data Release (SPDRv1, Katz et al. 2023a), each galaxy has been post-processed with `RASCAS` (Michel-Dansac et al. 2020) to simulate the self-consistent generation and propagation of an SED consisting of the stellar continuum, nebular continuum and nebular emission lines. A peeling algorithm (e.g. Yusef-Zadeh et al. 1984; Zheng & Miralda-Escudé 2002; Dijkstra 2017) was used to mock observe these dust-attenuated SEDs along ten consistent lines-of-sight, producing photometric images and magnitudes in *JWST* NIRCam filters. Comparisons between mock SPHINX[20] and JADES photometry and colour have been carried out, confirming that this is a representative sample (see Figs. 15 and 16 of Katz et al. 2023a). A complete description of the methods used to generate this data-set are provided in Katz et al. (2023a) and Choustikov et al. (2024b).

[4] `https://github.com/maho3/ltu-ili`

Hyperparameter	Optimized	Value
Number of hidden features	✓	21
Number of mixture components	✓	3
Optimizer learning rate	✓	8.932×10^{-4}
Training batch size	✓	45
Early stopping criterion	✓	13
Validation fraction	✗	0.2
Gradient norm clipping	✗	5

Table 1. Selected hyperparameters of the ILI model predicting $\log \dot{n}_{\rm ion}$ from JADES filters and source redshift. The hyperparameter naming follows the LTU-ILI interface and we outline the hyperparameter optimization routine in Section 2.

We train the model to predict $\log_{10} \dot{n}_{\rm ion}$, apply standard scaling to both the features and targets, and opt for a 20-80% test-train split by galaxies, not by individual lines-of-sight, to ensure that a single galaxy is not present in both splits. Furthermore, to make training more robust, we only use galaxies with $f_{\rm esc} \geqslant 10^{-6}$ to remove a small tail of outliers[5] and ensure that the full distribution of $\dot{n}_{\rm ion}$ values are represented in the training set. We use `Optuna` (Akiba et al. 2019) to optimize the following hyperparameters: number of hidden features in the network, number of mixture components, optimizer learning rate, training batch size and the early stopping criterion. We run `Optuna` for 1,000 trials to find the best hyperparameters and optimize the mean of Eq. (3) in a 10-fold cross-validation across galaxies. We list the selected hyperparameters to predict $\log \dot{n}_{\rm ion}$ from the JADES filters and redshift in Table 1.

Having trained the model, we can draw samples from $\hat{\mathcal{P}}(\boldsymbol{\theta}|\boldsymbol{X}, I)$. When testing the model on simulated data without uncertainties, we either draw 1,000 samples from the learnt posterior or summarize those draws with the maximum posterior value and an asymmetric 1σ uncertainty around it. On the other hand, when applying the model to observational data with uncertainties, we assume the uncertainties to be Gaussian such that $\boldsymbol{X} \pm \Delta\boldsymbol{X}$ and re-sample $\boldsymbol{X}$ 500 times, each time sampling 1,000 draws from the posterior. In doing so, we propagate both the model and photometric uncertainties into the prediction of $\boldsymbol{\theta}$.

In Fig. 1 we compare the predicted $\dot{n}_{\rm ion}$ with the true $\dot{n}_{\rm ion}$ of SPHINX[20] galaxies, isolating the sample of mock observations at each redshift in our sample. We note that we train a single model with redshift as a feature as opposed to training a separate model for each redshift bin. In all cases, we find that the running mean of the distribution matches the one-to-one line well, with the complete sample having a median absolute error (MAE) of 0.31 dex. The model performs particularly well for sources with $\log_{10}(\dot{n}_{\rm ion}/[{\rm photon/s}]) > 51$, struggling more with the LyC-dimmest sources at the highest redshifts for which training data is limited. For completeness, we perform a variety of other benchmark tests on the model. These are discussed in Appendix A.

Finally, we highlight that this method allows us to predict the *global* escaped ionizing luminosity of high-redshift galaxies without having to dust-correct observations or assume some model for the LyC escape fraction. As a result, this method is completely self-consistent, simple and efficient; as compared to traditional SED-fitting methods.

[5] Doing so improves the general performance of the model, as machine learning methods can struggle to reproduce outliers.

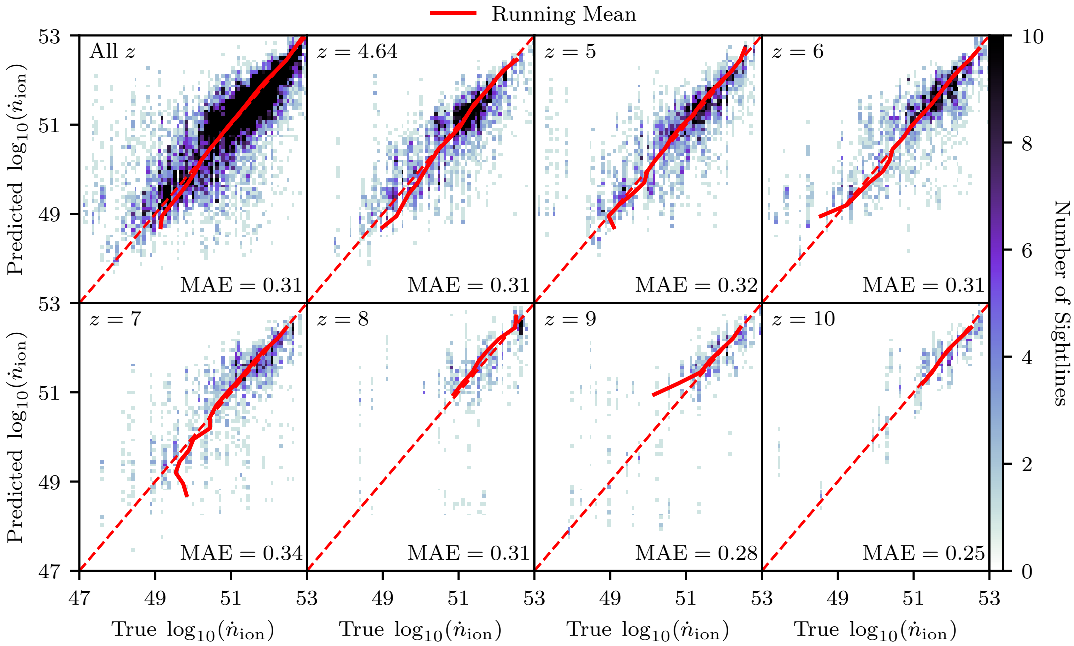

Figure 1. Histogram of $\dot{n}_{\rm ion}$ predicted using the ILI pipeline applied to mock dust-attenuated photometry of SPHINX^{20} galaxies as well as the true values computed using RASCAS, broken down by redshift bins. We include the running means in red as well as the median absolute error (MAE) for each redshift bin, showing how well the model performs in this validation experiment.

3 BENCHMARKING THE MODEL

3.1 Comparison with BAGPIPES

While we have shown that our model is accurate with well-behaved uncertainties, it is important to compare the efficacy of this method to that of a traditional SED-fitting code. To this end, we use BAGPIPES (Carnall et al. 2018) with priors given in Table 2. For a random sample of 30 galaxies from the SPHINX^{20} database we use BAGPIPES to find the best-fit model spectrum for each line-of-sight mock photometry. We note that to ensure that this is a fair test, we only sample galaxies from the test set of the ILI model introduced in Section 2.

The model spectrum is converted to a photon flux and integrated over rest-frame wavelengths $\leqslant 912$ Å to compute $\dot{n}_{\rm ion}$. To model the escape fraction, we repeat the process described in Chisholm et al. (2022) and Mascia et al. (2023b). For the former, we fit the UV spectral index (β, such that $f_\lambda \propto \lambda^\beta$) in the range $\lambda \in [1300, 1800]$ Å. These UV continuum slopes are then converted into an approximate escape fraction ($f_{\rm esc}^{\rm C22}$) using Equation 11 of Chisholm et al. (2022). For the latter, we use the effective half-light radius (R_e) of the galaxy mock observed in the F115W filter and measured by PHOTUTILS (Bradley et al. 2016; see also Katz et al. 2023a; Choustikov et al. 2024a), the ratio of [O III] λ5007/[O II] $\lambda\lambda$3726,3728 (O_{32}[6]) predicted by BAGPIPES as well as β as discussed above. These are then combined using the multivariate estimator given by Equation 1 of Mascia et al. (2023b) to produce the estimate of $f_{\rm esc}^{\rm M23}$. The product of the calculated $\dot{n}_{\rm ion}$ and one of the inferred escape fractions is taken as the final prediction. Finally, we repeat this process by also predicting $\dot{n}_{\rm ion}$ using the ILI model described in Section 2.

This entire process is demonstrated in Fig. 2, where we display all of the necessary information for all 10 lines-of-sight for a randomly selected SPHINX^{20} galaxy at redshift $z = 6$. On the *top*, we show the full mock SPHINX^{20} SED (*in colour*), the mock *JWST* NIRCam photometry in the JADES filters (*green*), as well as the best-fit BAGPIPES SED (*black*), confirming that it is a good match. On the *bottom*, we show the full ILI posterior distribution for each sight line (with matched colours). In each case, we include the true value of $\dot{n}_{\rm ion}$, computed directly from RASCAS ($\dot{n}_{\rm ion}^{\rm SPHINX^{20}}$, *solid*), the best-

Description	Parameter	Prior
SFH: Double Power Law		
Time of SFR Peak	τ / Gyr	$\mathcal{U}(0, 10)$
Falling Slope	$\log_{10}\alpha$	$\mathcal{U}(-2, 3)$
Rising Slope	$\log_{10}\beta$	$\mathcal{U}(-2, 3)$
Stellar Mass	$\log_{10}(M_*/M_\odot)$	$\mathcal{U}(6, 10)$
Metallicity	$\log_{10}(Z/Z_\odot)$	$\mathcal{U}(-1, 1)$
Dust: SMC (Gordon et al. 2003)		
V-band Attenuation	A_V	$\mathcal{U}(0, 2)$
Birth Cloud Reduction	η	2
Absorption Exponent	n	$\mathcal{G}(0.7, 0.3)$
Nebular		
Ionization Parameter	$\log_{10} U$	-2
Redshift		
Source Redshift	z	$z^{\rm SPHINX^{20}}$

Table 2. BAGPIPES priors used to fit SPHINX^{20} photometry in the JADES filters, inspired by those used in Carnall et al. (2018). $\mathcal{U}(u_{\rm min}, u_{\rm max})$ denotes a uniform distribution between $u_{\rm min}$ and $u_{\rm max}$, $\mathcal{G}(\mu, \sigma)$ a Gaussian distribution with mean μ and standard deviation σ, and if a single value is given then the parameter is fixed.

[6] We note that BAGPIPES is not designed to predict emission line fluxes based on wide and medium band photometry. Therefore, these values of O_{32} are likely to be a large source of uncertainty.

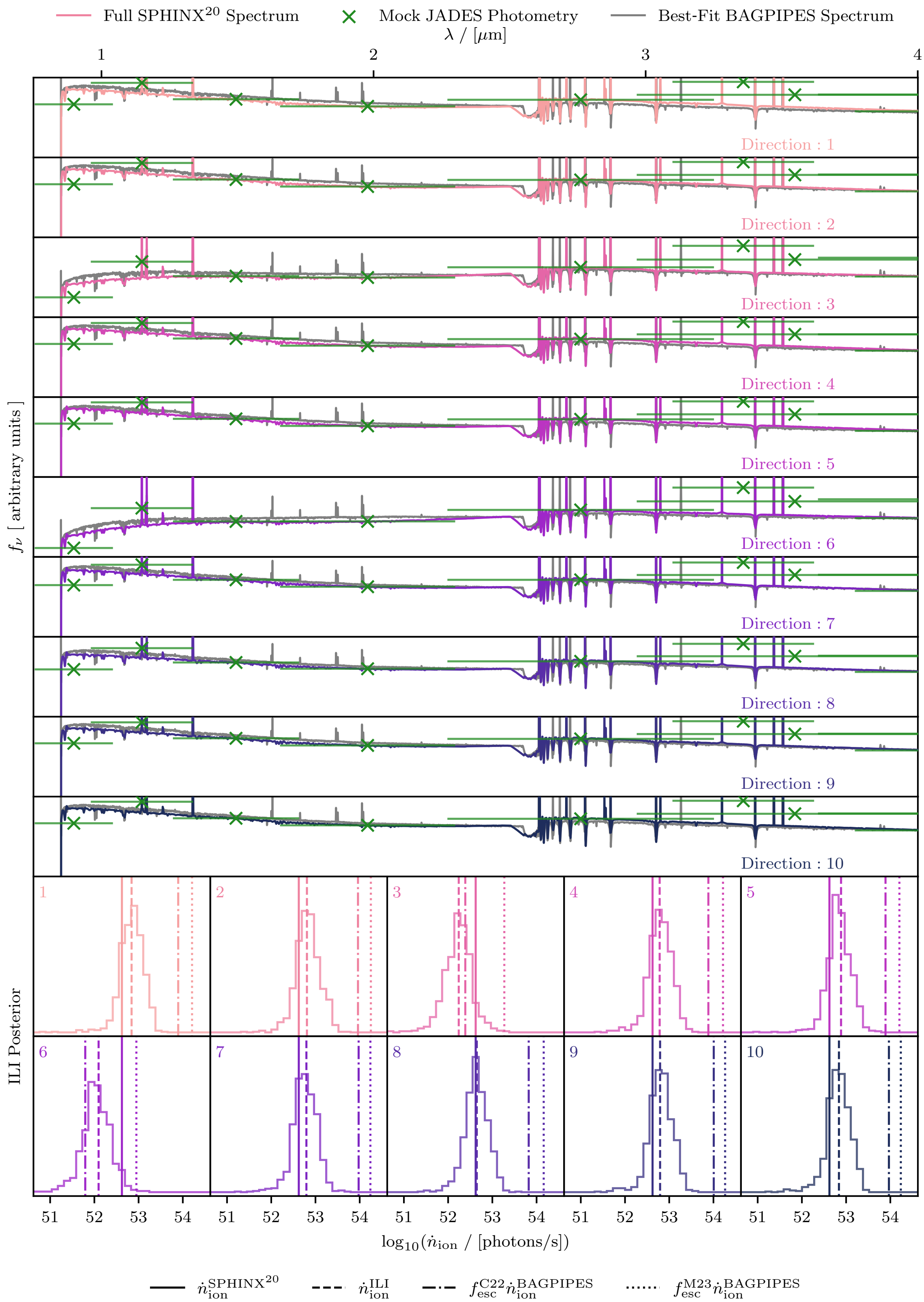

Figure 2. Comparison between the implicit likelihood inference (ILI) and `BAGPIPES` methods to inferring $\dot{n}_{\rm ion}$ for a random SPHINX^{20} galaxy at redshift $z = 6$. In the *long* panels, we include the mock SED (colour), mock *JWST* NIRCam photometry in the JADES filters (green), and best-fit `BAGPIPES` SED (black) for each line of sight. On the *bottom*, we include the full posterior distributions for $\dot{n}_{\rm ion}$ as sampled by the ILI pipeline, along with the true (*solid*) value, best-fit ILI (*dashed*) and best-fit `BAGPIPES` predictions (*dot-dashed*) using $f_{\rm esc}$ models from Chisholm et al. (2022) and Mascia et al. (2023b).

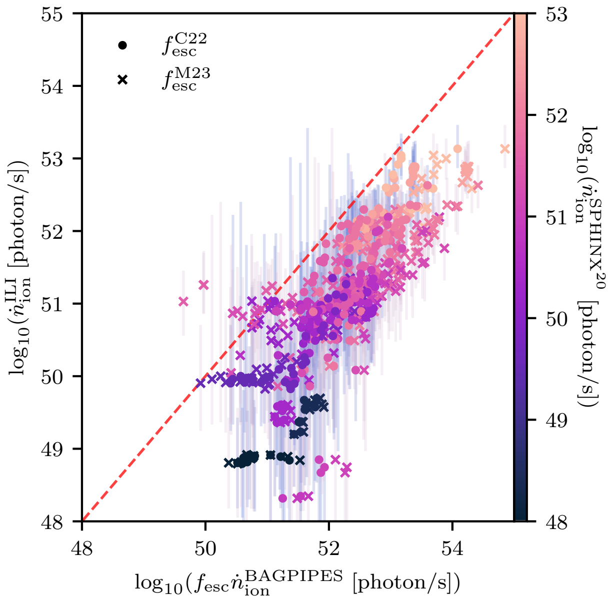

Figure 3. Comparison between the escaped ionizing luminosity of 30 SPHINX^{20} galaxies as predicted using the ILI pipeline ($\dot{n}_{\rm ion}^{\rm ILI}$) and BAGPIPES ($f_{\rm esc}\dot{n}_{\rm ion}^{\rm BAGPIPES}$), coloured by the true value computed using RASCAS ($\dot{n}_{\rm ion}^{\rm SPHINX^{20}}$). Because this quantity needs to also be inferred, we show BAGPIPES values computed using two different predictions for the LyC escape fraction, using the method from Chisholm et al. 2022 (*points*) and Mascia et al. 2023b (*crosses*).

fit ILI prediction ($\dot{n}_{\rm ion}^{\rm ILI}$, *dashed*), along with the best-fit BAGPIPES predictions ($\dot{n}_{\rm ion}^{\rm BAGPIPES}$) with $f_{\rm esc}^{\rm C22}$ (*dot-dashed*) and $f_{\rm esc}^{\rm M23}$ (*dotted*). Here, we find that the ILI-inferred values are typically much more accurate and consistent than those inferred from BAGPIPES model SEDs, despite the fact that BAGPIPES is inferring the SED well. Furthermore, it is clear to see that the lines-of-sight for which this is not the case (3 and 6) are significantly dustier than the others. In these cases, the ILI pipeline performs as expected and produces a more uncertain, broader posterior that tends to under-predict $\dot{n}_{\rm ion}$. However, the larger error bars confirm that the model is behaving as required. In practise, a similar behaviour was found for lines-of-sight with extremely blue spectral slopes, in which case the model would over-predict $\dot{n}_{\rm ion}$. Finally, it is interesting to compare the relative success of the two BAGPIPES results. Here, we find that these methods tend to over-predict $\dot{n}_{\rm ion}$ for a given source. However, in all cases, the inferred $\dot{n}_{\rm ion}$ based on the LyC escape fractions predicted by Chisholm et al. (2022) is more accurate than those predicted by Mascia et al. (2023b).

While this is reassuring, it is crucial to confirm that these trends exist for a larger sample of galaxies. To this end, Fig. 3 shows a comparison between $\dot{n}_{\rm ion}^{\rm ILI}$ and $\dot{n}_{\rm ion}^{\rm BAGPIPES}$, with points coloured by the true value of $\dot{n}_{\rm ion}^{\rm SPHINX^{20}}$. We find that the BAGPIPES methods tend to consistently over-predict the escaped ionizing luminosity, as compared to ILI-predicted values. Next, we find that the β slope method of Chisholm et al. (2022) performs slightly better than the multivariate method given by Mascia et al. (2023b). We note that in trials, $f_{\rm esc}^{\rm M23}$ was more accurate for galaxies with the greatest LyC escape fractions, with issues otherwise being driven by the dependence on O_{32} and R_e. Furthermore, the method proposed by Chisholm et al. (2022) was previously shown to fit SPHINX^{20} galaxies well (see Figure 8 of Choustikov et al. 2024b), suggesting that the over-estimates are driven by BAGPIPES itself.

It is now necessary to understand where the disagreements highlighted in Fig. 3 come from. To explore this further, we proceed to compare both prediction methods directly with the true value of $\dot{n}_{\rm ion}$ in Fig. 4. In the *left* panel, we plot the true value from SPHINX^{20} versus the ILI prediction, along with associated error bars, while on the *right*, we plot the true value against the BAGPIPES values, using $f_{\rm esc}^{\rm M23}$ (*purple*), $f_{\rm esc}^{\rm C22}$ (*pink*), as well as the escape fraction computed using RASCAS, $f_{\rm esc}^{\rm SPHINX^{20}}$ (*salmon*). Here, we learn several things. First, the ILI method is the most accurate, consistently agreeing across the whole sample. Moreover, in the rare cases where there are strong disagreements, we find significantly wider uncertainties, showing that the model is performing as expected. Next, we find that the UV slope-corrected value from Chisholm et al. (2022) is reasonably similar to that computed using the true value of $f_{\rm esc}$, albeit with more scatter. In contrast, the values predicted using the multivariate model of Mascia et al. (2023b) seem to systematically over-predict $\dot{n}_{\rm ion}$ by around an order of magnitude on average. Together, this highlights the fact that a large part of the uncertainty seen in Fig. 2 is due to the assumed escape fraction models. Finally, we note that while the ILI-inferred values have no apparent bias, the use of BAGPIPES along with the 'true' value of $f_{\rm esc}$ from SPHINX^{20} tends to over-predict $\dot{n}_{\rm ion}$ for galaxies releasing the most ionizing photons, confirming that this particular over-estimate arises due to imperfect fitting with BAGPIPES. This is consistent with previous work which has shown that SED-fitting methods tend to over-predict $\xi_{\rm ion}$ compared to emission line-based techniques, e.g. using PROSPECTOR (Johnson et al. 2021) (see Figure 7 of Simmonds et al. 2024b). This may be significant for discussions of whether reionization is driven by the brightest or faintest leakers (e.g. Finkelstein et al. 2019; Naidu et al. 2020), by artificially boosting the impact of the strongest sources.

Recently, Muñoz et al. (2024) discussed the possibility that present-day *JWST* observational constraints on $\rho_{\rm UV}$ (e.g. Donnan et al. 2024) and $\xi_{\rm ion}$ (e.g. Simmonds et al. 2024b) suggest that there might be *too* many LyC photons, reionizing the Universe too early for alternative probes of the IGM. Muñoz et al. (2024) extrapolated the $f_{\rm esc}$ model of Chisholm et al. (2022). However, the results of Fig. 4 suggest that doing so (alongside SED-fitted estimates of $\xi_{\rm ion}$) is likely to over-estimate the number of escaping LyC photons. These effects will compound to contribute to this conclusion by artificially boosting the ionizing contribution of high-redshift galaxies.

3.2 Why is there a Difference Between ILI and BAGPIPES?

As we have shown in the previous Section, the predicted values of $\dot{n}_{\rm ion}$ using the ILI pipeline and BAGPIPES can be dramatically different. As such, it is crucial to discuss why this might be the case, especially given the fact that both methods work by effectively searching for the best-fit SED from a reference sample.

The similarity between these two methods comes from the fact that ILI is trained on a library of mock photometry (based on underlying SEDs, see discussion in Section 4.3 of Katz et al. 2023a) and then interpolates between them to infer $\dot{n}_{\rm ion}$, while BAGPIPES generates a large sample of model SEDs, processes them to compute photometry and iterates this process to find a best-fitting model, which is then integrated to compute $\dot{n}_{\rm ion}$ (see also Carnall et al. 2018). These are effectively the same operations, albeit in different orders. However, while both the SPHINX^{20} and BAGPIPES SEDs account for nebular contributions with CLOUDY, the primary difference is that they use different stellar population synthesis models. Namely, SPHINX^{20} makes

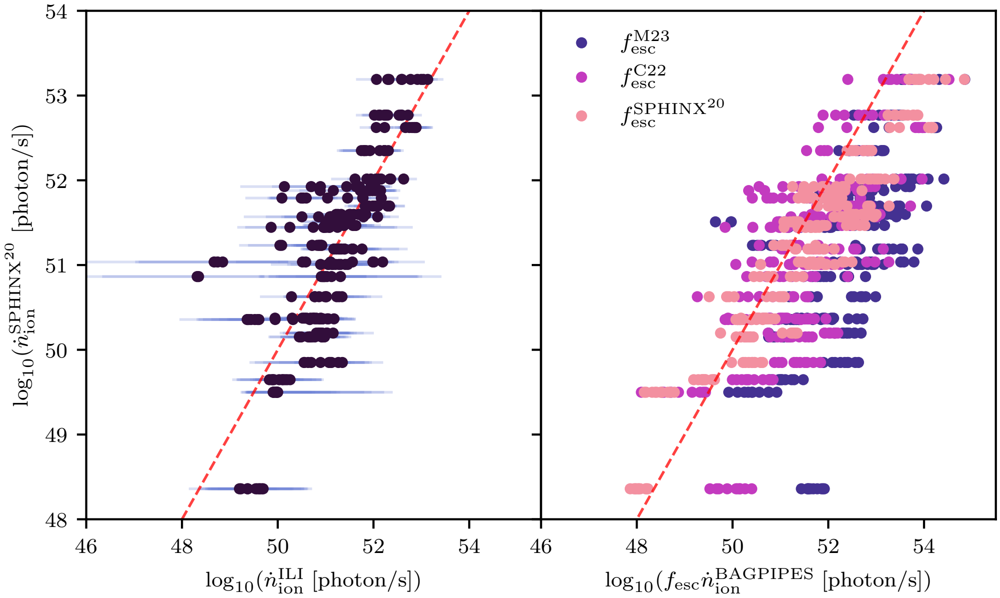

Figure 4. (*Left*): Comparison between the true value of escaped ionizing luminosity ($\dot{n}_{\rm ion}^{\rm SPHINX^{20}}$) of training-set SPHINX[20] galaxies computed using RASCAS with predictions from the ILI pipeline. We include asymmetric 1σ error bars based on the ILI posteriors, as discussed in Section 2. *Right*: Comparison between the true value of $\dot{n}_{\rm ion}$ with predictions using the best-fit BAGPIPES SED. Here, we use LyC escape fractions computed using methods proposed by Mascia et al. (2023b) (*purple*), Chisholm et al. (2022) (*pink*), as well as the true values computed by RASCAS (*salmon*).

use of BPASS SEDs (Stanway & Eldridge 2018) while BAGPIPES uses the 2016 version of Bruzual & Charlot (2003) models. This may introduce a systematic bias in the production of ionizing flux, with binary evolution increasing the hardness of the spectrum blueward of 912 Å, particularly at low metallicities (Stanway et al. 2016; Eldridge et al. 2017). This compounds with the results of Figs. 3, 4, confirming that the BAGPIPES method significantly over-predicts $\dot{n}_{\rm ion}$.

The second difference is that galaxies are not made up of a single H II region with some intrinsic metallicity and ionization parameter. As such, the fact that galaxies in our SPHINX[20] sample have resolved multi-phase ISMs, as well as contributions from multiple H II regions with non-zero escape fractions (see also Nakajima et al. 2013; Law et al. 2018; Ramambason et al. 2022) and diffused ionised gas (see also Zhang et al. 2017; Kewley et al. 2019) all contribute to producing more realistic nebular emission. However, one place where SPHINX[20] falters similarly to BAGPIPES is in the fact that it uses equilibrium abundances for non-hydrogenic species. It has been shown that galaxy-scale simulations with non-equilibrium thermochemistry produce dramatically different abundance distributions (Katz 2022; Katz et al. 2022a), however we leave explorations of how this effects the complete forward-modelling of SEDs to future work. Finally, our analysis has also taken into account the impact of the nebular continuum on the escaped ionizing luminosity, which has recently been shown to be important (Simmonds et al. 2024a).

The third difference is that the star-formation histories (SFHs) of SPHINX[20] galaxies (see Figure 7 of Katz et al. 2023a) are governed by a realistic environment, capturing effects such as mergers (e.g. Kaviraj et al. 2015; Fensch et al. 2017; Pearson et al. 2019) and gas accretion from the circumgalactic medium (Conselice et al. 2013; Putman 2017; Luo et al. 2021). In turn, simple SFH prescriptions employed by codes like BAGPIPES do not capture this effect, leading to discrepancies in inferred quantities (Lower et al. 2020; Narayanan et al. 2024). Using non-parametric SFHs it may, however, be possible to mimic the more realistic SFHs that simulations are able to generate (e.g. Wan et al. 2024), though such exercises are beyond the scope of the present work.

Next, we must consider the fact that in this exercise, we are fundamentally trying to predict an angle-averaged property by using line-of-sight observables. While our ILI model has been trained on multiple sight-lines for each object and therefore produces fairly consistent outputs (see Fig. 2), this is particularly a challenge for indirect models for the LyC escape fraction (Choustikov et al. 2024a,b). While we do not discuss this further, another potential option here is to use a framework such as ILI to predict angle-averaged $f_{\rm esc}$ based on the same photometric information to then combine with traditional SED fitting.

Another important issue is that SPHINX[20] makes use of a Kroupa-like initial mass function (IMF, Kroupa 2001), whereas BPASS and BAGPIPES can be tweaked to account for differing IMFs (e.g. Stanway & Eldridge 2023). This is particularly important in the context of recent work suggesting deviations from a standard IMF at high redshifts and low metallicities (Cameron et al. 2023), that in this context would particularly influence the ionizing radiation field[7].

A further consideration to note is that of dust attenuation. The assumptions taken in the post-processing of dust in SPHINX[20] have already been discussed extensively (Katz et al. 2023a; Choustikov et al. 2024b), however the key issue for our work is the fact that the attenuation law used is fixed to scale with metallicity. In principle, this can be overcome by reprocessing each SPHINX[20] galaxy with RASCAS using a variety of different attenuation laws, thus expanding the training set and building this flexibility into the ILI model. In contrast, SED modelling with BAGPIPES is able to self-consistently modify and fit the levels and kinds of dust attenuation. This likely

[7] Naturally, a self-consistent treatment of this would also affect the distribution and feedback of stars in the SPHINX[20] simulation. We leave explorations of this to future work.

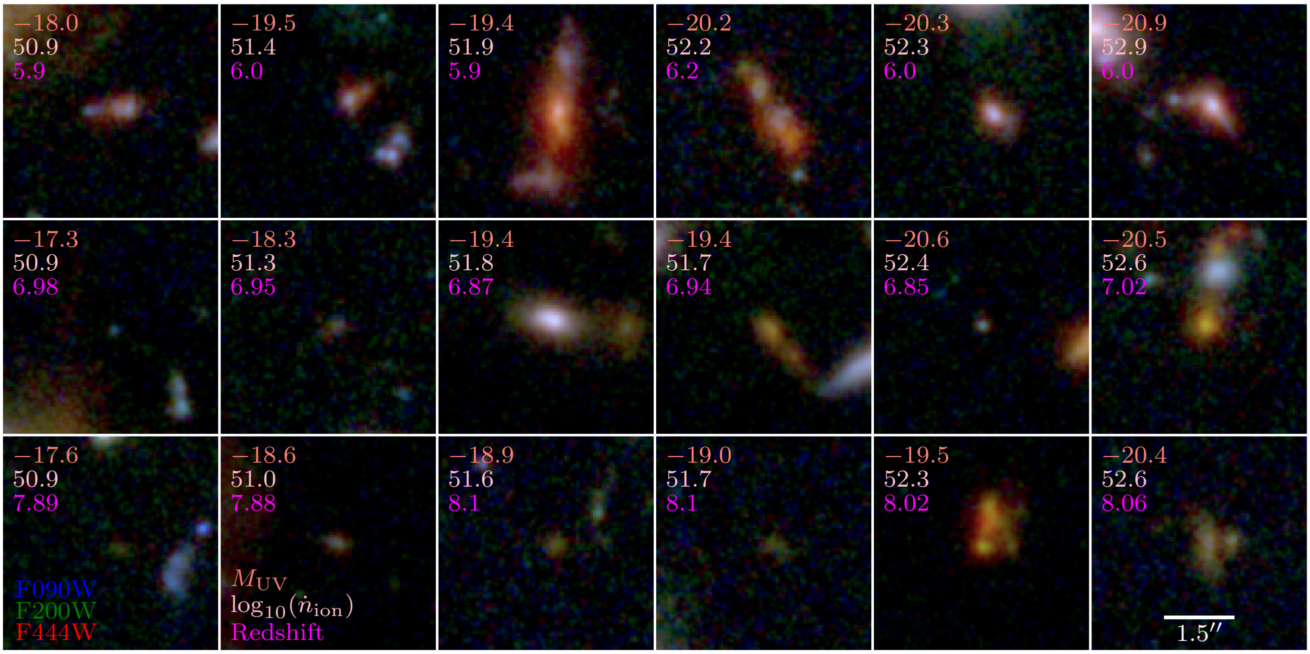

Figure 5. Thumbnail images of 18 galaxies in GOODS-S imaged by *JWST* NIRCam as part of JADES (Eisenstein et al. 2023a). RGB images are made using F444W in the red channel, F200W in the green and F090W in the blue channel. For each galaxy, we provide the absolute UV luminosity, $\dot{n}_{\rm ion}$ predicted by our ILI pipeline as well as photometric redshift. Galaxies are shown in order of their ionizing photon contributions at each redshift.

explains why it is better able to infer $\dot{n}_{\rm ion}$ for particularly dusty sight lines[8] (see sight-lines 3 and 6 of Fig. 2).

The final important difference is that the SPHINX^{20} simulation does not include AGN. As such, the model presented in this paper will only be able to account for the star-forming component of galaxies. This should not be a significant issue for our purposes, as AGN are expected to be a subdominant portion of the sources, whose contribution to reionization is expected to be small (e.g. Dayal et al. 2020; Trebitsch et al. 2021). For a complete discussion of the caveats implicit in the modelling and post-processing of our mock dataset of SPHINX^{20} galaxy photometry, the reader is invited to consult Katz et al. (2023a) and Choustikov et al. (2024b).

To conclude, we reiterate that the ILI method presented in this paper is generally more accurate than an SED-fitting alternative and produces self-consistent uncertainties while also taking significantly less time to run (with a speed-up of $\sim 10^2 - 10^3$). This makes it a comparable, if not superior option for investigating the contribution of a large number of galaxies imaged using photometry toward reionization.

4 PREDICTING THE ESCAPED IONIZING LUMINOSITIES FOR A POPULATION OF JADES GALAXIES

4.1 Application to JADES NIRCam Data

We now apply our ILI pipeline to real data to infer the ionizing photon luminosity of photometrically-observed galaxies. To do so, we use NIRCam Deep imaging (Rieke et al. 2023), taken and publicly released[9] as part of the *JWST* Advanced Deep Extragalactic Survey (JADES: Eisenstein et al. 2023a). These data are taken in the GOODS-S field, covering an area of ~ 25 arcmin^2. Specifically, we make use of magnitudes in the F115W, F150W, F200W, F277W, F335M, F356W, F410M, and F444W filters, computed using a Kron parameter of $K = 2.5$, which has been point spread function-convolved to the resolution in the F444W filter, as recommended in the data release. To complete our feature set, we also use photometric redshifts derived using `EAZY` (Brammer et al. 2008), as included in the JADES catalogue (Hainline et al. 2024). Apparent UV magnitudes, $m_{\rm AB}^{1500}$, are computed by fitting a power law ($f_\lambda \propto \lambda^\beta$) to the three filters nearest to rest-1500Å, selected for each redshift. For a full discussion of the approach as well as comparisons to spectroscopic redshifts the reader is directed to Hainline et al. (2024) and Rieke et al. (2023). Before proceeding, we make the following cuts to reduce our sample:

- We require a signal-to-noise ratio (S/N) in all filters redward of F200W to be greater than or equal to 3.
- We remove any sources that have been flagged as stars or that are affected by diffraction spikes.
- We remove any sources with $M_{\rm UV} \leqslant -23$ at $z > 6$ as these are likely to be dominated by AGN.

Following this process, we are left with a sample of 4,559 galaxies.

For each object, we use the ILI pipeline to predict $\dot{n}_{\rm ion}$ based on observed magnitudes in each filter normalised by $m_{\rm AB}^{1500}$, three colours (F115W-F150W, F150W-F277W, and F277W-F444W), and $m_{\rm AB}^{1500}$. In each case, we account for the model, photometric magnitude, and

[8] This is due to the fact that there is a limited number of such sight-lines in SPHINX^{20}.

[9] All of the JADES data used in this paper can be found on the MAST data-base at https://doi.org/10.17909/z2gw-mk31.

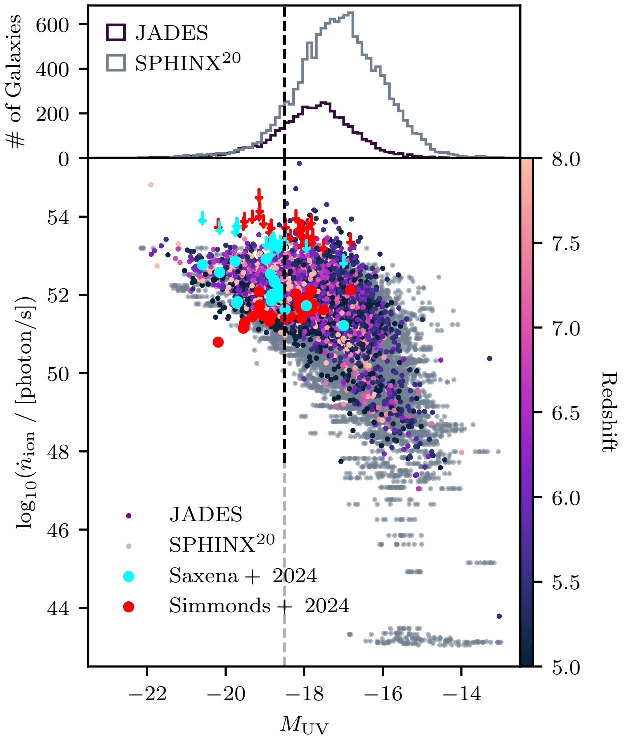

Figure 6. Escaping ionizing luminosity as a function of observed absolute UV magnitude for our sample of JADES (*coloured by redshift*) and SPHINX^{20} (*gray*) galaxies. Other observational data from Saxena et al. (2024) and Simmonds et al. (2024b) are included for comparison in *cyan* and *red*. In each case, we follow each papers' method to predict $f_{\rm esc}$. For the former, we use the multivariate model from Choustikov et al. (2024b), while in the latter we infer $f_{\rm esc}$ from the absolute UV magnitude based on the relation from Anderson et al. (2017). We also include a histogram of the observed absolute UV magnitudes for our sample of JADES (*black*) compared to SPHINX^{20} (*gray*) galaxies. Finally, the cut of UV-bright and UV-dim galaxies ($M_{\rm UV} = -18.5$) used elsewhere in this paper is also shown as a dashed line. Galaxies brighter than this value account for 20% of the sample.

redshift uncertainties by resampling as described in Section 2. As an overview, Fig. 5 shows 18 example galaxies from the JADES catalogue in redshift bins of $z \in \{6, 7, 8\}$. Here, we compile RGB images composed of F444W in the red channel, F200W in the green and F090W in the blue. For each object, we also list their observed absolute UV luminosity, predicted value of $\dot{n}_{\rm ion}$ from the ILI pipeline, and photometric redshift. Galaxies are shown in order of their ionizing photon contributions in each given redshift bin.

Next, in Fig. 6, we show the escaped ionizing luminosity of JADES galaxies as a function of their observed absolute UV magnitude, coloured by redshift. For comparison, we include spectroscopic measurements from Saxena et al. (2024) (*cyan*) as well as SED fitted predictions using PROSPECTOR from Simmonds et al. (2024b) (*red*). In both cases, we follow the reported methods of predicting $f_{\rm esc}$. In the first case, we use the multivariate model proposed by Choustikov et al. (2024b), while in the latter we use escape fractions inferred from the absolute UV magnitude ($M_{\rm UV}$), based on the VULCAN simulation (Anderson et al. 2017). However, we caution that the relation between $M_{\rm UV}$ and $f_{\rm esc}$ has been shown to be very dependent on stellar mass (see Figures 12 and 13 of Choustikov et al. 2024b) and is in general not a good predictor for $f_{\rm esc}$ (e.g. Flury et al. 2022b; Saxena et al. 2024; Choustikov et al. 2024b). Both sets of values are included, with intrinsic $\dot{n}_{\rm ion}$ shown as an *arrow* and escaped $\dot{n}_{\rm ion}$ given as *points*.

Here, we can see that there is some correlation between $\dot{n}_{\rm ion}$ and $M_{\rm UV}$. Galaxies with $M_{\rm UV} < -20$ are rare, but all have large escaped ionizing luminosities ($\dot{n}_{\rm ion} \gtrsim 10^{52}$ photons/s). We find that UV-dim galaxies with $M_{\rm UV} > -17$ are much more common but have much smaller values of $\dot{n}_{\rm ion}$, with all of these galaxies having $\dot{n}_{\rm ion} \lesssim 10^{53}$ photons/s. To illustrate the distribution of absolute UV magnitudes, we include a histogram (*top*) comparing the distribution of JADES galaxies to those from SPHINX^{20}. Beyond confirming that SPHINX^{20} galaxies are suitable analogues, this shows the sheer number of UV-dim galaxies in our sample. We define an absolute UV magnitude cut at $M_{\rm UV} = -18.5$ (shown as a *dashed* line), which is used to explore whether faint galaxies are the dominant contributors of ionizing photons during the epoch of reionization (Finkelstein et al. 2019, cf. Naidu et al. 2020).

Fig. 7 shows the inferred values of $\dot{n}_{\rm ion}$ as a function of redshift for all JADES galaxies in our sample, along with associated error bars. We colour points by their redshift uncertainty, to highlight the objects whose ionizing luminosity uncertainties are dominated by photometric redshift uncertainties. For comparison, we include other observational data from Saxena et al. (2024) (corrected using the $f_{\rm esc}$ relation of Choustikov et al. 2024b; *cyan*) and Simmonds et al. (2024b) (corrected with the $M_{\rm UV}$ relation of Anderson et al. 2017; *red*). Given we have uncertainties in both $\dot{n}_{\rm ion}$ and z, we use ROXY (Bartlett & Desmond 2023)[10], which provides an unbiased linear fit accounting for uncertainties in both x and y. We find a weak evolution with redshift, given by:

$$\log_{10}(\dot{n}_{\rm ion} \text{ / [photons/s]}) = (0.08 \pm 0.01)z + (51.60 \pm 0.06), \quad (4)$$

that we also plot (*lime*) with associated 3 σ uncertainties. We find that this matches our running mean (*blue*) well. Such a slow evolution with z is in agreement with previous works, which suggest little change in $\xi_{\rm ion}$ (e.g. Saxena et al. 2024; Simmonds et al. 2024b) as well as the LyC escape fraction (Mascia et al. 2023b). Next, we see that there is a small secondary population present, with $\dot{n}_{\rm ion}$ lower by about 2 dex. These are likely to be galaxies with particularly dusty sight lines, for which the model tends to struggle and under-estimates $\dot{n}_{\rm ion}$[11]. This can also be seen, due to the particularly large error bars of these objects (see also the discussion of sight lines three and six in Figure 2).

Next, we use the intrinsic ionizing luminosities ($\dot{n}_{\rm ion}^{\rm R22}$) and LyC escape fractions ($f_{\rm esc}^{\rm R22}$) from the SPHINX^{20} simulation to compute average values for each redshift bin. In doing so, we show the average of the product of these two quantities ($\langle f_{\rm esc}^{\rm R22} \dot{n}_{\rm ion}^{\rm R22} \rangle$, representing the ILI approach) as well as the product of their respective averages ($\langle f_{\rm esc}^{\rm R22} \rangle \langle \dot{n}_{\rm ion}^{\rm R22} \rangle$, representing the use of population-averaged statistics) in *gold*. We find that in general these two values do not agree, with the latter method over-predicting the average escaped ionizing luminosity by 0.5 dex toward the end of reionization. This emphasizes the fact that it is the angle-averaged product of these two quantities which is important to measure in order to accurately investigate galaxy contributions to reionization.

Finally, we can also see the fundamental UV magnitude limit derived from the JADES NIRCam depths (Eisenstein et al. 2023a). This leads to a reduction in the number of sources with redshift, as *JWST* is able to see fewer sources with the given S/N in each

[10] https://github.com/DeaglanBartlett/roxy
[11] In practise, it was found that the presence of this group of outliers did not affect the line of best-fit given in Equation 4.

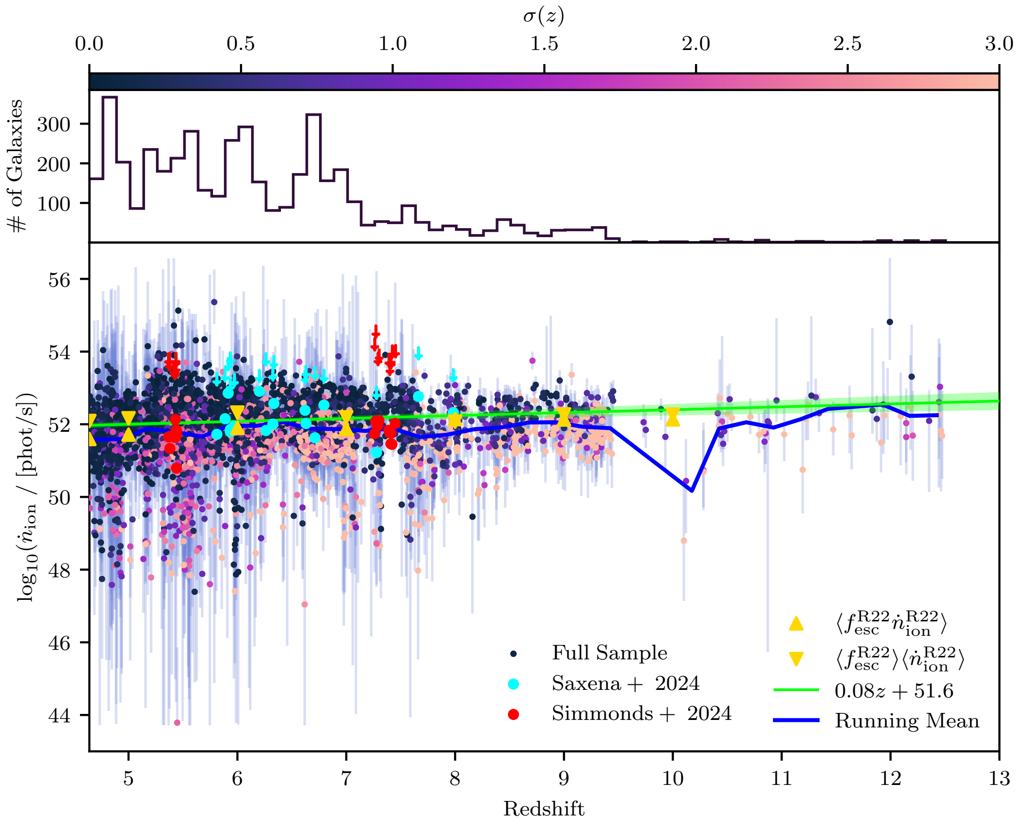

Figure 7. Predicted ionizing luminosities of *JWST* galaxies coloured by their photometric redshift uncertainties. Error bars are produced by resampling the model and photometric uncertainties, as described in the text. We include a running mean (*blue*) as well as a line of best-fit (*lime*), computed using ROXY (Bartlett & Desmond 2023). For comparison, we include data from Saxena et al. (2024) and Simmonds et al. (2024b), as described in Fig. 6. Finally, we also include global averages for escaped $\dot{n}_{\rm ion}$, computed at each redshift in SPHINX[20] (*gold*). This is to demonstrate the over-prediction which comes from studying these two quantities in isolation.

filter. We note that in practise, the trend seen in Fig. 7 remains fairly unchanged with respect to signal-to-noise cuts.

4.2 Implication for Reionization in GOODS-S

Now that we have predictions for the ionizing luminosity of a large number of galaxies in the GOODS-S field, we can reconstruct a reionization history for the survey volume. To do this, we integrate the ionizing luminosity contributions in numerous redshift bins, while also integrating the comoving volume of each bin, as follows:

$$\dot{N}_{\rm ion}(z) = \rho_{\rm UV}(z)\xi_{\rm ion}(z)f_{\rm esc} = \frac{\int_{z-\delta z}^{z+\delta z} {\rm d}z\, \dot{n}_{\rm ion}(z)}{\int_{z-\delta z}^{z+\delta z} {\rm d}V(z)}. \quad (5)$$

This tells us how many ionizing photons are being emitted by galaxies per Mpc^3 in a given redshift bin. This value can then be compared to various models of reionization. It is instructive to use Equation 26 from Madau et al. (1999):

$$\dot{N}_{\rm ion} = (10^{51.2}[{\rm photons/s/Mpc^3}])\, C\left(\frac{1+z}{6}\right)^3\left(\frac{\Omega_{\rm b}h_{50}^2}{0.08}\right)^2, \quad (6)$$

where $\Omega_{\rm b}$ is the baryonic density fraction of the Universe and C is the ionized hydrogen clumping factor, accounting for the fact that baryons are not uniformly distributed through the IGM. In particular, this model depends on a time-dependent clumping factor that is typically calibrated with large-scale simulations (e.g. So et al. 2014, see also Gnedin & Madau 2022 for a review).

Fig. 8 shows the integrated redshift evolution of $\dot{N}_{\rm ion}$ for all galaxies in our sample, as compared to the theoretical models from Finkelstein et al. (2019); Kulkarni et al. (2019a), Bayesian-inferred history from (Mason et al. 2019a) as well as various observational data (Simmonds et al. 2024b; Rinaldi et al. 2023; Mascia et al. 2023a). Finally, we include curves showing the number of ionizing photons required to ionize the neutral IGM for various clumping factors $C \in \{1, 3, 10\}$ given by Eq. (6). We find that our data is consistent with all of the observations, and predicted histories for the evolution of ionizing photon sources. It is also interesting to explore the question of whether reionization is driven by a small number of UV-bright sources or by a large number of UV-dim sources. To test this, we make a further cut in our data, computing $\dot{N}_{\rm ion}$ for all galaxies in our sample with $M_{\rm UV} < -18.5$ (*magenta*) and $M_{\rm UV} \geqslant -18.5$ (*purple*), accounting for the two groups respectively. We find that at late times ($z \lesssim 8$) the cohort of UV-dimmer galaxies (that account for 80% of the population) release more ionizing photons into the IGM overall, agreeing with previous work (e.g. Finkelstein et al. 2019). It is difficult to constrain the two groups' relative importance beyond this redshift due to the difficulty in observing dim galaxies with such a selection function at these distances.

However, there are several key points to discuss. The first is that, as noted previously, our model does not specifically include AGN.

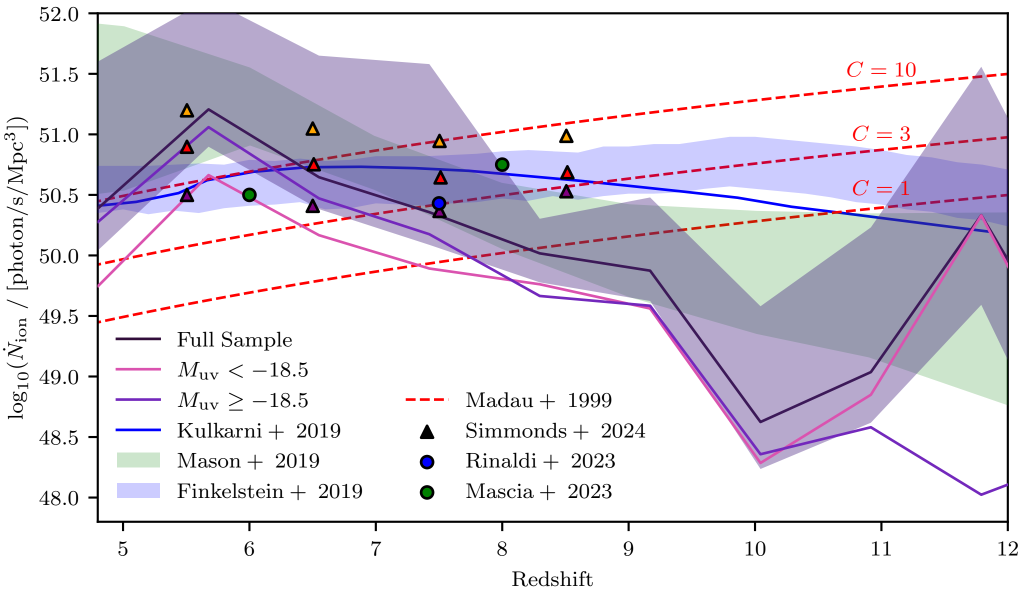

Figure 8. Number density of ionising photons produced and emitted into the IGM within the GOODS-S field as a function of redshift, based on ILI predictions of individual galaxies. We include lines for the entire sample (*black*) as well as for UV-bright galaxies ($M_{\rm UV} < -18.5$; *magenta*) and UV-dim galaxies ($M_{\rm UV} \geqslant -18.5$; *purple*). For the full sample, we also include uncertainties computed by resampling both photometric and model uncertainties (*dark purple*). Comparisons with a Bayesian-inferred history (Mason et al. 2019a), theoretical model (Finkelstein et al. 2019), and observational data (Simmonds et al. 2024b; Rinaldi et al. 2023; Mascia et al. 2023a) are also provided, as well as an analytical estimate of the number of photons required to ionize the neutral IGM for various clumping factors $C \in \{1, 3, 10\}$ (Madau et al. 1999).

Therefore, while we do include AGN hosts as sources (as we do not make any AGN-related selection cuts apart from removing exceptionally bright sources), our model does not account for any changes in the production or escape of ionized photons induced by the presence of an AGN (e.g. Grazian et al. 2018). Therefore, we do not observe, for instance, the late-time bump in ionizing luminosity that AGN cause (Kulkarni et al. 2019b; Dayal et al. 2020; Trebitsch et al. 2021). The second is that at the highest redshifts, our prediction of $\dot{N}_{\rm ion}$ become under-estimated due to the UV magnitude limit imposed by JADES being a flux-limited survey, thus effectively reducing the completeness of the sample at $z \gtrsim 8$ (see also the discussion in Robertson et al. 2023).

Now, we aim to use this sample to produce an explicit reionization history, tracing the evolution of the ionized fraction ($Q_{\rm HII}$) based on only our sample of galaxies. To do this, we make use of the modified "reionization equation" of Madau (2017):

$$\frac{dQ_{\rm HII}}{dt} = \frac{\dot{N}_{\rm ion}}{\langle n_{\rm H}\rangle(1+\langle\kappa^{\rm LLS}_{\nu_L}\rangle/\langle\kappa^{\rm IGM}_{\nu_L}\rangle)} - \frac{Q_{\rm HII}}{\bar{t}_{\rm rec}}, \tag{7}$$

where $\langle n_{\rm H}\rangle = 1.9\times10^{-7}\,{\rm cm}^{-3}$ is the comoving number density of hydrogen in the IGM (Gnedin & Madau 2022) and $\langle\kappa^{\rm LLS}_{\nu_L}\rangle$ and $\langle\kappa^{\rm IGM}_{\nu_L}\rangle$) are absorption coefficients due to high-density clumps known as Lyman-limit systems (Crighton et al. 2019; Becker et al. 2021; Zhu et al. 2023; Georgiev et al. 2024) as well as the IGM itself. This term is proportional to $1-Q_{\rm HII}$ and becomes important as ionized bubbles begin to merge and overlap (at $z \sim 6$), accounting for the presence of optically thick absorbers that ensure that the mean-free path of LyC photons remains small once overlap begins to occur (Gnedin & Fan 2006; Furlanetto & Mesinger 2009; Worseck et al. 2014). The ratio of these two quantities is given as a function of redshift in Equation 32 of Madau (2017) and is taken as 0 for $z > 6$. Finally, $\bar{t}_{\rm rec}$ is an "effective" recombination timescale in the IGM. For our purposes, we use the following fitting formula:

$$\bar{t}_{\rm rec} = 2.3\left(\frac{1+z}{6}\right)^{-4.35}\ {\rm Gyr}, \tag{8}$$

based on analysis of a radiation hydrodynamical simulation by So et al. (2014). We choose this expression because it does not require an estimate of the clumping factor C, although much work has been carried out to estimate redshift-dependent values of C using cosmological hydrodynamic simulations (Kohler et al. 2007; Pawlik et al. 2009; Finlator et al. 2012; Shull et al. 2012; So et al. 2014; Kaurov & Gnedin 2014). We note, however, that there is evidence for a large galaxy over-density in GOODS-S at $z \sim 5.4$ (Helton et al. 2024), which may further stress the effectiveness of this approximation at low redshifts, towards the end of reionization. In fact, as expected, we also find a slight bump in $\dot{N}_{\rm ion}$ (Figure 8) at this redshift.

Another key quantity to compute is the Thompson optical depth to the microwave background, τ. This can be computed as (Kuhlen & Faucher-Giguère 2012; Robertson et al. 2015; Robertson 2022):

$$\tau(z) = c\sigma_T\langle n_{\rm H}\rangle \int_0^z dz' \, \frac{(1+z')^2}{H(z')}\left[1+\frac{\eta Y}{4X}\right]Q_{\rm HII}(z'), \tag{9}$$

where c is the speed of light, σ_T is the Thompson cross section, and we assume that helium is fully ionized ($\eta = 2$) at redshifts $z < 4$ and singly ionized ($\eta = 1$) before this.

Using these expressions, as well as our results from Fig. 8, in Fig. 9 we show the computed evolution histories and associated uncertainties for $Q_{\rm HII}$[12] (*left*) and τ (*right*). In the case of $Q_{\rm HII}$, we compare

[12] Due to the flux limits of the survey, we solve Equation 7 from $Q_{\rm HII} = 0$ at $z = 13$. We also artificially set $Q_{\rm HII} = 1$ once reionization is complete. This is due to the fact that Eq. (7) is only valid until a given patch is nearing

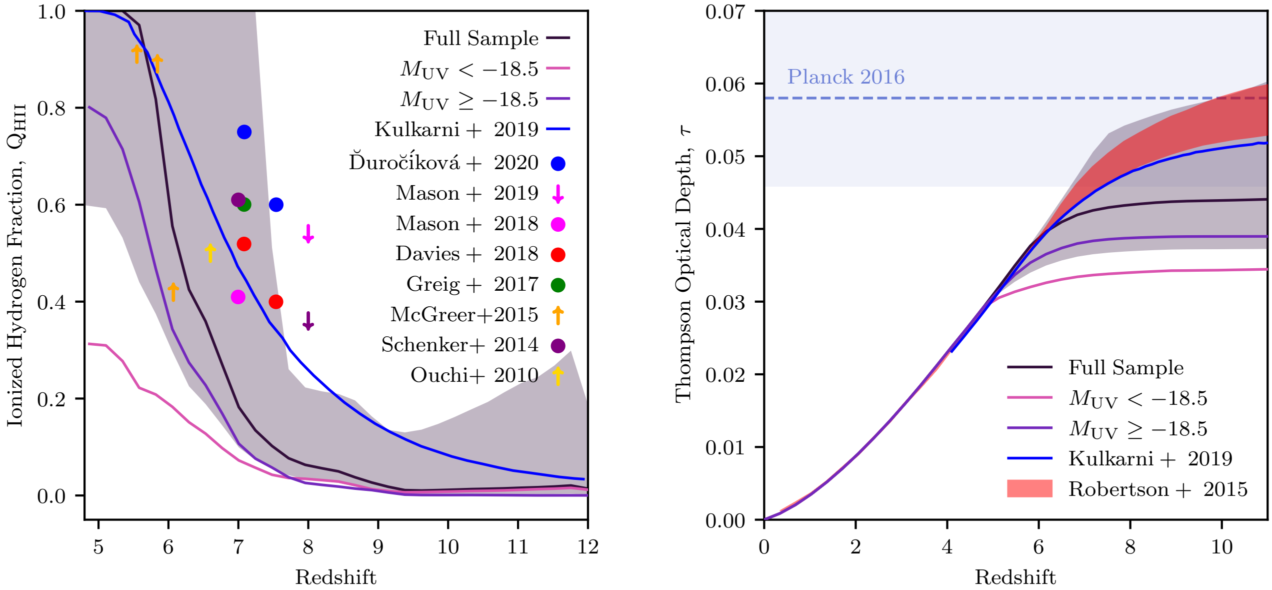

Figure 9. Evolution of the volume-averaged ionized fraction of hydrogen (*left*) as well as the Thompson optical depth (*right*) as a function of redshift. We include curves for our full sample (*black*) as well as for only UV-bright galaxies ($M_{\rm UV} < -18.5$; *magenta*) and only the UV-dim galaxies ($M_{\rm UV} \geqslant -18.5$; *purple*). For the full sample, we also include uncertainties computed by resampling both photometric and model uncertainties (*dark purple*). UV-bright galaxies are particularly important at high redshift ($z \gtrsim 8$), but can only reionize 35% of the volume by themselves, despite accounting for the brightest 20% of the sample. The large number of remaining UV-dim galaxies dominate at lower redshifts, reionizing 85% of the survey volume. Thus, neither group are solely responsible, but together are able to drive reionization to completion by $z \sim 5.3$. For comparison to $Q_{\rm HII}$, we include simulation results from Kulkarni et al. (2019a) as well as a number of observational results (Ouchi et al. 2010; Schenker et al. 2014; McGreer et al. 2015; Greig et al. 2017; Davies et al. 2018; Mason et al. 2018, 2019b; Ďurovčíková et al. 2020). Likewise for τ we compare to results from Robertson et al. (2015); Kulkarni et al. (2019a) as well as constraints from *Planck* (Planck Collaboration et al. 2016).

to results from Kulkarni et al. (2019a) as well as observational constraints (Ouchi et al. 2010; Schenker et al. 2014; McGreer et al. 2015; Greig et al. 2017; Davies et al. 2018; Mason et al. 2018, 2019b; Ďurovčíková et al. 2020). For τ, we compare to results from Kulkarni et al. (2019a); Robertson et al. (2015) as well as constraints from Planck Collaboration et al. (2016).

We find that the galaxies considered within this sample are able to complete reionization within the GOODS-S region by $z \sim 5.3$. Furthermore, the rapid evolution in $Q_{\rm HII}$ for the full sample begins very late, being only $\sim 20\%$ complete at $z = 7$, in agreement with various observational probes favouring a relatively late reionization (Ouchi et al. 2010; Schroeder et al. 2013; Schenker et al. 2014; Oñorbe et al. 2017; Bañados et al. 2018; Villasenor et al. 2022) as well as *Planck* (Planck Collaboration et al. 2016). Finally, as before, we repeat this calculation for the UV-bright and UV-dim galaxies defined in Fig. 6. Here, we find that UV-bright galaxies are only able to reionize $\sim 30\%$ of the volume by themselves, despite accounting for the brightest 20% of the population. On the other hand, the larger number of UV-dim galaxies become completely dominant at $z < 7.5$, managing to ionize $\sim 80\%$ of the volume by themselves. As such, we conclude that neither group of sources is able to reionize the Universe on time solely by themselves, but that the complete set of star-forming galaxies are able to complete reionization without the help of AGN or more exotic sources of ionizing photons (Furlanetto & Oh 2008; Robertson et al. 2015; Liu et al. 2016; Kulkarni et al. 2019b; Dayal et al. 2020; Ma et al. 2021; Trebitsch et al. 2021, 2023).

An important caveat is the fact that in this analysis we are only integrating so far down the UV luminosity function, owing to the flux limited sample of JADES[13] (Eisenstein et al. 2023a) as well as our selection function. In doing so, we are not completely sampling galaxies at fainter magnitudes (particularly at $z \gtrsim 7$), with no meaningful representation at $M_{\rm UV} \geqslant -15$. In turn, these sources may have comparable $\dot{n}_{\rm ion}$ contributions, despite the potential turnover at the faint end of the UV luminosity function (Bouwens et al. 2022; Williams et al. 2024). For example, recent work by Wu & Kravtsov (2024) suggests that for a constant $f_{\rm esc}$, dwarf galaxies with $M_{\rm UV} > -14$ might contribute $\approx$40-60% of the ionizing photon budget at $z > 7$, reducing to $\approx$20% at $z = 6$, highlighting the need to account for these objects. In practise, including these sources will increase $\dot{N}_{\rm ion}$ (particularly at higher redshifts), thus particularly modifying the intermediate reionization history and making it conclude slightly earlier, potentially in line with other observational constraints. As such, it would be interesting to repeat this exercise with other deep surveys (e.g. JADES Origin Field; Eisenstein et al. 2023b, NGDEEP; Bagley et al. 2024, GLASS; Treu et al. 2022), wider surveys (e.g. CEERS; Bagley et al. 2023, PRIMER; Dunlop et al. 2021, COSMOS-WEB; Casey et al. 2023) and particularly those which are targeted at lensing clusters which can push to even fainter UV luminosities (e.g. UNCOVER; Bezanson et al. 2022). We leave such explorations to future work, though note that our model can also be trained on other sets of *JWST* filters and is therefore suitable for these applications.

In the case of the Thompson optical depth, we recover a redshift evolution in agreement with previous results from Robertson et al.

complete reionization. The interested reader is directed to discussions in Robertson et al. (2013); Madau (2017); Gnedin & Madau (2022).

13 For a complete discussion, the reader is directed to Robertson et al. (2023).

(2015); Kulkarni et al. (2019a) up to $z \sim 6$. However, at redshifts beyond this, we similarly find that the reduced number of sources in our sample at higher redshift leads to a value of $\tau = 0.043$, falling below the constraints from *Planck* (Planck Collaboration et al. 2016). In agreement with the evolution of $Q_{\rm HII}$, the majority of optical depth evolution is driven by UV-dim galaxies, confirming their importance.

It is important to note that we do not suggest that the curve shown in Fig. 9 is the definitive history of reionization in GOODS-S, particularly given that we do not have a complete sample by definition (see the selection described in Section 4.1). Instead, the purpose of this work has been to show that galaxy properties such as (but not limited to[14]) $\dot{n}_{\rm ion}$ can be self-consistently derived from photometry. It is, however, particularly interesting that the sample studied here is able to drive reionization to completion on a realistic time-scale, leaving space for ever dimmer galaxies to make their mark. In conclusion, this work further accentuates the fact that while *JWST* has certainly ushered in a new era for the study of reionization, it is necessary to use deep surveys with well-defined selection functions and self-consistent models to build a complete picture of cosmic dawn.

5 CONCLUSIONS

We have implemented an implicit likelihood inference (ILI) model based on the LTU-ILI pipeline (Ho et al. 2024) to predict the angle-averaged escaped ionizing luminosity, $\dot{n}_{\rm ion}$, of Epoch of Reionization galaxies based on observed photometric magnitudes and redshifts. Trained on 13,800 mock dust-attenuated photometric line-of-sight measurements of *JWST* analogues from the SPHINX[20] simulation (Katz et al. 2023a), this model has been validated and shown to perform better than estimates computed using SED-fitting techniques with `BAGPIPES`, including better performance across multiple lines-of-sight for the same object. One of the key novelties of our model compared to previous analyses is that rather than treating the ionizing photon production efficiency and LyC escape fraction as separate quantities, they are inferred together. Hence, our method does not require a separate prescription for the LyC escape fraction, $f_{\rm esc}$, or for any dust-correction, as these are handled self-consistently by the model.

This ILI model was then deployed on a sample of 4,559 photometrically observed galaxies in the GOODS-S field as part of the JADES programme (Eisenstein et al. 2023a), allowing us to explore the redshift evolution of $\dot{n}_{\rm ion}$, the number density of ionizing photons released into the intergalactic medium (IGM), $\dot{N}_{\rm ion}$, as well as the volume-averaged ionized fraction of hydrogen, $Q_{\rm HII}$.

Our conclusions are as follows:

- Our ILI method for the inference of $\dot{n}_{\rm ion}$ from photometry is more accurate (and produces self-consistent uncertainties) than traditional SED-fitting methods. Additionally, it is orders of magnitude faster, allowing for easy application to large datasets.
- In a comparative test against an SED-fitting method on a sample of SPHINX[20] galaxies, `BAGPIPES` was found to perform worse, often over-estimating $\dot{n}_{\rm ion}$ for galaxies which release the most ionising photons. Furthermore, estimates from `BAGPIPES` required the use of a model for the LyC escape fraction, which further inflated this over-estimate.
- For our sample of photometric galaxies, $\dot{n}_{\rm ion}$ evolves slowly with redshift, as: $\log_{10}(\dot{n}_{\rm ion}) = (0.08 \pm 0.01)z + (51.60 \pm 0.06)$.
- Star-forming galaxies observed within this sample are capable of producing a reionization history that begins late and completes at $z \sim 5.3$.
- UV-dim galaxies (with $M_{\rm UV} \geqslant -18.5$, accounting for 80% of the sample) are able to reionize $\sim 80\%$ of the survey volume, while UV-bright galaxies (with $M_{\rm UV} < -18.5$, 20% of the sample) reionize $\sim 30\%$ of the volume. Thus, neither subgroup is capable of driving reionization by themselves but faint galaxies appear to be crucial.

We have utilised the synergy of photometric *JWST* observations and cosmological radiation hydrodynamic simulations with a resolved multi-phase interstellar medium to build an inference pipeline for the luminosity of ionizing photons released into the IGM by galaxies during the Epoch of Reionization. Beyond providing valuable insight into the contributions of star-forming galaxies to the evolution of reionization, this work further highlights the necessity for observers and simulators to work together as we continue to explore the cosmic dawn.

ACKNOWLEDGEMENTS

N.C. thanks Brant Robertson for insightful discussions. The authors thank Jonathan Patterson for smoothly running the Glamdring Cluster hosted by the University of Oxford, where part of the data processing was performed.

N.C. acknowledges support from the Science and Technology Facilities Council (STFC) for a Ph.D. studentship, as well as the hospitality of Princeton's Department of Astrophysical Sciences where part of this work was completed. RS acknowledges financial support from STFC Grant No. ST/X508664/1. AS acknowledges funding from the "FirstGalaxies" Advanced Grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 789056)

This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1 and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure.This work was performed using the DiRAC Data Intensive service at Leicester, operated by the University of Leicester IT Services, which forms part of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K000373/1 and ST/R002363/1 and STFC DiRAC Operations grant ST/R001014/1. DiRAC is part of the National e-Infrastructure.

AUTHOR CONTRIBUTIONS

The main roles of the authors were, using the CRediT (Contribution Roles Taxonomy) system[15]:

Nicholas Choustikov: Conceptualization; Formal analysis; Methodology; Software; Visualisation; Writing - original draft. **Richard Stiskalek**: Conceptualization; Methodology; Software; Writing - original draft. **Aayush Saxena**: Conceptualization; Writing

[14] In principal such a method (using ILI applied to photometry) can be leveraged to predict any galaxy property included in the SPHINX[20] public data release.

[15] `https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/credit.html`

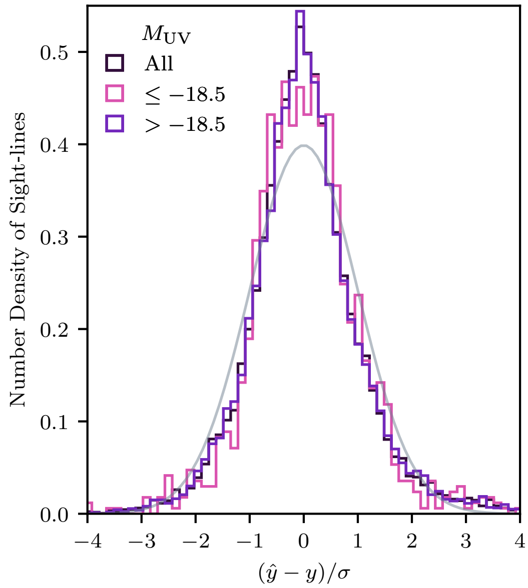

Figure A1. Standardised residuals (as defined in Eq. A1 for ILI predictions, for the full sample of (*black*), UV-bright ($M_{\rm UV} \leqslant -18.5$; *magenta*), and UV-dim ($M_{\rm UV} > -18.5$; *purple*) sight-lines. We include the means (**bold**) and standard deviations (*dashed*), as well as a standard Gaussian ($\mathcal{G}(0,1)$; *gray*) for comparison

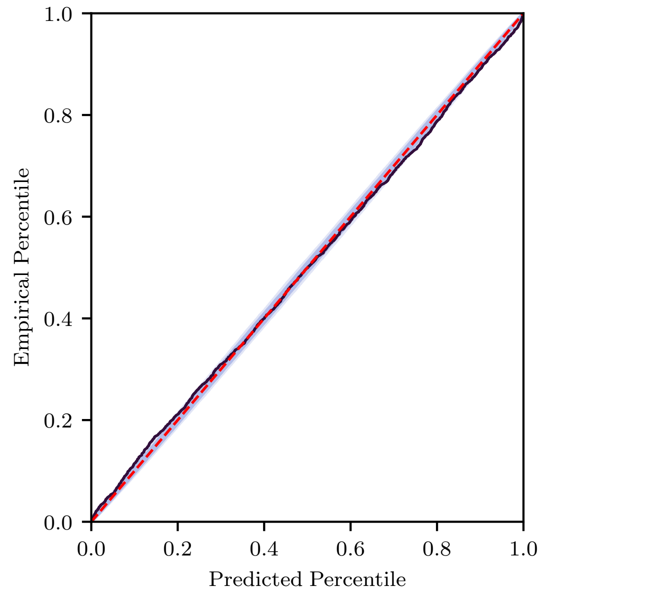

Figure A2. Probability integral transform diagnostic for the ILI model, quantifying the proportion of posterior samples that are below the true value.

- review and editing. **Harley Katz**: Conceptualization; Writing - review and editing. **Julien Devriendt**: Supervision; Resources; Writing - review and editing. **Adrianne Slyz**: Supervision; Resources; Writing - review and editing.

DATA AVAILABILITY

The SPHINX[20] data used in this work is available as part of the SPHINX Public Data Release v1 (SPDRv1, Katz et al. 2023a), available at `https://github.com/HarleyKatz/SPHINX-20-data`. The JADES photometric catalogue used in this work is available at `https://archive.stsci.edu/hlsp/jades`. The code and model developed for this work will be made available upon acceptance. All other data will be shared upon reasonable request to the corresponding author.

APPENDIX A: VALIDATING THE MODEL ON SPHINX[20]

Here, we proceed to complete a variety of benchmark tests on the model described in Section 2.

First, in order to further confirm that the uncertainties produced by the model are self-consistent, in Fig. A1 we show histograms of the standardised residuals given by:

$$x \equiv \frac{\dot{n}_{\rm ion}^{\rm predicted} - \dot{n}_{\rm ion}^{\rm true}}{\langle {\rm unc}(\dot{n}_{\rm ion}^{\rm predicted}) \rangle}, \qquad (A1)$$

where $\langle {\rm unc}(\dot{n}_{\rm ion}^{\rm predicted}) \rangle$ is the average of the asymmetric 1σ uncertainties of the ILI posterior. We include histograms for the full sample (*black*) as well as for the observed UV-bright ($M_{\rm UV} \leqslant -18.5$; *magenta*) and UV-dim ($M_{\rm UV} > -18.5$; *purple*) sight-lines of SPHINX[20] galaxies. For completeness, we also show the means (**bold**) and standard deviations (*dashed*), as well as the standard Gaussian distribution, $\mathcal{G}(0,1)$, as a comparison. We find that in all three cases our ILI model performs very well, without significant outliers. Interestingly, the model performs better in the case of UV-bright galaxies.

To further reinforce this point, we also inspect the Probability Integral Transform (PIT, Cook et al. 2006) diagnostic, shown in Fig. A2. The PIT quantifies the proportion of posterior samples $\boldsymbol{\theta}$ that are below the true value. If the distribution of PIT values is uniformly distributed, then the predicted posterior distributions are consistent with the true values (Zhao et al. 2021). The PIT distribution is typically assessed on a percentile-percentile plot, which compares the cumulative density function of PIT values to that of a uniform random variable. If the learnt posterior is well-calibrated, then the two cumulative density functions should agree. If not, the PIT plot is a useful probe of a global bias or over- and under-dispersion. We verify that the test-set PIT distribution of our ILI model predicting $\log \dot{n}_{\rm ion}$ passes this test.

Lastly, we also verify that the predictions of our ILI model agree with those of an Extra-Trees regressor (ET, Geurts et al. 2006) as implemented in `scikit-learn` (Pedregosa et al. 2011). While this is not a validation of the ILI pipeline, it is nevertheless a useful sanity check. We similarly optimize the hyperparameters of the ET model and find that the maximum posterior ILI and ET predictions are correlated with a Spearman correlation coefficient of 0.98 and that with respect to the true values the ILI model marginally outperforms the ET, while also providing self-consistent uncertainties.

REFERENCES

Akiba T., Sano S., Yanase T., Ohta T., Koyama M., 2019, arXiv e-prints, p. arXiv:1907.10902

Alsing J., Wandelt B., Feeney S., 2018, MNRAS, 477, 2874

Anderson L., Governato F., Karcher M., Quinn T., Wadsley J., 2017, MNRAS, 468, 4077

Bañados E., et al., 2018, Nature, 553, 473

Bagley M. B., et al., 2023, ApJ, 946, L12
Bagley M. B., et al., 2024, ApJ, 965, L6
Bartlett D. J., Desmond H., 2023, arXiv e-prints, p. arXiv:2309.00948
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Becker G. D., D'Aloisio A., Christenson H. M., Zhu Y., Worseck G., Bolton J. S., 2021, MNRAS, 508, 1853
Bezanson R., et al., 2022, arXiv e-prints, p. arXiv:2212.04026
Bishop C., 1994, Workingpaper, Mixture density networks. Aston University
Bosman S. E. I., et al., 2022, MNRAS, 514, 55
Bouwens R. J., et al., 2021, AJ, 162, 47
Bouwens R. J., Illingworth G., Ellis R. S., Oesch P., Stefanon M., 2022, ApJ, 940, 55
Bowler R. A. A., Jarvis M. J., Dunlop J. S., McLure R. J., McLeod D. J., Adams N. J., Milvang-Jensen B., McCracken H. J., 2020, MNRAS, 493, 2059
Boyett K. N. K., Stark D. P., Bunker A. J., Tang M., Maseda M. V., 2022, MNRAS, 513, 4451
Bradley L., et al., 2016, Photutils: Photometry tools, Astrophysics Source Code Library, record ascl:1609.011
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brinchmann J., 2023, MNRAS, 525, 2087
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cameron A. J., Katz H., Witten C., Saxena A., Laporte N., Bunker A. J., 2023, arXiv e-prints, p. arXiv:2311.02051
Carnall A. C., McLure R. J., Dunlop J. S., Davé R., 2018, MNRAS, 480, 4379
Casey C. M., et al., 2023, ApJ, 954, 31
Chevallard J., et al., 2018, MNRAS, 479, 3264
Chisholm J., et al., 2018, A&A, 616, A30
Chisholm J., Prochaska J. X., Schaerer D., Gazagnes S., Henry A., 2020, MNRAS, 498, 2554
Chisholm J., et al., 2022, MNRAS, 517, 5104
Choustikov N., et al., 2024a, arXiv e-prints, p. arXiv:2401.09557
Choustikov N., et al., 2024b, MNRAS, 529, 3751
Conselice C. J., Mortlock A., Bluck A. F. L., Grützbauch R., Duncan K., 2013, MNRAS, 430, 1051
Cook S. R., Gelman A., Rubin D. B., 2006, Journal of Computational and Graphical Statistics, 15, 675
Cranmer K., Brehmer J., Louppe G., 2020, Proceedings of the National Academy of Science, 117, 30055
Crighton N. H. M., Prochaska J. X., Murphy M. T., O'Meara J. M., Worseck G., Smith B. D., 2019, MNRAS, 482, 1456
Davies F. B., et al., 2018, ApJ, 864, 142
Dayal P., et al., 2020, MNRAS, 495, 3065
Dijkstra M., 2017, Saas-Fee Lecture Notes: Physics of Lyman Alpha Radiative Transfer, doi:10.48550/ARXIV.1704.03416, `https://arxiv.org/abs/1704.03416`
Donnan C. T., et al., 2023, MNRAS, 518, 6011
Donnan C. T., et al., 2024, arXiv e-prints, p. arXiv:2403.03171
Dunlop J. S., et al., 2021, PRIMER: Public Release IMaging for Extragalactic Research, JWST Proposal. Cycle 1, ID. #1837
Eisenstein D. J., et al., 2023a, arXiv e-prints, p. arXiv:2306.02465
Eisenstein D. J., et al., 2023b, arXiv e-prints, p. arXiv:2310.12340
Eldridge J. J., Stanway E. R., Xiao L., McClelland L. A. S., Taylor G., Ng M., Greis S. M. L., Bray J. C., 2017, Publ. Astron. Soc. Australia, 34, e058
Fan X., Carilli C. L., Keating B., 2006, ARA&A, 44, 415
Fensch J., et al., 2017, MNRAS, 465, 1934
Finkelstein S. L., et al., 2019, ApJ, 879, 36
Finlator K., Oh S. P., Özel F., Davé R., 2012, MNRAS, 427, 2464
Fletcher T. J., Tang M., Robertson B. E., Nakajima K., Ellis R. S., Stark D. P., Inoue A., 2019, ApJ, 878, 87
Flury S. R., et al., 2022a, ApJS, 260, 1
Flury S. R., et al., 2022b, ApJ, 930, 126
Furlanetto S. R., Mesinger A., 2009, MNRAS, 394, 1667
Furlanetto S. R., Oh S. P., 2008, ApJ, 681, 1
Georgiev I., Mellema G., Giri S. K., 2024, arXiv e-prints, p. arXiv:2405.04273
Geurts P., Ernst D., Wehenkel L., 2006, Machine Learning, 63, 3
Giallongo E., et al., 2015, A&A, 578, A83
Giallongo E., et al., 2019, ApJ, 884, 19
Giovinazzo E., Trebitsch M., Mauerhofer V., Dayal P., Oesch P. A., 2024, arXiv e-prints, p. arXiv:2402.17635
Gnedin N. Y., Fan X., 2006, ApJ, 648, 1
Gnedin N. Y., Madau P., 2022, Living Reviews in Computational Astrophysics, 8, 3
Gordon K. D., Clayton G. C., Misselt K. A., Landolt A. U., Wolff M. J., 2003, ApJ, 594, 279
Grazian A., et al., 2018, A&A, 613, A44
Greenberg D. S., Nonnenmacher M., Macke J. H., 2019, arXiv e-prints, p. arXiv:1905.07488
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017, MNRAS, 466, 4239
Hainline K. N., et al., 2024, ApJ, 964, 71
Harikane Y., et al., 2022, ApJS, 259, 20
Hassan S., Davé R., Mitra S., Finlator K., Ciardi B., Santos M. G., 2018, MNRAS, 473, 227
Helton J. M., et al., 2024, ApJ, 962, 124
Henry A., Scarlata C., Martin C. L., Erb D., 2015, ApJ, 809, 19
Hermans J., Begy V., Louppe G., 2019, arXiv e-prints, p. arXiv:1903.04057
Ho M., et al., 2024, arXiv e-prints, p. arXiv:2402.05137
Iliev I. T., Mellema G., Pen U. L., Merz H., Shapiro P. R., Alvarez M. A., 2006, MNRAS, 369, 1625
Inoue A. K., Shimizu I., Iwata I., Tanaka M., 2014, MNRAS, 442, 1805
Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018a, MNRAS, 474, 4514
Izotov Y. I., Worseck G., Schaerer D., Guseva N. G., Thuan T. X., Fricke Verhamme A., Orlitová I., 2018b, MNRAS, 478, 4851
Jaskot A. E., Oey M. S., 2014, ApJ, 791, L19
Johnson B. D., Leja J., Conroy C., Speagle J. S., 2021, ApJS, 254, 22
Kakiichi K., Gronke M., 2021, ApJ, 908, 30
Katz H., 2022, MNRAS, 512, 348
Katz H., et al., 2022a, arXiv e-prints, p. arXiv:2211.04626
Katz H., et al., 2022b, MNRAS, 515, 4265
Katz H., et al., 2023a, The Open Journal of Astrophysics, 6, 44
Katz H., et al., 2023b, MNRAS, 518, 592
Kaurov A. A., Gnedin N. Y., 2014, ApJ, 787, 146
Kaviraj S., Devriendt J., Dubois Y., Slyz A., Welker C., Pichon C., Peirani S., Le Borgne D., 2015, MNRAS, 452, 2845
Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511
Kimm T., Cen R., 2014, ApJ, 788, 121
Kimm T., Blaizot J., Garel T., Michel-Dansac L., Katz H., Rosdahl J., Verhamme A., Haehnelt M., 2019, MNRAS, 486, 2215
Kimm T., Bieri R., Geen S., Rosdahl J., Blaizot J., Michel-Dansac L., Garel T., 2022, ApJS, 259, 21
Kohler K., Gnedin N. Y., Hamilton A. J. S., 2007, ApJ, 657, 15
Kriss G. A., et al., 2001, Science, 293, 1112
Kroupa P., 2001, MNRAS, 322, 231
Kuhlen M., Faucher-Giguère C.-A., 2012, MNRAS, 423, 862
Kulkarni G., Keating L. C., Haehnelt M. G., Bosman S. E. I., Puchwein E., Chardin J., Aubert D., 2019a, MNRAS, 485, L24
Kulkarni G., Worseck G., Hennawi J. F., 2019b, MNRAS, 488, 1035
Law K.-H., Gordon K. D., Misselt K. A., 2018, ApJS, 236, 32
Leclercq F., et al., 2024, arXiv e-prints, p. arXiv:2401.14981
Leitherer C., et al., 1999, ApJS, 123, 3
Leitherer C., Hernandez S., Lee J. C., Oey M. S., 2016, ApJ, 823, 64
Li J., et al., 2024, arXiv e-prints, p. arXiv:2403.00074
Liu H., Slatyer T. R., Zavala J., 2016, Phys. Rev. D, 94, 063507
Livermore R. C., Finkelstein S. L., Lotz J. M., 2017, ApJ, 835, 113
Lower S., Narayanan D., Leja J., Johnson B. D., Conroy C., Davé R., 2020, ApJ, 904, 33
Luo Y., et al., 2021, ApJ, 908, 183
Ma X., Quataert E., Wetzel A., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2020, MNRAS, 498, 2001
Ma X., Quataert E., Wetzel A., Faucher-Giguère C.-A., Boylan-Kolchin M., 2021, MNRAS, 504, 4062
Madau P., 2017, ApJ, 851, 50
Madau P., Haardt F., Rees M. J., 1999, ApJ, 514, 648

Marin J.-M., Pudlo P., Robert C. P., Ryder R., 2011, arXiv e-prints, p. arXiv:1101.0955

Mascia S., et al., 2023a, arXiv e-prints, p. arXiv:2309.02219

Mascia S., et al., 2023b, A&A, 672, A155

Maseda M. V., et al., 2020, MNRAS, 493, 5120

Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2

Mason C. A., et al., 2019a, MNRAS, 485, 3947

Mason C. A., et al., 2019b, MNRAS, 485, 3947

McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499

Michel-Dansac L., Blaizot J., Garel T., Verhamme A., Kimm T., Trebitsch M., 2020, Astronomy & Astrophysics, 635, A154

Muñoz J. B., Mirocha J., Chisholm J., Furlanetto S. R., Mason C., 2024, arXiv e-prints, p. arXiv:2404.07250

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, MNRAS, 510, 4582

Nakajima K., Ouchi M., 2014, MNRAS, 442, 900

Nakajima K., Ouchi M., Shimasaku K., Hashimoto T., Ono Y., Lee J. C., 2013, ApJ, 769, 3

Narayanan D., et al., 2024, ApJ, 961, 73

Oñorbe J., Hennawi J. F., Lukić Z., Walther M., 2017, ApJ, 847, 63

Ouchi M., et al., 2010, ApJ, 723, 869

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

Papamakarios G., Murray I., 2016, arXiv e-prints, p. arXiv:1605.06376

Papamakarios G., Sterratt D. C., Murray I., 2018, arXiv e-prints, p. arXiv:1805.07226

Papamakarios G., Nalisnick E., Jimenez Rezende D., Mohamed S., Lakshminarayanan B., 2019, arXiv e-prints, p. arXiv:1912.02762

Pawlik A. H., Schaye J., van Scherpenzeel E., 2009, MNRAS, 394, 1812

Pearson W. J., et al., 2019, A&A, 631, A51

Pedregosa F., et al., 2011, Journal of Machine Learning Research, 12, 2825

Planck Collaboration et al., 2014, A&A, 571, A1

Planck Collaboration et al., 2016, A&A, 596, A108

Puchwein E., et al., 2023, MNRAS, 519, 6162

Putman M. E., 2017, in Fox A., Davé R., eds, Astrophysics and Space Science Library Vol. 430, Gas Accretion onto Galaxies. p. 1 (`arXiv:1612.00461`), doi:10.1007/978-3-319-52512-9_1

Ramambason L., et al., 2022, A&A, 667, A35

Rieke M. J., et al., 2023, ApJS, 269, 16

Rinaldi P., et al., 2023, arXiv e-prints, p. arXiv:2309.15671

Robertson B. E., 2022, ARA&A, 60, 121

Robertson B. E., et al., 2013, ApJ, 768, 71

Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, ApJ, 802, L19

Robertson B., et al., 2023, arXiv e-prints, p. arXiv:2312.10033

Rosdahl J., et al., 2018, MNRAS, 479, 994

Rosdahl J., et al., 2022, MNRAS, 515, 2386

Saldana-Lopez A., et al., 2022, A&A, 663, A59

Saxena A., et al., 2022a, MNRAS, 511, 120

Saxena A., et al., 2022b, MNRAS, 517, 1098

Saxena A., et al., 2024, A&A, 684, A84

Schaerer D., Izotov Y. I., Verhamme A., Orlitová I., Thuan T. X., Worseck G., Guseva N. G., 2016, A&A, 591, L8

Schaerer D., Marques-Chaves R., Barrufet L., Oesch P., Izotov Y. I., Naidu R., Guseva N. G., Brammer G., 2022, A&A, 665, L4

Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, ApJ, 795, 20

Schroeder J., Mesinger A., Haiman Z., 2013, MNRAS, 428, 3058

Shapiro P. R., Giroux M. L., 1987, ApJ, 321, L107

Shull J. M., Tumlinson J., Giroux M. L., Kriss G. A., Reimers D., 2004, ApJ, 600, 570

Shull J. M., France K., Danforth C. W., Smith B., Tumlinson J., 2010, ApJ, 722, 1312

Shull J. M., Harness A., Trenti M., Smith B. D., 2012, ApJ, 747, 100

Simmonds C., et al., 2023, MNRAS, 523, 5468

Simmonds C., Verhamme A., Inoue A. K., Katz H., Garel T., De Barros S., 2024a, arXiv e-prints, p. arXiv:2402.04052

Simmonds C., et al., 2024b, MNRAS, 527, 6139

So G. C., Norman M. L., Reynolds D. R., Wise J. H., 2014, ApJ, 789, 149

Stanway E. R., Eldridge J. J., 2018, Monthly Notices of the Royal Astronomical Society, 479, 75

Stanway E. R., Eldridge J. J., 2023, MNRAS, 522, 4430

Stanway E. R., Eldridge J. J., Becker G. D., 2016, MNRAS, 456, 485

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

Tang M., Stark D. P., Chevallard J., Charlot S., 2019, MNRAS, 489, 2572

Trebitsch M., Blaizot J., Rosdahl J., Devriendt J., Slyz A., 2017, MNRAS, 470, 224

Trebitsch M., et al., 2021, A&A, 653, A154

Trebitsch M., Hutter A., Dayal P., Gottlöber S., Legrand L., Yepes G., 2023, MNRAS, 518, 3576

Treu T., et al., 2022, ApJ, 935, 110

Varadaraj R. G., Bowler R. A. A., Jarvis M. J., Adams N. J., Häußler B., 2023, MNRAS, 524, 4586

Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, A&A, 578, A7

Verhamme A., Orlitová I., Schaerer D., Izotov Y., Worseck G., Thuan T. X., Guseva N., 2017, A&A, 597, A13

Villasenor B., Robertson B., Madau P., Schneider E., 2022, ApJ, 933, 59

Wan J. T., Tacchella S., Johnson B. D., Iyer K. G., Speagle J. S., Maiolino R., 2024, arXiv e-prints, p. arXiv:2404.14494

Wang B., et al., 2021, ApJ, 916, 3

Williams C. E., et al., 2024, ApJ, 960, L16

Worseck G., et al., 2014, MNRAS, 445, 1745

Worseck G., Prochaska J. X., Hennawi J. F., McQuinn M., 2016, ApJ, 825, 144

Wu Z., Kravtsov A., 2024, On the contribution of dwarf galaxies to reionization of the Universe (`arXiv:2405.08066`)

Xu H., Wise J. H., Norman M. L., Ahn K., O'Shea B. W., 2016, ApJ, 833, 84

Yuan Y., Martin-Alvarez S., Haehnelt M. G., Garel T., Sijacki D., 2024, arXiv e-prints, p. arXiv:2401.02572

Yusef-Zadeh F., Morris M., White R. L., 1984, ApJ, 278, 186

Zhang K., et al., 2017, MNRAS, 466, 3217

Zhao D., Dalmasso N., Izbicki R., Lee A. B., 2021, in de Campos C., Maathuis M. H., eds, Proceedings of Machine Learning Research Vol. 161, Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence. PMLR, pp 1830–1840, `https://proceedings.mlr.press/v161/zhao21b.html`

Zheng Z., Miralda-Escudé J., 2002, The Astrophysical Journal, 578, 33

Zheng W., et al., 2004, ApJ, 605, 631

Zhu Y., et al., 2023, ApJ, 955, 115

Ďurovčíková D., Katz H., Bosman S. E. I., Davies F. B., Devriendt J., Slyz A., 2020, MNRAS, 493, 4256

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